Improved Characterization of Non-Stationary Flows Using a Regularized Spectral Analysis of Ultrasound Doppler Signals

A. Herment ([1,*]), J.-F. Giovannelli ([2]), G. Demoment ([2]),
B. Diebold ([1]) and A. Delouche ([1])

([1]) INSERM, Unité 66, Imagerie Biomédicale, Morphologique et Fonctionnelle,
CHU Pitié Salpétrière, 91 boulevard de l'Hôpital, 75634 Paris Cedex 13, France
([2]) Laboratoire des Signaux et Systèmes, CNRS - ESE - UPS ESE, Plateau de Moulon,
91192 Gif-Sur-Yvette Cedex, France

Abstract. — This paper addresses the problem of ultrasound Doppler spectral estimation when only a short observation set is available. Following the work of Kitagawa and Gersch, the spectra are described by a long autoregressive model whose coefficients are estimated in a Bayesian regularized least squares framework accounting for spectral smoothness in order to avoid too spiky spectra. The critical computation of the tradeoff parameters is addressed using both maximum likelihood and generalized cross validation criteria in order to automatically tune the smoothness constraint. The practical potential of the method is demonstrated using both simulated and *in vitro* signals. In a Monte-Carlo simulation study, investigation of quantitative indices such as quadratic distances shows interesting improvements with respect to the usual least squares method whatever the window data length and the signal to noise ratio. When applied to actual Doppler signals, the proposed method offers better description of the Doppler spectrum morphology than the usual least squares one.

1. Introduction

Spectral analysis of ultrasound Doppler signals is an efficient tool for blood flow analysis. Actually, in ideal conditions such as stationary flow, first order statistics for the backscattering of ultrasound by blood, and flow insonification by an infinite plane wave, the Doppler signal is shown to be a zero-mean, stationary, Gaussian random process which can be completely specified by its autocorrelation function or equivalently by its Power Spectral Density (PSD) [3]. The Doppler signal PSD, commonly referred to as the "Doppler spectrum", provides a useful image of the velocity histogram of red blood cells, which statistically describes the flow as a distribution of the number of red cells traveling through the sample volume at a given velocity [34].

([*]) Author for correspondence (e-mail: herment@imed.jussieu.fr)

However, in practical applications, these ideal conditions are not fulfilled [8] and three main factors distort the image of the velocity histogram of red blood cells:

- Physical factors. Backscattering of ultrasound by blood is a highly complex non linear phenomenon [3] due in particular to the aggregation of red blood cells which can vary even during the same cardiac cycle. The limited size of the sample volume induces, among other effects, an artificial increase in the Doppler spectrum bandwidth [7, 26]. Finally, the wavefront is distorted [4], especially by phase cancellation effects [34] induced by intervening media located between the ultrasound transducer and the analyzed flow.

- Physiological factors. Physiological flows are essentially non-stationary ones, mainly during the ascending and descending systolic phases of the cardiac cycle.

- Estimation factors. Computation of the Doppler spectrum itself can induce severe distortions in the histogram image.

The physical factors are numerous and are described by the acoustic physics. However, the quality of flow description depends directly on the ability of the spectral analysis method to take into account the physiological factors (*i.e.* the non-stationarity of the Doppler signal) and to provide a more reliable spectrum than with the currently used Fourier methods.

1.1. NON-STATIONARY SPECTRAL ANALYSIS. — The problem of non-stationary analysis of Doppler signals was addressed recently by Guo *et al.* [16] using a signal modelization corresponding to the Doppler data recorded in a normal femoral artery. Time-frequency spectral analysis based on modern Cohen's class distributions [11] were used and compared to parametric AR stationary spectral analysis for Signal to Noise Ratios (SNR) ranging from infinity to 10 dB. The spectral analysis methods were classified with respect to their Integrated Mean Square Error (IMSE). The best results were obtained using the Bessel Distribution while the stationary AutoregRessive (AR) approach and the Choi-Williams distribution instead of a slightly 30% higher IMSE were considered by the author as valuable alternatives. However, a significant gain was obtained for the computation times (up to 7.5) using the parametric methods.

Aiming at practical implementation of the spectral analysis of Doppler signal, we readdress the spectral analysis in a parametric framework, assuming the pseudo-stationarity of the Doppler signal in short enough windows. The following survey of the relevant literature is used as a basis for our hypotheses.

1.2. PARAMETRIC SPECTRAL ESTIMATION OF STATIONARY SIGNALS. — Parametric spectral analysis has gained interest for Doppler signal analysis, as an alternative to Fourier analysis. This method consists of i) defining specifically a time model for the signal which provides a correct compromise between a good signal description and a stable spectral estimate, ii) computing the parameters of this model, and finally, iii) introducing these parameters into the theoretical expression of the model spectrum.

1.2.1. Model of the Doppler Signal. — Guo *et al.* [18] suggested that the quadrature Doppler signal recorded at peak systole nearby the aortic valve can be modeled as a complex AR process arguing that the modeling residual has a white noise characteristic when the Akaike Information Criterion (AIC) is used for optimizing the model order. However, this assumption which is true for long windows, can be debatable when using shorter windows for which the model order is often underestimated by the AIC criterion, as shown by Kaluzinski [20]. Conversely, Vaitkus and Cobbold [44] indicate that due to the transit time effect of the red blood cells

within the ultrasound beam, the Doppler signal is totally decorrelated beyond a time lag equal to this transit time because the distribution of red blood cells has completely changed within the sample volume. They concluded that a Moving Average (MA) part must be taken into account in the Doppler signal model. Finally, Kay and Marple showed in [21] that an AR process converts into an AutoRegressive Moving Average (ARMA) one when it is corrupted by a white Gaussian additive noise. From this standpoint, because of the possible variations of the SNR during the cardiac cycle (the SNR sometimes falls down during diastole, due to the clutter rejection filter and the flow characteristics), the Doppler signal should rather be modeled as an AR process in systole and either as an AR or an ARMA process in diastole.

1.2.2. Model Order . — Numerous criteria have been proposed to estimate the AR model order. These criteria are generally effective for long data sets. However, their performance degrades differently when the Doppler window is reduced. Among the eventual criteria, Schwindlein and Evans [39] showed that the Final Prediction Error (FPE) criterion [1], the AIC [2], and the AR transfer function criterion (CAT) [36] give coherent results and a correct order selection for long data windows and AR model; they also showed that the First Zero-Crossing of the autocorrelation function (FZC) is less effective. Yamaguchi [45] emphasized the fact that the best criterion is the first local minimum of the AIC rather than the FPE criterion whose minimum cannot always be clearly identified on short data sets. Vaitkus *et al.* [44] preferred using the Pao and Lee criterion [35] which studies the conditioning of the system as a function of the model order, and the Cadzow criterion [9] which examines the eigen structure of an over determined system of equations using a Singular Value Decomposition (SVD) of the corresponding matrix. However, for short Doppler windows, all these criteria tend to underestimate the model order when compared to the "best order" defined as the one providing the most convenient Doppler spectrum for clinical use [20]. Kitney *et al.* [25] pointed out that an AR order much larger than the one given by the FPE is necessary to find out deterministic structures in poststenotic flow disturbances. Barbaresco [5] proposed a regularized test derived from the criterion described by Ligget [27]. This new approach should improve consistency of the test on small sets of data, but no quantitative performance has yet been provided.

1.2.3. Parameter Estimation. — Schlindwein and Evans [39] used the Maximum Entropy Method (MEM) (Levinson-Durbin algorithm applied to the biased autocorrelation function estimator). Yamaguchi [45] proposed a modified AR-MEM method derived from the original Burg method [21] and integrating ensemble averaging concept. Kaluzinski [20] used the PAR-COR algorithm [29]. Vaitkus *et al.* [43] showed that the Yule-Walker algorithm provides better results than the MEM and the Least Squares (LS) approach [31] with respect to integrated mean square error and noise robustness.

   These methods are derived in a LS framework and related to usual empirical estimation of the correlation function. Moreover, they rely on the resolution of a linear system of equations. From this stand point, they are adapted to low computational complexity algorithms. Recently, regularized versions of Levinson and Burg algorithms [6] have become available that should still improve the efficiency of these methods.

1.2.4. Comparison of the Different Estimation Methods. — Vaitkus and Cobbold [44] addressed the choice of the best parametric estimation method for Doppler signals. They used for that purpose the stationary model of Mo and Cobbold [32] corresponding to the CW Doppler signal recorded at peak systole in a carotid artery. The comparison between ARMA, AR and MA models was made using 10 ms data sets (256 Doppler samples) and different SNR ranging from infinity to 0 dB. Numerous algorithms were compared for the determination of the model order

and for the estimation of the model parameters. The authors concluded that the best results were obtained using:

- The ARMA model with SVD algorithm described by Cadzow [9] for estimating the AR and the MA parameters. As for the determination of the order of the AR part, again, the best results were given by the Cadzow's procedure described in [9] and, for the MA part of the model, by a non linear approximate likelihood optimization consisting in selecting successive values of the MA order and testing the whiteness of the residues [8] for the MA part of the model.

- The AR parameters calculated through the Yule-Walker system of equations [23] and the Pao and Lee [35] or Cadzow [9] criteria for estimating the model order, which was estimated around 8 [43].

Comparison of the IMSE for these methods indicates a slight superiority for the ARMA model for high SNR signals and better results for the AR model when the SNR becomes lower than 10 dB. However, estimation of the AR model is four times faster than computation of the ARMA one.

1.3. Pseudo-Stationary Spectral Analysis. — The previous results by Guo *et al.* [16] and Vaitkus *et al.* [43] and [44] lead us to investigate the possibility of estimating the Doppler spectrum in very short windows in which the Doppler signal can be assumed stationary. Guo *et al.* [17], evaluating the normality and stationarity of the Doppler signal recorded behind aortic valves concluded that using a 10 ms window was acceptable in practice because 82% of the recorded segments were considered as stationary but they also indicated that using shorter windows would be preferable. Talhamy and Kitney [40] stressed that 5 or even 2.5 ms windows could be best fitted to resolving detailed flow structures. In such conditions, using a conventional 5 kHz PRF velocimeter, 25 samples (for 5 ms windows) or even only 12 samples (for 2.5 ms windows) are available for spectral estimation.

This paper describes a method for spectral analysis of Doppler signals achieving a compromise between a short analysis window for assessing the non-stationarity of the Doppler signal and a reasonable computation complexity for providing practical applications.

- As already stated in Section 1.2.1 and 1.2.2, since the choice between an ARMA and an AR model appears very difficult, since the criteria proposed for objective order selection seem to underestimate the model order and, finally, stating (Wold decomposition theorem) that any ARMA process can be represented by a high order AR process, an AR model with the highest possible order will be used for spectral analysis. Namely, the number of parameters of the AR model will be set to the number of Doppler data in the analysis window. At this step we assume that this high model order should be able to describe a wide variety of Doppler signals and that it should enable the Doppler signal to be accurately modeled even when using short windows.

- The strict LS framework on which the classical estimation of the model parameters is grounded (see Sect. 1.2.3 and 1.2.4) will be reformulated in order to introduce a regularization component to enhance the estimate stability while keeping a linear system of equations in order to cope with computational efficiency.

Section 2 introduces smoothness priors used for regularization of the spectral estimates obtained with long AR models; Section 3 is devoted to solutions for estimating the trade off parameters of this model. Section 4 presents the methodology of the simulation study and Section 5 comments the obtained results. Finally, Section 6 describes the development of a suboptimal algorithm aiming at practical applications.

2. Smoothness Prior and Long AR Models

2.1. LIMITATION OF CONVENTIONAL AR ESTIMATION . — As mentioned in the introduction, the Doppler signal $\mathbf{x} = [x_1, x_1, \ldots, x_N]^{\mathrm{t}}$, where superscript t denotes a transpose, is assumed to be a zero-mean stationary Gaussian N dimensional vector. It is well known that the statistical characteristics of such a vector are entirely summarized in its covariance matrix R, and that its probability density function is:

$$f_{X|R}(\mathbf{x}|R) = (2\pi)^{-N/2}\det R^{-1/2}\exp\left[-\frac{1}{2}(\mathbf{x}^{\mathrm{t}}R^{-1}\mathbf{x})\right]. \tag{1}$$

Since under our assumptions the $N\times N$ covariance matrix R is Toeplitz symmetric, it can clearly be parameterized by a N dimensional parameter vector $\boldsymbol{\theta}$. This vector can be equivalently chosen as: N correlation lags, N reflection coefficients, N cepstral parameters, N eigen values of the Toeplitz matrix R or $(N-1)$ AR parameters and noise power.

Thus, the previous probability density $f_{X|R}(\mathbf{x}|R)$, given equation (1), can be rewritten simply $f_{X|\Theta}(\mathbf{x}|\boldsymbol{\theta})$. Hence, it is to be emphasized that the spectral estimation problem is the one of $\boldsymbol{\theta}$ estimation *i.e.* the estimation of N parameters from N observations.

A common approach to the parameter estimation problem is maximum likelihood. Such an approach is usually justified on the basis of its asymptotic unbiased and minimum variance properties but may fail in a non asymptotic situation. Nevertheless the following part is devoted to ML estimation since this approach seems to be often desired and since the likelihood is one of the two elements of our approach.

As the density for $\mathbf{x}$ given equation (1), seen as a function of $\boldsymbol{\theta}$, is also the $\boldsymbol{\theta}$ likelihood

$$\mathrm{L}(\boldsymbol{\theta}) = f_{X|\theta}(\mathbf{x}|\boldsymbol{\theta}), \tag{2}$$

it is possible to estimate $\boldsymbol{\theta}$ through a maximum likelihood technique. This technique consists of choosing $\boldsymbol{\theta}$ as the maximizer of the probability of the observed data:

$$\theta_{\mathrm{ML}} = \arg\max \mathrm{L}(\boldsymbol{\theta}). \tag{3}$$

Such an expression for the likelihood is highly non linear with respect to $\boldsymbol{\theta}$ due to the presence of the determinant and the inverse of R. Several techniques [22, 38] were proposed in order to maximize $\mathrm{L}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$, but this very difficult non linear optimization problem is untractable for real time applications.

In order to avoid this problem, a particular choice for $\boldsymbol{\theta}$ and an approximation of the likelihood are made, leading to a computationally efficient estimator. Generally, $\boldsymbol{\theta}$ is chosen as the vector of AR coefficients

$$\boldsymbol{\theta} = \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ . \\ a_{N-1} \end{bmatrix}, \tag{4}$$

and as long as the PSD is not sharply peaked an approximate expression for the likelihood is derived in [23]:

$$f_{X|R}(\mathbf{x}|\mathbf{a}) = (2\pi\sigma_u^2)^{-N/2}\exp\left[-\frac{1}{2\sigma_u^2}(\mathbf{x} - X\mathbf{a})^{\mathrm{t}}(\mathbf{x} - X\mathbf{a})\right]. \tag{5}$$

Where σ_u^2 represents the variance of the generating noise. The data vector $\mathbf{x}$ and the data matrix X (also called observation vector $\mathbf{x}$ and observation matrix X) are designed in the

classical manner [23, 30] as follows:

$$
\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ . \\ x_N \\ . \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad X = \begin{bmatrix} 0 & 0 & . & . & 0 \\ x_1 & 0 & . & . & 0 \\ x_2 & x_1 & . & . & 0 \\ x_3 & x_2 & . & . & 0 \\ . & . & . & . & . \\ x_{N-1} & x_{N-2} & . & x_2 & x_1 \\ . & . & . & . & . \\ 0 & . & . & x_{N-1} & x_{N-2} \\ 0 & . & . & x_N & x_{N-1} \\ 0 & . & 0 & 0 & x_N \end{bmatrix}. \tag{6}
$$

Equation (5) clearly shows that the maximization of the approximated likelihood is equivalent to the minimization of the usual LS criterion [23, 30]:

$$
Q_0(\mathbf{a}) = (\mathbf{x} - X\mathbf{a})^{\mathrm{t}}(\mathbf{x} - X\mathbf{a}). \tag{7}
$$

The explicit expression for the minimizer of this criterion, the LS estimate, is well known:

$$
\begin{aligned}
\hat{a}_{\mathrm{LS}} &= \arg\min Q_0(\mathbf{a}) \\
&= (X^{\mathrm{t}}X)^{-1}X^{\mathrm{t}}\mathbf{x}
\end{aligned} \tag{8}
$$

Comments on the LS and the maximum likelihood estimates are now necessary. It has already been pointed out that the maximum likelihood and the LS estimates have interesting asymptotic properties. It means for a fixed AR model order and an infinite number of observation, or at least for a large enough observation vector with respect to a small model order. Here, as mentioned earlier, the signal is parameterized by N parameters, so estimation of N coefficients from N observations is desired. In such a situation the LS and the maximum likelihood estimation techniques leads to large deviations from the mean solution, *i.e.* large variance. Hence, even if this solution is weakly biased its strong variance leads to unacceptable estimates. Our approach consists of modifying the LS criterion. Such a problem of estimation of N parameters from N observations is an ill-posed one *i.e.* it suffers from a lack of information provided by N observations in order to infer about N parameters. Literature on the subject is abundant and various since this kind of problem arises in almost every branch of engineering and applied physics [13]. The resolution of ill-posed problems requires the use of the regularization concept, *i.e.* the introduction of prior information in the solution. The proposed Bayesian framework provides an attractive and coherent framework for that.

The ill-posedness can also be related to the ill-conditioning of the matrix $X^{\mathrm{t}}X$, for high order AR model with respect to the sample number (see Fig. 1). In such a case it can be shown [12] that any small variation on the data $\mathbf{x}$ will produce large variations on the solution $\hat{a}_{\mathrm{LS}} = (X^{\mathrm{t}}X)^{-1}X^{\mathrm{t}}\mathbf{x}$ and consequently large variations on the PSD estimate $\hat{S}_{\mathbf{x}}(f)$. Therefore this solution cannot be accepted as a physical one. The ill-conditioning of the matrix X is caused by a lack of information provided by the measured data, thus the only way of building a physically acceptable solution is to introduce some prior information in the solution and not just build the solution from the information provided by the data. This idea is developed in the next Section.

2.2. BAYESIAN APPROACH. — Introducing prior information in the solution is equivalent to changing the estimation criterion: instead of maximizing a likelihood of the parameter vector, we maximize the posterior likelihood which collects information provided by the observations
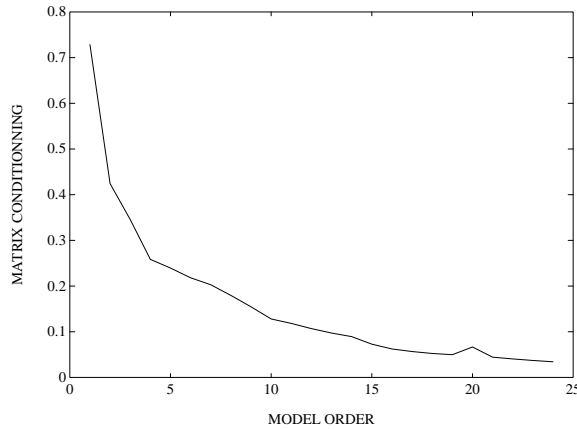
Fig. 1. — Normalized conditioning number of the observation matrix X as a function of AR model order and for a particular realization (24 samples) of the simulated Doppler signal. The matrix is well-conditioned when the number is near 1, and ill-conditioned when the number is near 0. Vertical axis, conditioning number. Horizontal axis, model order.

(*via* the likelihood) and by the *a priori* idea about the expected solution (*via* the prior law). This compound criterion incorporates both "prior knowledge" and "data-based knowledge" about $\boldsymbol{\theta}$ [13, 14].

The first problem to solve is the choice of a prior probability density $f_\Theta(\boldsymbol{\theta})$ for the vector $\boldsymbol{\theta}$. This density is supposed to contain the prior knowledge about the expected solution. This point is addressed in the next subsection 2.3. The information provided by the data is introduced through the conditional density for the observations, $f_{X|\Theta}(\mathbf{x}|\boldsymbol{\theta})$. Finally, the Bayes rule combines information provided by the observation (*via* the likelihood) and the prior information (*via* the prior law) in the posterior density for $\boldsymbol{\theta}$:

$$f_{\Theta|X}(\mathbf{x}|\boldsymbol{\theta}) = \frac{f_{X|\Theta}(\mathbf{x}|\boldsymbol{\theta})f_\Theta(\boldsymbol{\theta})}{f_X(\mathbf{x})}. \tag{9}$$

In a strict Bayesian sense, equation (9) yields the solution to the problem since it gathers all the information about the AR coefficients. However, the need of a spectrum implies the choice of a punctual estimator *i.e.* the selection of one representative point. A popular choice as a punctual estimator is the Maximum *A Posteriori* (MAP), defined as the maximizer of the posterior density:

$$\hat{a}_{\mathrm{MAP}} = \arg\max f_{A|X}(\mathbf{a}|\mathbf{x}). \tag{10}$$

The computation of this solution requires values for the remaining parameters, called the hyperparameters, (parameters of the prior law and noise variance). This is the crucial problem of tuning the compromise between the fidelity to the data and the fidelity to the prior knowledge. The problem has never been studied in the field of ultrasound spectral analysis and will be addressed in Section 3.

2.3. SPECTRUM SMOOTHNESS PRIOR. — The objective of this Section, is to choose a prior law for the AR coefficients, introducing a smoothness constraint on the PSD and leading to an easily computable solution.

The choice of a particular class of prior law is essentially constrained by the need of rapidly computable estimation and leads to deal with the class of Gaussian prior:

$$f_A(\mathbf{a}) = (2\pi)^{[(N-1)/2]}\det R_a^{-1/2}\exp\left[-\frac{1}{2}(\mathbf{a}^{\mathrm{t}}R_a^{-1}\mathbf{a})\right],\tag{11}$$

where the smoothness information about the PSD is introduced *via* the prior correlation R_a. The aim of the following is to find such a prior correlation matrix and we follow the idea of Kitagawa and Gersch [24].

The PSD $S_x(f)$ for an AR process is:

$$S_x(f) = \frac{\sigma_u^2}{\left|1 - \sum_{k=1}^{p} a_k\mathrm{e}^{2j\pi fk}\right|^2} = \frac{\sigma_u^2}{|1 - A(f)|^2},\tag{12}$$

with:

$$A(f) = \sum_{k=1}^{p} a_k\mathrm{e}^{2j\pi fk}.\tag{13}$$

When the PSD is known to be smooth, Kitagawa and Gersch [24] have proposed to constrain the PSD variations, *i.e.* to penalize its strong variations. They consider the kth derivative of $A(f)$ defined in equation (13) in order to measure the PSD variations. Averaging the squared modulus of this derivative over the whole reduced frequency domain, they define the PSD k^{th} smoothness by:

$$D_k = \int_{-1/2}^{1/2}\left|\frac{\partial^k}{\partial f^k}A(f)\right|^2 \mathrm{d}f.\tag{14}$$

After elementary algebra they show in [24] that:

$$D_k \propto \mathbf{a}^{\mathrm{t}}\Delta_k\mathbf{a},\tag{15}$$

where the Δ_k matrix, called the kth smoothness matrix is defined by:

$$\Delta_k = \begin{bmatrix} 1^{2k} & 0 & . & 0 \\ 0 & 2^{2k} & . & 0 \\ . & . & . & . \\ 0 & 0 & . & p^{2k} \end{bmatrix}.\tag{16}$$

A small value of D_k means a small value of the averaged k^{th} derivative of $A(f)$ hence a rather smooth PSD. In the limit when $D_k = 0$, $\mathbf{a} = 0$ and $\hat{S}_x(f) = \mathrm{C}^{\mathrm{te}}$ over the whole frequency domain, *i.e.* the PSD estimate is completely flat. On the contrary, a large value of D_k implies strong variations of $A(f)$, hence a peaky PSD.

In order to introduce this smoothness prior on the AR PSD, the covariance matrix R_a is designed from the kth order smoothness matrix Δ_k through the following equation:

$$R_a^{-1} = \frac{\lambda}{\sigma_u^2}\Delta_k.\tag{17}$$

The regularization parameter λ balances the regularized solution between the prior one and the data-based one. The question of its estimation is addressed in Section 3.

2.4. POSTERIOR LAW AND MAP ESTIMATE. — On the one hand, the previous section gives a prior law for the AR coefficients in equation (11) and on the other hand, the approximated likelihood for the AR coefficients is in the form given in equation (5). Application of the Bayes rule yields the posterior probability density function:

$$f_{A|X}(\mathbf{a}|\mathbf{x}) = \frac{f_{X|A}(\mathbf{x}|\mathbf{a}) f_A(\mathbf{a})}{f_X(\mathbf{x})}. \tag{18}$$

The denominator of equation (18), $f_X(\mathbf{x})$, is independent of $\mathbf{a}$ then its contribution is no more than a normalizing constant K; elementary algebra leads to the posterior law:

$$f_{A|X}(\mathbf{a}|\mathbf{x}) = K \exp\left[-\frac{1}{2\sigma_u^2} Q_{\mathrm{MAP}}(\mathbf{a})\right], \tag{19}$$

where $Q_{\mathrm{MAP}}(\mathbf{a})$ is given by:

$$Q_{\mathrm{MAP}}(\mathbf{a}) = (\mathbf{x} - X\mathbf{a})^{\mathrm{t}}(\mathbf{x} - X\mathbf{a}) + \lambda \mathbf{a}^{\mathrm{t}} \Delta_k^{-1} \mathbf{a}. \tag{20}$$

where the new hyperparameter λ balances the solution between the prior one and the data-based one. Since both the prior law $f_A(\mathbf{a})$ and the conditional law $f_{X|A}(\mathbf{x}|\mathbf{a})$ are Gaussian, the posterior density is also Gaussian. Hence the choice of a punctual estimator is no longer crucial: the MAP, the posterior mean, *etc.* are strictly equivalent and we will deal with the MAP. This estimator $\hat{a}_{\mathrm{MAP}}$ is defined as the posterior probability density maximizer or equivalently as the minimizer of the Regularized Least Squares (RLS) criterion $Q_{\mathrm{MAP}}(\mathbf{a})$ of equation (20):

$$\begin{aligned} \hat{a} &= \hat{a}_{\mathrm{MAP}} \\ &= \arg\max f_{A|X}(\mathbf{a}|\mathbf{x}) \\ &= \arg\min Q_{\mathrm{MAP}}(\mathbf{a}). \end{aligned} \tag{21}$$

Since the problem is linear and Gaussian, we have an explicit expression for its minimum:

$$\hat{a} = (X^{\mathrm{t}} X + \lambda \Delta_k)^{-1} X^{\mathrm{t}} \mathbf{x}. \tag{22}$$

2.5. QUADRATIC REGULARIZATION INTERPRETATION. — This part is devoted to the interpretation of this method in terms of quadratic regularization, out of the Bayesian framework.

The regularized criterion of equation (20) is composite: on the one hand, its first term $Q_\infty(\mathbf{a}) = (\mathbf{a}^{\mathrm{t}} \Delta_k \mathbf{a})$ is a prior criterion; on the other hand, its second term is the data-based criterion $Q_0(\mathbf{a}) = (\mathbf{x} - X\mathbf{a})^{\mathrm{t}}(\mathbf{x} - X\mathbf{a})$. Combining these two criteria, the proposed criterion incorporates both the prior and the data knowledge:

$$\begin{aligned} Q_{\mathrm{RLS}}(\mathbf{a}) &= Q_0(\mathbf{a}) + \lambda Q_\infty(\mathbf{a}) \\ &= (\mathbf{x} - X\mathbf{a})^{\mathrm{t}}(\mathbf{x} - X\mathbf{a}) + \lambda \mathbf{a}^{\mathrm{t}} \Delta_k \mathbf{a} \end{aligned} \tag{23}$$

As in the usual LS case, an explicit solution is available for $\hat{a}_{\mathrm{RLS}}$, the minimizer of the regularized criterion:

$$\hat{a}_{\mathrm{RLS}} = \arg\min Q_{\mathrm{RLS}}(\mathbf{a}) \tag{24}$$

$$\hat{a}_{\mathrm{RLS}} = (X^{\mathrm{t}} X + \lambda \Delta_k)^{-1} X^{\mathrm{t}} \mathbf{x}, \tag{25}$$

called the regularized least squares solution.

For λ small enough ($\lambda = 0$ at the limit) the criterion reduces to the usual LS one, $Q_{\mathrm{RLS}}(\mathbf{a}) = Q_0(\mathbf{a})$ and the usual LS solution is found again: $\hat{\mathbf{a}} = \mathbf{a}_0 = (X^{\mathrm{t}} X)^{-1} X^{\mathrm{t}} \mathbf{x}$. For λ large enough ($\lambda = \infty$ at the limit) the criterion becomes the prior one $Q_{\mathrm{RLS}}(\mathbf{a}) = Q_\infty(\mathbf{a})$ and the prior

solution is found again: $\hat{\mathbf{a}} = \hat{\mathbf{a}}_\infty = 0$ so that the PSD estimate is constant over the whole frequency domain.

Between these two extreme solutions, the "prior-based" one and the "data-based" one, an acceptable one still remains to be found. The aim is then to estimate, from the data, the crucial parameter λ in order to build a physically acceptable solution. The following part is devoted to this fundamental problem.

3. Hyperparameter Estimation

The method described above requires values for three hyperparameters σ_u^2, λ and k. The parameter σ_u^2 is a scaling factor while the two other parameters design the spectral shape. The parameter λ is of major importance for the spectral shape, so our study is focused on its estimation. On the contrary, the smoothness order is of lower influence and moreover it is practically possible, given a value of k, to find a value for λ giving a presuposed smoothness degree. Thus in our study the parameter is usually fixed to $k = 1$ while its influence is nevertheless evaluated in Section 5.2.

The problem of hyperparameter estimation is the most delicate one in regularization approaches, and has been extensively studied [14, 19, 41, 42]. Numerous techniques have been proposed and compared in these papers and two approaches seem to be of great interest. The first applied strategy, derived from the Bayesian framework, is maximum likelihood and allows estimation of both σ_u^2 and λ. The second strategy, derived from the deterministic quadratic regularization framework and called generalized cross validation provides an alternative to λ estimation, (while σ_u^2 estimation will be kept identical with in maximum likelihood approach).

3.1. MAXIMUM LIKELIHOOD. — One of the interests of the Bayesian framework is to provide coherent techniques to estimate the hyperparameters: Maximum Likelihood (ML). The most commonly employed technique consists of maximizing the marginal likelihood obtained by integrating the AR parameters out of the problem:

$$f_X(\mathbf{x}|\lambda, \sigma_u^2) = \int_{\mathbf{a}} f_{X,A}(\mathbf{x}, \mathbf{a})d\mathbf{a} \tag{26}$$

$$= \int_{\mathbf{a}} f_{X|A}(\mathbf{x}, \mathbf{a})f_A(\mathbf{a})d\mathbf{a}. \tag{27}$$

Kitagawa *et al.* [24] derived the hyperparameter Anti-Log-Likelihood (ALL):

$$\text{ALL}(\lambda, \sigma_u^2) = -2 \log f_X(\mathbf{x}|\lambda, \sigma_u^2) \tag{28}$$

$$= N \log \sigma_u^2 + \frac{1}{\sigma_u^2}\mathbf{x}^{\text{t}}(1 - M_2)\mathbf{x} - N\lambda - \det(M_1) \tag{29}$$

where the M_1 and M_2 matrices are defined by

$$M_1 = X^{\text{t}}X + \lambda\Delta_k \tag{30}$$

$$M_2 = X^{\text{t}}(X^{\text{t}}X + \lambda\Delta_k)^{-1}X = X^{\text{t}}M_1X. \tag{31}$$

The hyperparameters σ_u^2 and λ are chosen as the minimizer $\hat{\sigma}_u^2$ and $\hat{\lambda}$ of ALL with respect to σ_u^2 and λ:

$$\left(\hat{\lambda}, \hat{\sigma}_u^2\right) = \arg \min \text{ALL}(\lambda, \sigma_u^2). \tag{32}$$

This problem is a two dimensional optimization problem, but can be explicitly optimized with respect to $\hat{\sigma}_u^2$ as shown below.

3.1.1. σ_u^2 Estimation. — The ML estimate for σ_u^2 is deduced from equation (29) by nullifying its derivative with respect to σ_u^2 and gives the usual empirical estimate for the noise variance:

$$\sigma_u^2 = \frac{1}{N}\mathbf{x}^{\mathrm{t}}(1 - M_2)\mathbf{x}. \tag{33}$$

It is also to be remarked that this expression is an explicit λ function through the M_2 dependence with λ.

3.1.2. λ Estimation. — Replacing this expression in the likelihood obtained in equation (29), Kitagawa and Gersch derived in [24] the following expression for the anti-log-likelihood of λ:

$$\mathrm{ALL}(\lambda) = -\log(\det(M_1)) - N\log(\lambda) + \log(\mathbf{x}^{\mathrm{t}}(1 - M_2)\mathbf{x}). \tag{34}$$

The hyperparameter λ is then chosen as its minimizer:

$$\hat{\lambda}_{\mathrm{ML}} = \arg\min\ \mathrm{ALL}(\lambda). \tag{35}$$

No explicit expression is available for $\hat{\lambda}_{\mathrm{ML}}$ and numerical optimization is performed by calculating the ALL on a discrete logarithmic grid for λ and selecting the value which minimizes ALL.

3.2. GENERALIZED CROSS VALIDATION. — The previous method is derived in the Bayesian framework while this part is devoted to the generalized cross validation derived from the quadratic regularization interpretation presented in Section 2.5.

The Cross Validation (CV) criterion is an estimate of the mean square error, calculated from the data only [15]. The basic principle is very simple: removing one observation x_i from the data and predicting it on the basis of the regularized solution obtained from the remaining data gives a prediction error. Averaging the prediction errors over all the removed data yields an approximate mean square error, as a function of λ, called the CV criterion. The minimum of this criterion should give a good value of λ. In fact we shall minimize not CV but a modified version of CV called Generalized Cross Validation (GCV). This criterion does not differ greatly from the CV one and present more pleasant properties [15]. This criterion takes the form:

$$\mathrm{GCV}(\lambda) = \frac{\|(I - M_2)\mathbf{x}\|}{\mathrm{tr}(I - M_2)} \tag{36}$$

and λ is chosen as $\hat{\lambda}_{\mathrm{GCV}}$, the minimizer of this criterion with respect to λ:

$$\hat{\lambda}_{\mathrm{GCV}} = \arg\min\ \mathrm{GCV}(\lambda). \tag{37}$$

As in the ML case, no explicit expression is available for the minimizer of GCV and optimization must be numerically achieved.

4. Methodology for the Simulation Study

A theoretical comparative study between the performance of the different PSD estimation methods is strongly desirable, but is a very difficult task especially for finite length observation vectors. Therefore a comparative simulated study is required: Monte-Carlo experiments yields empirical statistical indices in the following way. A large number of signals have been simulated, and for each the PSD has been estimated using different methods and conditions. Indices were then calculated and averaged over a large number of realizations.
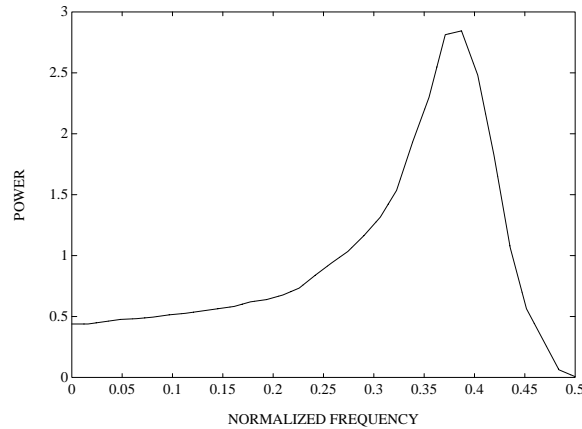
Fig. 2. — Theoretical PSD approximating the one typically seen at the level of the carotid artery at around peak systole, proposed by Mo and Cobold [33] and used in our simulation study. Vertical axis, arbitrary units. Horizontal axis, normalized frequency.

4.1. Measurement of Error and Assessment of Performance . — Following Vaitkus and Cobbold [44] two kinds of quality estimation measurement have been used. First the usual statistical estimation characteristics: bias, variance and mean square error are presented. Then three scalar indices integrating the frequency over the whole frequency domain are proposed.

The first statistical characteristic is the bias $B(f)$ which characterizes the mean difference between the estimation expectation $E(f)$ and the true value $S_x(f)$. The second one is the standard deviation $SD(f)$, which quantifies the mean variability of the estimate around $E(f)$. Finally, mean square error integrates both bias and variance and is here defined via $MSE(f)^2 = B(f)^2 + SD(f)^2$ in order to quantify an index homogeneous to the PSD.

Since the three previous indices are functions of frequency, qualitative and visual criterion are the only possible ways for comparing the different estimation method, over the whole frequency domain. This difficulty is avoided by integrating the variations of each index over the frequency range as proposed by Vaitkus et $al.$ [44] yelding: the Integrated Bias (IB), the integrated variance (ISD) and the Integrated Mean Square Error (IMSE). They are calculated by integrating respectively the bias, variance and mean square error over the whole frequency domain.

4.2. Theoretical Doppler Spectrum and Simulated Signals . — Mo and Cobbold have proposed in [33] both a particular spectral shape and a simulation method in order to generate zero-mean stationary Gaussian processes having a given PSD. The given PSD $S_x(f)$, shown in Figure 2 closely approximates the one recorded with a CW Doppler velocimeter at around peak systole under normal flow conditions. The PSD $S_x(f)$ is sampled on a fine frequency grid, each sample is multiplied by a χ^2 random variable, and the square root of the obtained samples is multiplied by a uniformly distributed complex phase. Finally the inverse Fourier transform of these samples yields the simulated signal $x(n)$. It can be shown that $x(n)$ is a second order stationary, zero-mean, Gaussian process, having the given PSD $S_x(f)$.

This spectral shape and this simulation model has been used to generate 500 signals of 256 samples. The following study was made from 16 to 256 of these samples in order to compare the methods for different data window lengths and the effect of an additive Gaussian white noise will also be investigated, for a SNR from -30 dB to 30 dB.
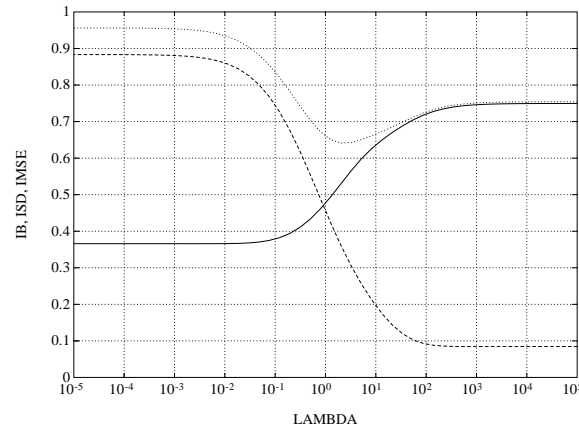
Fig. 3. — AR-RLS solution: integrated bias (IB, solid line), integrated standard deviation, (ISD, dashed line) and integrated mean square error (IMSE, dotted line) as a function of the regularising coefficient λ. Number of samples 16, SNR = 20 dB. Vertical axis, normalized. Horizontal axis, λ value.

From simulated signals, the AR coefficient vector $\mathbf{a}$ is calculated from equation (8) in the LS framework and from equation (25) in the RLS one. The PSD estimates $S_x(f)$ were finally computed from equation (12). Since the true PSD function $S_x(f)$ is known, empirical averaging gave the characteristics of interest: $B(f)$, $SD(f)$ and $MSE(f)$. Finally numerical integration yielded the three indices of interest IB, ISD and IMSE.

5. Simulation Study

Section 5.1, presents a comparison of the PSD estimation properties obtained *via* the LS method with the best model order *vs.* the RLS method with the best regularization parameter. In a second step, Section 5.2 investigates the RLS method with automatic tuning of the hyperparameter, still compared to the best RLS method. These previous studies were both achieved using a 16 samples window and a SNR of 20 dB. Section 5.3 investigates influence of window data length and noise power.

5.1. Best LS *versus* Best RLS Solution. — The first step is a study of IB, ISD and IMSE as functions of λ for the AR-RLS solution. Simulations were made on a logarithmic grid of 100 values of λ from $\lambda = 10^{-5}$ to $\lambda = 10^5$ and results are presented in Figure 3. On the one hand, for λ small enough ($\lambda < 10^{-3}$), the solution is not regularized *i.e.* no prior knowledge is introduced. So, the LS solution is entirely data-based. Therefore, as expected, the statistical properties are still those pointed out in Section 2.1: even though the bias is small, the large variance leads to unreliable results. On the other hand, for λ high enough ($\lambda > 10^3$), the solution is infinitely regularized *i.e.* data are no more taken into account and the solution is entirely prior-based, *i.e.* the estimated spectrum is constant over the whole frequency domain. Therefore, as expected, statistical properties are on the opposite: the estimation has a very low variance, but the large bias makes the results unreliable. Between these two extreme situations, IMSE shows, for $\lambda_{\min} = 2.02$, a minimum $\text{IMSE}_{\text{RLS}} = 0.64$ representing in this sense the best RLS performances (see Fig. 3).
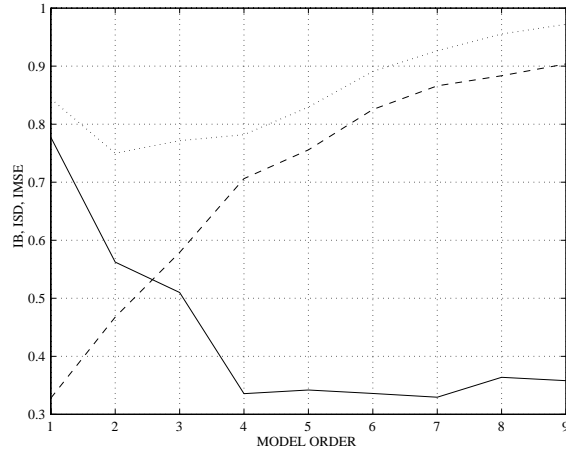
Fig. 4. — AR-LS solution: integrated bias (IB, solid line), integrated standard deviation, (ISD, dashed line) and Integrated mean square error (IMSE, dotted line) as a function of the AR model order p. Number of samples 16, SNR = 20 dB. Vertical axis, normalized. Horizontal axis, model order.

Table I. — *Integrated indices using the best LS and RLS solutions, with $\hat{\lambda}_{ML}$ and $\hat{\lambda}_{GCV}$.*

	LS	RLS-ML	RLS-GCV
IB	0.47	0.61	0.56
ISD	0.56	0.30	0.38
IMSE	0.75	0.69	0.68

In a second step, in a similar way to that just described, we studied the AR-LS solution with respect to the model order p. For each possible order ($p = 1$ to $p = 15$), the three indices, IB, ISD, and IMSE have been computed using the usual AR-LS criterion and results are given in Figure 4. Among all the possible order p, the order $p_{\min} = 2$ minimizes the IMSE and the corresponding IMSE is $\text{IMSE}_{\text{LS}} = 0.75$, which is in this sense the best LS performances.

Comparison of the above results shows that the "best" value for the LS gives an IMSE (0.75) which is 17% higher than the "best" value of IMSE for the RLS method.

5.2. COMPARISON WITH AUTOMATICALLY TUNED AR-RLS METHOD. — In the present section, the LS solution is left in this advantageous situation ($p = 2$) and compared with RLS one without knowing the best λ. For each simulated signal, the two λ estimation criteria (ML and GCV) have been computed on a logarithmic grid of 100 values of λ between 10^{-3} and 10^{3}, minima have been selected, and RLS solutions have been calculated.

For instance, from a particular signal, Figure 5 shows ML and GCV criteria, Figure 6 shows the estimated PSD with the best LS approach and with the RLS one using $\hat{\lambda}_{\text{ML}}$ and $\hat{\lambda}_{\text{GCV}}$.

Figures 7, 8 and 9 show averaged results for the best LS solution and for the RLS with automatic tuning of the spectral smoothness, *i.e.* the LS solution knowing the best model order and the RLS method estimating the regularization parameter. This comparison shows clear improvement with respect to the LS solution. More precise numerical results are given in Table I. The IMSE obtained from the regularized method, $\text{IMSE}_{\text{ML}} = 0.69$ with $\hat{\lambda}_{\text{ML}}$ and $\text{IMSE}_{\text{GCV}} = 0.68$ with $\hat{\lambda}_{\text{GCV}}$ are both smaller than IMSE_{LS} obtained for the best LS solution.
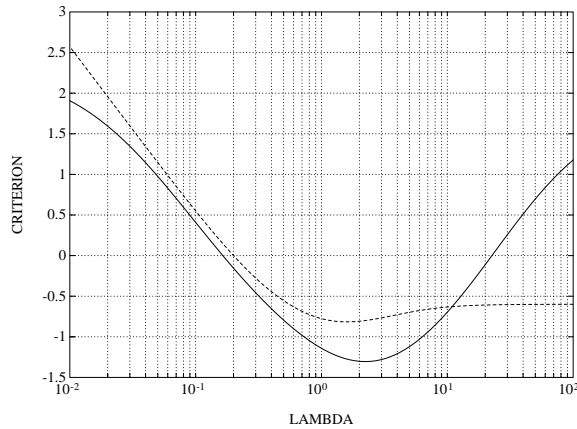
Fig. 5. — The two λ estimation criteria computed from one particular realization of the Doppler signal. Number of samples 16, SNR = 20 dB. Generalized cross validation (GCV, solid line) and maximum likelihood (ML, dashed line). Vertical axis, criterion value. Horizontal axis, λ value.
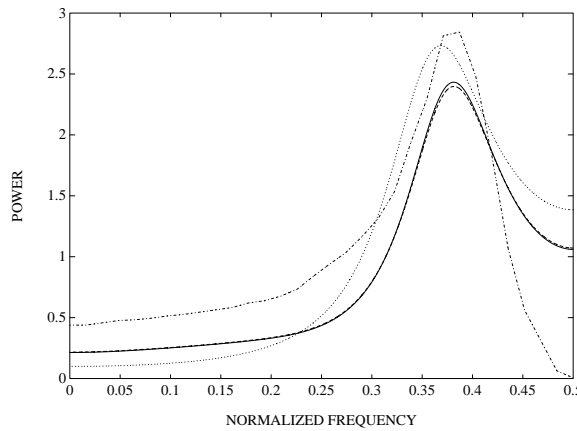

Fig. 6. — Regularized solution with $\hat{\lambda}_{\mathrm{GCV}}$ (solid line) and $\hat{\lambda}_{\mathrm{ML}}$ (dashed line) and LS one with $p = 2$ (dotted line) from the same signal as in Figure 5. The theoretical PSD is also plotted (dashdot line). Number of samples 16, SNR = 20 dB. Vertical axis, Power spectral density. Horizontal axis, normalized frequency.

Hence, despite the strong advantage given to the LS solution, the proposed one still shows a gain of 10% in terms of IMSE.

Moreover, comparing GCV and ML estimation method, one can see from Table I and from Figures 8 and 9 that the GCV method for estimating the regularization parameter yields slightly better PSD estimation quality.

Influence of the smoothness order has been investigated and Table II gives results for smoothness orders $k = 0$, $k = 1$ and $k = 2$. On the one hand, this table shows that the first order smoothness prior gives very slightly better performances in terms of IMSE, and on the other
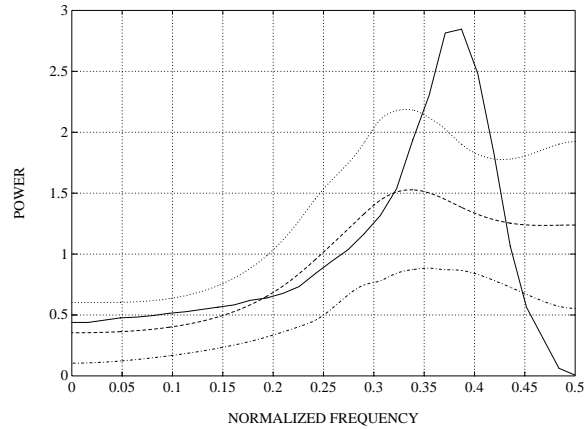
Fig. 7. — Best LS solution ($p = 2$). Actual PSD (solid line), estimator expectation $E(f)$ (dashed line), $E(f) + SD(f)$ (dotted line), and $E(f) - SD(f)$ (dashdot line). Number of samples 16, SNR = 20 dB. Vertical axis, Power spectral density. Horizontal axis, normalized frequency.
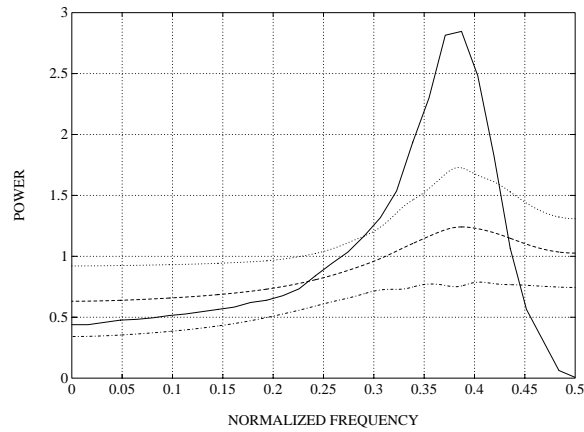

Fig. 8. — RLS solution with automatic estimation of the regularization parameter *via* ML. Actual PSD (solid line), estimator expectation $E(f)$ (dashed line), $E(f) + SD(f)$ (dotted line), and $E(f) - SD(f)$ (dashdot line). Number of samples 16, SNR = 20 dB. Vertical axis, Power spectral density. Horizontal axis, normalized frequency.

Table II. — *Integrated indices using the regularized solution for three smoothness orders $k = 0$, $k = 1$ and $k = 2$ and with the two $\hat{\lambda}$ estimation methods.*

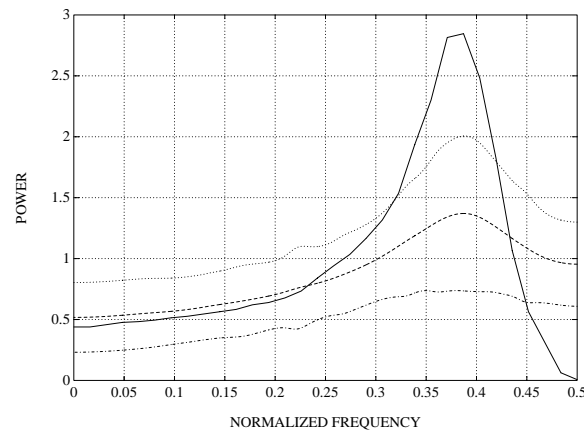|	$k = 0$	$k = 0$	$k = 1$	$k = 1$	$k = 2$	$k = 2$
	ML	GCV	ML	GCV	ML	GCV
IB	0.68	0.59	0.61	0.56	0.31	0.55
ISD	0.24	0.38	0.30	0.38	0.66	0.43
IMSE	0.71	0.70	0.69	0.68	0.72	0.70

Fig. 9. — RLS solution with automatic estimation of the regularization parameter *via* GCV. Actual PSD (solid line), estimator expectation $E(f)$ (dashed line), $E(f)+SD(f)$ (dotted line), and $E(f)-SD(f)$ (dashdot line). Number of samples 16, SNR = 20 dB. Vertical axis, Power spectral density. Horizontal axis, normalized frequency.

hand, performances show low sensivity to this parameter. Anyway, the error index always remains smaller than the best LS one.

5.3. INFLUENCE OF THE EXPERIMENTAL CONDITIONS. — Thus far the simulation results have been presented for SNR = 20 dB and 16 samples. In the following, the effects of varying these parameters are described.

5.3.1. Data Window Length. — The PSD estimate properties for different window data length ($N = 16$, 32,..., 256) are now investigated still using a SNR of 20 dB. In a first step, IMSE are computed for each window data lengths ($N = 16$, 32,..., 256), using the usual LS solution with several AR orders $p = N/2$, $N/4$,..., $N/16$. In a second step, from the same simulated signals and the same window data lengths, the IMSE are calculated using the RLS solution with both $\hat{\lambda} = \hat{\lambda}_{GCV}$ and $\hat{\lambda} = \hat{\lambda}_{ML}$ and with the first smoothness order. Figure 10 shows the IMSE curves as a function of the window data length. From these results it can be seen that the RLS procedure using either method for optimizing λ gives the smallest IMSE.

5.3.2. Signal to Noise Ratio. — This section investigates the influence of noise power from -30 dB to 30 dB, using a window with the same number of samples (16) as previously. On the one hand, using the LS solution for the best model order, IMSE are calculated for several signal to noise ratio from -30 dB to 30 dB. On the other hand, in the same conditions, but using the RLS solution with first smoothness order and both $\hat{\lambda} = \hat{\lambda}_{GCV}$ and $\hat{\lambda} = \hat{\lambda}_{ML}$, the IMSE are computed. IMSE curves as a function of signal to noise ratio are plotted in Figure 11. One can notice slightly better performance for the ML method for a negative SNR, and better performances for the GCV method for a positive one. Anyway, simulations results show that for all SNR the regularized solution yields better estimation qualities in terms of IMSE.
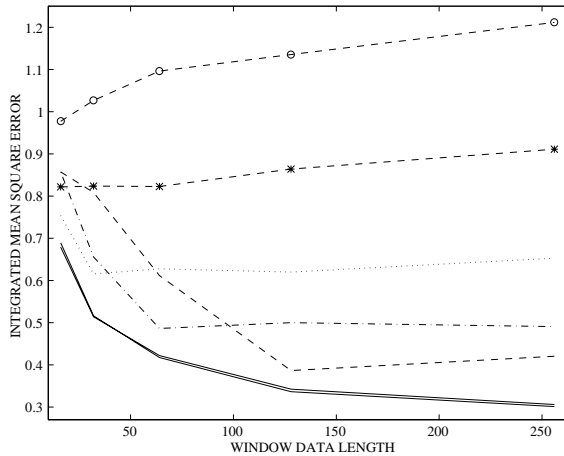
Fig. 10. — IMSE for different model orders as a function of window data length, in the LS framework. Solution with $p = N/2$ (circle-line) $p = N/4$ (star line) $p = N/8$ (dotted line) $p = N/16$ (dash-dotted line) $p = N/32$ (dashed line). RLS solution with $\hat{\lambda} = \hat{\lambda}_{\mathrm{ML}}$ and $\hat{\lambda} = \hat{\lambda}_{\mathrm{GCV}}$ (solid lines almost coincident). Number of samples 16 - 256, SNR = 20 dB. Vertical axis, IMSE. Horizontal axis, window length.
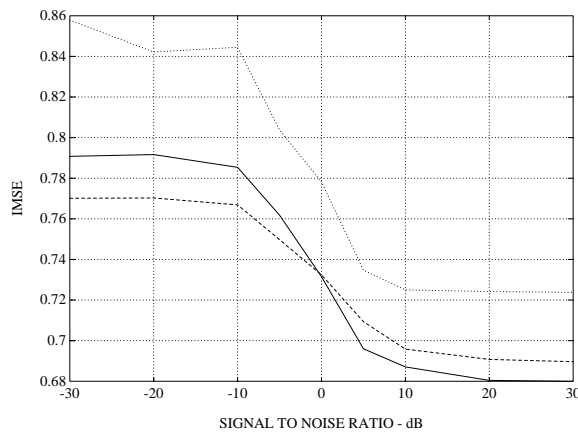

Fig. 11. — IMSE as a function of signal to noise ratio in dB. LS solution (dotted line), RLS solution with $\hat{\lambda}_{\mathrm{GCV}}$ (solid line), and RLS solution with $\hat{\lambda}_{\mathrm{ML}}$ (dashed line). Number of samples 16. Vertical axis, IMSE. Horizontal axis, SNR (dB).

6. Fast Suboptimal Spectral Analysis

The proposed estimation methods for the hyperparameters σ_u^2, k, and λ are not suitable for practical applications. In addition, the matrix inversion in equation (25) in turn can still be time consuming. The possible improvements of these two problems are now addressed.

6.1. Hyperparameter Determination

6.1.1. Smoothness Order. — Estimation of the parameters of the regularized model requires that the matrix of equation (25) be inverted. In the general case, this matrix has no specific properties and the inversion algorithm requires about $O(N^3)$ multiplications. We propose to choose a zero-order smoothness constraint ($k = 0$) for which the resulting Δ_k matrix becomes the identity matrix, thus the parameters of the AR model are now given by:

$$\hat{\mathbf{a}} = (X^{\mathrm{t}}X + \lambda I)^{-1} X^{\mathrm{t}}\mathbf{x}. \tag{38}$$

The matrix $X^{\mathrm{t}}X + \lambda I$ has a Toeplitz structure. It can then be inverted using fast $O(N^2)$ algorithms based on Levinson recursions. Practical applications can then be contemplated. Demoment *et al.* [12] derived an efficient algorithm providing a possible solution based on a fast Kalman filter. A complete description of this algorithm applied to ultrasound signals can be found in [28].

6.1.2. Noise Variance. — The value σ_u^2 only modifies the spectrum amplitude as indicated in equation (12) but has no influence on the spectrum shape. Because, in medical spectrum presentation, a scaling of the Doppler spectra is generally applied before presentation (the maximum value of each spectrum is set to the maximum white level), we decided to use, subsequently, a constant value: $\sigma_u^2 = 1$. An individual scaling of the amplitude of each spectrum in the range [0, 256] was then used for presentation. It accounts for the vertical stripes in Figures 12 and 13 corresponding to low energy spectra artificially enhanced by this mode of data presentation.

6.1.3. Regularization Parameter. — Once these two hyperparameters are set, the problem of the best balance between data and smoothness priors needs to be addressed. From equation (23), the value of $Q_0(\mathbf{a})$ is a quadratic function of $\mathbf{x}$, while that of $Q_\infty(\mathbf{a})$ is independent of $\mathbf{x}$. Consequently, the value of λ (Eqs. (17), (22)) must take into account the signal energy in order to maintain a constant spectral smoothness, whatever the Doppler signal amplitude (which is highly variable, even within the same cardiac cycle). A simple scaling factor can be obtained from the energy of the Doppler signal within the window:

$$\hat{\lambda} = \mathbf{x}^{\mathrm{t}}\mathbf{x}\lambda_0 \tag{39}$$

λ_0 can be defined before hand and once for all, for a class of Doppler signals using for instance the methods proposed in Sections 3.1 and 3.2.

6.1.4. Computation Time. — Using such estimation of the λ value, a rapid computation of the parameter vector is made possible. The spectrum itself (Eq. (12)) can also be rapidly calculated using an FFT algorithm. Using a DEC alpha 3000 the overall spectrum calculation written in MATLAB code needed less than 25 and 100 ms for 16 and 32 sample windows respectively. This results indicates that with specialized hardware it would be possible to design a real time Doppler spectrum analyser.

6.2. Results of Fast Spectral Estimation.
— Pulsatile flow was obtained using an oscillating pump whose frequency was adjusted at 70 cycles per minute. This pump was connected to a discharge tank through an inflow ball valve connected to the discharge tank and to the test circuit through an outflow ball valve. The ejected fluid flowed through a divergent-convergent structure in order to dissipate the upstream disturbances before entering the 30 mm diameter nozzle. The stenose orifice was circular with a diameter of 7.1 mm. This design generated a flat velocity profile at the origin of the jet. This tube was connected
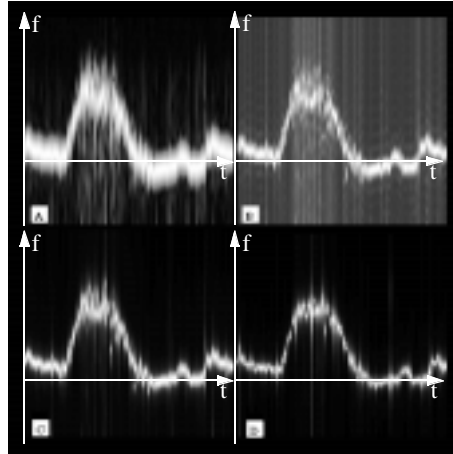
Fig. 12. — Spectra of the *in vitro* laminar flow in the jet core downstream a stenosis (7.1 mm in diameter). Linear grey scale. PRF = 10 kHz, probe frequency 3 Mhz, focus 16 mm, sample volume length < 1 mm, sample volume width < 5 mm. A) periodogram (FFT and Hanning window); B) fast RLS estimation ($p = 32$, $k = 0$); C) RLS estimation ($p = 32$, $k = 1$); D) LS estimation ($p = 4$). Horizontal axis, time (0.7 s from left to right). Vertical axis, Doppler frequency. Grey scale power spectral density.
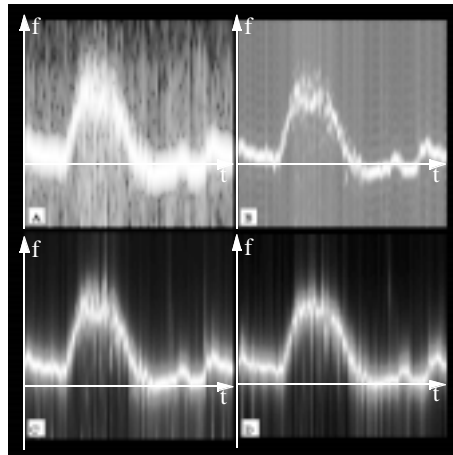

Fig. 13. — Spectra of the *in vitro* laminar flow in the jet core downstream a stenosis (7.1 mm in diameter). Log grey scale. PRF = 10 kHz, probe frequency 3 MHz, focus 16 mm, sample volume length < 1 mm, sample volume width < 5 mm. A) periodogram (FFT and Hanning window); B) fast RLS estimation ($p = 32$, $k = 0$); C) RLS estimation ($p = 32$, $k = 1$); D) LS estimation ($p = 4$). Horizontal axis, time (0.7 s from left to right). Vertical axis, Doppler frequency. Grey scale log power spectral density.

to adjustable downstream resistance and compliance which allowed, together with upstream resistance and compliance, the adjustment of the velocity waveform. The fluid was a mixture of water, glycerol and corn starch. The 30% concentration of glycerol provided a viscosity of 0.04 P at room temperature, while the 5% corn starch provided scatterers for the Doppler ultrasound measurements.

The velocity measurements were performed using a commercial ultrasound Doppler velocimeter modified for *in-vitro* studies. The probe angle was 70 degrees with respect to flow axis and the ultrasound beam was focused at 1.6 mm *i.e.* on the flow axis. A 3 MHz r.f. frequency was selected and the pulse repetition frequency was set to 10 kHz in order to allow the measurement of velocity without aliasing and to preserve both reasonable longitudinal and lateral resolutions with respect to the axis of the ultrasonic beam. Using this set up, the sample volume dimensions were less than one millimeter along the tube diameter direction and less than 5 mm along the tube axis. The sample volume was thus entirely located inside the laminar jet core so that the dispersion of velocities was extremely small in the sample volume.

A 32 samples window (3.2 ms) was used to comply with the stationary hypothesis in the window for spectral analysis. Results are presented in Figures 12 and 13 using the usual spectrum presentation, where the horizontal axis corresponds to the time along the flow cycle, the vertical one to positive and negative flow velocities and the grey level to the spectrum amplitude.

Figures 12 and 13 show in A, B, C and D the results respectively given by the FFT based periodogram associated with a Hanning window, the zero order smoothness AR-RLS, the first order smoothness AR-RLS. and the classical AR-LS estimation for $p = 4$ (lowest IMSE), The proposed AR-RLS method presents a good compromise between the smeared, low resolution spectrum of the Fourier approach and the spiky spectrum of the LS method which cannot give correct information on the spectrum bandwidth. The first order smoothness offers the best description of the Doppler spectrum morphology. The smearing effect of the Fourier approach has been removed while the spiky aspect of pure LS spectra has disappeared. This reduction of spectrum bandwidth can be considered, in the present application, as an improvement of the spectrum quality. Because the sample volume, was entirely embedded in the laminar core of the jet on the flow axis, the flow dispersion was very small. Because the sample volume was located at the focus of the ultrasound beam, the spectrum distorsion due to geometrical effects was reduced. Furthermore, because the AR approach identifies the model from the avaible samples within the window but does not assume that the Doppler signal must be zero outside the Doppler window, as does the FFT approach, the spectrum distorsion due to the transit time effect was removed. The zero order smoothness offers comparable improvements. The above developments indicate in this situation, more spiky spectra. However, using the usual time frequency representation, help the observer to compensate for this effect and still provides very legible results.

7. Conclusion

We have addressed the problem of ultrasound Doppler spectral estimation when only a short span of data is available for analysis (down to 16 observations). In such a situation usual autoregressive estimation strategies, such as empirical correlation estimation or least squares approach, enforce the estimation of a very small number of parameters which precludes the description of a large class of power spectral densities. The Bayesian approach presented here and initially proposed by Kitagawa and Gersch [24] alleviates this limitation. It admits the robust estimation of long autoregressive parameter vectors (typically 16 parameters from 16 observations) and avoids too spiky spectra. The proposed comparative study is grounded on both simulated and *in vitro* data and shows improvement with respect to the usual approach.

In the simulation study, we have compared two estimation methods of the fundamental regularization parameter: maximum likelihood and generalized cross validation methods in terms of recovering a known power spectral density. The conclusion of this study is that the generalized cross validation method performs slightly better than the maximum likelihood

method. Nevertheless, this result should be taken with caution since the observed difference is very slender.

While the first part of the simulation study deals with hyperparameter estimation, in the second part, comparison of the presented Bayesian method *versus* the usual one has been achieved in different situations, varying the data span length and the signal-to-noise ratio. For that purpose, bias, standard deviation and mean square error, have been estimated using a Doppler signal model. These results confirm and extend those of Kitagawa and Gersch. It was found that the Bayesian method with automatic tuning of the hyperparameters yields better indices than the usual least squares method whatever the model order.

Finally, the *in vitro* data processing also shows interesting improvements, especially better legibility of the spectra, hence offers better description of the Doppler spectrum morphology than the usual least squares one.

Acknowledgments

References

[1] Akaïke H., Statistical predictor identification, *Ann. Inst. Statist. Math.* **22** (1970) 203-217.

[2] Akaïke H, A new look at the statistical model identification, *IEEE Trans. Autom. Contr.* **19** (1974) 716-723.

[3] Angelsen B.A.J., A theoretical study of the scattering of ultrasound from blood, *IEEE Trans. Biomed. Eng.* **27** (1980) 61-67.

[4] Ata O.W. and Fish P.J., Effect of deviation from plane wave conditions on the Doppler spectrum from an ultrasonic blood flow detector, *Ultrasonics* **29** (1991) 395-403.

[5] Barbaresco F., Super Resolution Spectrum Analysis Regularisation: Burg, Capon and Ago-antagonistic Algorithms, EUSIPCO-96 (1995) pp. 2005-2008.

[6] Barbaresco F., Algorithme de Burg régularisé FSDS, comparaison avec l'algorithme de Burg MFE, *GRETSI* **1** (1995) 29-32.

[7] Bascom P.A.J. and Cobbold R.S.C., Effects of transducer beam geometry and flow velocity profile on the Doppler power spectrum: A theoretical study, *Ultrasound. Med. Biol.* **16** (1990) 279-295.

[8] Box G.E.P. and Jenkins G.M., Time series analysis: Forecasting and control (San Francisco: Holden Day, 1976).

[9] Cadzow J.A., Spectral estimation: An overdetermined rational model equation approach, *Proc. IEEE* **70** (1982) 907-939.

[10] Censor D., Newhouse V.L., Vontz T. and Ortega H.V., Theory of ultrasound Doppler-spectra velocimetry for arbitrary beam and flow configurations, *IEEE Trans. Biomed. Eng.* **35** (1988) 740-751.

[11] Cohen L., Time-Frequency Distribution A Review, *Proc. IEEE (77)* **7** (1989) 941-981.

[12] Demoment G., Houacine A., Herment A. and Mouttapa I., Adaptive Bayesian spectrum estimation. Fourth IEEE ASSP Workshop on Spectrum Estimation and Modeling (1988) pp. 33-38.

[13] Demoment G., Image reconstruction and restoration: overview of common estimation structures and problems, *IEEE Trans. Acoust. Speech Signal Processing* **37** (1989) 2024-2036.

[14] Fortier N., Demoment G. and Goussard Y., Comparison of GCV and ML methods of determining parameters in image restoration by regularisation, *J. Vis. Comm. Image Represent.* **4** (1993) 157-170.

[15] Golub G.H., Heath M. and Wahba G., Generalized cross validation as a method for choosing a good ridge parameter, *Technometrics* **21** (1979) 215-223.

[16] Guo Z., Durand L.G., Allard L., Cloutier G., Lee H.C. and Langlois Y.E., Cardiac Doppler blood flow signal analysis. Part 1 Evaluation of the normality and stationarity of the temporal signal, *Med. and Biol. Eng. Comput.* **31** (1993) 237-242.

[17] Guo Z., Durand L.G., Allard L., Cloutier G., Lee H.C. and Langlois Y.E., Cardiac Doppler blood flow signal analysis. Part 2 Time-frequency representation based on autoregressive modeling, *Med. and Biol. Eng. Comput.* **31** (1993) 243-248.

[18] Guo. Z., Durand L.G. and Lee H.C., Comparison of Time-Frequency distribution Techniques for Analysis of Simulated Doppler Ultrasound Signals of the Femoral Artery, *IEEE Trans. Biomed. Eng.* **41** (1994) 332-342.

[19] Hall P. and Titterigton D.M., Common structures for choosing smoothing parameters in regression problems, *Journal of Royal Statistical Society* **49** (1987) 184-198.

[20] Kaluzynski K., Analysis and application possibilities of autoregressive modeling to Doppler blood flow signal spectral analysis, *Med. Biol. Eng. Comp.* **25** (1987) 73-76.

[21] Kay S.M. and Marple S.L., Spectrum analysis - a modern perspective, *Proc. IEEE* **69** (1981) 380-1419.

[22] Kay S.M., Recursive maximum likelihood estimation of autoregressive processes, *IEEE Trans. Acoust. Speech Signal Processing* **21** (1983) 56-65.

[23] Kay S.M., Modern spectral estimation, theory and applications (Prentice Hall, 1988).

[24] Kitagawa G. and Gersch W., A smoothness priors long AR model method for spectral estimation, *IEEE Trans. Autom. Contr.* **30** (1985) 57-65.

[25] Kitney R.I. and Giddens D.P., Analysis of blood flow velocity waveforms by phase shift averaging and autoregressive spectral estimation, *J. Biom. Eng.* **105** (1983) 398-401.

[26] Law Y.F., Bascom P.A.J., Johnston K.W.P. and Cobbold R.S.C., Experimental study of the effects of pulsed Doppler sample volume size and position on the Doppler spectrum, *Ultrasonics* **29** (1991) 404-410.

[27] Ligget W.S., Passive sonar: Fitting models to multiple time series, NATO ASI Signal Processing (1972) pp. 327-345.

[28] Lymberis A., Herment A., Demoment G. and Fric C., Estimation of frequency dependent attenuation based on parametric spectral analysis and correlation lags of the demodulated echo signal, *Ultrason. Imag.* **13** (1991) 1-26.

[29] Markel J.D. and Gray A.H., Linear prediction of speech (Springer-Verlag, 1976).

[30] Marple S.L., Digital Spectral Analysis with Applications (Prentice Hall, 1987).

[31] Marple L., A new autoregressive spectrum analysis algorithm, *IEEE Trans. Acoust. Speech Signal Processing* **28** (1980) 429-433.

[32] Mo L.Y.L. and Cobbold R.S.C., A stochastic model of the backscattered Doppler ultrasound from blood, *IEEE Trans. Biomed. Eng.* **33** (1986) 20-27.

[33] Mo L.Y.L. and Cobbold R.S.C., Speckle in continuous wave Doppler ultrasound spectra: A simulation study, *IEEE Trans. Ultras. Ferro. and Freq. Control* **33** (1986) 747-753.

[34] O'Donnell M. and Flax S.W., Phase aberration measurements in medical ultrasound: Human studies, *Ultrason. Imag.* **10** (1988) 1-11.

[35] Pao Y. and Lee D.T., Performance characteristics of the Cadzow modified direct ARMA method for spectrum estimation. First IEEE ASSP Workshop on Spectrum Estimation (1981) 2.5.1-2.5.10.

[36] Parzen E., An approach to times series modeling and forecasting illustrated by hourly electricity demands. Tech. Rep. 37, Statistical Science Division, State University of New York (1976).

[37] Peronneau P., Herment A., Moutet J.-P., Xhaard M., Piechocki M. and Guglielmi J.-P., Doppler signal analysis. Cardiovascular applications of Doppler echography, P. Peronneau and B. Diebold, Eds. (Paris, INSERM Pub., 1982) pp. 81-112.

[38] Pham Dinh Tuan, Maximum likelihood estimation of the autoregressive model by relaxation on the reflection coefficients, *IEEE Trans. Acoust. Speech Signal Processing* **36** (1988) 1363-1367.

[39] Schlindwein F.S. and Evans D.H., Selection of the order of autoregressive models for spectral analysis of Doppler ultrasound signals, *Ultrasound Med. Biol.* **16** (1990) 81-91.

[40] Talhami H.E. and Kitney R.I., Maximum likelihood frequency tracking of the audio pulsed Doppler ultrasound signal using a Kalman filter, *Ultrasound Med. Biol.* **14** (1988) 599-609.

[41] Titterington D.M., Common structure of smoothing techniques in statistics, *International Statistical Review* **53** (1985) 141-170.

[42] Thompson A.M., Brown J.C., Kay J.W. and Titterington D.M., A study of choosing the smoothing parameter in image reconstruction by regularization, *IEEE Trans. Pattern Anal. Mach. Intel.* **13** (1991) 326-339.

[43] Vaitkus P.J. and Cobbold R.S.C., A comparative study and assessment of Doppler ultrasound spectral estimation techniques. Part I: Estimation methods, *Ultrasound Med. Biol.* **14** (1988) 661-672.

[44] Vaitkus P.J., Cobbold R.S.C. and Johnston K.W., A comparative study and assessment of Doppler ultrasound spectral estimation techniques. Part II: Methods and results, *Ultrasound Med. Biol.* **14** (1988) 673-687.

[45] Yamaguchi T., Kikkawa S. and Parker K.H., Simulation of non-stationary spectral analysis of turbulence in the aorta using a modified autoregressive or maximum entropy method, *Med. Biol. Eng. Comp.* **25** (1987) 533-542.