

Analytical and Numerical Investigations of Probabilistic Monochromatic Problem

Stefan Schmuck^{1*}, Jakob Svensson²

¹Istituto di Fisica del Plasma, CNR, Milan, Italy

²Max-Planck-Institut für Plasmaphysik Teilinstitut Greifswald, Greifswald, Germany

Email: *schmuck@ifp.cnr.it

How to cite this paper: Schmuck, S. and Svensson, J. (2019) Analytical and Numerical Investigations of Probabilistic Monochromatic Problem. *Journal of Applied Mathematics and Physics*, 7, 793-808.

<https://doi.org/10.4236/jamp.2019.74054>

Received: February 28, 2019

Accepted: April 13, 2019

Published: April 16, 2019

Copyright © 2019 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

A probabilistic formalism, relying on Bayes' theorem and linear Gaussian inversion, is adapted, so that a monochromatic problem can be investigated. The formalism enables an objective test in probabilistic terms of the quantities and model concepts involved in the problem at hand. With this formalism, an amplitude (linear parameter), a frequency (non-linear parameter) and a hyperparameter of the Gaussian amplitude prior are inferred jointly given simulated data sets with Gaussian noise contributions. For the amplitude, an analytical normal posterior follows which is conditional on the frequency and the hyperparameter. The remaining posterior estimates the frequency with an uncertainty of MHz, while the convolution of a standard approach would achieve an uncertainty of some GHz. This improvement in the estimation is investigated analytically and numerically, revealing for instance the positive effect of a high signal-to-noise ratio and/or a large number of data points. As a fixed choice of the hyperparameter imposes certain results on the amplitude and frequency, this parameter is estimated and, thus, tested for plausibility as well. From abstract point of view, the model posterior is investigated as well.

Keywords

Bayes' Theorem, Linear Inversion, Uncertainty, Monochromatic, Spectral Limitation

1. Introduction

Fourier transformation tools are used to obtain information about spectra for a given data set. As any data has an uncertainty, Fourier transformation techniques can be supported by probabilistic theory captured by Bayes' theorem [1] to improve scientific results and conclusions. The works [2] [3] demonstrate some

advantages of the probabilistic ansatz over a conventional approach.

A common and misleading assumption in the field is that the Nyquist theorem determines the spectral band limitation. However, as the Nyquist frequency follows from finite data sampling, it can only be an upper limit but cannot be an estimator for the limit caused actually by a source and/or diagnostic throughput. A similar reasoning applies to the lower spectral limit. After a basic formalism has been derived, it could be shown in Ref. [4] that the band limitation can be well inferred from experimental data. Thereby, a linear Gaussian inversion technique was used to infer spectral amplitudes. Furthermore, a settings posterior was introduced which estimates non-linear parameters and hyperparameters of a problem. For instance, the spectral band limits and the uncertainty on the derived quasi-continuous spectrum, originating in non-probed Fourier coefficients, have been inferred jointly.

For many applied analysis schemes, certain model assumptions and their implications are not tested but assumed to be valid with infinite trust. From scientific point of view, any analysis scheme should be tested given simulated noisy data for which all model assumptions are clear. Then, analysis results and model assumptions can be investigated which is achieved objectively by a probabilistic ansatz. If this can be carried out analytically, valuable information is available when only actual measured data are given for a scientific problem.

The problem of a monochromatic source is a good example to show the powerfulness of Bayesian inference and to test model assumptions. In this work, the formalism derived in Ref. [4] is adapted to a monochromatic problem in Section 2 and applied to simulated data with different noise levels in Section 3. In addition, analytical and numerical investigations are carried out, so that the probabilistic findings can be understood in more detail. For example, a better signal-to-noise ratio improves the estimation of the frequency, while more data points compensate a worsened signal-to-noise ratio. The conclusion section can be found at the end.

2. Adapted Probabilistic Formalism

The formalism derived in Ref. [4] is adapted below to a monochromatic problem, involving one frequency parameter f_1 and one amplitude parameter S_1 . In addition, a hyperparameter is at hand, entering in the normal prior for S_1 . Abstractly, Bayes' theorem reads then

$$p(S_1, f_1, \sigma_{S_1, Pr} | \mathbf{D}) = \frac{p(\mathbf{D} | S_1, f_1)}{p(\mathbf{D})} p(S_1 | \sigma_{S_1, Pr}) p(\sigma_{S_1, Pr}) p(f_1) \quad (1)$$

with the joint posterior on the left-hand, the likelihood $p(\mathbf{D} | S_1, f_1)$ divided by the evidence, the amplitude prior $p(S_1 | \sigma_{S_1, Pr})$, the prior for the hyperparameter, and the prior for the frequency $p(f_1)$.

The amplitude S_1 maps linearly to the data domain via the vector $\mathbf{M}_1 = \mathbf{M}(f_1)$ dependent on the frequency in a non-linear manner. The data \mathbf{D}

with N_D entries are assumed to be acquired independently, and the measurement uncertainty of each data point follows a normal distribution with standard deviation σ_D . With these assumptions, the Gaussian likelihood becomes

$$p(\mathbf{D} | S_1, f_1) = \frac{\exp\left[-\frac{1}{2}(\mathbf{D} - \mathbf{M}_1 S_1)^\top \boldsymbol{\Sigma}_D^{-1} (\mathbf{D} - \mathbf{M}_1 S_1)\right]}{(2\pi)^{N_D/2} |\boldsymbol{\Sigma}_D|^{1/2}} \quad (2)$$

with the covariance matrix $\boldsymbol{\Sigma}_D = \sigma_D^2 \delta_{ij}$. The amplitude prior takes the form

$$p(S_1 | \sigma_{S_1, Pr}) = \frac{\exp\left[-\frac{1}{2} S_1^2 / \sigma_{S_1, Pr}^2\right]}{(2\pi)^{1/2} \sigma_{S_1, Pr}} \quad (3)$$

for vanishing prior mean and variance $\sigma_{S_1, Pr}^2$. After some analytical operations, one obtains the Gaussian amplitude posterior

$$p(S_1 | f_1, \sigma_{S_1, Pr}, \mathbf{D}) = \frac{\exp\left[-\frac{1}{2} (S_1 - S_{1, Po}) \sigma_{S_1, Po}^{-2} (S_1 - S_{1, Po})\right]}{(2\pi)^{1/2} \sigma_{S_1, Po}} \quad (4)$$

conditional on f_1 and $\sigma_{S_1, Pr}$. In the above equation, the posterior variance is given by $\sigma_{S_1, Po}^2 = (\mathbf{M}_1^\top \mathbf{M}_1 \sigma_D^{-2} + \sigma_{S_1, Pr}^{-2})^{-1}$, and the posterior mean reads $S_{1, Po} = \sigma_{S_1, Po}^2 \mathbf{M}_1^\top \mathbf{D} / \sigma_D^2$.

From the remaining terms, the so-called settings posterior

$$p(f_1, \sigma_{S_1, Pr} | \mathbf{D}) = \frac{1}{K'} \frac{\sigma_{S_1, Po}}{\sigma_{S_1, Pr}} \exp\left[\frac{1}{2} S_{1, Po}^2 / \sigma_{S_1, Po}^2\right] p(\sigma_{S_1, Pr}) p(f_1) \quad (5)$$

is derived which can be interpreted as Ockham's razor with respect to S_1 . This settings posterior has no general analytical solution, and, thus, the normalisation constant K' cannot be stated further. For the remainder of the paper, $p(\sigma_{S_1, Pr}) = 1/\Delta_\sigma$ and $p(f_1) = 1/\Delta_f$ are taken as uniform distributions. By this choice, $K' = K p(\sigma_{S_1, Pr}) p(f_1)$ can be used to investigate analytically and numerically the quantity

$$p_{\times K} = \frac{\sigma_{S_1, Po}}{\sigma_{S_1, Pr}} \exp\left[\frac{1}{2} S_{1, Po}^2 / \sigma_{S_1, Po}^2\right]. \quad (6)$$

Finally, the evidence can be identified as

$$p(\mathbf{D}) = \frac{K}{\Delta_\sigma \Delta_f} \frac{\exp\left[-\frac{1}{2} \mathbf{D}^\top \boldsymbol{\Sigma}_D^{-1} \mathbf{D}\right]}{(2\pi)^{N_D/2} |\boldsymbol{\Sigma}_D|^{1/2}}. \quad (7)$$

The constant $K/(\Delta_\sigma \Delta_f)$ varies with the chosen model H , including likelihood, priors, and the prior domains for f_1 and $\sigma_{S_1, Pr}$. Hence, this constant is linked to the model posterior $p(H | \mathbf{D})$ which is of importance, when the model is further abstracted, or a comparison with an alternative model is carried out.

3. Monochromatic Problem

For a monochromatic even source, data sets with different noise levels are simulated, and the formalism derived in the previous section is applied. To explain the results found, analytical and numerical investigations are carried out. While this problem is treated abstractly in the following, the main results will be presented for two examples with low and high noise contribution to ease the presentation.

(A) Simulated Data: Two Examples

Data sets are modelled for the real-world interferometer found in Ref. [5] which achieves an optical path difference to obtain an interferogram. Formally, the interferogram $V_i = 2S \cos(2\pi f x_i / c)$ is simulated with the amplitude $S = 0.5$ and the frequency $f = 50$ GHz. The optical path difference grid x_i is chosen as $x_i = (i - 129)\Delta x$ with $i = 1, \dots, N_D = 789$ with constant increment $\Delta x = 40 \mu\text{m}$. Accordingly, the total length $L = 31.56$ mm and the length $L_{SS} = 21.28$ mm of the single-sided region follow. Two noisy data sets $D_i = V_i + \mathcal{N}_i(0, \sigma_D^2)$ are generated with the noise levels σ_D at 0.05 and 1.0 (see Figure 1).

(B) Application of Formalism

The linear mapping of the amplitude parameter is given by $M_{1,i} = 2 \cos(2\pi f_1 x_i / c)$, yielding M_1 . For the two simulated data sets, Figure 2 shows $p(S_1 | f_1, \sigma_{S_1, Pr}, \mathbf{D})$, varying in f_1 and for three values of $\sigma_{S_1, Pr}$. Even for the high noise level case, $S_{1, Po}$ peaks close to 50 GHz, and the uncertainty $\sigma_{S_1, Po}$ remains reasonably small. The peak has the full width at half maximum (FWHM) of about 8 GHz which is roughly twice as large as the classical spectral resolution stated by $c/(2L) \approx 4.75$ GHz. High and low values of $\sigma_{S_1, Pr}$ seem to have little effect on the results.

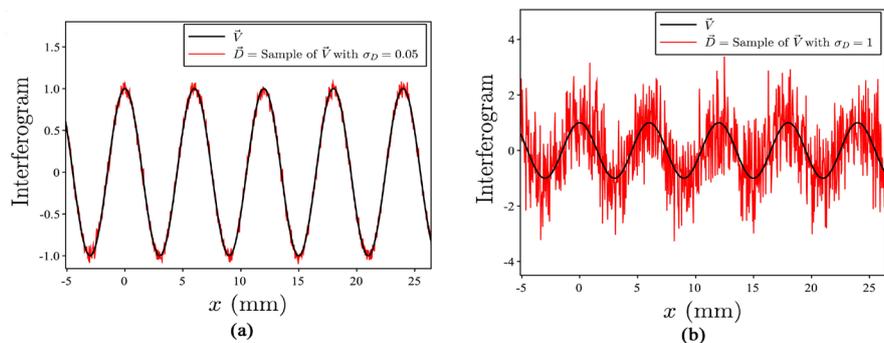


Figure 1. Simulated interferograms V (without noise, black) and D (with noise, red). Formally, $V = 2S \cos(2\pi f x / c)$ was used with amplitude $2S = 1$, frequency $f = 50$ GHz, and optical path difference grid x with increment $\Delta x = 40 \mu\text{m}$ (number of data points $N_D = 789$). The data-like quantity $D = V + \mathcal{N}(0, \sigma_D^2)$ resembles a sample of V , assuming a Gaussian noise component with variance σ_D^2 . Furthermore, the noise is chosen to be independent of x . (a) $\sigma_D = 0.05$ (low noise); (b) $\sigma_D = 1.0$ (high noise).

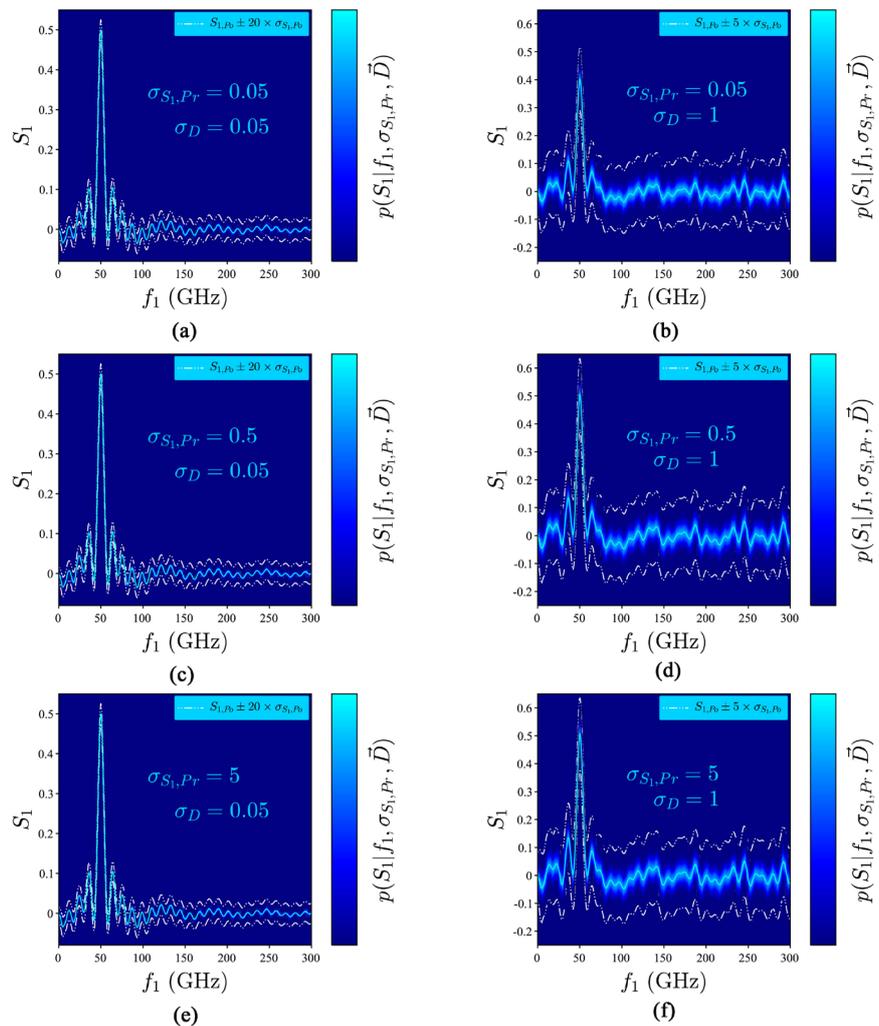


Figure 2. Conditional Gaussian amplitude posterior $p(S_1|f_1, \sigma_{S_1,Pr}, \mathbf{D})$ with mean $S_{1,Pr}$ and standard deviation $\sigma_{S_1,Pr}$. This posterior depends on the frequency f_1 , prior standard deviation $\sigma_{S_1,Pr}$ for S_1 , and noisy interferogram data \mathbf{D} . The conditional posterior is obtained for two sets \mathbf{D} (see **Figure 1**) modelled by a monochromatic source with the frequency $f = 50$ GHz, the amplitude $S = 0.5$ and the noise levels $\sigma_D = 0.05$ (left column) and 1 (right column). The posterior mean (cyan) peaks at values of about 0.5 close to 50 GHz. The spectral FWHM of this peak accounts for about 8 GHz. The posterior standard deviation (dashed-white for its multiples) increases with the noise level, as one would expect. Only small values of $\sigma_{S_1,Pr}$ affect $S_{1,Pr}$ when the noise level is high.

The settings posterior $p(f_1, \sigma_{S_1,Pr} | \mathbf{D})$ is proportional to the quantity $p_{\times K}$ which takes very large values even on logarithmic scale and, thus, $\ln p_{\times K}$ is used in the following. For the two cases, $\ln p_{\times K}$ is shown in **Figure 3(a)** and **Figure 3(b)** versus the spectral parameter and for three prior values of $\sigma_{S_1,Pr}$. All three cross-sections look similar, but close to 50 GHz the peak is highest for values near $\sigma_{S_1,Pr} = 0.5$. Identifying this global maximum by $\max(p_{\times K})$, the ratio $p_{\times K} / \max(p_{\times K})$ takes reasonable numbers and is presented for the

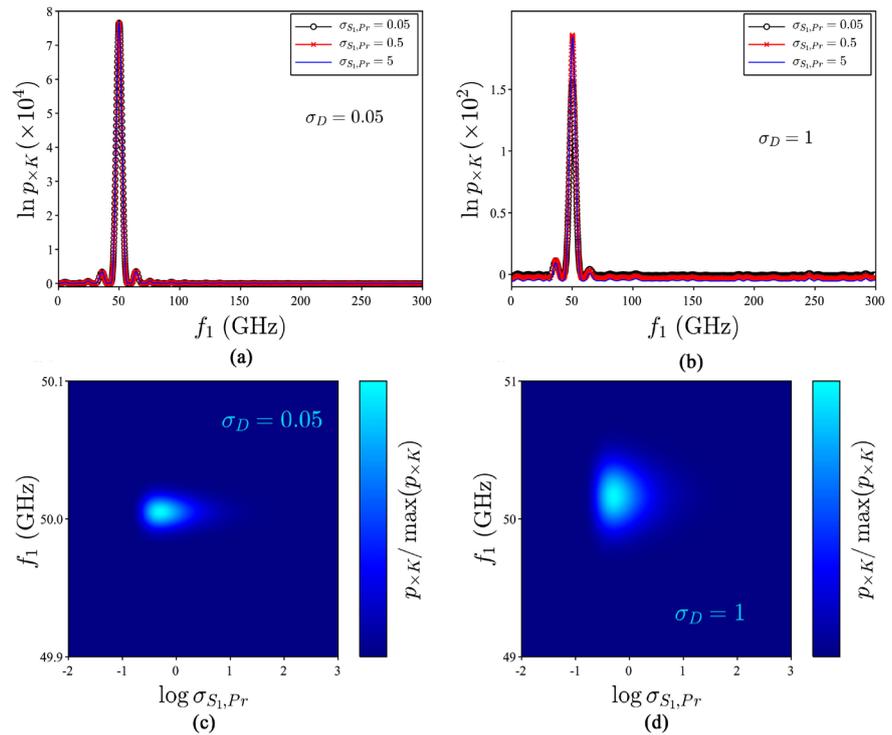


Figure 3. $p_{\times K}$, being proportional to joint posterior $p(f_1, \sigma_{S_1,Pr} | \mathbf{D})$, dependent on frequency f_1 and prior standard deviation $\sigma_{S_1,Pr}$ given interferogram data set \mathbf{D} (see **Figure 1**) with low and high noise levels $\sigma_D = 0.05$ (left column) and 1 (right column), respectively. Since $p_{\times K}$ takes large values, $\ln p_{\times K}$ is presented in (a) and (b). Deviding by the global maximum, the ratio $p_{\times K}/\max(p_{\times K})$ becomes manageable from numerical point of view as shown in (c) and (d). This ratio peaks in the vicinity of 50 GHz and $\sigma_{S_1,Pr} = 0.5$ ($\log \sigma_{S_1,Pr} = -0.3$). The spectral width of $p_{\times K}/\max(p_{\times K})$ increases with the noise level from about 10 to 200 MHz but remains well below the classical spectral resolution of about 4.75 GHz. The width with respect to $\sigma_{S_1,Pr}$ remains constant and covers one order of magnitude. (a) $\ln p_{\times K}$ (low noise case) for broad spectral domain at three values of $\sigma_{S_1,Pr}$; (b) $\ln p_{\times K}$ (high noise case) for broad spectral domain at three values of $\sigma_{S_1,Pr}$; (c) $p_{\times K}/\max(p_{\times K})$ (low noise case) for spectral domain 49.9 - 50.1 GHz; (d) $p_{\times K}/\max(p_{\times K})$ (high noise case) for spectral domain 49 - 51 GHz.

relevant domains of f_1 and $\sigma_{S_1,Pr}$ by **Figure 3(c)** and **Figure 3(d)** for both data sets. This ratio is narrow in the vicinity of 50 GHz and in the domain $0.3 \leq \sigma_{S_1,Pr} \leq 3$. Furthermore, the peaking has a spectral width of about 10 MHz and 200 MHz for the low and high noise case, respectively. When interpreted as spectral resolution, this width is smaller by several orders of magnitude with respect to its classical counterpart of about 4.75 GHz.

(C) Analytical and Numerical Investigations

To have further inside in the results, the conditional amplitude posterior and $p_{\times K}$ are analytically investigated. In order to do so, the noise contribution in the data is neglected, meaning that D_i is replaced with the noise-less interferogram $V_i = 2S \cos(2\pi f x_i/c)$. This allows the tracking of the dependencies of the

posteriors on the original amplitude S and frequency f noise level σ_D , and number of data points N_D . Furthermore, a second factorisation $p(f_1, \sigma_{S_1, Pr} | \mathbf{D}) = p(f_1 | \sigma_{S_1, Pr}, \mathbf{D}) p(\sigma_{S_1, Pr} | \mathbf{D})$ can be justified. Parallel to this investigation, numerics and conditional posteriors are presented.

1) Conditional Amplitude Posterior

Starting point is the variance $\sigma_{S_1, Po}^2$ of the conditional posterior $p(S_1 | f_1, \sigma_{S_1, Pr}, \mathbf{D})$ (see Equation (4)). Formally, one finds for the inverse

$$\sigma_{S_1, Po}^{-2} = \frac{\mathbf{M}_1^T \mathbf{M}_1}{\sigma_D^2} + \frac{1}{\sigma_{S_1, Pr}^2}. \tag{8}$$

As long as the spatial sampling is well enough, one can use the approximation

$$\begin{aligned} \mathbf{M}_1^T \mathbf{M}_1 &= 4 \sum_{i=1}^{N_D} \cos\left(2\pi \frac{f_1}{c} x_i\right) \cos\left(2\pi \frac{f_1}{c} x_i\right) \\ &\approx 2N_D \left[1 + \cos\left(2\pi \frac{f_1}{c} L_{SS}\right) \text{sinc}\left(2\pi \frac{f_1}{c} L\right) \right], \end{aligned} \tag{9}$$

as shown by Equation (A1), to get the posterior variance

$$\sigma_{S_1, Po}^2 = \frac{\sigma_D^2 / (2N_D)}{1 + \cos\left(2\pi \frac{f_1}{c} L_{SS}\right) \text{sinc}\left(2\pi \frac{f_1}{c} L\right) + \frac{\sigma_D^2}{2N_D \sigma_{S_1, Pr}^2}}. \tag{10}$$

This variance increases quadratically with the noise level and reduces with the number of data samples. Thus, more data points per spatial unit can compensate the noise contribution, at least partly. However, this assumes that the noise is independent of the data sampling.

The modulated sinc function in the denominator of Equation (10) depends on f_1 , L and L_{SS} for which the interferogram is available. The sinc function becomes unity and vanishes for small and large frequencies, respectively. For instance, since both data examples share the same spatial domain, the sinc function is close to 0 above about 20 GHz. The amplitude prior influences $\sigma_{S_1, Po}^2$, when the ratio $\sigma_D^2 / (2N_D \sigma_{S_1, Pr}^2)$ becomes larger than 1 which is obtained for a large noise level, a small value of $\sigma_{S_1, Pr}$ and/or few data points.

For the posterior mean,

$$S_{1, Po}(f_1) = \frac{\sigma_{S_1, Po}^2}{\sigma_D^2} \mathbf{M}_1^T \mathbf{V} = \frac{2 \sum_{i=1}^{N_D} V_i \cos\left(2\pi \frac{f_1}{c} x_i\right) / (2N_D)}{1 + \cos\left(2\pi \frac{f_1}{c} L_{SS}\right) \text{sinc}\left(2\pi \frac{f_1}{c} L\right) + \frac{\sigma_D^2}{2N_D \sigma_{S_1, Pr}^2}} \tag{11}$$

follows, leaving aside the noise contribution but the uncertainty on the mean is still captured by $\sigma_{S_1, Po}$. With the approximation (A1) and the trigonometric identity (B1), one can resolve

$$S_{1, Po}(f_1) = S \frac{\cos\left(\pi \frac{f-f_1}{c} L_{SS}\right) \text{sinc}\left(\pi \frac{f-f_1}{c} L\right) + \cos\left(\pi \frac{f+f_1}{c} L_{SS}\right) \text{sinc}\left(\pi \frac{f+f_1}{c} L\right)}{1 + \cos\left(2\pi \frac{f_1}{c} L_{SS}\right) \text{sinc}\left(2\pi \frac{f_1}{c} L\right) + \frac{\sigma_D^2}{2N_D \sigma_{S_1, Pr}^2}}. \tag{12}$$

Indeed, one finds the original amplitude $S_{1,Po} = S$ for $f_1 = f$ and prior values with the condition $\sigma_D^2 / (2N_D) \ll \sigma_{S_1,Pr}^2$. Furthermore, due to

$\cos\left(\pi \frac{f-f_1}{c} L\right) \text{sinc}\left(\pi \frac{f-f_1}{c} L_{SS}\right)$ there is a peaking of $S_{1,Po}(f_1)$ in the vicinity of $f_1 \approx f$ with the FWHM determined by both trigonometric arguments. Away from f , the oscillating amplitude drops like $1/(f-f_1)$ due to the sinc function which can be seen in **Figure 2**.

2) Conditional Posterior for f_1

Now the information is available to investigate $p(f_1, \sigma_{S_1,Pr} | \mathbf{D})$. Neglecting the noise, the triple product in the exponent of the exponential in $p_{\times K}$ reads

$$\begin{aligned} \frac{1}{2} S_{1,Po} \sigma_{S_1,Po}^{-2} S_{1,Po} &\approx \frac{1}{2} \frac{\mathbf{M}_1^T \mathbf{V}}{\sigma_D^2} \sigma_{S_1,Po}^2 \frac{\mathbf{M}_1^T \mathbf{V}}{\sigma_D^2} \\ &= \frac{1}{2} \sigma_{S_1,Po}^2 \frac{4N_D^2 S^2}{\sigma_D^4} \left[\cos\left(\pi \frac{f-f_1}{c} L_{SS}\right) \text{sinc}\left(\pi \frac{f-f_1}{c} L\right) \right. \\ &\quad \left. + \cos\left(\pi \frac{f+f_1}{c} L_{SS}\right) \text{sinc}\left(\pi \frac{f+f_1}{c} L\right) \right]^2. \end{aligned} \tag{13}$$

The global maximum is close to $f_1 = f$, and, assuming a sufficiently large frequency f , one gets the approximation

$$\begin{aligned} &\left[\cos\left(\pi \frac{f-f_1}{c} L_{SS}\right) \text{sinc}\left(\pi \frac{f-f_1}{c} L\right) \right]^2 \\ &\approx \left[1 - \left(\pi \frac{f-f_1}{c} L_{SS}\right)^2 - \frac{1}{3} \left(\pi \frac{f-f_1}{c} L\right)^2 \right] \end{aligned} \tag{14}$$

when the Taylor series expansions $\text{sinc}(z) \approx 1 - z^2/3!$ and $\cos(z) \approx 1 - z^2/2$ are used. Then,

$$\frac{1}{2} S_{1,Po} \sigma_{S_1,Po}^{-2} S_{1,Po} \approx \frac{1}{2} \sigma_{S_1,Po}^2 \frac{4N_D^2 S^2}{\sigma_D^4} \left[1 - (f-f_1)^2 \left(\frac{\pi}{c}\right)^2 \left(L_{SS}^2 + \frac{1}{3}L^2\right) \right] \tag{15}$$

remains, and the term which is independent on $f-f_1$ takes very large numerical values and is treated in more detail in the next subsection. The remaining term can be rewritten by a quadratic exponent of a Gaussian with posterior mean $f_{1,Po} = f$ and square root of the variance

$$\begin{aligned} \sigma_{f_1,Po} &= \frac{1}{\sigma_{S_1,Po}} \frac{\sigma_D^2}{2N_D S} \frac{1}{\pi L} \left(\frac{3}{3L_{SS}^2/L^2 + 1} \right)^{1/2} \\ &\approx \frac{\sigma_D}{SN_D^{1/2}} \frac{1}{\pi L} \left(\frac{3}{2} \frac{1}{3L_{SS}^2/L^2 + 1} \right)^{1/2}, \end{aligned} \tag{16}$$

using Equation (10) for large frequencies and $\sigma_{S_1,Pr}^2 \gg \sigma_D^2 / (2N_D)$. This uncertainty increases with the noise but decreases with the square root of the number of data points per spatial unit, the signal level and the spatial domain covered. For the data set examples, one inserts $L = 31.56$ mm, $L_{SS} = 21.28$ mm, $N_D = 789$, $\sigma_D/S = 0.1$ and 2 to find the astonishing numbers $\sigma_{f_1,Po} = 8.6$

and 171.9 MHz, respectively. This explains the narrowness of $p(f_1, \sigma_{S_1, Pr} | \mathbf{D})$ with respect to f_1 as shown in **Figure 3(c)** and **Figure 3(d)**.

In fact, $\sigma_{S_1, Po}$ changes with f_1 on a GHz, but the peaking in f_1 is of the order of MHz. Hence, the factorisation

$p(f_1, \sigma_{S_1, Pr} | \mathbf{D}) \approx p(f_1 | \sigma_{S_1, Pr}, \mathbf{D}) p(\sigma_{S_1, Pr} | \mathbf{D})$ with the conditional Gaussian posterior

$$p(f_1 | \sigma_{S_1, Pr}, \mathbf{D}) = \frac{\exp\left[-\frac{1}{2} \frac{(f_1 - f_{1, Po})^2}{\sigma_{f_1, Po}^2}\right]}{(2\pi)^{1/2} \sigma_{f_1, Po}} \quad (17)$$

and

$$p(\sigma_{S_1, Pr} | \mathbf{D}) = \frac{\sigma_{S_1, Po}}{\sigma_{S_1, Pr}} \frac{\exp\left[-\frac{1}{2} \frac{\sigma_{S_1, Po}^2}{\sigma_D^4} \frac{4N_D^2 S^2}{\sigma_D^4}\right]}{K} (2\pi)^{1/2} \sigma_{f_1, Po} \quad (18)$$

is justified. Thereby, $\sigma_{S_1, Po}$ needs to be taken at $f_1 = f_{1, Po}$. The remaining probability $p(\sigma_{S_1, Pr} | \mathbf{D})$ is investigated in the next subsection.

Numerics

Because the variance $\sigma_{S_1, Po}^2$ changes in f_1 on GHz scale but $\exp\left[\frac{1}{2} S_{1, Po} \sigma_{S_1, Po}^{-2} S_{1, Po}\right]$ on MHz scale, only the exponential needs to be taken into account for the determination of $f_{1, Po}$ and $\sigma_{f_1, Po}$. Furthermore, one faces large numbers, and, hence, the exponent is dealt with directly.

At a given $\sigma_{S_1, Pr}$, the frequency at which the maximum of the exponent occurs is identified as the posterior mean $f_{1, Po}(\sigma_{S_1, Pr})$. The posterior standard deviation $\sigma_{f_1, Po}(\sigma_{S_1, Pr})$ is obtained when the ratio of the exponent to its maximum at $f_{1, Po}$ reads $\frac{1}{2} S_{1, Po} \sigma_{S_1, Po}^{-2} S_{1, Po} / \max\left(\frac{1}{2} S_{1, Po} \sigma_{S_1, Po}^{-2} S_{1, Po}\right) = -1/2$. For both data sets, the ratios are shown in **Figure 4(a)** and **Figure 4(b)** in the vicinity of the peak. While the uncertainty of $f_{1, Po}$ and, hence, the width of the Gaussian augment with the noise level (see also **Figure 5(a)** and **Figure 5(b)**) as described by Equation (16), f remains included in the $2\text{-}\sigma_{f_1, Po}$ band centred at $f_{1, Po}$. Furthermore, the choice of $\sigma_{S_1, Pr}$ affects $f_{1, Po}$ and $\sigma_{f_1, Po}$ only for small values (see **Figure 5(a)** and **Figure 5(b)**).

3) Posterior for $\sigma_{1, Pr}$

After the factorisation, one combines Equations (16) and (18) like

$$\begin{aligned} p(\sigma_{S_1, Pr} | \mathbf{D}) &= \frac{P'}{K} = \frac{\sigma_{S_1, Po}}{\sigma_{S_1, Pr}} \frac{\exp\left[-\frac{1}{2} \frac{\sigma_{S_1, Po}^2}{\sigma_D^4} \frac{4N_D^2 S^2}{\sigma_D^4}\right]}{K} (2\pi)^{1/2} \sigma_{f_1, Po} \quad (19) \\ &= \frac{P}{K} \frac{\sigma_D^2}{N_D S L} \sqrt{\frac{3}{2\pi} \frac{L^2}{3L_{SS}^2 + L^2}} \end{aligned}$$

with the important term

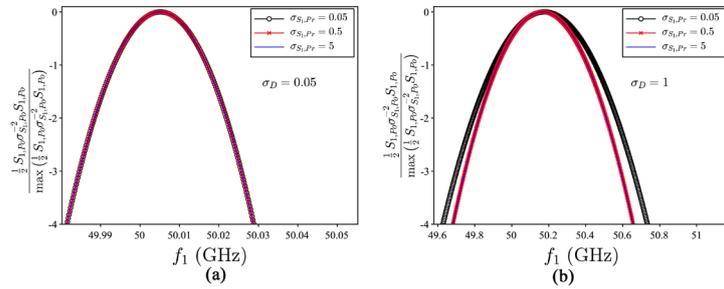


Figure 4. Ratio of exponent to its maximum dependent on frequency f_1 and keeping constant prior standard deviation $\sigma_{S_1,Pr}$. For the data set with low noise level ($\sigma_D = 0.05$, left), the peaking is more narrow than the one for the high noise level ($\sigma_D = 1$, right). The maximum location and the width of the ratio are used to estimate posterior mean $f_{1,Po}(\sigma_{S_1,Pr})$ and standard deviation $\sigma_{f_1,Po}(\sigma_{S_1,Pr})$. (a) Ratio for low noise data set at three values of $\sigma_{S_1,Pr}$. For $\sigma_{S_1,Pr} = 0.5$, $f_{1,Po} = 50.016$ GHz and $\sigma_{f_1,Po} = 8.5$ MHz are extracted which are consistent with 50 GHz used to simulate the data; (b) Ratio for high noise data set at three values of $\sigma_{S_1,Pr}$. For $\sigma_{S_1,Pr} = 0.5$, $f_{1,Po} = 50.167$ GHz and $\sigma_{f_1,Po} = 173$ MHz are estimated. $\sigma_{f_1,Po}$ and the deviation of $f_{1,Po}$ from 50 GHz increase with the noise level. Furthermore, small prior values for $\sigma_{S_1,Pr}$ affect $\sigma_{f_1,Po}$.

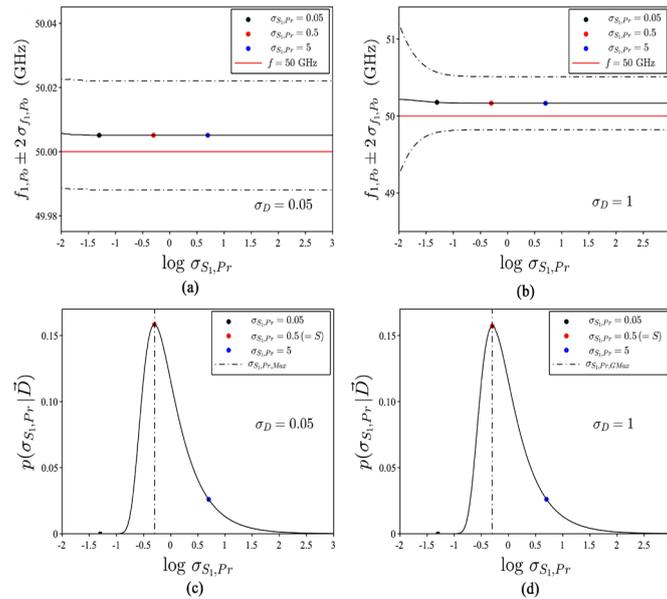


Figure 5. Posterior mean $f_{1,Po}$ and standard deviation $\sigma_{f_1,Po}$ for frequency f_1 dependent on prior standard deviation $\sigma_{S_1,Pr}$ for noise levels $\sigma_D = 0.05$ (left) and 1 (right). The original frequency $f = 50$ GHz (red line) is included usually in the band $f_{1,Po} \pm 2\sigma_{f_1,Po}$. However, $\sigma_{f_1,Po}$ increases from about 10 MHz to 200 MHz given both noise levels. The posterior $p(\sigma_{S_1,Pr} | \mathbf{D})$ (bottom row) changes little for the different noise levels and has a global maximum close to $S = 0.5$ (dashed-black). (a) $f_{1,Po} \pm 2\sigma_{f_1,Po}$ for $\sigma_D = 0.05$; (b) $f_{1,Po} \pm 2\sigma_{f_1,Po}$ for $\sigma_D = 1$; (c) $p(\sigma_{S_1,Pr} | \mathbf{D})$ for $\sigma_D = 0.05$; (d) $p(\sigma_{S_1,Pr} | \mathbf{D})$ for $\sigma_D = 1$.

$$P = \frac{\exp\left[\frac{1}{2}\sigma_{S_1,Pr}^2 \frac{4N_D^2 S^2}{\sigma_D^4}\right]}{\sigma_{S_1,Pr}}. \quad (20)$$

The exponent of P becomes

$$\frac{\sigma_{S_1,Pr}^2 \frac{N_D S^2}{\sigma_D^2}}{\sigma_{S_1,Pr}^2 \left[1 + \cos\left(2\pi \frac{f_{1,Pr}}{c} L_{SS}\right) \operatorname{sinc}\left(2\pi \frac{f_{1,Pr}}{c} L\right)\right] + \frac{\sigma_D^2}{2N_D}} \quad (21)$$

which is monotonically rising in $\sigma_{S_1,Pr}$ from 0 to about $N_D S^2 / \sigma_D^2$. In case $\sigma_{S_1,Pr}^2 > \sigma_D^2 / (2N_D)$ and $S/\sigma_D = 1$, the number of data points dominates. For the two examples with $N_D = 789$, large numbers $\exp(78900)$ and $\exp(197)$ are at hand as could be seen in **Figure 3(a)** and **Figure 3(b)**.

As the exponent vanishes for $\sigma_{S_1,Pr} = 0$, $1/\sigma_{S_1,Pr}$ and, thus, P and P' diverge. Hence, this global maximum should be excluded by a proper choice for the prior of $\sigma_{S_1,Pr}$. Doing so, P has the global maximum P_{GMax} at

$$\sigma_{S_1,Pr,GMax}^2 = \frac{S^2}{2} - \frac{\sigma_D^2}{2N_D} + \frac{S^2}{2} \left(1 - 2 \frac{\sigma_D^2}{N_D S^2}\right)^{1/2} \quad (22)$$

supposed a large $f_{1,Pr}$. Interestingly, N_D reduces the noise impact drastically, so that even for elevated noise levels $\sigma_D \approx S$, the maximum is still given by S with minor corrections.

The uniform prior $p(\sigma_{S_1,Pr})$ may be finite for the domain $\sigma_{S_1,Pr,L} \leq \sigma_{S_1,Pr,U}$, and the maximum of P in this domain is denoted by P_{Max} . Then, the peaking can be made more obvious by

$$\frac{P}{P_{Max}} = \frac{\sigma_{S_1,Pr,Max}}{\sigma_{S_1,Pr}} \exp\left[-\frac{1}{2} \frac{\sigma_{S_1,Pr,Max}^2 - \sigma_{S_1,Pr}^2}{\sigma_{S_1,Pr}^2 + \frac{\sigma_D^2}{2N_D}} \frac{S^2}{\sigma_{S_1,Pr,Max}^2 + \frac{\sigma_D^2}{2N_D}}\right] \quad (23)$$

after some algebra for a large $f_{1,Pr}$. With the normalisation of the above equation

$$\int_{\sigma_{S_1,Pr,L}}^{\sigma_{S_1,Pr,U}} \frac{P}{P_{Max}} d\sigma_{S_1,Pr} = K_{\sigma_{S_1,Pr}}(\sigma_{S_1,Pr,L}, \sigma_{S_1,Pr,U}), \quad (24)$$

the posterior becomes

$$p(\sigma_{S_1,Pr} | \mathbf{D}) = \frac{1}{K_{\sigma_{S_1,Pr}}} \frac{P}{P_{Max}}. \quad (25)$$

In case, the prior is finite for the domain $\sigma_{S_1,Pr,L} \leq \sigma_{S_1,Pr,GMax} \leq \sigma_{S_1,Pr,U}$, so that $P_{Max} = P_{GMax}$, one can set $\sigma_{S_1,Pr,Max}^2 = S^2 - \sigma_D^2 / (2N_D)$. This reveals the simple forms

$$P_{GMax} = \frac{\exp\left[\frac{N_D S^2}{\sigma_D^2} - \frac{1}{2}\right]}{\sqrt{S^2 - \frac{\sigma_D^2}{2N_D}}} \quad (26)$$

and

$$\frac{P}{P_{GMax}} = \frac{\sqrt{S^2 - \frac{\sigma_D^2}{2N_D}}}{\sigma_{S_1,Pr}} \exp \left[\frac{1}{2} \left(1 - \frac{S^2}{\sigma_{S_1,Pr}^2 + \frac{\sigma_D^2}{2N_D}} \right) \right] \tag{27}$$

of Equation (23). With the above expression and Equation (19), the constant

$$K = P'_{GMax} K_{\sigma_{S_1,Pr}} = \frac{\sigma_D^2}{N_D} \frac{c}{S L} \sqrt{\frac{3}{2\pi}} \frac{L^2}{3L_{SS}^2 + L^2} \frac{\exp \left[\frac{N_D S^2}{\sigma_D^2} - \frac{1}{2} \right]}{\sqrt{S^2 - \frac{\sigma_D^2}{2N_D}}} K_{\sigma_{S_1,Pr}} \tag{28}$$

follows with proper unit Hz. K depends in a complicated way on the noise level, the signal, the number of data points, the spatial domain for which the data is acquired, and the limits $\sigma_{S_1,Pr,L}$ and $\sigma_{S_1,Pr,U}$. For the two examples, one gets $\ln(P'_{GMax}/\text{Hz}) = 78910$ ($\sigma_D = 0.05$) and $\ln(P'_{GMax}/\text{Hz}) = 214$ ($\sigma_D = 1$).

Numerics

Starting point is Equation (19) which is rewritten like

$$\begin{aligned} \ln p(\sigma_{S_1,Pr} | \mathbf{D}) + \ln K &= \ln P' \\ &= \ln \left(\frac{\sigma_{S_1,Po}}{\sigma_{S_1,Pr}} \max \left(\exp \left[\frac{1}{2} S_{1,Po} \sigma_{S_1,Po}^{-2} S_{1,Po} \right] \right) (2\pi)^{1/2} \sigma_{f_1,Po} \right) \end{aligned} \tag{29}$$

to handle large numerical values. Thereby, $\sigma_{S_1,Po}$ is evaluated at $f_1 = f_{1,Po}$. The far right-hand side of the above equation is numerically available, and the maximum $\ln P'_{Max}$ at $\sigma_{S_1,Pr,Max}$ and the normalisation $K_{\sigma_{S_1,Pr}}$ can be determined for the domain from $\sigma_{S_1,Pr,L}$ to $\sigma_{S_1,Pr,U}$.

Choosing for the moment $\sigma_{S_1,Pr,L} = 10^{-2}$ and $\sigma_{S_1,Pr,U} = 10^3$, so that the global maximum P'_{GMax} is included, the above formalism yields $\ln(P'_{GMax}/\text{Hz}) = 76861$ (213) and $K_{\sigma_{S_1,Pr}} = 6.32$ (6.36) for the data set simulated with $\sigma_D = 0.05$ (1). For both examples, the order of magnitude of $\ln(P'_{GMax}/\text{Hz})$ is well described by Equation (26).

For the data sets, both P'_{GMax} are located at $\sigma_{S_1,Pr,GMax} = 0.499$ and 0.501 which is in the vicinity of S (see **Figure 5(c)** and **Figure 5(d)**) and is in agreement with expression (22). For the different noise levels, $p(\sigma_{S_1,Pr} | \mathbf{D})$ has a similar shape; the probability reduces drastically below $\sigma_{S_1,Pr,GMax}$ but remains finite above up to $\sigma_{S_1,Pr} = 10$, at least.

4) Model Posterior

The posterior of the model with uniform priors for σ_{S_1} and f_1 is given by

$$p(H | \mathbf{D}) = \frac{P'_{Max} K_{\sigma_{S_1,Pr}}}{(\sigma_{S_1,Pr,U} - \sigma_{S_1,Pr,L})(f_{1,Pr,U} - f_{1,Pr,L})} \tag{30}$$

This posterior weighs the constants obtained with respect to the spectral and amplitude domains. Hence, $p(H | \mathbf{D})$ increases with the constants in the numerator and smaller domains in the denominator. When one investigates

Figure 3(a) and **Figure 3(b)**, it is obvious that $p(H | \mathbf{D})$ increases, when $f_{1,Pr,L}$ and $f_{1,Pr,U}$ are chosen to include $f_{1,Po}$ and are separated by few GHz or even some 100s MHz. Similarly, choosing $\sigma_{S_1,Pr,L}$ and $\sigma_{S_1,Pr,U}$ to include the global maximum P'_{GMax} at $\sigma_{S_1,Pr,GMax}$ (see **Figure 5(c)** and **Figure 5(d)**) elevates $p(H | \mathbf{D})$. Furthermore, $p(H | \mathbf{D})$ has a global maximum when $\sigma_{S_1,Pr,L}$ and $\sigma_{S_1,Pr,U}$ approach $\sigma_{S_1,Pr,GMax} \approx S$ from either side. This is numerically confirmed by scanning $\sigma_{S_1,Pr,L}$ and $\sigma_{S_1,Pr,U}$ over several orders of magnitude and evaluating the model posterior for the two data sets (see **Figure 6(a)** and **Figure 6(b)**).

(D) Joint Posterior and Marginal Posteriors for S_1 and f_1

Because the model posterior peaks at $\sigma_{S_1,Pr,GMax}$, the marginal posterior for f_1 is simply given by the Gaussian $p(f_1 | \sigma_{S_1,Pr} = \sigma_{S_1,Pr,GMax}, \mathbf{D})$ with posterior mean $f_{1,Po}$ and standard deviation $\sigma_{f_1,Po}$ (see Section III C2 and **Figure 5(a)** and **Figure 5(b)**). Furthermore, the joint posterior

$$p(S_1, f_1 | \sigma_{S_1,Pr,GMax}, \mathbf{D}) = p(S_1 | f_1, \sigma_{S_1,Pr,GMax}, \mathbf{D}) p(f_1 | \sigma_{S_1,Pr,GMax}, \mathbf{D}) \quad (31)$$

can be evaluated and is shown for both data sets in **Figure 7(a)** and **Figure 7(b)**. This joint posterior peaks close to S and f used to simulate the data, and its width increases with the noise level. However, S and f are usually located inside the peak, and, hence, one can trust the results within the posterior uncertainties.

The marginal posterior for S_1 is obtained by the marginalisation

$$p(S_1 | \sigma_{S_1,Pr,GMax}, \mathbf{D}) = \int p(S_1, f_1 | \sigma_{S_1,Pr,GMax}, \mathbf{D}) df_1. \quad (32)$$

The above integration is performed numerically, and the results are shown in **Figure 7(c)** and **Figure 7(d)** for both data sets. Another way to achieve the marginalisation is to draw samples for f_1 from $p(f_1 | \sigma_{S_1,Pr,GMax}, \mathbf{D})$, and for each of those samples draw samples for S_1 from $p(S_1 | f_1, \sigma_{S_1,Pr,GMax}, \mathbf{D})$. Histograms of these posterior samples are given in **Figure 7(c)** and **Figure 7(d)**.

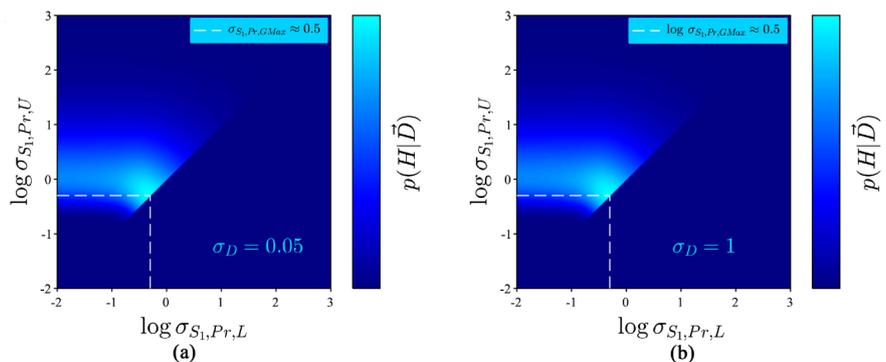


Figure 6. Model posterior $p(H | \mathbf{D})$. The stated model uses a uniform prior for the prior standard deviation $\sigma_{S_1,Pr}$ in the domain from $\sigma_{S_1,Pr,L}$ to $\sigma_{S_1,Pr,U}$. As these limits approach $\sigma_{S_1,Pr,GMax} \approx 0.499$ for $\sigma_D = 0.05$ and 0.501 for $\sigma_D = 1$ (dashed-white), the posterior increases. Thereby, $\sigma_{S_1,Pr,GMax} \approx \sqrt{S^2 - \sigma_D^2 / (2N_D)}$ is closely linked with the original amplitude $S = 0.5$. (a) Low noise case $\sigma_D = 0.05$; (b) High noise case $\sigma_D = 1$.

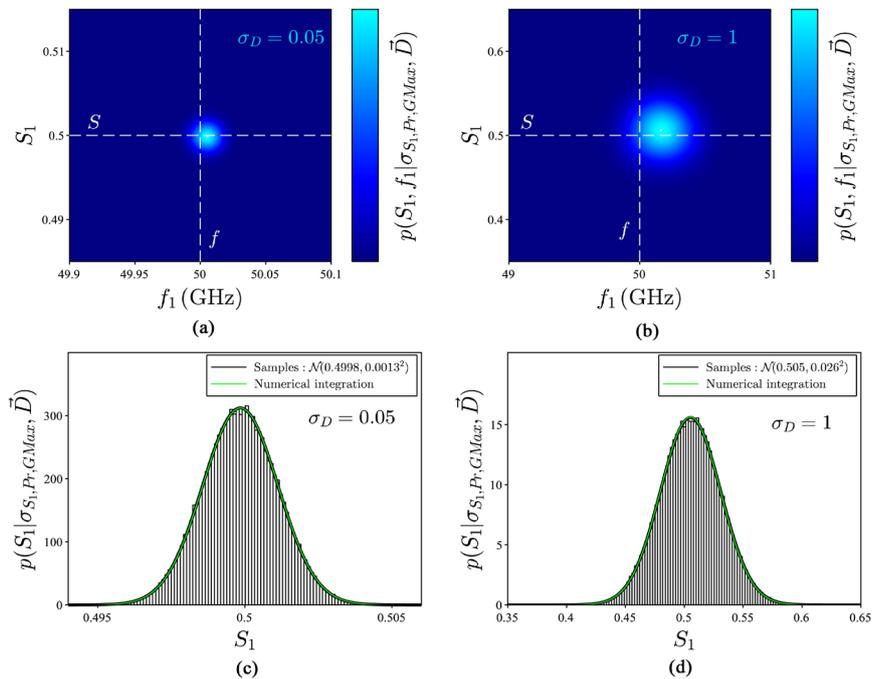


Figure 7. Joint posterior $p(S_1, f_1 | \sigma_{S_1, Pr, GMax}, \mathbf{D})$ and marginal posterior $p(S_1 | \sigma_{S_1, Pr, GMax}, \mathbf{D})$ of frequency f_1 and amplitude S_1 for low noise ($\sigma_D = 0.05$, left) and high noise data set ($\sigma_D = 1$, right). The joint posterior peaks close to the original values $S = 0.5$ and $f = 50$ GHz used to simulate the data sets. For the higher noise level, the posterior width increases, and the location of the global maximum separates further from S and f . By numerical integration, the marginal $p(S_1 | \sigma_{S_1, Pr, GMax}, \mathbf{D})$ is evaluated, and its shape is close to a normal distribution. This is confirmed by a second approach which uses 10^5 samples from the joint posterior to estimate the marginal posterior mean and standard deviation for S_1 . The histogram of the samples is given as well. (a) $p(S_1, f_1 | \sigma_{S_1, Pr, GMax}, \mathbf{D})$ for low noise level; (b) $p(S_1, f_1 | \sigma_{S_1, Pr, GMax}, \mathbf{D})$ for high noise level; (c) $p(S_1 | \sigma_{S_1, Pr, GMax}, \mathbf{D})$ for low noise level. The marginal mean and standard deviation read 0.4998 and 0.0013, respectively; (d) $p(S_1 | \sigma_{S_1, Pr, GMax}, \mathbf{D})$ for high noise level. One finds that the marginal mean and standard deviation account for 0.505 and 0.026, respectively.

Evaluating the mean and standard deviation from these samples approximates $p(S_1 | \sigma_{S_1, Pr, GMax}, \mathbf{D})$ by a Gaussian. In doing so for the data set with $\sigma_D = 0.05$ (1), the marginal posterior mean $S_{1, Po, Ma} = 0.4998$ (0.505) and standard deviation 0.0013 (0.026) are obtained. Both normal distributions are good approximations for $p(S_1 | \sigma_{S_1, Pr, GMax}, \mathbf{D})$ (see **Figure 7(c)** and **Figure 7(d)**).

4. Conclusions

The presented work investigates a monochromatic problem with an adapted version of Bayes' theorem to infer an amplitude parameter, a frequency parameter and a hyperparameter which acts on the amplitude prior only. The amplitude at a given frequency is a parameter which has a linear mapping to the

data domain. If the measurements have a normal noise contribution (Gaussian likelihood) and the amplitude prior is chosen to be Gaussian, an analytical linear inversion technique is enabled. The amplitude posterior, being conditional on the frequency and prior information, has similarities with the result of a conventional Fourier transformation analysis approach. However, the estimation of the frequency reveals a well localised domain which is in accordance with the measured data. The corresponding posterior for the frequency is orders of magnitude narrower than the width of the convolution function used by conventional approaches to estimate the spectral resolution. The prior information captured by one hyperparameter is tested as well, and its posterior peaks over a one order of magnitude and is very robust for significantly different noise levels on the data.

The findings for the monochromatic problem are investigated analytically. This reveals, for instance, the influence of the signal-to-noise ratio, the amount of data points and the spatial domains covered by the measurements on the results. The presented approach can be followed to examine more complex problems which involve more than one frequency and one amplitude and other diagnostic imperfections like a variable offset. Since this offset has a linear correspondence to the data domain, it would enter in the conditional amplitude posterior. Any non-linear parameter and additional hyperparameter would be estimated by the joint settings posterior specific for a certain problem. In doing so, a profound understanding of model implications could be established in analytical terms for applications relying on Fourier transformations. This is essential for designing a diagnostic for a given problem with somewhat known signal-to-noise ratio and hardware limitations, or for comparing results of different models in an objective way when a data set is given and, for example, the number of contributing frequencies is unknown.

Acknowledgements

This work has been carried out within the framework of the EURO fusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No. 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Sivia, D. and Skilling, J. (2006) *Data Analysis: A Bayesian Tutorial*. Oxford University Press, Oxford.
- [2] Jaynes, E.T. (1987) *Maximum Entropy and Bayesian Spectral Analysis and Estima-*

tion Problems. *Fundamental Theories of Physics*, **21**, 1-37.

- [3] Bretthorst, G.L. (1988). Bayesian Spectrum Analysis and Parameter Estimation. Springer-Verlag, Berlin, Heidelberg. <https://doi.org/10.1007/978-1-4684-9399-3>
- [4] Schmuck, S. and Svensson, J. (2017) Fourier Spectroscopy: A Bayesian Way. *International Journal of Spectroscopy*, 2017, Article ID: 9265084. <https://doi.org/10.1155/2017/9265084>
- [5] Schmuck, S., Fessey, J., Boom, J.E., Meneses, L., Abreu, P., Belonohy, E. and Lupelli, I. (2016) Electron Cyclotron Emission Spectra in X- and O-Mode Polarisation at JET: Martin-Puplett Interferometer, Absolute Calibration, Revised Uncertainties, Inboard/Outboard Temperature Profile, and Wall Properties. *Review of Scientific Instruments*, **87**, Article ID: 093506. <https://doi.org/10.1063/1.4962809>

Appendix A: Integral Approximation

The spatial grid x_i with $i \in [1, N]$ may be given by the set $x_i = i\Delta x + x_0$ of uniformly distributed locations separated by the constant increment Δx , so that $x_1 = -L_{DS}$ and $x_N = -L_{DS} + L$. Then, the sum

$$\begin{aligned} & \sum_{i=1}^N \cos\left(2\pi \frac{f_1}{c} x_i\right) \cos\left(2\pi \frac{f_2}{c} x_i\right) \\ & \approx \frac{1}{2\Delta x} \int_{-L_{DS}}^{-L_{DS}+L} \left[\cos\left(2\pi \frac{f_-}{c} x\right) + \cos\left(2\pi \frac{f_+}{c} x\right) \right] dx \\ & = \frac{1}{2\Delta x} \left[\frac{c}{2\pi f_-} \sin\left(2\pi \frac{f_-}{c} x\right) \right]_{-L_{DS}}^{-L_{DS}+L} + \frac{1}{2\Delta x} \left[\frac{c}{2\pi f_+} \sin\left(2\pi \frac{f_+}{c} x\right) \right]_{-L_{DS}}^{-L_{DS}+L} \quad (\text{A1}) \\ & = \frac{N}{2} \cos\left(\pi \frac{f_-}{c} L_{SS}\right) \text{sinc}\left(\pi \frac{f_-}{c} L\right) + \frac{N}{2} \cos\left(\pi \frac{f_+}{c} L_{SS}\right) \text{sinc}\left(\pi \frac{f_+}{c} L\right) \end{aligned}$$

is approximated by an integral over the spatial domain using the trigonometric identities (B1) and (B2). The approximation reveals two modulated sinc functions with a spatial and a spectral dependence. The spectral dependence is due to the sum and difference frequencies $f_- = f_1 - f_2$ and $f_+ = f_1 + f_2$, respectively. Spatially, the modulation depends on the length $L_{SS} = L - 2L_{DS}$ of single-sided domain which is present on side only. The spatial dependence of the sinc function is given by the length L of the total domain covered.

Appendix B: Trigonometric Identities

The trigonometric identities are

$$\cos(x)\cos(y) = \frac{1}{2} [\cos(x-y) + \cos(x+y)], \quad (\text{B1})$$

$$\sin(x) - \sin(y) = 2 \cos\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right). \quad (\text{B2})$$