Metrologia 53 (2016) S125–S135 doi:10.1088/0026-1394/53/4/S125

On challenges in the uncertainty evaluation for time-dependent measurements

S Eichstädt[1], V Wilkens[1], A Dienstfrey[2], P Hale[2], B Hughes[3] and C Jarvis[3]

[1] Physikalisch-Technische Bundesanstalt, Braunschweig and Berlin, Germany
[2] National Institute of Standards and Technology, Boulder, Colorado, USA
[3] National Physical Laboratory, Teddington, Middlesex TW11 0LW, UK

E-mail: sascha.eichstaedt@ptb.de

Received 1 October 2015, revised 24 February 2016
Accepted for publication 26 February 2016
Published 14 June 2016

Abstract

The measurement of quantities with time-dependent values is a common task in many areas of metrology. Although well established techniques are available for the analysis of such measurements, serious scientific challenges remain to be solved to enable their routine use in metrology. In this paper we focus on the challenge of estimating a time-dependent measurand when the relationship between the value of the measurand and the indication is modeled by a convolution. Mathematically, deconvolution is an ill-posed inverse problem, requiring regularization to stabilize the inversion in the presence of noise. We present and discuss deconvolution in three practical applications: thrust-balance, ultra-fast sampling oscilloscopes and hydrophones. Each case study takes a different approach to modeling the convolution process and regularizing its inversion. Critically, all three examples lack the assignment of an uncertainty to the influence of the regularization on the estimation accuracy. This is a grand challenge for dynamic metrology, for which to date no generic solution exists. The case studies presented here cover a wide range of time scales and prior knowledge about the measurand, and they can thus serve as starting points for future developments in metrology. The aim of this work is to present the case studies and demonstrate the challenges they pose for metrology.

Keywords: dynamic metrology, dynamic measurement, deconvolution, linear system, regularization, uncertainty

(Some figures may appear in colour only in the online journal)

1. Introduction

Measurements of quantities with time-dependent values are becoming increasingly relevant in all areas of metrology[4]. We define a *dynamic quantity* as any quantity whose value varies with time in such a way that this time dependence must be taken into account to characterize the quantity to the accuracy desired. Measurements in which at least one of the involved quantities is dynamic are called *dynamic measurements*. Dynamic measurements include a range of applications from single sensors to large networks, such as gas grids or electrical power grids; apply to a wide variety of physical quantities, such as pressure, force, or voltage; and cover timescales of variation from picoseconds to several minutes. For instance, measurements of mechanical quantities such as pressure, force, torque, acceleration and vibration in automotive and aerospace industries are usually taken under dynamic conditions [1, 2]. The same holds for measurements in the characterization of high-speed electronics [3, 4] or ultrasound devices [5, 6]. Despite the wide variety of applications, dynamic measurements share common mathematical structures in their measurement model. It follows that the challenges associated with their analysis are likewise overlapping. Nevertheless, current approaches for estimating measurands and their associated uncertainties differ significantly. A harmonized treatment with uniform guidance capable of cutting across application domains is needed.

[4] Official contribution of the National Institute of Standards and Technology; not subject to copyright in the United States.

The analysis of dynamic measurements is closely related to signal processing and system theory [7]. Thus, foundational elements of this theory are well-understood. The challenge for metrology lies in the adaptation of these methods, and more pointedly, in the development of statistical methods for the evaluation of measurement uncertainties [8]. In particular, the concept of uncertainty and its evaluation in signal processing differs from that in metrology outlined in the 'guide to the expression of uncertainty in measurement' (GUM) [9] and its supplements [10, 11]. For instance, in signal processing users are typically more interested in *robustified* methods than the evaluation of reliable statements of precision and accuracy, see, e.g. [12].

A common problem in the analysis of dynamic measurements is that the measurement model is ill-posed. It follows that the estimation of the desired quantity is unstable in the presence of noise. Thus, in order to obtain a reasonable estimate, one must introduce bias or prior knowledge into the estimation process, see sections 3 and 4. In mathematical and statistical analysis, the trade-off in which noise amplification is decreased at the expense of increased systematic error is known as *regularization* [13, 14]. The challenge for metrology lies in the corresponding assessment of the uncertainty contribution of this regularization. There are many approaches available in the literature that focus on carrying out regularization, but there is yet a significant lack of guidance regarding the evaluation of associated uncertainty.

In this article we focus on differing approaches to the estimation of the measurand. We consider three practical case studies: dynamic calibration of micro-thrusters, measurement of electrical pulses with dynamically calibrated sampling oscilloscopes and hydrophone measurements of pressure signals as generated by medical ultrasound devices. The three case studies cover a wide range of time scales and types of prior knowledge about the measurand. Thus, these examples can be considered as benchmark candidates for the development of a generic metrological treatment of regularized deconvolution. In one case study the regularization is treated implicitly, one utilizes explicit statistical approaches and in the third example the aim is to explicitly take advantage of prior knowledge about the measurand. The examples highlight that the treatment of regularization is strongly application dependent; an undesirable situation for generic guidance and the development of generic standard documents. In this regard, this paper represents an early step by gathering diverse applications to ensure that a suggested way forward is usable in a wide range of a metrological applications.

This paper is outlined as follows. Section 2 introduces the mathematical models considered and defines the estimation problem. In section 3 the three case studies are presented together with their individual approaches to the regularization problem. Section 4 discusses some existing approaches to accounting for the regularization bias, and it delineates a research framework to address regularization in an application-independent manner. Finally, section 5 discusses several further challenges in metrology for dynamic measurements.

2. Deconvolution—an introduction

The analysis of dynamic measurements is closely related to analysis of measurement problems in signal processing. As a consequence, the vocabulary in this area of metrology is that of system theory and signal processing. For instance, the measurement device is considered as *system* and the quantities with time-varying values as *signals*. For many dynamic measurement applications in metrology the system can be considered linear and time-invariant (LTI) in its working range. That is, the relationship between system $\mathcal{H}$, input $Y(t)$ and output $X(t)$ satisfy

$$\begin{aligned}\mathcal{H}(\alpha Y_1+\beta Y_2)&=\alpha\mathcal{H}(Y_1)+\beta\mathcal{H}(Y_2),\\ \mathcal{H}(Y(\bullet-\tau))&=X(\bullet-\tau)\end{aligned} \tag{1}$$

where α and β are scale factors and τ is an arbitrary time-shift. The first of the above equations represents linearity, and the second, time-invariance.

The system function $\mathcal{H}$ characterizes the dynamic behavior of the measurement system, and it can be given either as a transfer function in the Laplace domain

$$\boldsymbol{H}(s)=K_0\frac{\sum_{n=0}^{N_b}b_ns^n}{1+\sum_{n=1}^{N_a}a_ns^n}, \tag{2}$$

a complex valued frequency response function $H(f)$ in the frequency domain, or as an impulse response function $h(t)$ in the time domain[5]. For discrete-time systems, the z-transform is applied to the Laplace domain model (2) or an appropriate sampling applied to the frequency response or the impulse response function, respectively.

For LTI systems the mathematical relation between the measurand y and the (noise-free) measurement x is given by a *convolution*

$$x(t)=\int_{-\infty}^{\infty}h(t-s)y(s)\mathrm{d}s=:(h*y)(t). \tag{3}$$

For *causal* LTI systems the upper-limit of integration is replaced by t as, in this case, $h(t)=0$ for $t<0$. In any event, by the Convolution Theorem of Fourier analysis, the representation of (3) in the frequency domain is given by multiplication

$$X(f)=H(f)Y(f). \tag{4}$$

This holds true for continuous as well as for discrete time LTI systems.

An example of convolution is shown in figure 1. The figure shows an input waveform or measurand (blue) and the output (red). In this example, the output is given by convolution of the input with a proposed input response function (not shown). One observes that the relatively featureless impulse-like input is recorded as a damped, oscillatory waveform on the output of the measurement device. In engineering terms this is considered to be a consequence of the *finite bandwidth* of the measurement device.

[5] Note, throughout we use small letters for time domain functions and capital letters for frequency domain functions, in line with standard signal processing literature, see e.g. [7]. This is in contrast to the GUM, which uses capital letters for random variables and small letters for their corresponding realizations.

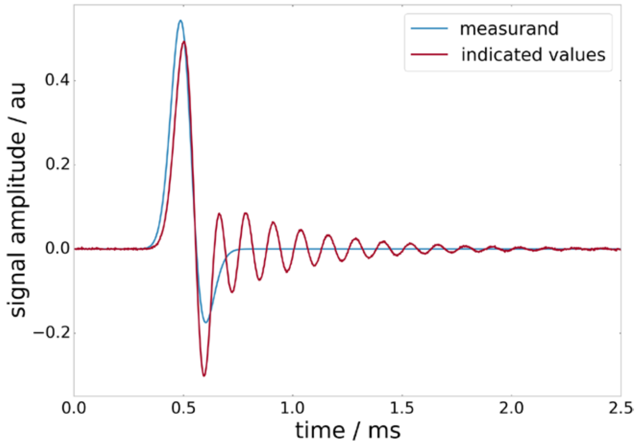

Figure 1. Simulated example for a typical dynamic measurand with time-dependent input and output. The time-dependent deviations of the output are owing to the non-ideal dynamic behaviour of the measurement system.

Estimation of the measurand requires deconvolution, which is well known to be an ill-posed inverse problem [13]. Analytically, this ill-posedness can be illustrated by considering the procedure in the frequency domain. Given $H(f)$ obtained by, for example, an independent calibration experiment, and the measured response $X(f)$, one solves (4) for $Y(f)$ by simple division

$$Y(f) = \frac{X(f)}{H(f)}. \quad (5)$$

The difficulty with (5) is that, for physical reasons, $|H(f)|$ decays to zero for large f[6]. By contrast, noise in the measurement entails that the spectrum of $X(f)$ is very broad. For example, in the case of additive white noise $|X(f)|$ tends to a constant for large f. The consequence is that the quotient (5) diverges at any frequency f for which $|H(f)| \approx 0$ and $|X(f)| > 0$. Furthermore and unfortunately, the above-mentioned physical considerations effectively guarantee that such f will exist in abundance. That is, measurement noise and numerical noise in the measured signal $x(t)$ are amplified strongly. In the language of statistical estimation theory, one has that while deconvolution (5) is the best linear unbiased estimate for the desired measurand, the variance is so large as to make $Y(f)$ unusable[7]. To rectify the situation, unbiased estimation has to be augmented in a reasonable way. In the literature this is known as regularization of the ill-posed inverse problem [14]. Considering the following case studies it is worth pointing out that, mathematically speaking, the role of y and h in (3) and (4) can be interchanged. For instance, in a calibration the system's impulse response h can be determined from measurements of x and y by a deconvolution.

[6] Physically, the impulse response function is continuous as a function of time; therefore, its Fourier transform vanishes for large $|f|$.
[7] As equation (3) is a linear operator, the best linear unbiased estimator is given by the pseudo-inverse applied to the measurement. This pseudo-inverse is equivalent to the division shown in (5) restricted to f such that $H(f) \neq 0$.

Table 1. Comparison of the estimation approaches for the three case studies.

Example	Frequency range	Deconvolution	Regularization
Thrust balance	≈ 5 Hz	Digital filter	Visual inspection
Oscilloscope	≈ 40 GHz	Linear estimation model	Data dependent
Hydrophone	≈ 40 MHz	Division in frequation domain	Prior knowledge

3. Case studies

We consider three practical case studies to illustrate application of the convolution model (3). The case studies take different approaches to carry out deconvolution for estimation of the measurand, see table 1. The first case study relies on designing a digital filter in the frequency domain; the second case study considers deconvolution as a linear estimation problem; whereas the third example aims at deconvolution in the frequency domain. Similarly, each case study takes a different approach for the regularization of the ill-posed inverse problem of deconvolution. For each case study we briefly introduce the technical background and present the deconvolution and regularization approach.

Despite the diverse background of the applications, there is the underlying generic estimation problem of determining a trade-off between reduced variance and increased systematic error. That is, in all case studies the ill-posed estimation problem requires some kind of regularization which in turn introduces a systematic error. The generic problem for the evaluation of uncertainty is thus the quantification of the uncertainty contribution of the systematic error introduced. An estimate of this error requires some prior knowledge about the measurand, which is still unusual for measurement analysis in metrology, see section 4.

3.1. Calibration of micro-Newton thrusters

So called micro-thrusters operate in the range from 0.1 μN up to 500 mN and are typically applied for spacecraft altitude control, drag compensation and precise flight control [1]. In the application, a command signal is issued to the thruster which then generates a corresponding force signal. For precise flight control, the actual generated force signal for a known command signal has to be determined in a calibration beforehand. The measurement principle for a traceable calibration of the forces generated by the thruster is based on a calibrated force actuator. This actuator aims at keeping the thruster in place on the thrust-balance assembly by generating forces itself. The force exerted by the thruster is the measurand and an unknown system input and the control signal issued to the force actuator is the system output, see figure 2. The control signal is modeled as a convolution of the thrust signal and the balance assembly response. A deconvolution then yields an estimate of the thruster force from the control signal.

Owing to the small forces generated by the thruster the calibration measurements are strongly affected by environmental

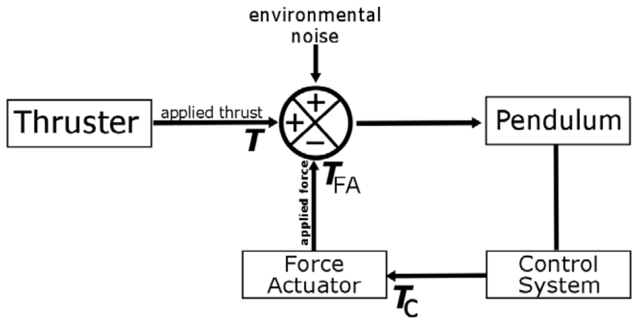

Figure 2. Working principle of the employed thrust balance. The force actuator acts as a compensation of the force generated by the thruster, and thereby provides an indication of the force generated.

noise. A reduction of measurement noise by simple low-pass filtering is not sufficient, because the required bandwidth of the filter would also deteriorate the measured signal. To this end, two parallel measurements are utilized as follows. The thrust-balance is equipped with two almost identical mechanical assemblies. The measurement balance assembly (MBA) carries the thruster and a force actuator, whereas the tilt compensation assembly (TCA) is equipped with a dummy and a force actuator. Thereby, the TCA is affected by the exact same environmental noise as the MBA.

The goal is then to utilize the measurement at the TCA to remove the environmental noise from the MBA measurements. However, both measurements are affected by the dynamic behavior of the respective devices. That is, assuming that both setups are sensing the identical environmental noise process $n(t)$, the measured output signal of the TCA setup is modeled as

$$x_{\mathrm{TCA}}(t) = \int_{-\infty}^{t} h_{\mathrm{TCA}}(t-s)n(s)\mathrm{d}s = (h_{\mathrm{TCA}} * n)(t). \quad (6)$$

Similarly, with $y(t)$ being the force generated by the thruster, the measured output at the MBA is given by

$$\begin{aligned} x_{\mathrm{MBA}}(t) &= \int_{-\infty}^{t} h_{\mathrm{MBA}}(t-s)(y(s)+n(s))\mathrm{d}s \\ &= (h_{\mathrm{MBA}} * (y+n))(t). \end{aligned} \quad (7)$$

For both assemblies the measured output is thus a convolution of the signal of interest and the device's impulse response. Although both, the TCA and the MBA, are mostly identically constructed, their dynamic behavior differs. That is, their impulse response and thus also their response to the noise process $n(t)$ are not identical. A utilization of the TCA for noise reduction in the MBA thus requires a deconvolution for both measurements beforehand.

The deconvolution approach utilized in this example is to design a digital filter

$$G_{\mathrm{inv}}(z) = \frac{\sum_{k=0}^{N_b} b_k z^{-k}}{1+\sum_{k=1}^{N_a} a_k z^{-k}}, \quad (8)$$

whose frequency response $G(\mathrm{e}^{j\omega})$ ideally equals the reciprocal of the system's frequency response $H(j\omega f_s)$ with f_s the sampling frequency. Therefore, the system's frequency response is determined using system identification, and the sought digital filter is then obtained by means of non-linear least-squares adjustment of the filter coefficients [15]. Uncertainties associated with the frequency response values are propagated to an uncertainty associated with the filter coefficients. Uncertainties associated with the deconvolution filter coefficients can be propagated to the corresponding estimate of the measurand either using closed formulas [16, 17] or a Monte Carlo method [18].

In an ideal noise-free scenario, the application of the obtained filter to the measured system output signal then produces the (discretized) system input signal

$$\hat{y}_0[n] = ((h*y)*g_{\mathrm{inv}})(nT_s) = (y*(h*g_{\mathrm{inv}}))(nT_s) = y(nT_s), \quad (9)$$

with T_s the time domain sampling period. However, in practice the ideal inverse filter results in a strong noise amplification, see figure 3. The compensation result has a flat frequency response amplitude, corresponding to an ideal dynamic behavior, whereas the result in the time domain does contain almost only noise. Thus, the ideal inverse filter cannot be utilized to obtain a reasonable estimate of the system input signal.

In order to render the ill-posed estimation problem stable, some kind of regularization is required. In this example an additional low-pass filter is designed which aims at suppressing high-frequency noise in the estimation process. The low-pass filter here was chosen as FIR-type filter with cut-off at 2 Hz, designed using a Kaiser window with length of 100 samples and with a scaling factor of $\beta = 8$. That is, the ideal inverse filter is replaced by the deconvolution filter

$$G_{\mathrm{dec}}(z) = G_{\mathrm{inv}}(z)G_{\mathrm{low}}(z), \quad (10)$$

and the application of the combined filter results in

$$\hat{y}[n] = ((h*y)*g_{\mathrm{dec}})(nT_s). \quad (11)$$

In the frequency domain the compensation result has a flat frequency amplitude up to a certain frequency and goes to zero from there on owing to the low-pass filter g_{low}. In the time domain the noise amplification is reduced significantly, see figure 4.

The design of the low-pass filter requires one to choose a suitable filter type, filter order and cut-off frequency. In particular the choice of the low-pass filter cut-off frequency introduces a bias to the estimation and controls the suppression of the frequency content of the estimated signal. Any choice $f_{\mathrm{cut}} > 0$ necessarily results in a systematic error in the estimate as it represents a deviation from the ideal inverse of the measurement equation, see figure 5. If this bias cannot be considered to be negligibly small, then the uncertainty contribution of the low-pass filter must be accounted for in the total uncertainty of the measurand. In this example the low-pass filter cut-off frequency is typically chosen such that the obtained estimate of the measurand is close to that expected by the experimenter. The effect of the systematic error is then assumed to be negligible. However, an assessment of the systematic error requires some prior knowledge about the measurand itself is known, see section 4.

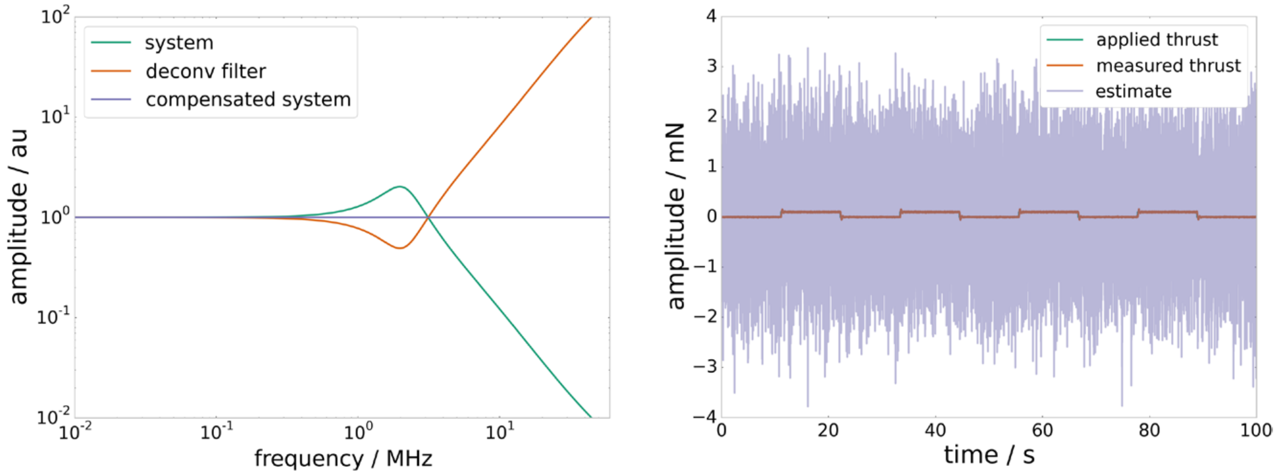

Figure 3. Result of the application of the ideal inverse filter in the frequency domain (left) and the time domain (right).

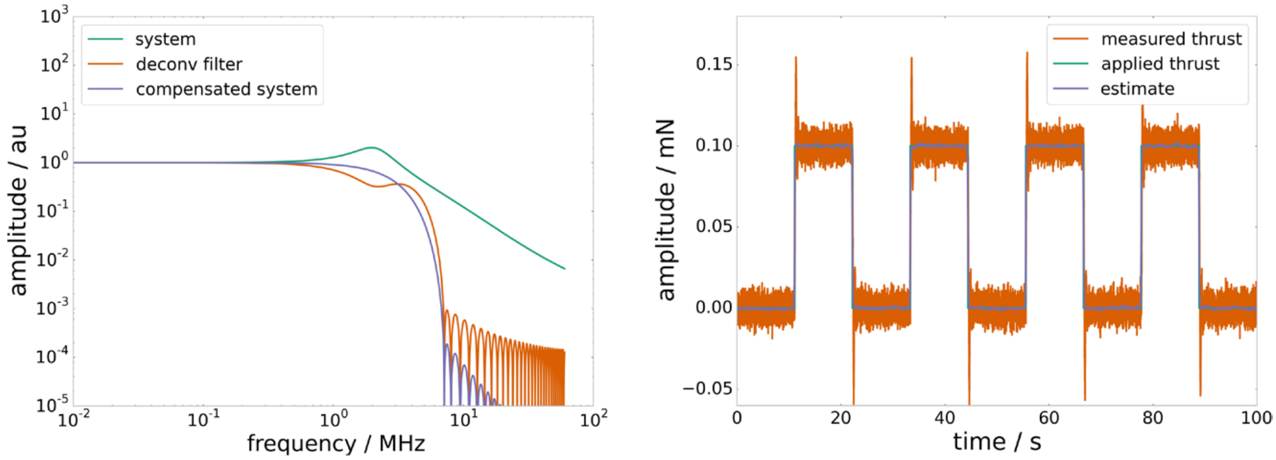

Figure 4. Result of the application of the inverse filter together with the additional low-pass filter in the frequency domain (left) and the time domain (right).

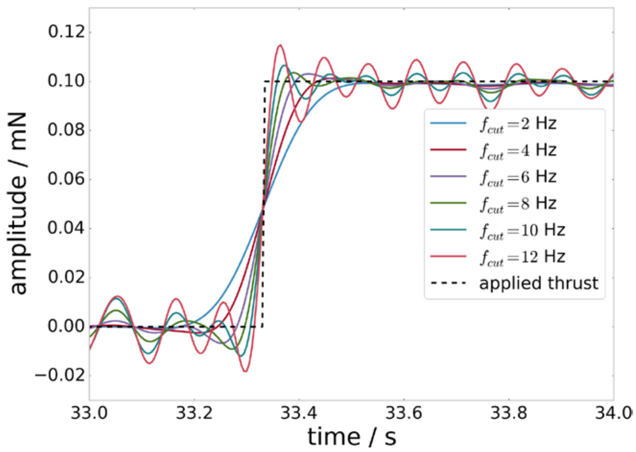

Figure 5. Effect of the low-pass filter cut-off frequency on the estimation result.

3.2. Sampling oscilloscope

Ultra-fast sampling oscilloscopes with a nominal bandwidth of over 100 GHz are ideal instruments for characterizing high-speed electrical equipment, such as pulse generators, and for the measurement of high-speed waveforms. A quasi-dynamic approach to take the oscilloscope's reaction time into account is based on the oscilloscope's rise time. However, many examples have shown that such a single parameter device characterization is insufficient as it does not take into account the non-ideal dynamic behavior. Moreover, despite their small rise time of about 4–7 ps, ultra-fast sampling oscilloscopes are sometimes only marginally faster than the signals to be measured, making some kind of correction necessary [19]. When the influence of the oscilloscope is not taken into account, the obtained measurement result can lead to a wrong characterization of the device under test. For instance, an eye-diagram test for a pulse generator may give a false rejection of the generator owing to the dynamic behavior of the oscilloscope not taken into account. To this end, a full waveform approach is recommended which completely takes into account the non-ideal dynamic behavior of the measuring instrument. Therefore, the mathematical relation between the measurand $\boldsymbol{y} = (y(t_1), \ldots, y(t_N))^T$ and the measured signal $\boldsymbol{x} = (x(t_1), \ldots, x(t_N))^T$ is considered to be modeled by

$$\boldsymbol{x} = \boldsymbol{H}\boldsymbol{y} + \boldsymbol{n} \tag{12}$$

with $\boldsymbol{n} = (n(t_1), \ldots, n(t_N))^T$ the noise process and $\boldsymbol{H}$ the convolution matrix calculated from the (discretised) oscilloscope's

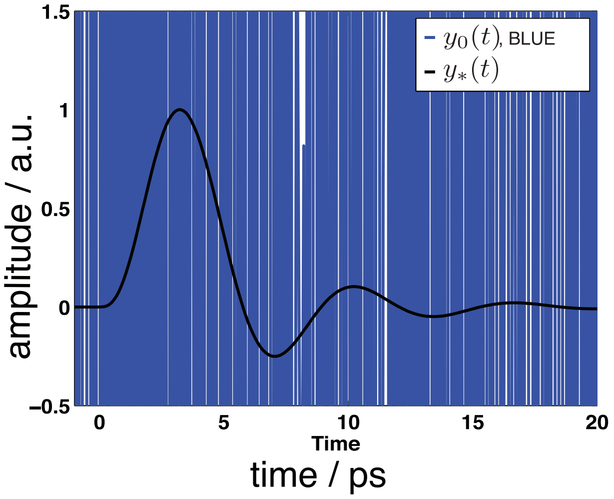

Figure 6. Deconvolution without regularization $y_0(t)$ as the result of the application of the best linear unbiased estimator (13). The black curve $y_*(t)$ shows the true input signal for comparison.

impulse response $h(t)$ [7]. The noise process $\boldsymbol{n}$ is assumed to be normally distributed $\boldsymbol{n} \sim N(0, \boldsymbol{\Sigma}_n)$ with known covariance matrix $\boldsymbol{\Sigma}_n$.

The mathematical relation (12) is a linear regression model with $\boldsymbol{y}$ being the vector of sought parameters. The best linear unbiased estimator (BLUE) for this regression problem is a weighted linear least-squares with solution given by [14]

$$\boldsymbol{y}_0 = (\boldsymbol{H}^T \boldsymbol{\Sigma}_n^{-1} \boldsymbol{H})^{-1} \boldsymbol{H}^T \boldsymbol{\Sigma}_n^{-1} \boldsymbol{x}. \tag{13}$$

However, similar to the first case study, the ideal estimator results in a strong noise amplification, see figure 6. As in the previous case study, this noise amplification reflects the ill-posedness of deconvolution [13, 14].

In order to render the estimation of the measurand $\boldsymbol{y}$ stable, a bias has to be introduced to the unbiased estimator (13) to reduce the variation in the estimated signal. One possible change to the estimator is to consider a smoothing penalty. For instance, with $\boldsymbol{L}$ the matrix representation of the second-order finite difference operator $(\Delta_L[x])(k) = (x(k-1) - 2\,x(k) + x(k+1))/T_s^2$, a regularized estimator can be derived as

$$\hat{\boldsymbol{y}}_\lambda = (\boldsymbol{H}^T \boldsymbol{\Sigma}_n^{-1} \boldsymbol{H} + \lambda^2 \boldsymbol{L}^T \boldsymbol{L})^{-1} \boldsymbol{H}^{-1} \boldsymbol{\Sigma}_n^{-1} \boldsymbol{x}, \tag{14}$$

with λ denoting the so called regularization parameter. The larger λ the stronger the influence of the penalty $\boldsymbol{L}$, and hence, λ controls the smoothness of the solution. For very small values of λ, the obtained estimate is still mostly covered by noise. In contrast, a large value of λ produces an overly smoothed estimate, see figure 7.

Many methods have been suggested in the literature to determine suitable values for λ based on an assumed statistical model for the measurement, see, e.g. [14, 20, 21]. Primarily these methods are heuristic and based on practical experience rather than a strict mathematical and statistical proof. For instance, a commonly applied approach is the so called *L*-curve method. It is based on the observation that when plotting the generalized norm ($\|\boldsymbol{Lx}\|^2$) of the regularized solution versus the norm of the corresponding residual for all valid regularization parameters, the resulting curve resembles the letter '*L*' and the optimal choice of the regularization parameter corresponds to the L-shape corner. The reasoning for this choice is that at the corner of the *L*-curve, the nature of the regularization changes from smoothing the estimate to reducing the residual and thus often provides a good compromise between the two parts. While many methods for choosing λ produce reasonably good estimates in specific scenarios, no method is provably 'optimal' for all cases. Moreover, all methods can fail completely when the underlying assumptions about the statistical model of the measurement are incorrect [21].

In any case, the regularization parameter introduces a systematic error to the estimation of the measurand whenever $\lambda > 0$. Thus, irrespective of the chosen method for the determination of an 'optimal' value for λ, the uncertainty contribution of the induced systematic error has to be taken into account. Currently, there is no generally applicable guidance for such analysis, see also section 4.

3.3. Hydrophones

Calibrated hydrophones are commonly utilized for precise measurements of technical and medical ultrasound fields. In particular for the characterization of medical ultrasound devices precise and reliable determination of positive and negative pressure peaks is required in order to ensure patient safety. The current standard practice in the analysis of hydrophone measurements is to consider the frequency response amplitude $M = H(f_{\text{awf}})$ at a single frequency, the so called working frequency f_{awf}. For hydrophones with sufficiently flat frequency response in the bandwidth of the measured ultrasound field this approach produces a reasonable estimate. However, for many practical ultrasound devices, hydrophones that satisfy this requirement are hardly available. Moreover, many hydrophones show non-flat frequency response amplitude values even in their stated working range, see figure 8. To this end, the full frequency response of the hydrophone has to be taken into account [6, 22].

In this case study, the convolution is considered in the frequency domain,

$$X(f) = H(f)Y(f) \tag{15}$$

with $Y(f)$, $X(f)$ denoting the discrete Fourier transform (DFT) of the measurand and the indication, respectively. The ideal deconvolution is therefore given by

$$Y(f) = \frac{X(f)}{H(f)}. \tag{16}$$

As observed previously, (16) results in significant noise amplification. To mitigate this issue, the deconvolution (16) is augmented with a low-pass filter with frequency response $L(f)$

$$\hat{Y}(f) = \frac{X(f)}{H(f)} L(f). \tag{17}$$

The associated time domain estimate of the measurand is calculated using the inverse DFT [6, 23]. Similarly to the first case study, the amount of regularization is controlled mainly by the

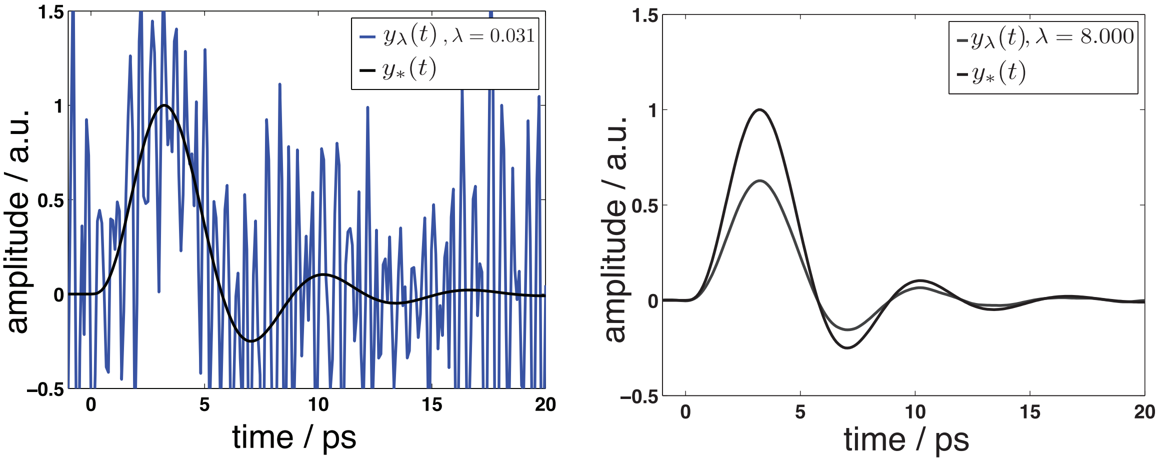

Figure 7. Application of the regularized estimator with regularization parameter being λ too small (left) and too large (right).

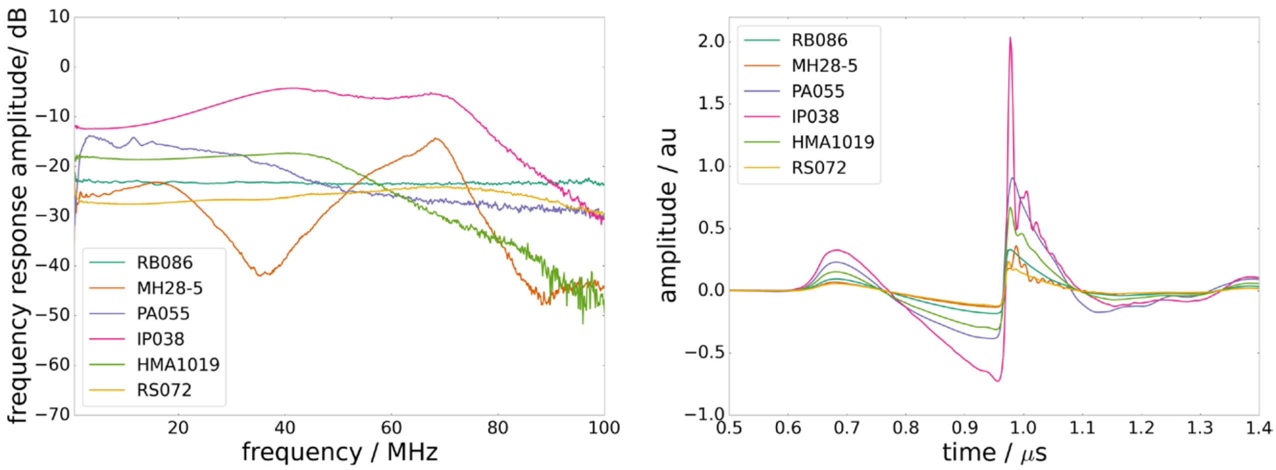

Figure 8. Left: Calibrated frequency response amplitude for various hydrophones Right: Simulated application of all hydrophones to the exact same input to demonstrate the impact of the individual sensor characteristic on the measured signal.

low-pass filter cut-off frequency. A small cut-off frequency results in overly smoothed estimates, whereas a large cut-off increases the noise content in the estimated waveform. The quantities of interest, the maximum and minimum in the estimated pressure versus time curve, depend nontrivially on the low-pass filter cut-off frequency, see figure 9. Thus, the possible systematic error introduced by the low-pass filter has to be taken into account.

For the example considered here, the shape of the measured pulses in the characterization of ultrasound devices is typically known to some extent owing to the nature of the signals generated. For instance, very strong high frequency oscillations and signal overshooting as can be seen in figure 8 (right) for some hydrophone signals is not realistic for ultrasonic pressure waveforms. This could be employed for the design of the low-pass filter. Moreover, such knowledge may be utilized to derive an estimate of the systematic error, see, section 4.

4. Regularization approaches for metrology

Since the deconvolution problem is an ill-posed inverse problem, a key task in the estimation is regularization and the evaluation of its uncertainty contribution. For metrology, a generic treatment of the regularization uncertainty at the level of the GUM would be favorable. A prerequisite of the application of the GUM is a model that accounts for all uncertainty sources. For ill-posed problems, and without incorporating additional information, such a model does not exist. Hence, GUM-compliant solutions always need to be application-dependent, making general guidance rather difficult. The here discussed examples provide three approaches to the regularization of the estimation problem, that differ in the kind of prior knowledge about the measurand and the corresponding potential treatment of regularization uncertainty.

For the thrust-balance signal in section 3.1, assumptions about the smoothness of the measurand guide the choice of the low-pass filter. The design of a low-pass filter as a means for regularization is a common approach in signal processing. In fact, many classical approaches, such as Wiener deconvolution and Tikhonov regularization can be interpreted as generalized low-pass filters in the following sense. Let us denote the frequency response of the regularized estimation kernel in a standardized, factored form by

$$G(f) = H^{-1}(f)R(f) \tag{18}$$

where $H(f)$ is the frequency response of the LTI system and $R(f)$ is that of the regularizer. Regularization approaches are distinguished by their choice of $R(f)$, which may be viewed as

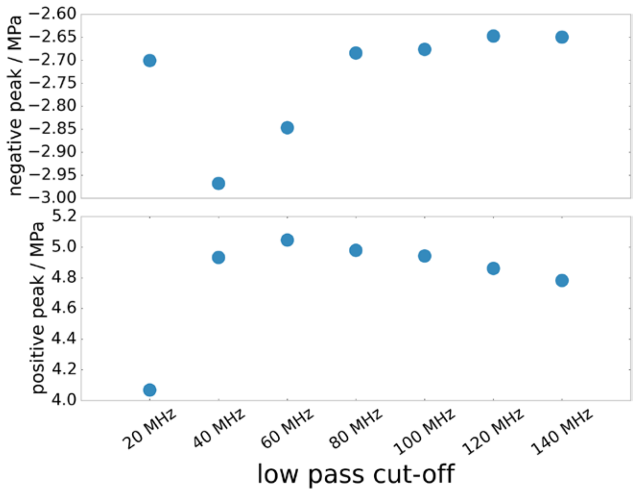

Figure 9. Illustration of the effect of the low-pass filter cut-off frequency on the determined pressure peak values in the deconvolution of a pulse measured with the IP038 hydrophone.

a generalized, low-pass filter. For example, the Wiener deconvolution is given by

$$R(f) = \frac{|H(f)^2|}{|H(f)|^2 + N(f)/Y(f)}$$

with $N(f)$ the power spectrum of the assumed noise process and $Y(f)$ that of the measurand when interpreted as wide-sense stationary noise process, see [24]. Tikhonov regularization replaces the reciprocal signal-to-noise ratio in the denominator with a less-specific form, for example

$$R(f) = \frac{|H(f)|^2}{|H(f)|^2 + \lambda^2|L(f)|^2}$$

with $L(f)$ being, for instance, the frequency domain representation of the smoothness penalty. For both classical methods the resulting frequency response $R(f)$ resembles that of a low-pass filter. In the thrust-balance example the regularizer $R(f)$ is chosen as a specific low-pass filter with cut-off frequency determined by expert knowledge about the smoothness, i.e. frequency content, of the measurand. Consequently, the resulting systematic bias introduced by the low-pass filter is assumed to be negligibly small compared to the other sources of uncertainty.

For the hydrophone example in section 3.3, the approach to regularization is also the design of a suitable low-pass filter. However, in this case, prior knowledge about the frequency content of the measurand is more specific than in the thruster application. This allows for a GUM-compliant evaluation of the uncertainty contribution of the regularization bias in the following way [25]. Assume that knowledge about the measurand is given by means of an upper bound $B(f)$ on the absolute value of its Fourier transform $Y(f)$, in other words, $|Y(f)| < |B(f)|$ for all real f. Then there exists an upper bound $\overline{\Delta}$ such that

$$|\hat{y}[n] - y(nT_s)| \leqslant \overline{\Delta} \tag{19}$$

independent of n, with $\hat{y}[n]$ the estimate of the measurand $y(t)$ at time instant nT_s with T_s the sampling interval length. That is, the upper bound for the magnitude of the Fourier spectrum is translated into an upper bound in the time domain, see [8]. In this way, knowledge about the measurand is applied to derive knowledge about the regularization error. According to GUM supplement 1 [10], this kind of knowledge is expressed in terms of a probability density function (PDF) by assigning a rectangular distribution to the regularization error

$$p_{\Delta[n]}(\delta) = \begin{cases} 1/(2\overline{\Delta}) & |\delta| \leqslant \overline{\Delta} \\ 0 & \text{otherwise} \end{cases}. \tag{20}$$

It is worth noting, that this approach can also be employed to derive a low-pass filter cut-off frequency that yields minimal uncertainty [25].

Whereas the hydrophone and the thruster example use explicit prior knowledge about the measurand, the oscilloscope example aims at an objective data-driven method for the determination of the regularization parameter. Many approaches for this task can be found in the literature, although no method yields mathematically proven optimality. To some extent, the regularization parameter chosen by a data-driven method aims at balancing the amount of noise reduction and the size of the regularization bias. Prior knowledge is incorporated implicitly by the choice of the regularization operator, e.g. a smoothness operator. Depending on the statistics of the data, then a regularization parameter value is determined to weigh the impact of the prior knowledge yielding a more objective result. However, some knowledge about the measurand is still required for the assignment of an uncertainty to the systematic error induced by the regularization process, unless its value can be assumed to be negligible compared to the other sources of uncertainty.

Bayesian inference allows for an incorporation of prior knowledge about the measurand in a probabilistic framework. The viewpoint on probabilities and the interpretation of probability density functions in the GUM is closely related to that in Bayesian statistics. However, the relation of a Bayesian inference to the framework of the GUM and its supplements is a topic of ongoing research in metrology. In a Bayesian inference, the prior distribution expresses one's subjective belief about the measurand, which translates to a PDF assigned to the estimate of the measurand which expresses the (subjective) belief about its value. For instance, the approach taken for the oscilloscope example in section 3.2 can similarly be derived in a Bayesian framework when the data is normally distributed. If one assigns a multi-variate normal with mean zero and covariance $\boldsymbol{L}^T\boldsymbol{L}$ as *a prior* distribution for the measurand, then Bayes' rule yields a multivariate normal distribution for the measurand posterior. Under this interpretation, equation (14) in section 3.2 amounts to estimation of the measurand by the maximum posterior. Several approaches towards the determination of prior distributions that incorporate assumptions about signal smoothness can be found in the literature, see [26, 27] and references therein. Whereas the approach for the oscilloscope example aims at determining a value of the regularization parameter λ from the observed data using heuristic methods, a Bayesian inference would employ prior elicitation methods. The resulting prior distribution then expresses the prior belief about the measurand, which propagates to a posterior belief using Bayes' rule. Consequently, there is no formal estimation bias to be determined. This is a conceptual advantage of a (subjective) Bayesian inference, but requires great care regarding the choice of the prior distribution associated with the measurand in order to obtain reliable results.

Bayesian inference is a common regression approach used when prior knowledge about the sought parameter is available, see [28]. The common use case for Bayesian inference in regression entails estimating a parameter vector of small dimension. In the context of dynamic measurements, one may consider a parametric model of the measurand and infer its parameter values from the measured data. However, this model choice requires considerable prior knowledge and imposes significant structure in the resulting measurand. In many metrological applications this is not feasible. As an alternative, the discrete-time values of the measurand itself can be viewed as the parameters of interest. In principle, this allows the application of Bayesian regression using standard tools. However, in practice the huge dimensionality of the measurand poses significant challenges. First of all, reasonable prior knowledge has to be formulated for each time instant at which the value of the measurand is considered. Secondly, the application of Bayesian regression in most cases requires numerical tools such as Markov Chain Monte Carlo (MCMC) [28]. These tools work very well for small dimensions, but require an increased effort for increased dimensionality of the sought parameter vector. To this end, in some cases so-called conjugate priors can be employed that allow for an analytical treatment, i.e. without the need of MCMC. There is an ongoing discussion in the statistical literature about the consistency of Bayesian estimates when the dimensionality increases [8]. That is, the estimates may become inconsistent when the sampling frequency is increased. Another alternative is the application of non-parametric Bayesian inference, for instance by using penalized spline regression, Dirichlet processes, wavelets or reproducing Kernel Hilbert space methods [29]. However, currently available methods are very challenging conceptually as well as numerically, making their widespread use in metrology laboratories unfeasible.

Summarizing the above discussion, Bayesian inference appears to be a viable approach for regularization in dynamic metrology, but further research is necessary in order to provide generic guidance for a broad class of users.

5. Further challenges

Beyond the challenge of regularized deconvolution and the corresponding evaluation of uncertainties discussed here, there are additional mathematical and statistical challenges in dynamic metrology.

5.1. Dimensionality

The previously discussed huge dimensionality of the discretised dynamic quantities pose a challenge for the evaluation of an uncertainty from measurements alone. For instance, the measured system output signal has to be accompanied with an uncertainty. For univariate quantities typically repeated measurements are utilized for this purpose. However, the number of repeated measurements required to reliably determine the covariance matrix of a multivariate quantity increases with its dimension. Thus, in most practical cases parametric approaches for the uncertainty have to be developed.

5.2. Mathematical modeling

In many cases mathematical models for dynamic systems are based on differential equations whereas static measurements typically rely on algebraic models. This results in mathematical challenges for the derivation of appropriate measurement models. Moreover, it poses conceptual challenges, because the uncertainty associated with parameters of the differential equations result in stochastic differential equations as measurement models, leading to stochastic process noise models.

5.3. Statistical modeling

In many applications so far only idealized noise structures are considered, partly for simplicity and owing to the above mentioned challenge in uncertainty evaluation. However, in practice measurement noise is typically non-ideal. For instance, an anti-aliasing filter or some other kind of filtering applied to the measured signal introduces correlation in the signal noise. For practical reasons, this can be dealt with only by utilizing parametrized models, such as, e.g. auto-regressive models, for the statistical modeling of measurement uncertainty.

5.4. Software tools

In recent years many software tools have been developed to carry out uncertainty evaluation in line with the GUM and its supplements. However, these tools are not applicable to dynamic measurements owing to the dimensionality and the ill-posedness of the estimation task. A prerequisite for the wide-spread implementation of dynamic metrology is the availability of harmonised standard approaches and software tools for common tasks, such as estimation of the measurand and the propagation of measurement uncertainty.

5.5. Calibration certificates

Often a non-parametric calibration of dynamic systems is carried out. For instance, sampling oscilloscopes are calibrated dynamically in terms of their discretised impulse response $\boldsymbol{h} = (h(t_1), \ldots, h(t_N))^T$. Reporting and transferring the calibration result requires transferring the associated uncertainties, i.e. the covariance matrix. Typically, the dimension N of the calibrated impulse response is on the order of several thousands, which makes traditional calibration certificates infeasible.

5.6. Incomplete calibration data

In many applications the proposed deconvolution relies on approximate information about the dynamic behavior of the measurement system. Hydrophones, for instance, are usually experimentally calibrated in a limited frequency range due to technical limitations and restrictions of time and costs. For deconvolution applications, however, the frequency response is needed from zero up to the Nyquist frequency of the measurement application. Therefore appropriate frequency response extrapolations are necessary in addition to the regularization procedure, and their implications on the uncertainty of the results need to be considered as well.

5.7. Key comparisons

An essential part of metrology is the execution of laboratory intercomparison studies. For static measurements a unique univariate value is considered and a key comparison reference value is sought. For dynamic metrology this concept cannot be adopted easily. One reason is the dimensionality of the dynamic measurand. First approaches, such as in dynamic calibration of accelerometers, mimic the static approach by considering individual values at defined frequencies as independent key comparisons, neglecting correlation.

6. Summary and outlook

Although industrial applications carry out dynamic measurements on a day-to-day basis, the underlying metrological foundation is still mostly based on static calibration. However, an increasing number of applications make the implementation of dynamic measurement analysis inevitable.

A typical task in dynamic metrology is deconvolution as a method to estimate the dynamic measurand. Three different approaches have been presented and discussed here. Despite their different physical quantities and time scales, all case studies have to deal with similar mathematical challenges. The key issue is the regularization of the otherwise unstable estimation, owing to its mathematical ill-posedness. As regularization unavoidably introduces a systematic error to the estimation, its uncertainty contribution has to be taken into account. In order to estimate this systematic error, some kind of prior knowledge about the measurand has to be available. Unfortunately, currently there is no generally applicable approach to accomplish this task. Methods available in the literature typically focus on statistical optimality rather than a metrologically sound evaluation of uncertainty. Hence, there is a great need for future metrological research and the development of harmonized guidance for end-users.

References

[1] Hughes B 2012 System identification and uncertainty analysis for challenging measurement applications: a case study in micro-Newton level force measurement *BIPM Workshop on Challenges in Metrology for Dynamic Measurements*

[2] Bartoli C, Beug M F, Bruns T, Eichstädt S, Esward T, Klaus L, Knott A, Kobusch M and Schlegel C 2014 Dynamic calibration of force, torque and pressure sensors *IMEKO 22nd TC3, 12th TC5 and 3rd TC22 Int. Conf. (Cape Town, Republic of South Africa, 2014)*

[3] Hale P D, Williams D F, Dienstfrey A, Wang J, Jargon J, Humphreys D, Harper M, Füser H and Bieler M 2012 Traceability of high-speed electrical waveforms at NIST, NPL and PTB *Conf. on Precision Electromagnetic Measurements* (IEEE) pp 522–3

[4] Füser H, Eichstädt S, Baaske K, Elster C, Kuhlmann K, Judaschke R, Pierz K and Bieler M 2012 Optoelectronic time-domain characterization of a 100 GHz sampling oscilloscope *Meas. Sci. Technol.* **23** 025201

[5] Harris G R 1996 Are current hydrophone low frequency response standards acceptable for measuring mechanical/cavitation indices? *Ultrasonics* **34** 649–54

[6] Wilkens V and Koch C 2004 Amplitude and phase calibration of hydrophones up to 70 MHz using broadband pulse excitation and an optical reference hydrophone *J. Acoust. Soc. Am.* **115** 2892

[7] Oppenheim A V and Schafer R W 1989 *Discrete-time Signal Processing* (Englewood Cliffs, NJ: Prentice-Hall)

[8] Eichstädt S 2012 Analysis of dynamic measurements: evaluation of dynamic measurement uncertainty *PhD Thesis* Technical University of Berlin

[9] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML 2008 *Evaluation of Measurement Data: Guide to the Expression of Uncertainty in Measurement* (Joint Committee for Guides in Metrology, JCGM 100)

[10] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML 2008 *Evaluation of Measurement Data–Supplement 1 to the 'Guide to the Expression of Uncertainty in Measurement'–Propagation of Distributions Using a Monte Carlo Method* (Joint Committee for Guides in Metrology, JCGM 101)

[11] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML 2011 *Evaluation of Measurement Data: Supplement 2 to the 'Guide to the Expression of Uncertainty in Measurement': Extension to Any Number of Output Quantities* (Joint Committee for Guides in Metrology, JCGM 102)

[12] Dabóczi T 1998 Uncertainty of signal reconstruction in the case of jittery and noisy measurements *IEEE Trans. Instrum. Meas.* **47** 1062–6

[13] Riad S M 1986 The deconvolution problem: an overview *Proc. IEEE* **74** 82–5

[14] Hansen P C and O'Leary D P 1993 The use of the *L*-curve in the regularization of discrete ill-posed problems *SIAM J. Sci. Comput.* **14** 1487–503

[15] Eichstädt S, Elster C, Esward T J and Hessling J P 2010 Deconvolution filters for the analysis of dynamic measurement processes: a tutorial *Metrologia* **47** 522–33

[16] Elster C and Link A 2008 Uncertainty evaluation for dynamic measurements modelled by a linear time-invariant system *Metrologia* **45** 464–73

[17] Link A and Elster C 2009 Uncertainty evaluation for IIR (infinite impulse response) filtering using a state-space approach *Meas. Sci. Technol.* **20** 055104

[18] Eichstädt S, Link A, Harris P and Elster C 2012 Efficient implementation of a Monte Carlo method for uncertainty evaluation in dynamic measurements *Metrologia* **49** 401–10

[19] Dienstfrey A and Hale P D 2014 Analysis for dynamic metrology *Meas. Sci. Technol.* **25** 035001

[20] Krawczyk-Stando D and Rudnicki M 2007 Regularization parameter selection in discrete ill-posed problems—the use of the *U*-Curve *Int. J. Appl. Math. Comput. Sci.* **17** 157–64

[21] Dienstfrey A and Hale P D 2014 Colored noise and regularization parameter selection for waveform metrology *IEEE Trans. Instrum. Meas.* **63** 1769–78

[22] Wear K A, Gammell P M, Maruvada S, Liu Y and Harris G R 2011 Time-delay spectrometry measurement of magnitude and phase of hydrophone response *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **58** 2325–33

[23] Wear K A, Liu Y, Gammell P M, Maruvada S and Harris G R 2015 Correction for frequency-dependent hydrophone response to nonlinear pressure waves using complex deconvolution and rarefactional filtering: application with fiber optic hydrophones *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **62** 152–64

[24] Wiener N 1952 Extrapolation, interpolation and smoothing of stationary time series with engineering applications *J. Amer. Stat. Assoc.* **47** 319–21

[25] Eichstädt S, Link A, Bruns T and Elster C 2010 On-line dynamic error compensation of accelerometers by uncertainty-optimal filtering *Measurement* **43** 708–13

[26] Kimeldorf G S and Wahba G 1970 A correspondence between bayesian estimation on stochastic processes and smoothing by splines *Ann. Math. Stat.* **41** 495–502

[27] Fitzpatrick B G 1991 Bayesian analysis in inverse problems *Inverse Problems* **7** 675–702

[28] Elster C *et al* 2015 A Guide to Bayesian Inference for Regression Problems *Technical Report* PTB, NPL, LNE, LGC Ltd., INRIM, VSL, SP

[29] Müller P and Quintana F A 2004 Nonparametric Bayesian data analysis *Stat. Sci.* **19** 95–110