

Article

Uncertainty Propagation for Vibrometry-Based Acoustic Predictions Using Gaussian Process Regression

Andreas Wurzinger [†]  and Stefan Schoder ^{*,†} 

Aeroacoustic and Vibroacoustic Group, Institute of Fundamentals and Theory in Electrical Engineering, Graz University of Technology, Inffeldgasse 16c, 8010 Graz, Austria; andreas.wurzinger@tugraz.at

* Correspondence: stefan.schoder@tugraz.at

[†] These authors contributed equally to this work.

Abstract

Shell-like housing structures for motors and compressors can be found in everyday products. Consumers significantly evaluate acoustic emissions during the first usage of products. Unpleasant sounds may raise concerns and cause complaints to be issued. A prevention strategy is a holistic acoustic design, which includes predicting the emitted sound power as part of end-of-line testing. The hybrid experimental-simulative sound power prediction based on laser scanning vibrometry (LSV) is ideal in acoustically harsh production environments. However, conducting vibroacoustic testing with laser scanning vibrometry is time-consuming, making it difficult to fit into the production cycle time. This contribution discusses how the time-consuming sampling process can be accelerated to estimate the radiated sound power, utilizing adaptive sampling. The goal is to predict the acoustic signature and its uncertainty from surface velocity data in seconds. Fulfilling this goal will enable integration into a product assembly unit and final acoustic quality control without the need for an acoustic chamber. The Gaussian process regression based on PyTorch 2.6.0 performed 60 times faster than the preliminary reference implementation, resulting in a regression estimation time of approximately one second for each frequency bin. In combination with the Equivalent Radiated Power prediction of the sound power, a statistical measure is available, indicating how the uncertainty of a limited number of surface velocity measurement points leads to predictions of the uncertainty inside the acoustical signal. An adaptive sampling algorithm reduces the prediction uncertainty in real-time during measurement. The method enables on-the-fly error analysis in production, assessing the risk of violating agreed-upon acoustic sound power thresholds, and thus provides valuable feedback to the product design units.

Keywords: Gaussian process regression; Vibro-Acoustics; production; measurements; laser scanning vibrometry; risk mitigation



Academic Editors: Kai Yang and John S. Allen

Received: 28 August 2025

Revised: 19 September 2025

Accepted: 29 September 2025

Published: 1 October 2025

Citation: Wurzinger, A.; Schoder, S. Uncertainty Propagation for Vibrometry-Based Acoustic Predictions Using Gaussian Process Regression. *Appl. Sci.* **2025**, *15*, 10652. <https://doi.org/10.3390/app151910652>

Copyright: © 2025 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

Quantifying the sound generation of highly integrated products is essential during design and production. The overall goal is to enhance product quality and meet contractual and legal requirements. For any device with direct user interaction, noise emissions can negatively affect both the user and the user's environment [1]. Therefore, tools [2] to predict and quantify sound emissions should be integrated early in the design process and aligned with production-ready methods for quality control and production-based design feedback [3]. A production-ready system for acoustic end-of-line testing to meet

vibroacoustic requirements is presented in [4]. Ref. [4] requires an acoustically treated environment for accurate sound measures, whereas the used acoustic measurement method [5] does not rely on a specific treatment of the acoustics in the production facility.

We identify the following specifications of the system [4]: firstly (specification 1), the system should be ready to perform measurements at different locations, contactless, and without any additional assembly cost, in a harsh acoustic environment such as a production facility. Secondly (specification 2), the system should select observation points based on the data prediction accuracy and uncertainty calculated by a self-learned data-driven method. Thirdly (specification 3), the technique should provide data that gives the maximum valuable insight into specific characteristics where the deviations occurred over time. Fourthly and finally (specification 4), the measurement time should align with the production cycle of a production line and the performance test of the end test before delivery.

The objective of the work is to establish an acoustic measurement technique for acoustic end-of-line testing that fulfills system specification 1 and 2, and establish the necessary pre-conditions for system specification 3 and 4. The development of experimental sound prediction methods with far-field prediction based on the boundary element method (BEM) was found to provide accurate results even in non-favorable acoustic environments [5] (realizing specification 1 through being contactless). This method is well-suited to fulfill specification two by utilizing data-driven mapping techniques, such as Gaussian process regression (GPR) [6], to minimize the overall uncertainty of the sound prediction method [7]. The benefit of using GPR is a direct estimate of the uncertainty quantification and risk estimation of exceeding limits is possible [8]. For instance, the next best training points are collected based on where the data is most likely to be the most informative. This Bayesian optimization strategy [9,10] is beneficial in situations where data acquisition is costly or slow, and can improve convergence. It is an iterative search (adaptive) of the next best sampling point based on previous samples. The sampling is updated to focus on areas of the input space that are expected to yield the most significant improvement of the model [11].

In the data acquisition process of surface motion data, GPR was used in various applications of robotics and biomechanics [12–14], including real-time [15,16]. In general, the main advantage of GPR over other existing interpolation methods is that it provides a maximum of information of the statistical description from the given data [11,17]. This maximum information gain is well-suited to the stated specifications in the introduction and can further benefit from the sequential design of experiments. However, specification 4 represents a significant obstacle in industrial integration for acoustic testing. Therefore, the contribution of this publication is that it presents the first application of GPR-based uncertainty predictions in vibroacoustics based on descriptive statistics and it discusses the potential speed-up of the Monte Carlo simulation method used in [7]. Firstly, the sampling will be discussed, starting with a minimum number of sampling points and then iteratively refining the number of sampling points. Secondly, the GPR routine of GPyTorch [18] will significantly improve the training and prediction speed of the GPR while continuously updating the model. Thirdly, the far-field prediction will be replaced by the low-fidelity and rapid Equivalent Radiated Power (ERP) method, which is known to be as accurate as the boundary element method (BEM) for high frequencies. Based on the uncertainty estimates from the GPR of the surface velocity, uncertainty in the resulting acoustic signal can be described by the ERP method presented in this work. The usage of BEM would require either the eigenvalues of the correlation matrix of the data points or a Monte Carlo Simulation [19].

In [20], the GPR was performed using scikit-learn 1.6.1 [21], considering 30 restarts to guarantee a robust minimum of the marginal log likelihood cost function. Therewith, at each frequency step (for 32 frequencies), an independent GPR was fitted separately to both the real and imaginary parts of the surface normal velocity frequency response. For each frequency step, 2 GPRs were trained, and restarted 30 times, for 128 sampling points. Subsequently, the two converged GPRs were evaluated using a Monte Carlo method 100 times to compute the mean of the radiated sound power at each vibration frequency. The computation of the radiated sound power was performed using three-dimensional conventional collocational BEM using the Burton–Miller method. The total sound power (BEM) estimation time of this computation was composed of 188.6 s for the GPR training (in total 2×32 , 64 GPR were executed), 1072 s for drawing the Monte Carlo samples (in total $2 \times 32 \times 100$, 6400 samples were drawn) and 387.9 s to evaluate the BEM (in total 32 BEM systems were solved and 100×32 BEM result calculations were performed) on the reference machine ($2 \times$ Intel Xeon G 5218-16 cores, Intel, Santa Clara, CA, USA). All three steps (GPR training, sample drawing, and the BEM simulations) consume a significant amount of time and consume a large fraction of the process time. In this article, statistical measures for the sound power uncertainty are derived, the speed-up potential of the workflow is discussed, and an adaptive refinement scheme to reduce the uncertainty in the predicted sound power is presented. To show the accuracy of the uncertainty propagation, the statistical sampling is compared to a Monte Carlo (BEM)-simulated reference acoustic prediction.

In summary, integrating GPR into the experimental sound field estimation provides a new perspective on assessing surface vibration measurements for acoustical quantification. In Section 2.1, the GPR details are discussed. Section 3 describes the experimental sound field estimation using laser scanning Doppler vibrometry (LSV). It is followed by the results Section 5, which reports the accuracy, performance, and adaptive refinement procedure of the proposed workflow. Concluding remarks are given in Section 6.

2. Methodology

Figure 1 depicts the proposed workflow schematically. It is based on the workflow presented in [5], where the boundary element method (BEM) was used to predict the radiated sound field based on the structural dynamic. At first, the surface velocity will be obtained from a region of interest (ROI) by data acquisition at a limited number of probing locations, by means of e.g., Laser Scanning Vibrometry (LSV) and Finite Element (FE) simulation. The data is then mapped by a standard interpolation technique [5] or Gaussian process regression (GPR) [20] onto the geometry and the evaluation points for the sound power prediction. Finally, the sound propagation is simulated. The radiated sound field can be predicted in various ways, such as using finite element method, BEM, or comparably fast low-fidelity methods. In this work, GPR is used on a dataset at given sample points to obtain a statistical description of the normal velocity field on the whole structure. Therewith, a statistical description of the sound emission based on the Equivalent Radiated Power (ERP) is computed based on either the Monte Carlo method or direct evaluation as proposed in Section 2.4. The statistical description of the ERP can then be used to assess the quality of the prediction and, finally, additional sample points can be defined in an adaptive refinement step based on an expected improvement function as described in Section 2.5.

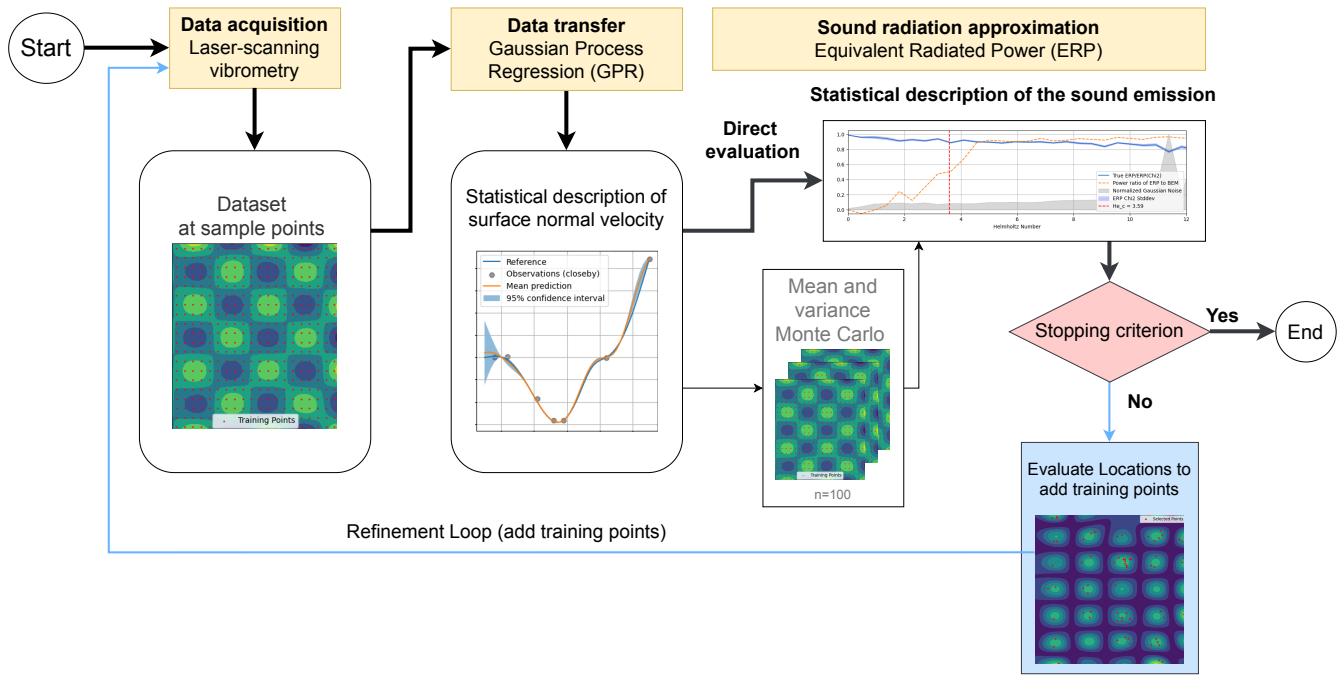


Figure 1. Schematic of the presented workflow predicting the radiated sound field based on surface velocity measurements, including an adaptive refinement procedure.

2.1. Interpolation Based on Gaussian Process Regression

In this article, the Gaussian process regression (GPR) is selected as an interpolation technique to provide both a mean and an uncertainty quantification [22] of the underlying physical process. If it cannot explain the data, the GPR has a noise level that increases, which may be used as an indication of such an event. It is known to reconstruct statistical distributions from noisy observations without overfitting [23], making it an ideal method for this experimental sound power estimate. The output is modeled at different input points with a mean and a covariance function. It has cubic complexity $O(n^3)$ in terms of the input data points n , which limits its application to large datasets. As a consequence, a variational Gaussian process is chosen to reduce the computational complexity [24,25] based on approximate inference [26]. The radial basis function (RBF) kernels are a valid choice for a spatially decreasing correlation and may improve the computational speed due to the sparsity of the correlation function. In [20], the complex-valued normal surface velocity $v_s \cdot n := v_n(x)$ was expressed as Gaussian processes (GP) $v_n \sim \mathcal{GP}\{\mu, k\}$ for both real and imaginary part. Sampling this distribution of functions, this results in the Gaussian vector

$$v_n \sim \mathcal{GP}\{\mu, K\}. \quad (1)$$

where $\mu := \mu(x)$ defines the GP's mean and $K := \text{cov}(x, x', \theta)$ the covariance matrix, with θ denoting hyper-parameters. The radial basis kernel function used to evaluate the covariance matrix reads

$$K = \sigma_x^2 \exp\left(-\frac{1}{2} \sum_{j=1}^d \frac{\|d\|_2^2}{l^2}\right). \quad (2)$$

Here, $d = x - x'$ is the distance vector, l length scale of the automatic relevance determination (ARD) for adjustable length scales for different input dimensions d [27], and σ_x^2 is a gain

parameter. The hyperparameters θ of the Gaussian process are adapted by optimization to minimize the negative log marginal likelihood objective function $L(\mathbf{y}, \theta)$ [28]

$$\underset{\theta}{\operatorname{argmin}}\{-\log L(\mathbf{v}, \theta)\} \quad (3)$$

from an initial prior distribution (zero-mean Gaussian process $\sim \mathcal{N}(\mathbf{0}, \mathbf{K})$). Thereafter, the interpolation operation using GPR can be performed by making a prediction $\mathbf{v}_n^*$ at location $\mathbf{x}^*$ based on the set of training data $\mathbf{v}_n$ at $\mathbf{x}$ as described in [28]. The joint distribution reads

$$\begin{bmatrix} \mathbf{v}_n \\ \mathbf{v}_n^* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}(\mathbf{x}) \\ \boldsymbol{\mu}(\mathbf{x}^*) \end{bmatrix}, \begin{bmatrix} \mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma_*^2 \mathbf{I} & \mathbf{K}(\mathbf{x}, \mathbf{x}^*) \\ \mathbf{K}(\mathbf{x}^*, \mathbf{x}) & \mathbf{K}(\mathbf{x}^*, \mathbf{x}^*) \end{bmatrix}\right) \quad (4)$$

with the predictive uncertainty σ_*^2 and leads to the predictive equation

$$\begin{aligned} \mathbf{v}_n^* \sim \mathcal{N}\left(\boldsymbol{\mu}(\mathbf{x}^*) + \mathbf{K}(\mathbf{x}^*, \mathbf{x}) \left[\mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma_*^2 \mathbf{I}\right]^{-1} [\mathbf{v}_n - \boldsymbol{\mu}(\mathbf{x})], \right. \\ \left. \mathbf{K}(\mathbf{x}^*, \mathbf{x}^*) - \mathbf{K}(\mathbf{x}^*, \mathbf{x}) \left[\mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma_*^2 \mathbf{I}\right]^{-1} \mathbf{K}(\mathbf{x}, \mathbf{x}^*)\right). \end{aligned} \quad (5)$$

Typically, to improve the condition of the numerical optimization, the input and output data are normalized independently. The mean of the training data μ_{scaler} is subtracted and it is scaled to unit variance by $X = \frac{X - \mu_{\text{scaler}}}{\sigma_{\text{scaler}}}$. Consequently, due to $E\left(\frac{X - \mu_{\text{scaler}}}{\sigma_{\text{scaler}}}\right) = \frac{E(X) - \mu_{\text{scaler}}}{\sigma_{\text{scaler}}}$ and $\text{COV}\left(\frac{X - \mu_{\text{scaler}}}{\sigma_{\text{scaler}}}\right) = \frac{\text{COV}(X)}{\sigma_{\text{scaler}}^2}$ the fitted Gaussian process has to be rescaled as

$$\mathcal{GP}_x\{\boldsymbol{\mu}, \mathbf{K}\} = \mathcal{GP}_X\left\{\sigma_{\text{scaler}}\boldsymbol{\mu} + \mu_{\text{scaler}}\mathbf{e}, \sigma_{\text{scaler}}^2\mathbf{K}\right\} \quad (6)$$

with $\mathbf{e}$ defining a vector of ones with the size of $\boldsymbol{\mu}$. In doing so, the predictive uncertainty of the GPR on the normalized dataset can be defined as the sum of the model variance and the noise variance.

2.2. Sound Power Estimates

As presented in [5], the sound propagation is solved using the open-source BEM library NiHu [29]. It solves the Helmholtz boundary integral equation for smooth boundaries [2]

$$p_a = \frac{1}{2} \int_{\Gamma_s} \left(p_a \frac{\partial G}{\partial y_i} - G \frac{\partial p_a}{\partial y_i} \right) \cdot \mathbf{n} \, d\Gamma, \quad (7)$$

where p_a is the fluctuating acoustic pressure and $G(\mathbf{x}|\mathbf{y}, k)$ the three-dimensional free-field Green's function. The Neumann boundary condition for the acoustic pressure (7) can be expressed by Euler's equation of motion

$$\Gamma_s : \frac{\partial p_a}{\partial y_i} n_i = \rho_0 j\omega v_n, \quad (8)$$

with angular frequency $\omega = 2\pi f$ and mean density of air at ambient conditions $\rho_0 = 1.225 \frac{\text{kg}}{\text{m}^3}$. The Burton–Miller method [30] was used to solve the Helmholtz boundary integral equation. The radiated sound power

$$P_a = \frac{1}{2} \int_{\Gamma_s} \Re\{p_a v_n^*\} d\Gamma, \quad (9)$$

with the complex conjugate being denoted by the superscript $*$ was computed as an integral metric of the acoustic far field. Whenever the radiation efficiency is one, typically

encountered at higher frequency vibrations, the BEM simulation can be approximated by the ERP accurately as mentioned, e.g., in [31]. In contrast to the BEM, no system of equations has to be solved, which makes it computationally efficient. The ERP sound power estimate is defined as

$$P_{\text{ERP}} = \frac{1}{2} c_0 \rho_0 \int_{\Gamma_s} \|v_n\|^2 d\Gamma \approx \frac{1}{2} c_0 \rho_0 \sum_i A_i v_{n,i} v_{n,i}^* \quad (10)$$

with A_i denoting the area of element i . The major benefit of the ERP estimate is its fast computation. Assuming constant-valued elements on the boundary Γ_s , the surface integral simplifies to a summation of the corresponding element values.

The ERP gives a good approximation of the radiated sound power for convex rigid bodies and at high frequencies [32]. This is due to modeling all sources with a constant radiation efficiency of 1, which gives an upper bound for the structure-induced acoustical field in most cases [32]. A good indicator for the appropriate frequency range to use the ERP is the coincidence frequency, which is sometimes referred to as the critical frequency. The coincidence frequency indicates the frequency at which the phase speed of bending and acoustic waves is equal [33]. Below the frequency, the radiation efficiency is expected to be low, and cancellation effects must be considered for predicting the radiated sound power, as mentioned in, e.g., [34]. For frequencies larger than f_c , the radiation efficiency is high, and the ERP gives a good approximation of the radiated sound power.

2.3. Confidence of the Sound Power Prediction

For each frequency ω , the real and the imaginary parts of the Fourier transformed normal velocity are fit to a Gaussian process

$$\Re\{v_n(\mathbf{x}, \omega)\} \sim \mathcal{GP}\{\boldsymbol{\mu}_{\omega, \Re}, \mathbf{K}_{\omega, \Re}\}, \quad (11)$$

$$\Im\{v_n(\mathbf{x}, \omega)\} \sim \mathcal{GP}\{\boldsymbol{\mu}_{\omega, \Im}, \mathbf{K}_{\omega, \Im}\}, \quad (12)$$

The Gaussian process can be used with Monte Carlo simulation to estimate the ERP of the surface velocity for every frequency and evaluation point, and then estimate the variance $\text{VAR}[ERP]$ and the expected value $E[ERP]$ of the acoustic signal. This brute force technique can also be performed for the sound power estimate of the BEM. In each sampling location $\mathbf{X} \in \mathbb{R}^{2 \times m} \rightarrow \Re\{V_n(\omega)\} \in \mathbb{R}^m$ and $\mathbf{X} \in \mathbb{R}^{2 \times m} \rightarrow \Im\{V_n(\omega)\} \in \mathbb{R}^m$. The evaluation is that of a multivariate normal distribution $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{K})$ with a symmetric, real-valued, and positive definite covariance matrix $\mathbf{K}$. A computationally efficient draw of samples can be performed by

$$\mathbf{x} = \boldsymbol{\mu} + \mathbf{L}\mathbf{z} \quad (13)$$

with the standard normal random vector $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and the lower triangular matrix $\mathbf{L}$ defined as $\mathbf{K} = \mathbf{L}\mathbf{L}^\top$. $\mathbf{L}$ can be computed efficiently by Cholesky decomposition.

This brute force technique is relatively inefficient, since a large number of GP realizations have to be computed.

Alternatively, the collective normal velocity can be described by the descriptive statistics of a multivariate normal distribution with a given mean and covariance $\Re\{V_n(\omega)\} \sim \mathcal{N}(\boldsymbol{\mu}_{\Re}, \mathbf{K}_{\Re})$, $\Im\{V_n(\omega)\} \sim \mathcal{N}(\boldsymbol{\mu}_{\Im}, \mathbf{K}_{\Im})$ connected to the Gaussian processes, respectively. The

$$ERP \sim (\Re\{V_n(\omega)\})^\top (\Re\{V_n(\omega)\}) + (\Im\{V_n(\omega)\})^\top (\Im\{V_n(\omega)\}) \sim \chi_{gen, \Re}^2 + \chi_{gen, \Im}^2 \quad (14)$$

which is proportional to the surface motion energy, being proportional to the square of the collective normal velocities. The individual distributions of the two terms are uncorrelated general chi-squared distributions. The collective expected value and variance are the sum of the individual expected value and variance

$$E[\chi_{gen,\mathfrak{R}}^2 + \chi_{gen,\mathfrak{S}}^2] = E[\chi_{gen,\mathfrak{R}}^2] + E[\chi_{gen,\mathfrak{S}}^2] \quad (15)$$

$$VAR[\chi_{gen,\mathfrak{R}}^2 + \chi_{gen,\mathfrak{S}}^2] = VAR[\chi_{gen,\mathfrak{R}}^2] + VAR[\chi_{gen,\mathfrak{S}}^2] \quad (16)$$

The individual terms are computed as follows [35]: given a column vector $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{K})$, which could be the normal velocity, find the probability of the squared sum of the individual samples

$$q(\mathbf{x}) = \mathbf{x}^T \mathbf{I} \mathbf{x} > 0. \quad (17)$$

With $\mathbf{I}$ being the identity matrix. This can be viewed as the multi-dimensional integral or sum of the energy of a multivariate normally distributed function, $\mathbf{x} = \sqrt{\mathbf{K}}\mathbf{z} + \boldsymbol{\mu}$. $\sqrt{\mathbf{K}} = \mathbf{V}^T \sqrt{\mathbf{D}} \mathbf{V}$, where $\mathbf{D}$ is the diagonal of the eigenvalues and $\mathbf{V}$ the respective eigenvectors of the eigenvalues. Now, we can transform the expression to $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ by the inverse $\mathbf{z} = \mathbf{K}^{-1/2}(\mathbf{x} - \boldsymbol{\mu})$. This decorrelates the variables $\mathbf{x}$, and transforms the integration domain to a different quadratic form (bilinear form):

$$\tilde{q}(\mathbf{z}) = \mathbf{z}^T \mathbf{K} \mathbf{z} + 2(\mathbf{K}^{1/2} \boldsymbol{\mu}) \mathbf{z} + \boldsymbol{\mu}^T \boldsymbol{\mu} > 0 \quad (18)$$

Now the problem is to find the probability of the standard normal $\mathbf{z}$ in this domain. We can start at the eigen-decomposition $\mathbf{K} = \mathbf{V}^T \mathbf{D} \mathbf{V}$ of $\mathbf{K}$. The transformed variable $\mathbf{y} = \mathbf{V} \mathbf{z}$ is also standard normal, and in this space the quadratic is given by the following:

$$\begin{aligned} \hat{q}(\mathbf{y}) &= \mathbf{y}^T \mathbf{D} \mathbf{y} + 2(\mathbf{K}^{1/2} \boldsymbol{\mu}) \mathbf{V} \mathbf{y} + \tilde{q}_0 \quad \left(\mathbf{b}^T = 2(\mathbf{K}^{1/2} \boldsymbol{\mu}) \mathbf{V} \right) \\ &= \sum_i \left(D_i y_i^2 + b_i y_i \right) + \sum_{i'} b_{i'} y_{i'} + \boldsymbol{\mu}^T \boldsymbol{\mu} \\ &= \sum_i D_i \left(y_i + \frac{b_i}{2D_i} \right)^2 + \sum_{i'} b_{i'} y_{i'} + \boldsymbol{\mu}^T \boldsymbol{\mu} - \sum_i D_i \left(\frac{b_i}{2D_i} \right)^2 \\ &= \sum_i D_i \chi_{1, (b_i/2D_i)}^2 + \mathcal{N}(m, s^2), \end{aligned}$$

here, i and i' index the nonzero and zero eigenvalues of the eigenvalue decomposition. Finally, we obtain a weighted sum of non-central chi-square variables χ^2 , each with 1 degree of freedom, and a normal variable $\mathcal{N}(m, s^2)$. So this is a generalized chi-square variable $\tilde{\chi}_{w,k,\lambda,s,m}$, where we merge the non-central chi-square variables with the same weights, so that the vector of their weights $\mathbf{w} = \text{diag}(\mathbf{D}) \quad \forall \quad D_i \neq 0$ are the unique eigenvalues among D_i , their degrees of freedom k are the numbers of times the eigenvalues occur, and their non-centralities and the normal parameters are as follows:

$$\lambda_j = \frac{1}{4w_j^2} \sum_{i=0:D_i=w_j}^k b_i^2, \quad s = \sqrt{\sum_{i'=0:D_i=0} b_{i'}^2}, \quad m = \boldsymbol{\mu}^T \boldsymbol{\mu} - \mathbf{w} \cdot \boldsymbol{\lambda}.$$

The : in the summation index describes that we iterate over the indices i from zero to k , only for the entries where the weight w_j and the iterated D_i are equal. The mean estimate can be described over a sum of j distinct eigenvalues occurring k_j times in the diagonal matrix $\mathbf{D}$

$$\mu_q = \sum_j w_j(k_j + \lambda_j) + m \quad (19)$$

and the variance estimate

$$\sigma_q^2 = 2 \sum_j w_j^2(k_j + 2\lambda_j) + s^2. \quad (20)$$

Note that this sound power estimate can be also performed using the system matrix of the BEM. The main drawback of this method is that based on the number of points, an eigenvalue problem (algorithmic complexity of $O(n^3)$) has to be solved that slows down the prediction of the sound power estimate drastically.

2.4. Approximation of the χ^2 Expected Value and χ^2 Variance to Describe the Sound Power

In general, the expected value and the variance of the bilinear form (18) can be calculated by

$$\mu_q = E[z^T K z + 2(K^{1/2} \mu) z + \mu^T \mu] = E[z^T K z] + \mu^T \mu = \sum_i K_{ii} + \mu^T \mu \quad (21)$$

and the variance by

$$\text{VAR}[z^T K z + 2(K^{1/2} \mu) z + \mu^T \mu] = E[(z^T K z + 2(K^{1/2} \mu) z + \mu^T \mu - \mu_q)^2]. \quad (22)$$

This expression can be further expanded by the expected value into

$$E[(z^T K z - \sum_i K_{ii} + 2(K^{1/2} \mu) z)^2] = \text{VAR}[z^T K z] + \text{VAR}[2(K^{1/2} \mu) z]. \quad (23)$$

where the covariance of the normally distributed variables and the square of it is zero. Furthermore, we can utilize the details of the variance of the first and last expression for the positive definite covariance matrix K

$$\text{VAR}[z^T K z] = 2 \sum_i (K^T K)_{ii} = 2 \sum_{i,j} K_{ij}^2. \quad (24)$$

$$\text{VAR}[2(K^{1/2} \mu) z] = 4 \mu^T K \mu. \quad (25)$$

Finally, we obtain a simple expression for the variance

$$\sigma_q^2 = 2 \sum_{i,j} K_{ij}^2 + 4 \mu^T K \mu. \quad (26)$$

The predictive uncertainty $\sigma_*^2 = \sigma_q^2 / \sigma_{\text{scaler}}^2$ of the ERP can be defined concerning the normalized GPR output data, such that it will have a value close to zero for low variance in the prediction and a value close to one for uncertain predictions.

2.5. Adaptive Refinement Procedure

To improve the regression quality, additional sampling points can be defined based on an adaptive refinement procedure. The procedure follows the schematic given in Figure 1. It starts by performing a variational GPR, and as long as a stopping criterion is not met, the adaptive refinement procedure is carried out. Therefore, a potential improvement function is defined to estimate the optimal locations for describing the GPR with reduced noise. The idea is to target the maxima of the function and accurately predict them using the GPR. We

define the improvement function $I(x)$ at a new point x as the probability that it is located at a maximum of the absolute value of the underlying process

$$I(x) = \max(|f(x)|).$$

The expectation of this improvement $EI(x)$ is defined by

$$EI(x) = \mathbb{E}[\max(|f(x)|)]$$

Based on the $Z = \frac{\mu(x)}{\sigma(x)}$, the $EI(x)$ is given by

$$EI(x) = \mu(x)(1 - 2\Phi(-Z)) + 2\sigma(x)\phi(Z)$$

where $\Phi(\cdot)$ is the Cumulative Distribution Function (CDF) and $\phi(\cdot)$ is the Probability Density Function (PDF) of the standard normal distribution. Now we select locations x that maximize this $EI(x)$ value. Finally, we select the best 500 points from all adaptive refinement samples that improve the GPR, and then randomly select 20 points. This process is carried out independently for both the real and imaginary parts. Both new locations are added to obtain new measurements, as with one measurement, we obtain both the real and the imaginary parts. With that procedure, we minimize the noise in the final ERP values.

3. Application

The method was tested using a vibrating plate benchmark test (see Figure 2a). It comprises a thin square-shaped steel plate with dimensions $(1\text{ m} \times 1\text{ m} \times 0.01\text{ m})$ and material properties (density $\rho = 7872\text{ kg m}^{-3}$, Young's modulus $E = 200\text{ GPa}$, and Poisson number $\nu = 0.29$) [5]. The plate was clamped at the edge Γ_{fix} and was free to move on the other edges. A time-harmonic point force of strength $F_z = 1\text{ N}$ is applied at position $x_0 = [0.1, 0.2, 0.01]^\top\text{ m}$ in the z -direction, exciting a vibrating motion in the structure.

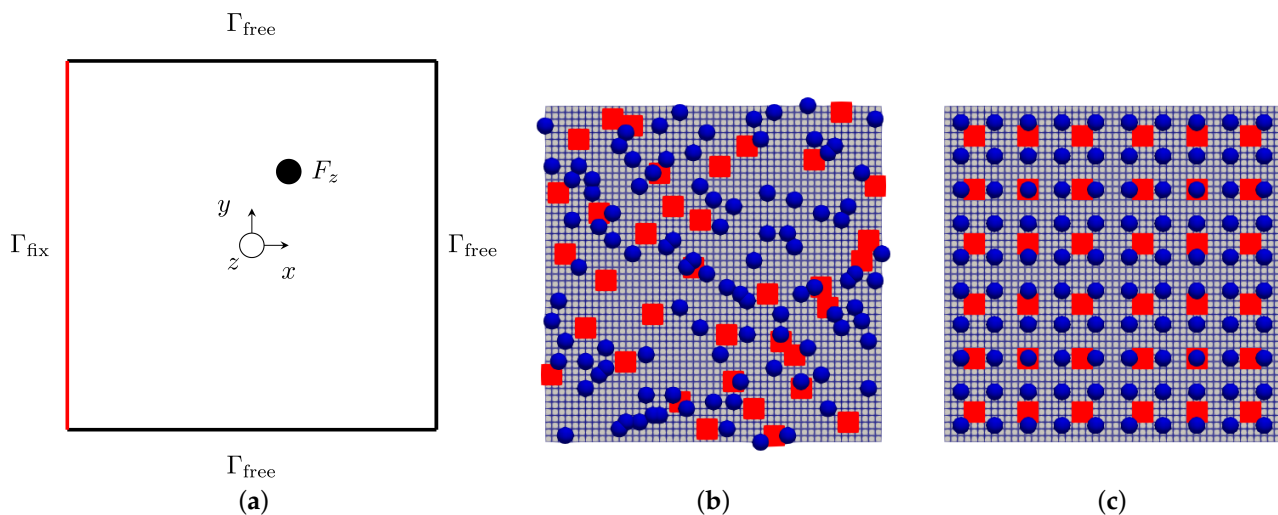


Figure 2. Top view of the investigated plate with an exemplary set of 36 (red) and 100 (blue) measurement points, based on Latin hypercube (b) and uniform (c) sampling, respectively. (a) Simulation setup, with force F_z and boundary conditions [5]. (b) Latin hypercube sampling. (c) Uniform sampling.

A time-harmonic finite element simulation was conducted using openCFS [36] to obtain the plate's surface vibration at each frequency. This finite element simulation served as pseudo-measurement data. At each vertex of the black grid depicted in Figure 2, the normal velocity is known and used for sampling the reference data as a training dataset. In

Figure 3a, some eigenmodes of the benchmark problem are visualized. The coincidence frequency is $f_c = 1233$ Hz. For frequencies larger than f_c , the radiation efficiency is high, and the ERP gives a good approximation of the radiated sound power as shown in Figure 3b.

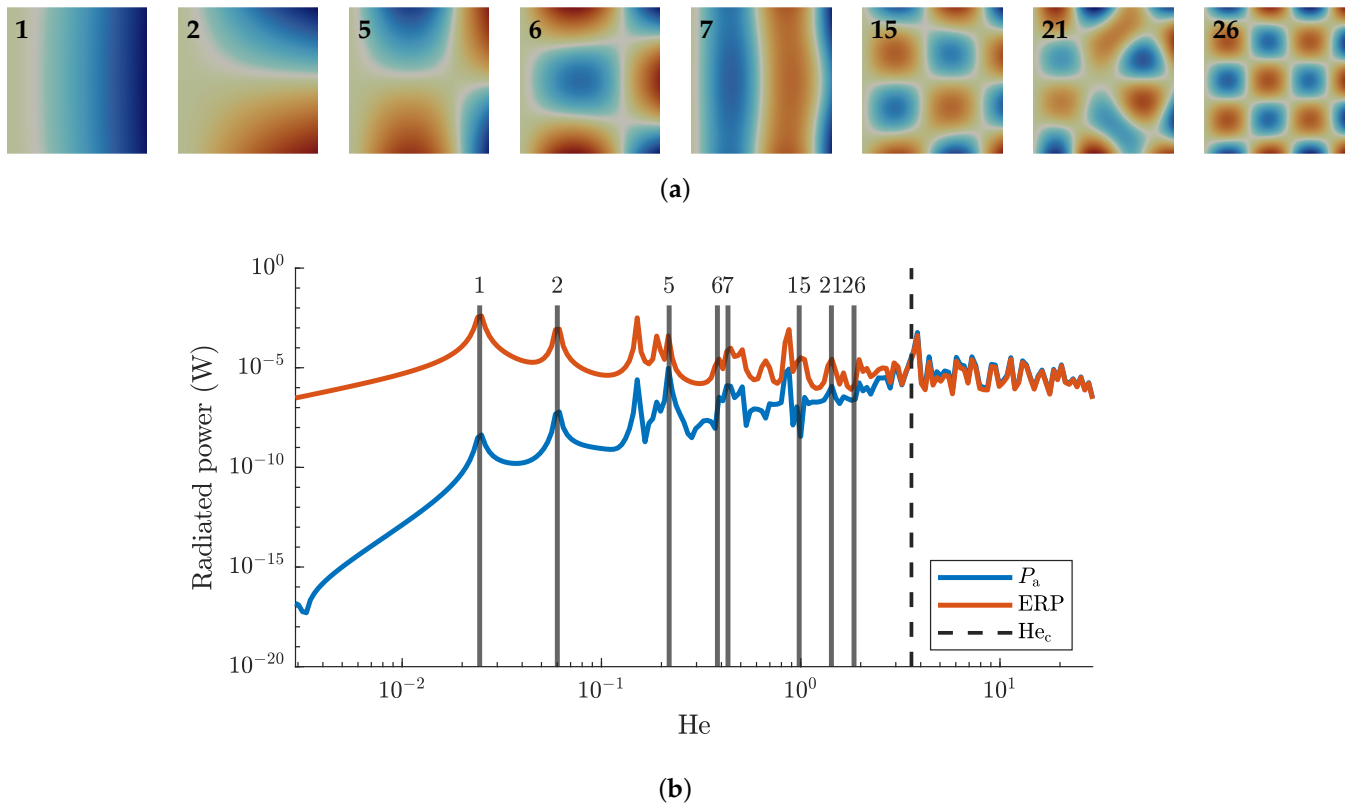


Figure 3. Solution of the benchmark problem. Indicated eigenmodes at $He = 0.024, 0.06, 0.22, 0.38, 0.43, 0.98, 1.43, 1.85$. The coincidence frequency for this setup is $He = 3.6$. (a) Real part of the mode shape for selected structural eigenmodes. (b) Radiated sound power based on BEM (P_a) and ERP.

Two sampling strategies were tested: (i) Latin hypercube (LHC) sampling and (ii) uniform grid sampling. In the case of the LHC sampling as depicted in Figure 2b, the closest node to each sample was used as a measurement point. For the uniform grid sampling depicted in Figure 2c, equally spaced nodes are selected. In Figure 2, two exemplary sets of measurement points are indicated in red and blue color, respectively. To emulate the effect of noisy observations, uncorrelated noise was added to the simulated pseudo-measurement data $\hat{v}_{n,\omega}(x) = v_{n,\omega}(x) + \epsilon_\omega(x)$. The added noise $\epsilon_\omega \sim N(0, \sigma_{\epsilon,\omega}^2)$ was defined as spatially normally distributed noise with scaled variance based on the mean velocity amplitude $\sigma_{\epsilon,\omega} = k_{\text{noise}} \overline{v_{n,\omega}(x)}$. The factor was chosen as $k_{\text{noise}} = 0.001$.

3.1. Exact Gaussian Process Regression

In [20], the prediction capabilities of exact Gaussian process regression are discussed. It is shown that the rule of thumb, that six points per wavelength are sufficient for a respective estimate [37], can be used to give a limit of the minimum number of sampling points required for the interpolation. As shown in [20], as soon as the vibration is under-resolved, it does not matter which interpolation method is used, and the result is poor. In contrast, as long as the experimental resolution is given, the radial basis function kernel GPR outperformed standard linear interpolation significantly. Similar to GPR, ordinary radial basis function interpolation [38] based on the same geometric length scale reaches a comparable level of accuracy.

At higher frequencies, the mode shapes become more complex, and the number of measurement points required to prevent under-resolution increases. Consequently, the dimensionality of the GPR gets high which first results to badly conditioned optimization problems and therefore significant numerical errors and secondly comes with substantial computational cost as the cost for inference scales at $\mathcal{O}(N^3)$ due to the need to invert the covariance matrix. Therefore, for the regression of high-frequency oscillation shapes, approximate inference can be the superior method.

3.2. Stochastic Variational Gaussian Process Regression

The exact GPRs complexity scales with the number of input data points N cubed $\mathcal{O}(N^3)$. To address this, we employ the stochastic variational Gaussian process (SVGP) framework introduced in [39]. This method combines variational inference with sparse GPs, reducing the cost from $\mathcal{O}(N^3)$ to $\mathcal{O}(NM^2)$, where $M \ll N$ is the number of inducing points.

Stochastic variational Gaussian process regression is based on variational Bayesian inference. An introduction to variational inference can be found in [40]. Bayesian inference derives the posterior distribution $p(f|x)$ using Bayes' rule

$$p(f|x) = \frac{p(x, f)}{p(x)} = \frac{p(x|f)p(f)}{p(x)} \quad (27)$$

from the prior distribution $p(f)$, the likelihood $p(x|f)$, and the evidence also referred to as marginal likelihood $p(x)$. The evidence is defined as

$$p(x) = \int p(x|f)p(f)df \quad (28)$$

To obtain computationally efficient inference, the integration in Equation (28) is typically approximated as direct integration is intractable in most cases. This can be achieved through Monte Carlo integration or by utilizing variational inference. In variational inference, an approximation

$$q(f) \approx p(f|x)$$

for most f, x is desired such that x can be inferred from f cheaply. The Kullback–Leibler divergence can be used to quantify the similarity between both distributions, the truth and its approximation. The according optimization problem describes finding the best approximation $q_*(f)$

$$q_*(f) = \operatorname{argmin}_{q(f)} \text{KL}(q(f)||p(f|x)) \quad (29)$$

Unfortunately, the Kullback–Leibler divergence [41] $\text{KL}(q(f)||p(f|x))$ comes with the same flaw as the evidence integral because it requires the evidence $\log p(x)$ in its evaluation. However, a lower bound of the objective can be stated

$$\text{ELBO} = E[\log p(f, x)] - E[\log q(f)] \leq \log p(x) \quad (30)$$

This function is called evidence lower bound (ELBO). Maximizing the ELBO is equivalent to minimizing the KL divergence [40] and ensures that the approximate posterior $q(f)$ is close to the true posterior $p(f|x)$ while avoiding the intractable computation of the evidence.

While variational inference provides a way to efficiently approximate the posterior distribution, sparse GPs alleviate the computational bottleneck of handling all training points. Combining these two approaches yields scalable inference. Sparse Gaussian

processes rely on inducing point methods [42] where the inducing variables $\mathbf{u}$ are used paired with the inducing inputs $\mathbf{Z}$. The distribution augmented with inducing points reads

$$p(\mathbf{f}, \mathbf{u}) = \mathcal{N}\left(\begin{bmatrix} \mathbf{f} \\ \mathbf{u} \end{bmatrix} \middle| \mathbf{0}, \begin{bmatrix} \mathbf{K}_{nn} & \mathbf{K}_{nm} \\ \mathbf{K}_{nm}^\top & \mathbf{K}_{mm} \end{bmatrix}\right) \quad (31)$$

with indices n representing the data points $x_n \in \mathbf{X}$ and m the inducing point indices $z_m \in \mathbf{Z}$. Again, Bayesian inference can be written as

$$p(\mathbf{u}|\mathbf{x}) = \frac{p(\mathbf{x}, \mathbf{f}, \mathbf{u})}{p(\mathbf{x})} = \frac{p(\mathbf{x}|\mathbf{f})p(\mathbf{f}|\mathbf{u})p(\mathbf{u})}{p(\mathbf{x})}$$

with $p(\mathbf{u}) \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_{mm})$, $p(\mathbf{f}|\mathbf{u}) \sim \mathcal{N}(\mathbf{K}_{nm}\mathbf{K}_{mm}^{-1}\mathbf{u}, \mathbf{K}_{nn} - \mathbf{Q}_{nn})$, and $\mathbf{Q}_{nn} = \mathbf{K}_{nm}\mathbf{K}_{mm}^{-1}\mathbf{K}_{nm}^\top$.

Using the Fully Independent Training Conditional (FITC) method to approximate the integration over $\mathbf{f}$, one arrives at a tractable bound on the marginal likelihood for the Gaussian case [24]

$$\begin{aligned} \log p(\mathbf{x}) &\geq \log \mathcal{N}\left(\mathbf{0}, \mathbf{K}_{nn}\mathbf{K}_{mm}^{-1}\mathbf{K}_{nm}^\top + \sigma^2\mathbf{I}\right) \\ &\quad - \frac{1}{2\sigma^2} \text{tr}(\mathbf{K}_{nn} - \mathbf{Q}_{nn}) \end{aligned} \quad (32)$$

This bound on the marginal likelihood, also referred to as *evidence*, can be used as an objective function in optimizing the covariance function parameters as well as the inducing input points $\mathbf{Z}$. The advantage of this approach is the reduced computational cost of Equation (32) to $\mathcal{O}(NM^2)$. Combining this approach of sparse GPs with the variational inference [43] yields

$$\begin{aligned} \log p(\mathbf{x}) &\geq \log \mathcal{N}\left(\mathbf{K}_{nm}\mathbf{K}_{mm}^{-1}\mathbf{m}, \sigma^2\mathbf{I}\right) \\ &\quad - \frac{1}{2\sigma^2} \text{tr}\left(\mathbf{K}_{nm}\mathbf{K}_{mm}^{-1}\mathbf{S}\mathbf{K}_{mm}^{-1}\mathbf{K}_{mn}\right) \\ &\quad - \frac{1}{2\sigma^2} \text{tr}(\mathbf{K}_{nn} - \mathbf{Q}_{nn}) - \text{KL}[q(\mathbf{u})\|p(\mathbf{u})] \end{aligned} \quad (33)$$

with the variational parameters $\mathbf{m}, \mathbf{S}$ such that $q(\mathbf{u}) \sim \mathcal{N}(\mathbf{m}, \mathbf{S})$.

In the following, this combined approach of sparse GPs and variational inference will be referred to as variational GPR. Its key advantage lies in enabling scalable optimization in both stochastic [43] and distributed settings [44,45].

4. Hypotheses

The core part of this article is the investigation of the predictive capabilities of Gaussian process regression for the interpolation task, specifically concerning the mean and variance of surface vibration fields. The primary interest is how the variability in terms of the mapped variance is translated into the sound emission predictions. In general, a significant variability results in a large variability in the predicted sound signal. There are two common methods to compute the standard deviation of the radiated sound power. First, a Monte Carlo simulation of the sound field using the boundary element method or the ERP. Another method, and a faster prediction method, is the estimate based on statistics of stochastic processes described in the previous sections. Both methods are suitable and can be used to update the number of sampling points and refine the underlying Gaussian process. Therefore, the following seven hypotheses are formulated to address the research questions regarding the usage of GPRs in acoustic end-of-line testing:

H1. *The algorithm derived can provide real-time capability regarding the sampling cycle of the laser scanning vibrometry system with fixed sample averaging and fixed measurement uncertainty thresholds.*

H2. *The number of training points for the Gaussian process regression positively correlates with the upper frequency limit at which the Gaussian process can accurately model the surface normal velocity and sound prediction.*

H3. *The interpolation error to the ground truth velocity correlates with the noise of the velocity prediction from the GPR.*

H4. *The Latin hypercube sampling with the same number of sampling points as regular grid sampling results in a lower level of error (of the mean compared to the ground truth) and variance of the predicted GP after the same number of training epochs.*

H5. *There exists an adaptive refinement method that reduces the (integrated) modeled noise of the velocity prediction and hence improves the accuracy in the sound power estimate.*

H6. *The ERP-based estimates based on the statistical theory have higher computational performance than the Monte Carlo with BEM, while maintaining the accuracy of the predicted sound power above a specific frequency where the radiation efficiency approaches one.*

H7. *The computational performance can be significantly increased using transfer learning, based on pre-trained models.*

H8. *The proposed method can handle different levels of noise contained in the measurement data.*

5. Results

The interpolation based on GPR has the unique advantage of quantifying uncertainty of a model, which is estimated based on experimental data. The model uncertainty is described by a mean, a covariance, and a likelihood variance estimate. This result provides the opportunity to model the variance of the surface velocity estimate, which can be translated into the variability of the radiated sound power. It serves as an informed approach to reducing uncertainty in the measurement procedure.

5.1. Computational Performance Improvement

One focus of this work was to reduce the time required for GPR training (Hypothesis 1). Thereby, the alterations from the preliminary setup [20] were twofold. First, the core framework for the training process was replaced by *GPyTorch 1.14* [18], a highly efficient and modular implementation of GPs. GPU acceleration may become attractive as we approach several tens of thousands of training points. The framework used is implemented in *PyTorch 2.6.0* [46]. On the reference setup using 3600 training points, an improvement of computational time by a factor of 60 was obtained comparing the variational GPR setup in *GPyTorch* compared to the exact GPR in *scikit-learn* (see Table 1). We tested two commonly used optimizers, the L-BFGS optimization algorithm and the Adam optimizer [47] in *GPyTorch*. The L-BFGS optimizer (learning rate of 0.01, with a maximum of 20 iterations and 10 previous gradients stored) took longer than the Adam optimizer (learning rate of 0.01) to reach a comparable accurate minimum. The exact GPR implementation was even faster when it converged. This was only the case at oscillations with very low frequency and using only a few training points (below 100). In other cases (when the number of training points increases), the exact GPR implementation in *GPyTorch* did not converge well.

Table 1. Comparison of computation time (in seconds) for $n = 3600$ uniformly distributed data points on the reference measurement grid. The mean time of an optimization run is presented in seconds. For the Monte Carlo Simulation 100 samples are drawn on a 60×60 evaluation grid. Computations were performed on a reference machine (Apple M4).

Task of Sound Prediction	BEM (Monte Carlo)		ERP (χ^2 -EV-Based)	(χ^2)
Exact GPR [20] Training (L-BFGS)			scikit-learn 535.0	
Variational GPR Training (L-BFGS)			GPyTorch 9.93	
Training (Adam)			7.96	
Draw samples		0.20	-	-
Evaluate radiated power distribution	5.49	0.004	3.15	0.04
Total time (Adam)	13.65	8.16	10.11	8.00

Inspecting the computation times of the preliminary investigations [20], the drawing of samples from the GP represents a significant contribution to the overall computation time. The comparably slow drawing of samples from the GP using the scikit-learn library was due to the used algorithm in the package, which computes the samples by singular value decomposition $K = U\Sigma V^T$ and $x = \mu + U\Sigma^{\frac{1}{2}}z$. As the covariance matrix of any multivariate normal distribution has to be positive definite, a computationally more efficient method to draw samples is the use of Cholesky decomposition $K = LL^T$ and Equation (13). By using this modified algorithm, we were able to reduce the computational time of the sample drawing in these cases (see Table 1) from over 16.7 s (scikit-learn [20]) to 0.20 s, which is a significant reduction.

The computation of the radiated sound field using the BEM represents a significant computational cost of 5.49 s (see Table 1) and 4.9 GB of RAM. As the number of required measurement points is dominated by the high-frequency oscillation patterns showing smaller length scale mode shapes, a high-frequency ERP approximation of the radiated sound power can be used instead. Compared to the radiated sound power of the BEM, the ERP-based method assumes a uniform radiation efficiency of one for any frequency of interest. These assumptions lead to significant deviations at low frequencies (below the coincidence frequency), while minor deviations occur for high frequencies of the signal (details will be provided in Section 5.6).

Therefore, the ERP was considered as an alternative for quality assessment using Monte Carlo simulation. It reduces the computational effort required to solve a system of algebraic equations (evaluation time of radiated sound power with BEM of 9.52 s) to a matrix-vector product (evaluation time of radiated sound power with ERP of 0.004 s). Furthermore, the explicit nature of the ERP makes it possible to reformulate it in a Bayesian sense to avoid the Monte Carlo simulation altogether, as shown in Section 2.4. There exist two estimates of the mean and variance, one being an exact estimate (χ^2 -EV-based) involving the eigenvalues of the covariance matrix and another approximate (χ^2) by directly acting on the covariance matrix and the mean vector. The inverse of the covariance matrix for a large number of sampling points is computationally costly and results in 3.15 s of computational time for the exact ERP estimate given in Section 2.3. In contrast, the ERP (χ^2) calculation using the covariance matrix and the mean vector directly outperforms all computations (resulting in a computational time of 0.04 s).

It then results in an overall time of 8.00 s for one prediction of a GPR and the evaluation of the sound power. In doing so, the computational time for a comparable number of

support points was reduced by a factor of 60 from the previous implementation [20] to 8.00 s on the same architecture.

5.2. Expected Value and Variance of the Sound Power

In this section, we investigate the predictive capability of the GPR for estimating the sound power as a function of the frequency (Hypothesis 2). For the exact GPR, when increasing the number of training points, the numerical implementation and data fitting became unstable. The expected trend, that a higher amount of training data allows for modeling plate vibrations at higher oscillation frequencies, was not provided. Therefore, we switched to a variational Gaussian process regression method (see Section 3.2) instead of the exact Gaussian process regression method. The overall setup remains unchanged. It was implemented in GPyTorch using the algorithms described in [39]. The number of inducing points was set to 1/10 of the total number of data points.

We selected a Gaussian prior for the length scales in each direction $\ell \sim \mathcal{N}(\mu_\ell, \sigma_\ell)$ (e.g., $\mu_\ell = (0.1, 0.1)^\top$, $\sigma_\ell = 10^{-3}\mathbf{I}$), which is updated during the training. The homoscedastic Gaussian noise on the scaled target vibration velocity of the real and imaginary parts is learned independently. For the sparse variational inference, we provided initial M inducing points (one tenth of the provided training points) $\mathbf{Z} = \{\mathbf{z}_m\}_{m=1}^M$ selected by k-means classification on the training points. The locations of the inducing points are learned during training. A Cholesky distribution parametrizes the variational posterior. The variational ELBO specifies the loss function, and mini-batch training is used with the Adam optimizer. The model and the likelihood parameters are trained jointly. Early stopping is implemented as soon as the noise is sufficiently small and no further improvement is visible on a validation set.

In Figure 4, the radiated sound power estimate compared to the ground truth is compared to the ERP estimation based on a generalized χ^2 distribution. Several dashed lines and shaded areas are shown in the plot, describing the results of different the uniform resolution of a total of 100, 400, 1600, 3600, and 6400 points. Firstly, in blue, the relative accuracy η_{χ^2} of the (true) ERP (estimated by the reference solution data) and the ERP predicted by χ^2 distribution descriptive statistics of the velocity data from the GPR

$$\eta_{\chi^2} = 1 - \left| \frac{P_{a,ERP}}{P_{a,ERP,\chi^2}} - 1 \right| = 1 - \epsilon_{\chi^2}.$$

Secondly, the blue-shaded band visualizes the predictive uncertainty σ_*^2 of the GPR. The mean and variance of statistical ERP prediction methods are in good agreement with the Monte Carlo simulation η_{MC} (mean deviations less than 10^{-5} for each study), suggesting that fast computations using descriptive statistics based on the χ^2 distribution are a valuable and computationally efficient alternative.

It is noticeable that the quality of the GPR interpolation is strongly governed by the predictive uncertainty σ_*^2 , which indicates how much can be described by the deterministic part of the GPR. In the case of a value of zero, the deterministic part can accurately describe the data. However, as the value increases towards 1, the descriptive capabilities of the GPR decrease, and one can see large deviations in the field plots compared to the ground truth. It means that the mean of the GPR is significantly different from the provided velocity data. In such a case, the modes of vibration cannot be explained by the GPR anymore. Figure 4 clearly shows that with an increasing number of training points, the limit where the variational GPR can still explain the surface vibration velocity rises monotonically. For a resolution of 100 points, vibration characteristics below a Helmholtz number of $He_{crit} = 0.25$ can be resolved. For 400 points, this limit increases to a Helmholtz number of $He_{crit} = 2.5$, for 1600 points to about a Helmholtz number of $He_{crit} = 13$, for 3600 and

6400 points above the maximum Helmholtz number of interest $He_{crit} > 29$. It is observed, that the predictive uncertainty is larger for 3600 in the whole frequency range compared to a 6400 points resolution.

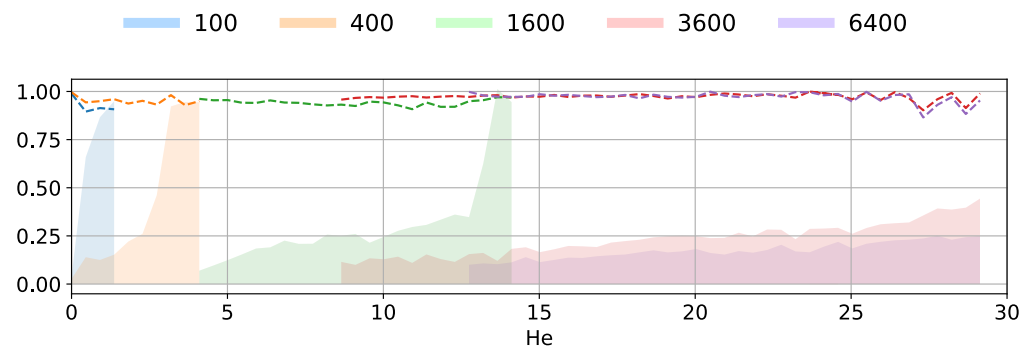


Figure 4. Evolution of the relative accuracies η_{χ^2} (dashed line) and predictive uncertainties σ_*^2 (shaded area) of the corresponding GPR as a function of the Helmholtz number, for different uniform resolutions of a total of 100 (blue), 400 (orange), 1600 (green), 3600 (magenta), 6400 (violet) points. The SNR level is 1000/1.

5.3. Correlation Between Sound Power Error ϵ and Predicted Uncertainty σ_*^2

In this section, we will investigate Hypothesis 3. According to Figure 4, there is no direct correlation between the interpolation error of the surface velocity and the predictive uncertainty σ_*^2 of the GPR visible. As soon as the GPR cannot explain the data by its deterministic part anymore, the predictive uncertainty σ_*^2 increases sharply, indicating that there are too few training points for the GPR. However, the error in the sound power (equal to the integrated surface velocity) increases only slightly.

Even at high GPR noise levels, the error of the predicted ERP is low. There is no clear trend whether the uniform or the Latin hypercube sampling performs better.

For a detailed analysis of the training time, the individual timings were separated into datasets, where the predictive uncertainty σ_*^2 is low and another where it is high. The separation was performed based on a critical Helmholtz number He_{crit} as a selector; everything below was judged as low noise, and everything above was regarded as a high noise region. In the respective ranges, the mean over all GPRs is computed. The means of the training times for each investigated set of are reported in Table 2, with training iterations over 1500 epochs to reach the low noise levels (Apple M4 Chip 10 Cores, 32 GB RAM, Apple Inc., Cupertino, CA, USA). The means are computed over all GPRs fulfilling the condition of the critical Helmholtz number He_{crit} . For a high number of training points and as soon as the GPR noise level increases, the training time increases substantially. There was no significant difference between the Latin hypercube sampling and the uniform sampling.

Table 2. GPR mean training times t_{tran} of the individual GPRs of the real and imaginary part.

Run	$\bar{t}_{tran,real} > He_{crit}$	$\bar{t}_{tran,real} < He_{crit}$	$\bar{t}_{tran,imag} > He_{crit}$	$\bar{t}_{tran,imag} < He_{crit}$
100, uni	2.626	0.428	2.589	0.256
400, uni	4.380	4.951	4.288	2.764
1600, uni	43.244	7.982	45.849	9.108
3600, uni	-	28.952	-	38.339
6400, uni	-	68.506	-	58.513

5.4. Uniform Sampling Versus Latin-Hypercube Sampling for the Surface Velocity

This section discusses two types of common (automated) data sampling strategies in connection with the underlying process specification (Hypothesis 4). The sampling strategies are as follows: first, a regular fit onto the geometry for uniform plate-like domains; second, a commonly used sampling strategy for arbitrary sampling spaces. The first strategy of uniform sampling is ideal in the case of uniform plate stiffness and the coordinate-sampling-aligned mode shapes. However, the results might deteriorate for inhomogeneous material, inhomogeneous plate thickness, non-rectangular geometries, non-uniform clamping at the boundaries, and other effects that produce modes with localized gradients and non-alignment with the sampling strategy. Therefore, we look into the details of the surface velocity predictions based on the two sampling strategies (using an exact GPR, which works for a low number of training points).

Figure 5 depicts the expected value and Figure 6 the standard deviation of the real part of the normal velocity for the respective sampling strategies for a frequency of 313 Hz.

The expected value is in close range to the ground truth of the reference data for both sampling strategies. For the standard deviations, uniform sampling leads to a significantly lower overall standard deviation between the sampling points compared to Latin hypercube sampling. For the Latin hypercube sampling, some blind spots on the geometry arise, where no support points of the GP are present, leading to a larger standard deviation in the prediction. These increased standard deviation spots could be used for adaptive refinement; however, they typically are located on the boundaries of the domain. This naive approach can be improved by using the variance-based refinement strategy presented in [48]. Another possibility is a modified version of the expected improvement [49].

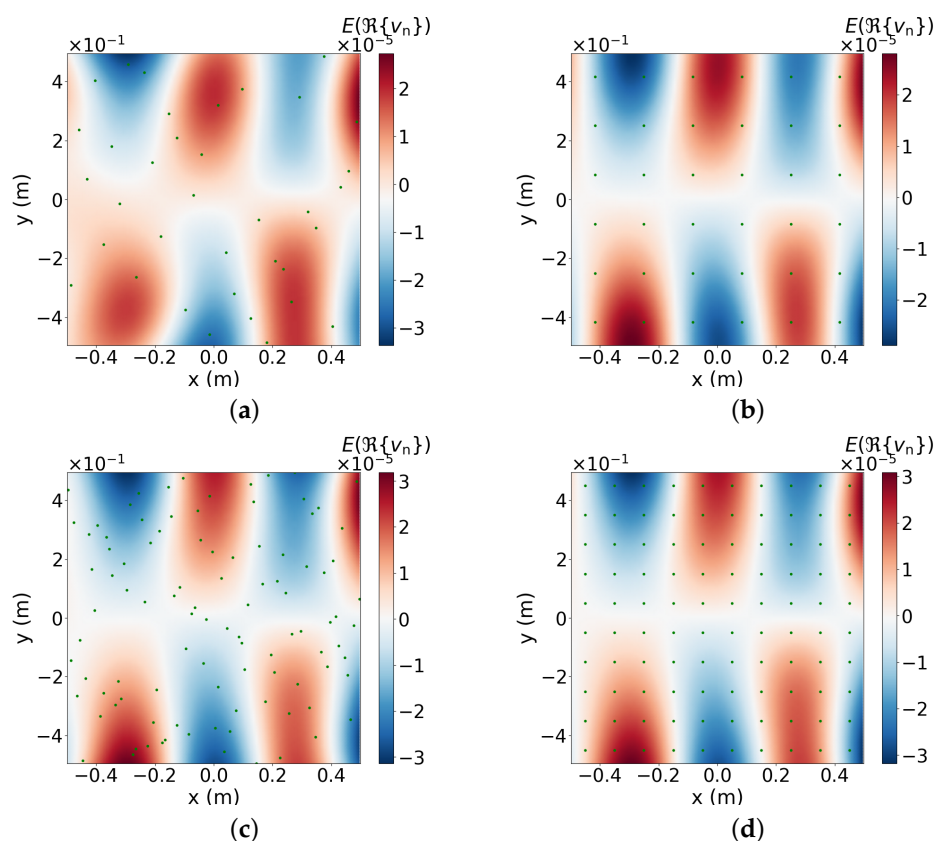


Figure 5. Mean of the real value of the normal velocity v_n at $f = 313$ Hz, Latin hypercube sampling, and uniform sampling for $n = 36$ and $n = 100$ pseudo measurement points indicated in green. (a) Latin hypercube sampling with $n = 36$. (b) Uniform sampling with $n = 36$. (c) Latin hypercube sampling with $n = 100$. (d) Uniform sampling with $n = 100$.

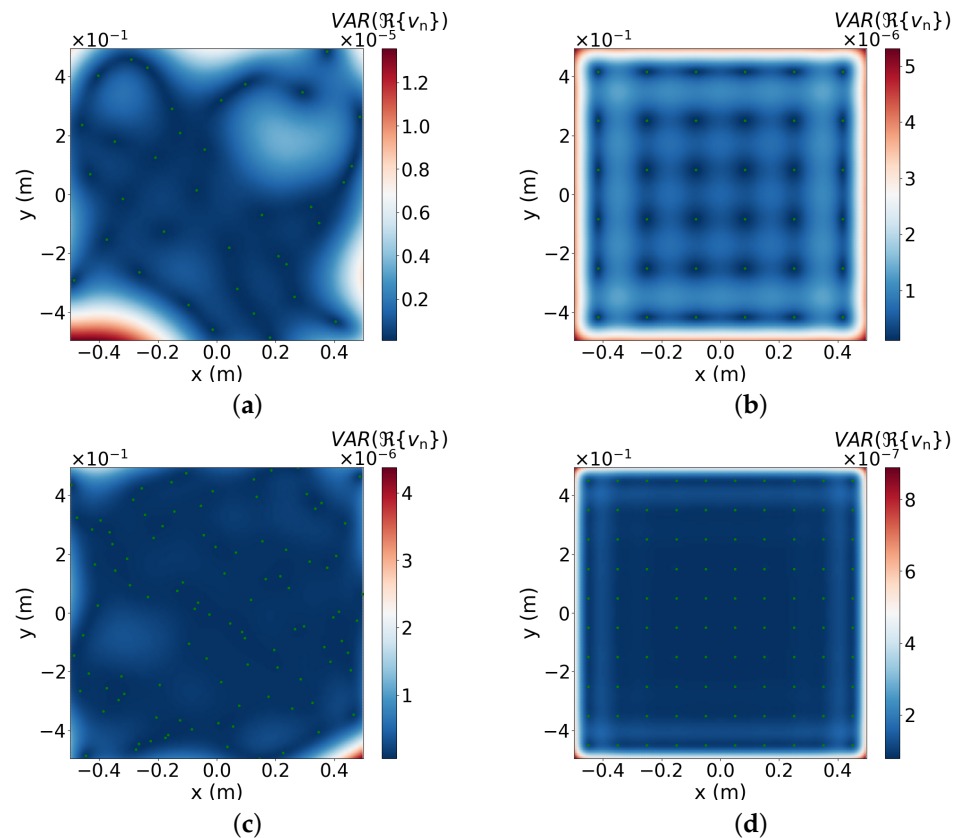


Figure 6. Standard deviation of the real value of the normal velocity v_n at $f = 313$ Hz, latin hypercube and uniform sampling for $n = 36$ and $n = 100$ pseudo measurement points indicated in green. (a) Latin hypercube sampling with $n = 36$. (b) Uniform sampling with $n = 36$. (c) Latin hypercube sampling with $n = 100$. (d) Uniform sampling with $n = 100$.

5.5. Design of Experiment Using Adaptive Refinement

To test Hypothesis 5, an adaptive refinement procedure as described in Section 2.5 was performed. As a stopping criterion, the normalized noise of the GPR exceeds was used with a threshold of 0.01. Figure 7 shows the frequency of 1250 Hz, the $EI(x)$ at the start of the adaptive refinement iterations, and at the third stage of the adaptive refinement iterations for the real part of the surface normal velocity field. At the first stage (Figure 7a), the GP is not able to resolve the velocity field adequately, and the EI suggests adding points at the domain corners. This is expected as at this stage the field is mostly described by noise everywhere. Due to the intended randomness in the point selection, the procedure remains robust and at the third stage (Figure 7b).

Figure 8 shows the effects of the adaptive refinement procedure on the sound power error and GPR predictive uncertainty as a function of the Helmholtz number. At the initial stage, before the refinement, 400 uniform training points were considered. As a comparison, this setup without adaptive refinement predicted a low noise in the function output until a Helmholtz number of about 2.5 (see Figure 4). In comparison, with the adaptive refinement procedure, a robust prediction was possible up to a Helmholtz number of 11. This accuracy is comparable to that of predictions provided by 1600 uniform training points, despite using only 1200 data points. It is observed that during the adaptive refinement, the predictive uncertainty is lower in the converged region ($He < 11$) compared to the predictions using 1600 uniform training points. However, the ERP estimate is slightly less accurate compared to the estimations based on 1600 uniform training points due to the early stopping criterion selected. The estimates were robust regarding the initial training distribution; when using a Latin hypercube sampling initially, comparable results were obtained.

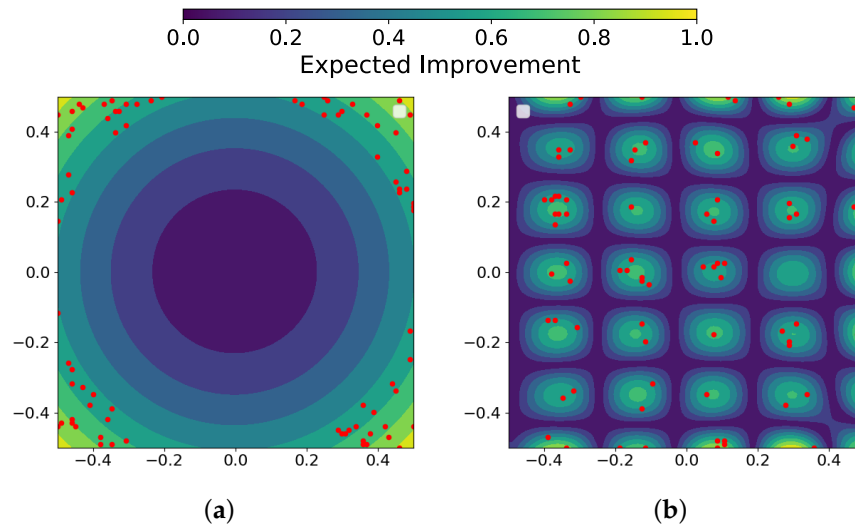


Figure 7. Two snapshots of $EI(x)$ during the adaptive refinement procedure for the real part of the surface normal velocity. The red dots indicate the added training pairs to the GPR training for both the real and imaginary parts. (a) First stage. (b) Third stage.

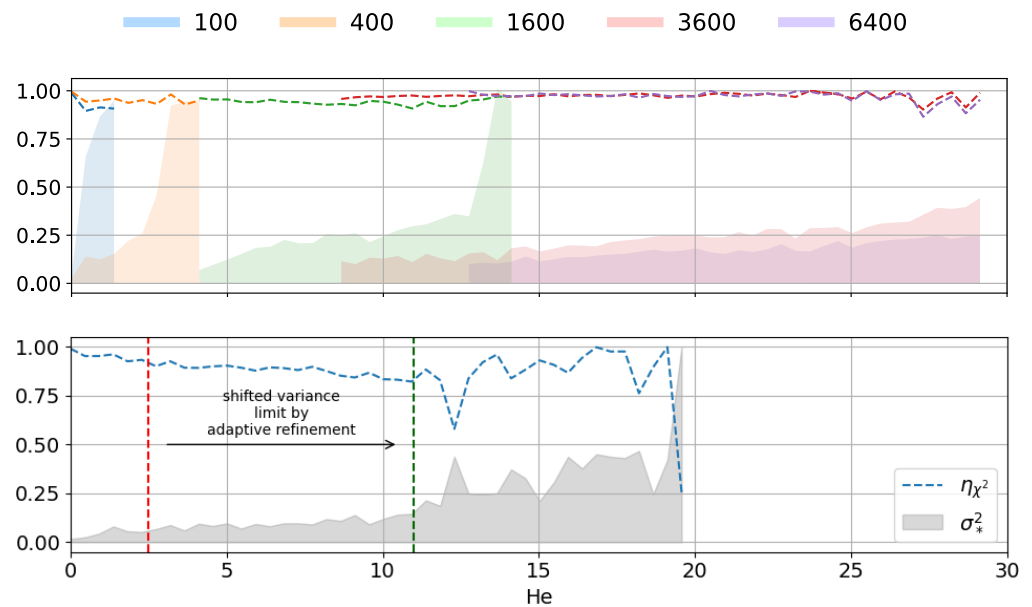


Figure 8. Evolution of the relative accuracy η_{χ^2} and the predictive uncertainty σ_*^2 as a function of the Helmholtz number. An initially uniform resolution of 400 was used, with random adaptive refinement. At the final stage, 1200 training points were considered.

5.6. Detailed Comparison of Sound Power Estimates

In Figure 9, the sound power estimate based on the ERP is compared to the one using BEM (Hypothesis 6). For the current case, below the coincidence frequency He_{co} (red line in Figure 9), the ERP estimate is not an accurate measure for the radiated sound power. Typically, with increasing frequency, the radiation efficiency increases, and as soon as it turns to one, the ERP and the BEM estimate coincide. Under this condition, the ERP measure is preferable, as the statistical estimation presented in this article enables the statistical description and estimates to be carried out. Regarding the simulation time, at low frequencies, a relatively coarse training point distribution is sufficient, allowing for the computation of the Monte Carlo simulation within a reasonable time frame. As indicated

by the results in Figure 4, the statistical values estimated by 100 simulations are sufficiently close to the analytically derived ones.

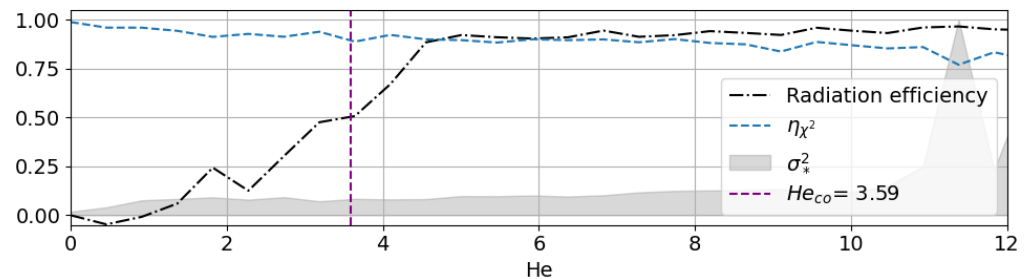


Figure 9. Comparison of the radiation efficiency, the relative ERP accuracy η_{χ^2} , and the predictive uncertainty σ_*^2 as a function of the Helmholtz number for the adaptive refinement case in Figure 8.

5.7. Computational Performance Using Pre-Trained Models

Finally, we tested if a significant speedup of the training time is possible using a pre-trained model (Hypothesis 7). Typically, during production, each product will not deviate much from the previous one (as long as it is of high quality); therefore, it might be interesting to use a pre-trained model. Table 3 compares the training times for several configurations. When using 1600 training points with the GPR setup, 1500 epochs of training, and a batch size of 1024, it took on average 5.57 s to train from scratch (in the region where convergence is achieved). When using a pre-trained GPR, the training time was reduced to 3.6 s (a nearly 36% reduction from 5.57 s to 3.6 s) with a standard deviation of 0.91 s. By relaxing the number of epochs (while still obtaining a converged GPR), the training time was reduced by approximately 68% from 3.6 s to 1.14 s. Overall, the mean training time was reduced from 5.57 s to 1.14 s, which yields an increase in computational performance of 79%.

Table 3. GPR mean training times below the critical Helmholtz number $He < He_{crit}$. The abbreviation uni stands for uniform sampling.

Run	$E[t_{tran}]$	$\sqrt{VAR[t_{tran}]}$	Training Epochs	Pre-Trained Model Used
1600, uni	5.57	1.94	1500	no
1600, uni	3.59	0.91	1500	yes
1600, uni	2.23	0.09	150	yes
1600, uni	1.14	0.12	75	yes

5.8. Influence of Measurement Noise

All previous investigations were based on a laser vibrometry measurement with a signal-to-noise ratio (SNR) of 1000/1, which represents a level of measurement quality that can be achieved for many applications. To test the effect of increased measurement noise on the proposed workflow (Hypothesis 8), the noise factor introduced in Section 3 was increased to $k_{noise} = 0.1$, representing a signal-to-noise ratio (SNR) of 10/1. Figure 10 depicts the evolution of the relative accuracy η_{χ^2} and the predictive uncertainty σ_*^2 as a function of the Helmholtz number for the increased noise level. One can observe that the method remains robust despite the increased measurement noise, and the GPR can interpolate the normal velocity field accurately for an adequate number of samples. Especially at a high number of samples, the results show a decreased critical Helmholtz number. Consequently, more samples are required to describe noisy measurement data in contrast to data with low SNR.

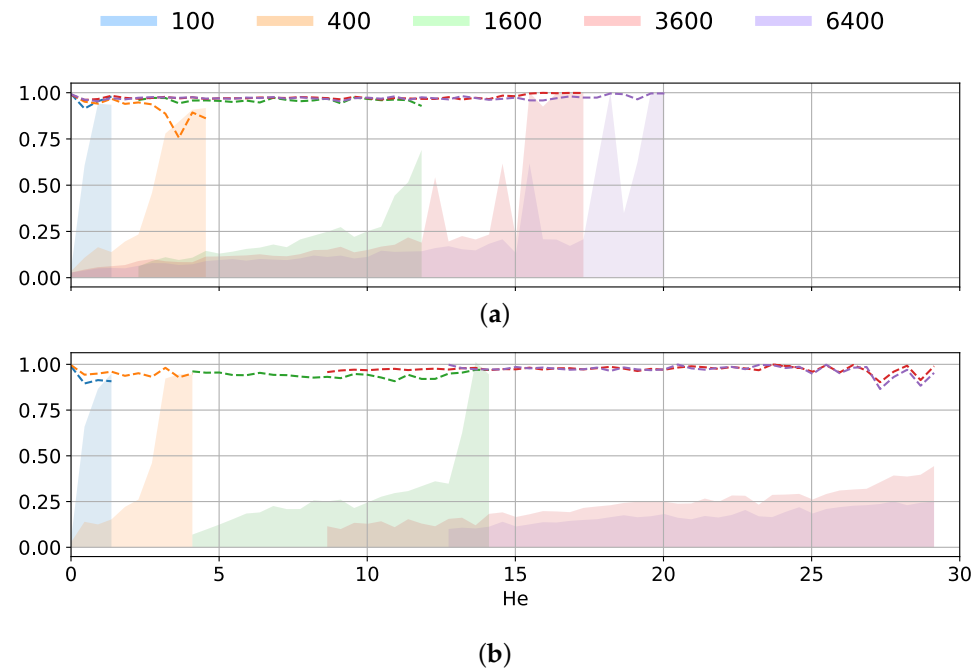


Figure 10. Evolution of the relative accuracy η_{χ^2} and the predictive uncertainty σ_*^2 as a function of the Helmholtz number with an increased SNR. (a) Increased SNR: $k_{\text{noise}} = 0.1$. (b) Reference SNR: $k_{\text{noise}} = 0.001$.

6. Conclusions

Initially, we identify four specifications for a measurement system for acoustic end-of-line testing to meet vibroacoustic requirements. Specification 1, to perform measurements at different locations, both contactless and without any additional assembly cost, was realized by a Laser scanning vibrometer setup. Specification 2 of the measurement system, which measures the accuracy and uncertainty of data prediction using a self-learned data-driven method, was achieved by the demonstration of the adaptive refinement method in combination with the underlying GPR. Specification 3, that the technique should provide data that give the maximum valuable insight into specific characteristics where deviations occurred over time, was not within the scope of the current investigation; however, it can be inferred through a comparative evaluation using the pre-trained models. The fourth specification that measurement and processing times should fit into the production cycle of a production line was demonstrated by the investigated hypotheses. The outcome of Hypothesis 1 is that the algorithm derived provides real-time capability regarding the sampling cycle of the laser scanning vibrometry system with fixed sample averaging and fixed measurement uncertainty thresholds. We demonstrated that the number of training points for Gaussian process regression positively correlates with the upper frequency limit at which the deterministic part of the Gaussian process can accurately model the surface normal velocity and sound prediction. We found that the ERP error with respect to the ground truth does not correlate with the noise in the velocity prediction from the GPR for the benchmark problem used. We found that Latin hypercube sampling, with the same number of sampling points, does not show better performance in training compared to uniform sampling. We also demonstrated the limitations of the ERP-based estimates by comparing them to sound power predictions using BEM. At higher oscillation frequencies, the ERP predicted sound power delivers accurate results. The computational performance was significantly increased (79% less time consumed) using transfer learning, based on pre-trained models. Overall, the use of GPyTorch over scikit-learn reduced the processing time by a factor of 60. Furthermore, we investigated the influence of uncorrelated noise on

the measurement data, where we observed an increased number of measurement points to describe noisy measurement data.

Finally, we want to point out perspectives to improve the GPR modeling, including a spatially varying measurement noise distribution and improvement of the adaptive refinement strategy. The measurement uncertainty based on the local measurement point of the experimental setup can be modeled within the GPR and, therefore, considered by the sound prediction method. A direct extension to measurement noise distribution in space (variation in input locations) can be undertaken via Variational Heteroscedastic Gaussian process regression [50]. Another direction would be to develop an improved adaptive refinement strategy that eliminates oscillatory behavior as soon as the predictive uncertainty increases. Ghosh et al. [51] propose a framework that partitions prediction uncertainty based on fidelity levels, allowing for a more nuanced approach to experimental design that considers both the cost and the expected information gain from various data sources. This strategy not only enhances the efficiency of the design process but also provides a systematic method for integrating information from different fidelity levels, thereby enriching the overall understanding of the system under investigation.

Author Contributions: Conceptualization, S.S.; Methodology, A.W. and S.S.; Software, A.W. and S.S.; Validation, A.W. and S.S.; Formal analysis, A.W. and S.S.; Investigation, A.W. and S.S.; Resources, S.S.; Data curation, A.W. and S.S.; Writing—original draft, A.W. and S.S.; Writing—review and editing, A.W. and S.S.; Visualization, A.W. and S.S.; Supervision, S.S.; Project administration, S.S.; Funding acquisition, S.S. All authors have read and agreed to the published version of the manuscript.

Funding: Stefan Schoder would like to thank the Austrian Federal Ministry for Innovation, Mobility and Infrastructure and the Austrian Research Promotion Agency (FFG) for funding the bio-COMP4acoustics project under the TakeOff programme line (project number 913972). Andreas Wurzinger and Stefan Schoder acknowledge partial funding support by the COMET project ECHODA (Energy Efficient Cooling and Heating of Domestic Appliances). ECHODA is funded within the framework of COMET - Competence Centers for Excellent Technologies by BMK, BMAW, the province of Styria, as well as SFG. The COMET programme is managed by FFG.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The models and the data presented in this study are available on request from the corresponding author.

Acknowledgments: Open Access Funding by the Graz University of Technology.

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

References

1. Kumar, S.; Forster, H.M.; Bailey, P.; Griffiths, T.D. Mapping Unpleasantness of Sounds to Their Auditory Representation. *J. Acoust. Soc. Am.* **2008**, *124*, 3810–3817. [\[CrossRef\]](#)
2. Marburg, S. Boundary Element Method for Time-Harmonic Acoustic Problems. In *Computational Acoustics*; Kaltenbacher, M., Ed.; CISM International Centre for Mechanical Sciences; Springer International Publishing: Cham, Switzerland, 2018; pp. 69–158. [\[CrossRef\]](#)
3. Vachálek, J.; Bartalský, L.; Rovný, O.; Šišmišová, D.; Morhác, M.; Lokšík, M. The digital twin of an industrial production line within the industry 4.0 concept. In Proceedings of the 2017 21st International Conference on Process Control (PC), Štrbské Pleso, Slovakia, 6–9 June 2017; pp. 258–262.
4. Klippel, W. End-of-line testing. In *Assembly Line-Theory and Practice*; Grzechca, W., Ed.; Books on Demand: Norderstedt, Germany, 2011; pp. 181–206.
5. Wurzinger, A.; Kraxberger, F.; Maurerlehner, P.; Mayr-Mittermüller, B.; Rucz, P.; Sima, H.; Kaltenbacher, M.; Schoder, S. Experimental Prediction Method of Free-Field Sound Emissions Using the Boundary Element Method and Laser Scanning Vibrometry. *Acoustics* **2024**, *6*, 65–82. [\[CrossRef\]](#)

6. Rasmussen, C.E. Gaussian processes in machine learning. In *Summer School on Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 63–71.
7. Wurzinger, A.; Kraxberger, F.; Mayr-Mittermüller, B.; Rucz, P.; Sima, H.; Kaltenbacher, M.; Schoder, S. Assessment of vibro-acoustic sound emissions based on structural dynamics. In Proceedings of the Proceedings of DAGA 2024, Hannover, Germany, 18–21 March 2024.
8. Li, J.; Wang, H. Gaussian Processes Regression for Uncertainty Quantification: An Introductory Tutorial. *arXiv* **2025**, arXiv:2502.03090. [[CrossRef](#)]
9. Schulz, E.; Speekenbrink, M.; Krause, A. A Tutorial on Gaussian Process Regression: Modelling, Exploring, and Exploiting Functions. *J. Math. Psychol.* **2018**, *85*, 1–16. [[CrossRef](#)]
10. Tuo, R.; Wang, W. Uncertainty quantification for Bayesian optimization. In Proceedings of the International Conference on Artificial Intelligence and Statistics, Virtual, 28–30 March 2022; pp. 2862–2884.
11. Bect, J.; Bachoc, F.; Ginsbourger, D. A Supermartingale Approach to Gaussian Process Based Sequential Design of Experiments. *Bernoulli* **2019**, *25*, 2883–2919. [[CrossRef](#)]
12. Jault, D.; Gillet, N.; Schaeffer, N.; Mandéa, M. Local Estimation of Quasi-Geostrophic Flows in Earth’s Core. *Geophys. J. Int.* **2023**, *234*, 494–511. [[CrossRef](#)]
13. Xia, G.; Han, Z.; Zhao, B.; Wang, X. Unmanned Surface Vehicle Collision Avoidance Trajectory Planning in an Uncertain Environment. *IEEE Access* **2020**, *8*, 207844–207857. [[CrossRef](#)]
14. Kocev, B.; Hahn, H.K.; Linsen, L.; Wells, W.M.; Kikinis, R. Uncertainty-Aware Asynchronous Scattered Motion Interpolation Using Gaussian Process Regression. *Comput. Med. Imaging Graph.* **2019**, *72*, 1–12 [[CrossRef](#)]
15. Jahanandish, M.H.; Fey, N.P.; Hoyt, K. Lower Limb Motion Estimation Using Ultrasound Imaging: A Framework for Assistive Device Control. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 2505–2514. [[CrossRef](#)]
16. Omainska, M. Visual Pursuit With Switched Motion Estimation and Rigid Body Gaussian Processes. *Trans. Inst. Syst. Control Inf. Eng.* **2023**, *36*, 327–335. [[CrossRef](#)]
17. Binois, M.; Huang, J.; Gramacy, R.B.; Ludkovski, M. Replication or Exploration? Sequential Design for Stochastic Simulation Experiments. *Technometrics* **2018**, *61*, 7–23. [[CrossRef](#)]
18. Gardner, J.; Pleiss, G.; Weinberger, K.Q.; Bindel, D.; Wilson, A.G. Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration. *Adv. Neural Inf. Process. Syst.* **2018**, *31*.
19. Raychaudhuri, S. Introduction to monte carlo simulation. In Proceedings of the 2008 Winter Simulation Conference, Miami, FL, USA, 7–10 December 2008; pp. 91–100.
20. Wurzinger, A.; Heidegger, P.; Kraxberger, F.; Mayr-Mittermüller, B.; Sima, H.; Kaltenbacher, M.; Schoder, S. Prediction of vibro-acoustic sound emissions based on mapped structural dynamics. In Proceedings of the INTER-NOISE and NOISE-CON Congress and Conference Proceedings, Nantes, France, 25–29 August 2024; Volume 270, pp. 8251–8260. [[CrossRef](#)]
21. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
22. Bilonis, I.; Zabarar, N. Multi-Output Local Gaussian Process Regression: Applications to Uncertainty Quantification. *J. Comput. Phys.* **2012**, *231*, 5718–5746. [[CrossRef](#)]
23. Mones, L.; Bernstein, N.; Csányi, G. Exploration, Sampling, And Reconstruction of Free Energy Surfaces with Gaussian Process Regression. *J. Chem. Theory Comput.* **2016**, *12*, 5100–5110. [[CrossRef](#)]
24. Titsias, M. Variational learning of inducing variables in sparse Gaussian processes. In Proceedings of the Artificial Intelligence and Statistics, Clearwater Beach, FL, USA, 16–18 April 2009; pp. 567–574.
25. Maddox, W.J.; Stanton, S.; Wilson, A.G. Conditioning Sparse Variational Gaussian Processes for Online Decision-making. In *Proceedings of the Advances in Neural Information Processing Systems*; Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., Vaughan, J.W., Eds.; Curran Associates, Inc.: New York, NY, USA, 2021; Volume 34, pp. 6365–6379.
26. Bauer, M.; Van der Wilk, M.; Rasmussen, C.E. Understanding probabilistic sparse Gaussian process approximations. *Adv. Neural Inf. Process. Syst.* **2016**, *29*.
27. Williams, C.; Rasmussen, C. Gaussian processes for regression. *Adv. Neural Inf. Process. Syst.* **1995**, *8*.
28. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*, 3rd ed.; Adaptive Computation and Machine Learning; MIT Press: Cambridge, MA, USA, 2008.
29. Fiala, P.; Rucz, P. NiHu: An Open Source C++ BEM Library. *Adv. Eng. Softw.* **2014**, *75*, 101–112. [[CrossRef](#)]
30. Burton, A.J.; Miller, G.F. The Application of Integral Equation Methods to the Numerical Solution of Some Exterior Boundary-Value Problems. *Proc. R. Soc. Lond.* **1971**, *323*, 201–210. [[CrossRef](#)]
31. Luegmair, M.; Münch, H. Advanced Equivalent Radiated Power (ERP) Calculation for Early Vibro-acoustic Product Optimization. In Proceedings of the The 22nd International Congress on Sound and Vibration, Florence, Italy, 12–16 July 2015.
32. Fritze, D.; Marburg, S.; Hardtke, H.J. Estimation of radiated sound power: A case study on common approximation methods. *Acta Acust. United Acust.* **2009**, *95*, 833–842. [[CrossRef](#)]

33. Cremer, L.; Heckl, M.; Petersson, B. Sound Radiation from Structures. In *Structure-Borne Sound: Structural Vibrations and Sound Radiation at Audio Frequencies*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 449–535. [\[CrossRef\]](#)
34. Kvist, K.; Sorokin, S.V.; Larsen, J.B. Radiation Efficiency Varying Equivalent Radiated Power. *J. Acoust. Soc. Am.* **2025**, *157*, 169–177. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Das, A.; Geisler, W.S. Methods to integrate multinormals and compute classification measures. *arXiv* **2020**, arXiv:2012.14331.
36. Schoder, S.; Roppert, K. openCFS: Open Source Finite Element Software for Coupled Field Simulation—Part Acoustics. *arXiv* **2022**, arXiv:2207.04443. [\[CrossRef\]](#)
37. Langer, P.; Maeder, M.; Guist, C.; Krause, M.; Marburg, S. More Than Six Elements Per Wavelength: The Practical Use of Structural Finite Element Models and Their Accuracy in Comparison with Experimental Results. *J. Comput. Acoust.* **2017**, *25*, 1750025. [\[CrossRef\]](#)
38. Schoder, S.; Roppert, K.; Weitz, M.; Junger, C.; Kaltenbacher, M. Aeroacoustic source term computation based on radial basis functions. *Int. J. Numer. Methods Eng.* **2020**, *121*, 2051–2067. [\[CrossRef\]](#)
39. Hensman, J.; Matthews, A.; Ghahramani, Z. Scalable Variational Gaussian Process Classification. *arXiv* **2014**, arXiv:1411.2005. <http://arxiv.org/abs/1411.2005>. [\[CrossRef\]](#)
40. Blei, D.M.; Kucukelbir, A.; McAuliffe, J.D. Variational Inference: A Review for Statisticians. *J. Am. Stat. Assoc.* **2017**, *112*, 859–877. [\[CrossRef\]](#)
41. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [\[CrossRef\]](#)
42. Snelson, E.; Ghahramani, Z. Sparse Gaussian processes using pseudo-inputs. *Adv. Neural Inf. Process. Syst.* **2005**, *18*.
43. Hensman, J.; Fusi, N.; Lawrence, N.D. Gaussian Processes for Big Data. *arXiv* **2013**, arXiv:1309.6835. <http://arxiv.org/abs/1309.6835>. [\[CrossRef\]](#)
44. Dai, Z.; Damianou, A.; Hensman, J.; Lawrence, N. Gaussian Process Models with Parallelization and GPU acceleration. *arXiv* **2014**, arXiv:1410.4984. <http://arxiv.org/abs/1410.4984>. [\[CrossRef\]](#)
45. Gal, Y.; van der Wilk, M.; Rasmussen, C.E. Distributed Variational Inference in Sparse Gaussian Process Regression and Latent Variable Models. In *Proceedings of the Advances in Neural Information Processing Systems*; Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., Weinberger, K., Eds.; Curran Associates, Inc.: New York, NY, USA, 2014; Volume 27.
46. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems 32*; Curran Associates, Inc.: New York, NY, USA, 2019; pp. 8024–8035.
47. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
48. Krause, A.; Singh, A.; Guestrin, C. Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. *J. Mach. Learn. Res.* **2008**, *9*, 235–284.
49. Jones, D.R.; Schonlau, M.; Welch, W.J. Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* **1998**, *13*, 455–492. [\[CrossRef\]](#)
50. Lázaro-Gredilla, M.; Titsias, M.K. Variational Heteroscedastic Gaussian Process Regression. In *Proceedings of the ICML*, Washington, DC, USA, 28 June–2 July 2011; pp. 841–848.
51. Ghosh, S.; Kristensen, J.; Zhang, Y.; Subber, W.; Wang, L. *A Strategy for Adaptive Sampling of Multi-Fidelity Gaussian Processes to Reduce Predictive Uncertainty*; American Society of Mechanical Engineers: New York, NY, USA, 2019. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.