SIGNAL PROCESSING III: Theories and Applications
I.T. Young et al. (editors)
Elsevier Science Publishers B.V. (North-Holland)

THE VARIATIONAL APPROACH IN SPECTRAL ESTIMATION

Miguel A. Lagunas

E.T.S. Ingenieros de Telecomunicación. U.P.C.
Procesado de Señal en Comunicaciones.
Apdo. 30002. 08071 BARCELONA - SPAIN.

SUMMARY

This is a very personal point of view of the underlying ideas that yields most of the currently reported spectral estimation techniques, procedures and algorithms. After a brief introduction on non-parametric spectral estimation, which includes a shutle aspect about how to use averaging in DFT based methods, the paper describes the potential of the variational approach in deriving already reported estimates and the way out to obtain new ones. Finally, and as the second approach of high interest in spectral estimation, the design and extensions of data-dependent filters for spectral estimation is reported.

I. INTRODUCTION

During the last decade a lot of work has been done in the field of spectral estimation. But it seems to me that there are not many reported works which can provide an adequate framework for the topic. In fact, some papers reported as tutorials or with some tutorial value look like as a collection of methods with a questionable rank for them; based in very concrect examples from concrect signals that, of course, does not make valuable the classification. For interested readers in spectral estimation usually each signal or signals under analysis represents a new problem, in such a way that a very expert answer for a spectral estimation problem will be something like "tell me what you need and I would tell you what is the best solution". There are two topics which are very important nowdays, one is a general framework to encompass most of the reported methods; and, in this way, providing certain knowledge about how to derive new estimates. The second one concerns with objetive measures of quality, different from correlation bias, resolution, etc., that will tell us the overall performance that the estimate yields.

This paper will address the first item by describing the variational approach for spectral estimation both in the frequency domain and in the time domain. A brief introduction about non-parametric procedures is included due to its interest in estimating the prior information in a spectral estimation procedure (i.e. the autocorrelation function or the averaged periodogram function). The variational approach in frequency and in the time domain will follow in sections III and IV.

No figures or examples are included and reference list has been minimized in accordence to the main objective of this work, which is not to report any procedure or any concrect reference to already reported methods.

II. NON-PARAMETRIC SPECTRAL ESTIMATION

Regardless this work was done thinking in parametric spectral estimation, non-parametric procedures deserves an important consideration, since they could be viewed as a prior-estimation step in all the techniques which starts from an estimate of the data autocorrelation function. It is important to notice that this first step, in general, should not be considered as a data reduction stage, in the sense that we could obtain the same number of data samples of the a.c.f. than we have in the data sample record. All the a.c.f. lags contain potential information to be used in the spectral estimation procedure. Of course, it is easy to conclude that the high variance, associated to the a.c.f.

values obtained in a N to N computation basis from the N data samples, will promote a low statistical stability of the resulting estimate; but, when the main concern is resolution, even as a cosmetic factor as in angle of arrival detection or Pisarenko like methods, long correlation sequences should be used at the expense of a poor quality in the stabily sense.

At his point is word to take in mind that the claimed superior quality of direct data methods over correlation methods refers only with linear systems models of the procedure. In other words, minimum phase properties of the model, analysis/synthesis problems, or any other constraint related with deterministic signal generation models will produce the difference between the two approaches. In a general contest, and having available the currently reported algorithms, correlation based methods can provide at least the same performance that direct data methods.

Furthermore, in may cases the average involved in computing the a.c.f. estimate could be the best candidate for a first step noise reduction.

Thus, focussing the problem of an estimate for the autocorrelation function in a spectral estimation problem, or the cross-correlation function in a cross-spectrum procedure, the WOSA estimate have to be mentioned. As it is well-known by the reader this procedure reported by Welch many years ago consist in a partitioning of the original data record x(n) (n=0, N-1) in records of length M, overlapped over D samples and weigthed by a data window $w_x(n)$ if desired ... Then we could define $\tilde{x}_q(n)$ and $\tilde{y}_q(n)$ as indicated in (1).

$$\tilde{x}_q(n)=x(q(M-D)+n)w_x(-\frac{M}{2}+n) \quad n=1,M$$

$$\tilde{y}_q(n)= y(q(M-D)+n)w_y(-\frac{M}{2}+n) \quad n=1,M \qquad (1)$$

o elsewhere.

The WOSA cross or (y equal to x) autospectrum estimate is computed as shown in (2).

$$P_{xy}(l) = \frac{1}{K_o} \sum_{q=1}^{Q} X_q^*(l).Y_q(l) \qquad (2)$$

Where, K_o is a factor, depending on the overlapping length D and windows $w_x(.)$ and $w_y(.)$, and it is set in order to remove the bias; $X_q(l)$ is the L points DFT of the original M points length sequence $\tilde{x}_q(n)$ padded with L-M zeros if desired; and Q the number of records obtained from the original records of N samples each.

At this point two things deserve more attention, one is a formal question, the other is a natural but not usual extension of (2) which produces additional averaging highly desired in parametric methods for very short data records. The formal question is about the constant K_o. This constant does not remove the bias of P(w) with respect the actual cross or autospectrum (let us say it is $S_x(w)$. P(w) is always biased since its expected value is a convolution of the effective lag window with the true power spectrum density. Constant K_o is set just to constraint that the inverse Fourier transform of $P_{xy}(l)$ will be the zero lag value of the initial the c.c.f. or the a.c.f.

The second question is the so-called STUSE method |1|. Bassically, it can be stated as that, there are not any problem in considering the crossproduct of $X_q^*(l)$ with $Y_p(l)$ in the averaging process, whenever a delay correction factor is used prior the term is included in (2). Thus a more general, and lower variance than (2), estimate is (3), where these new terms are under consideration.

$$P_{xy}(l)=\frac{1}{K_o} \sum_{p,q=1,Q} X_q^*(l)Y_p(l). \cdot \exp(-j\frac{2\pi l}{L}(p-q).(N-D)) \qquad (3)$$

Once $P_{xy}(l)$ is obtained, the IDFT of it will provide the desired estimate of the cross-correlation or the autocorrelation function. A quadratic window or lag-reshape methods could be used if desired in this first step of our parametric estimate algorithm. I would like to recall the importance of the STUSE method when compared with classical WOSA method chiefly when the number of samples for $r_{xy}(n)$ or $r_{xx}(n)$ to be used in the parametric procedure is high when compared with the data record length N.

To end with this section, I would like to insist in that all the information is in $P_{xy}(l)$ or $P_{xx}(l)$ l=0, L-1; and any procedure which do not use the overall L samples will miss something in the way out to obtain the final spectral estimate. Of course this statement does not mean that all the L samples of $P_{xy}(l)$ or the associated $r_{xy}(n)$ have to play the same role in the parametric spectral estimation technique. It is just to ask for the importance of functions, which constraining the spectral estimation procedure, describe the global behavior of the cross or autocorrelation sequence.

III. VARIATIONAL APPROACH FOR SPECTRAL ESTIMATION

One of the most powerfull tool to derive parametric models for power spectrum estimates is to set some objective function $\emptyset(S_x(w))$ or $\emptyset(S_x)$, and minimize or maximize this objective constrained by the relevant information we have available from the prior-estimate $P_x(w)$, in many cases the signal periodogram or the STUSE estimate.

It could be infered that two important decisions have to be taken under such approach. The first one deals with the objective function $\emptyset(S_x)$. The most familiar one is the entropy as shown in (3).

$$H_o = \frac{1}{2\pi} \cdot \int_{-\pi}^{\pi} Ln\ S_x(w)\ dw \qquad (3)$$

Which corresponds to the exact expression of it when x(n) (n=0,N-1) belongs to a gaussian and stationary random process. This is a nice justification with an adequate and theoretical mathematical background to the wide use of the entropy as objective in spectral estimation. Many words can be written in this topic about the important role of entropy in spectral estimation. Any case, I prefere to pay attention to its features when the zero lag correlation constraint is set. When H_o is maximized under the constraint (4), the resulting estimate is a flat estimate.

$$\frac{1}{2\pi} \cdot \int_{-\pi}^{\pi} S_x(w)dw = r_x(0) = \frac{1}{2\pi} \cdot \int_{-\pi}^{\pi} P_x(w)dw \qquad (4)$$

This is, as concerns with a general background, a desired feature of the objective, (i.e. it produces a maximum flatness estimate coherent with the constraints). In other words, maxima and minima in the estimate will depend on the constraints because the objective under the power constraint will try to relent a peaky behavior in the estimated power density. Once the entropy is viewed in this way, it is clear that other objectives could deserve attention. A brief list of them could be the following one:

$$H_1 = \int \exp\ S(w)\ dw \qquad (5\text{-}a)$$

$$H_2 = \int S\ (w)\ Ln\ S(w) \qquad (5\text{-}b)$$

$H_3 = \int S^n(w)\ dw$ for $n \neq 1,0$ not necessarily an integer number. An special case of interest is n=2 (5-c)

$$H_4 = \int \left(\frac{dS}{dw}\right)^2 dw \qquad (5\text{-}d)$$

Further discussion about these objective functions needs the constraints we desire to set to the desired spectral estimate in order to be included in finding a maximum or minimum of these objectives. To better report the role of the constraints is worthwhile to solve the general problem of an objective with some constraints.

objective $\int \emptyset(S)dw$ (6-a)

constraints $\int \Psi(S)e^{jqw}dw = \rho_q$; $|q| < Q$ (6-b)

Note that finding the constraints does not face any problem since they can be computed from the periodogram $P_x(w)$ as in (7) using the same function $\Psi(\cdot)$ we used in (6-b).

$$\rho_q = \int \emptyset(P(w))e^{jnw}dw \ ; \ |q| < Q \qquad (7)$$

Now, forming the Lagrangian $\mathcal{E}$ with a multiplier λ_q for every constraint, (8) results;

$$\mathcal{E} = \emptyset(S) - \sum_{|q|<Q} \lambda_q \cdot \Psi(S)e^{jqw}$$

seting derivative of the Lagrangian with respect to S equal to zero, the estimate is obtained as a function of the Lagrange parameters.

$$\frac{d\pounds}{dS} = \emptyset(S) - \sum_{|q|<Q} \lambda_q e^{jqw}\dot{\Psi}(S) = 0 \qquad (8)$$

or

$$\frac{\dot{\emptyset}(S)}{\dot{\Psi}(S)} = \sum_{|q|<Q} \lambda_q . e^{jqw} = \lambda(w) \qquad (9)$$

Before to go further in the discussion the reader should remind that (8) must be changed when the objective is as (5-d) or $\emptyset$ depends on high order (greather than zero) derivatives.

Thus (9) reveals that when correlation lags without missing points are set as constraints in the variational procedure, i.e. $\emptyset(S)=S(w)$ and $\dot{\Psi}(S)=1$ and $\rho_n=r_n$ the data autocorrelation lags, any of the previous objectives will produce a polynomial modeling or a linear model in the inverse Fourier domain, for the derivative of the objective function.

Case 1 (5-a) $S(w)=Ln(\lambda(w))$

the correlation is extrapolated from the cepstrum of $\lambda(w)$.

Case 2 (5-b) $S(w)=\exp(\lambda(w)-1)$

the cepstrum of the estimate is of finite length equal to the number of correlation constraints.

Case 3 (5-c) $S(w)=\left(\frac{\lambda(w)}{n}\right)^{1/n-1}$

for n=2 an all-zero or MA model results as a structural constraint for the desired estimate. This is the formal way to justify the B-T estimate, because $\rho_q=r_q$. This proves that the B-T estimate is a minimum energy for the extrapolated a.c.f. This is a trivial conclusion but alow to include the prior-estimate B-T in this framework.

Case 4 (3) Classical Burg estimate with an AR model as structural constraint for the resulting estimate. $S(w)=1/\lambda(w)$.

From the previous paragraphs it results clear that any claim, in a realistic or practical approach, about the adequateness or not of a given objective has no sense without knowing the set of constraints we are going to use.

The variety of models or structural constraints for the estimate provides support to the idea that constraints deserves more attention that the objective itself in a variational approach for parametric spectral estimation.

Thinking deeply in the role of constraints in these techniques, we can conclude that, when selecting then, we are facing the hard problem, or taking the hard decision, about wheter or not the kind and number of constraints, we set in the procedure, reflects the overall knowledge we have from the data record sample. $x(n)$; n=0, N-1.

It is worthwhile to note that the number of constraints we use, is usually less than the number of data samples N. It could be said that the number of independent samples in the data record is probably less than N. In other words, that Q data constraints summarize quite well all the information of the original data sample and, as a consequence, we can be confident with the selected data constraints. Thus there is an uncetainly that could be stated as follows: When the first Q data autocorrelation lags are all the information we need from the data sample in order to obtain the actual density estimate? or, taking in mind that the second order information r_q steems from an estimation procedure, a close estimate to the actual spectrum results.?

The answer for this question is just that we are doing rigth whenever the structural form of the estimated presented above as cases 1,...,4 are in fact valid for the actual estimate (i.e. selecting case 1; when a finite order MA model for exp(S(w)) obeys to a correct assumption for the actual power density spectrum of the random process to which data record x(n) belongs).

When there is not prior knowledge about the signal model it is clear that as much information we reflect in the constraints better the performance of the procedure will be. This last sentence points out that, once a given objective has been selected, correlation constraints are not the only choice to form them in a variational procedure. Other second order information like real cepstrum o third order functions can be set as constraints in addition to the correlation ones in order to improve the resulting quality of the associated spectral estimate. At the same time, it looks realistic to include more data constraints to the, lets say, a successive correlation lags, when it is clear that high correlation lags behavior play an important role in some interesting features of the spectral estimate as resolution or bias. Also the problem addressed in |2|, |3| are interesting at this point.

One of the most interesting cases under such philosophy is the inclusion of cepstrum and correlation constraints in maximum entropy problem |4|. The problem is stated as it is shown below:

Maximize $\int \mathrm{Ln}\ S(w)dw$ for $S(w)$

constrained to

$$\int S(w)\exp(jnw)dw = r(n);\quad |n|\leqslant Q \tag{10}$$

and

$$\int \mathrm{Ln}\ S(w)\exp(jmw)dw = c(m);\quad |m|\leqslant P,\ m\neq 0$$

The Lagrangian is

$$\mathrm{Ln}\ S.(1+\sum_{\substack{m\neq 0\\ |m|\leqslant P}} \mu_m.\exp(jmw)) + S(w)((\sum_{|n|\leqslant Q} \lambda_n.\exp(jnw))) \tag{11}$$

Seting derivation with respect S(w) equal to zero the new estimate results;

$$S(w) = \frac{1+\sum \mu_m.\exp(jmw)}{\sum \lambda_n.\exp(jnw)} \tag{12}$$

Where the reader can see that all the classical models for spectral estimation, i.e. AR, MA and ARMA, can be encompassed in this framework.

Anycase the important conclussion is up to what degree a free look of the variational procedure can open new ways to perform well supported parametric spectral estimation. It is worthwhile to note that parametric modelling is a concept that, as concerns with spectral estimation cannot be longer applied only to the zero-pole modelling of the power spectrum. Let see an example of this statement. Let us suppose the following problem.

Maximize $\int S(w)\mathrm{Ln}\ S(w)dw$ for $S(w)$

constrained to

$$\int S(w)\mathrm{Ln}\ S(w)e^{jnw}dw;\ |n|\neq 0, \leqslant P \qquad \int S(w)e^{jmw}dw;\ |m|\leqslant Q \tag{13}$$

The solution arises to an ARMA (P,Q) model for the real cepstrum of the resulting estimate. In general, an ARMA model results for $\emptyset(S(w))$ whathever we set P constraints on $\emptyset(S(w))$ and Q correlation constraints. In reference |4| it can be viewed that, in the same way that correlation constraints provides information about poles of the spectral density, cepstrum, constraints provides information about the zero location.

Finally it is interesting to remark that the solution of (10) faces a non-linear problem when solving for parameters λ_n and μ_m from the initial constraints. But, when the additional constraint of minimum phase for both denominator and numerator of the ARMA model involved is assumed, thus the parameters λ_m and μ_n can be found from an eigenvector formulation of mixed first and second order information of the initial data sample |5|.

IV. BANK FILTER APPROACH

In this section it will be described the bank filter approach for spectral estimation. In some sense this approach can be viewed as an extension of the variational approach reported in section III to the time domain problem.

To start with, let us suppose that we are interested in designing a FIR filter which its impulse response will be denoted as $\underline{w}$ in a vectorial notation. The components of such vector w(q) are the corresponding impulse response at time sample q.T, being T the sampling period. Thus, the filter output also named as the residual output $\varepsilon(n)$ obeys (14), where $\underline{x}_n$ is the current data vector with components x(n), x(n-1),...,x(n-Q) being Q+1 the filter length.

$$\varepsilon(n) = \underline{w}^T.\underline{x}_n \tag{14}$$

Facing the design of $\underline{w}$ for spectral estimation purposes, there is a prior decision concerning whether the residual $\varepsilon(n)$ or the filter $\underline{w}$ have to retain the information concerning the power density of the data input x(n) (n=0, N-1). To be more concrect in any linear, or even non-linear, processing of x(n) producing a quasi-white residual $\varepsilon(.)$ the transfer function of the DSP performed retains all the information, including the complete second order information, (i.e. the power spectrum), of the input sequence. Under this framework are included all the first order modelling procedures MA, AR or ARMA; because all of them form the spectral estimate from the transfer function of the associated DSP.

The second framework, which will deserve more attention herein than the previous one, consists in measuring the power of the input signal x(n) inside a frequency beam, with central frequency and bandwidth determined by the filter $\underline{w}$. Once the power of $\varepsilon(n)$ is measured as a function of the filter central frequency w_o and the frequency bandwidth $B_w(w_o)$, the power level and the power density spectrum $S(w_o)$ can be formulated as (15).

$$P(w_o) = E(\varepsilon^2(n)) \qquad (15\text{-a})$$

$$S(w_o) = P(w_o)/B_w(w_o) \qquad (15\text{-b})$$

It results clear that the quality in resolution and bias of the resulting estimate $S(w_o)$ will depend mainly in the leakage introduced by the bandwidth $B_w(w_o)$. In other words the filter $\underline{w}$ will be a pass-band filter with a bandwidth as narrow as possible and distorsionless in the frequency band where the filter is steered. To design the filter there are many non-data-dependent choices; these choices are listed below and they indlude the most classical procedures for spectral analyzers used in acoustics, vibration analysis, hi-fi design testing, etc.

a) Fixed bandwidth filter analysis bank.
b) Octave filter bank.
c) Sequential spectrum analyzers.

In c) a chirp signal is multiplied by the input sequence and a fixed high-quaity low pass filter provides the residual that, after a quadratic detector, gives the power level $P(w_o)$ which coincides with $S(w_o)$ without the constant bandwidth of the low pass filter. It is worthwhile to mention that the energy-time delay estimation technique could be encompassed under this kind of techniques |6|.

The design of filter $\underline{w}$ is obtained from minimizing the output power $P(w_o) = \underline{w}^T.\underline{\underline{R}}.\underline{w}$ under user selected constraints in order to guarantee that distorsion and leakage are under adequate levels always limited by the a priori constrained FIR length Q(i.e. $\underline{w}^T=(w(0),w(1),...,w(Q-1))$. The most familiar constraint is to set a given value for the filter response at the central frequency w_o. Using vector $\underline{S}^T$ as $(1, \exp(-jw_o),..., \exp(-j(Q-1)w_o))$, this constraint can be formulated as it is shown in (16),

$$\underline{S}^T\ \underline{w} = 1 \qquad (16)$$

a 0dB response at the central frequency w_o. When additional constraints are used, like first or high order derivations of the main lobe, the constraints can be formulated as (17) where $\underline{\underline{c}}$ reflects them in a matrix formulation. Of course the rank of $\underline{\underline{c}}$ have to be less than Q.

$$\underline{\underline{c}}.\underline{w} = \underline{f} \qquad (17)$$

After setting Lagrange multipliers for (17)

$$\underline{w}^T.\underline{\underline{R}}.\underline{w} \quad \text{minimum}$$
$$\underline{\underline{c}}.\underline{w} = \underline{f}$$
$$\underline{w}^T.\underline{\underline{R}}.\underline{w} - \underline{\lambda}^T(\underline{\underline{c}}.\underline{w}+\underline{f})$$

setting derivative with respect $\underline{w}$ equal to zero

$$\underline{\underline{R}}.\underline{w} = \underline{\underline{c}}^T.\underline{\lambda}$$
$$\underline{w} = \underline{\underline{R}}^{-1}.\underline{\underline{c}}^T.\underline{\lambda}$$

and using the constraint equation

$$\underline{\lambda} = (\underline{\underline{c}}.\underline{\underline{R}}^{-1}.\underline{\underline{c}}^T)^{-1}.\underline{f}$$

Thus the optimum vector filter is:

$$\underline{w}_{opt} = \underline{\underline{R}}^{-1}.\underline{\underline{c}}^T(\underline{\underline{c}}.\underline{\underline{R}}^{-1}.\underline{\underline{c}}^T)^{-1}.\underline{f} \qquad (18)$$

The residual power, i.e. the power level estimate is:

$$\underline{w}_{opt}^T.\underline{\underline{R}}.\underline{w}_{opt} = \underline{f}^T(\underline{\underline{c}}.\underline{\underline{R}}^{-1}.\underline{\underline{c}}^T)^{-1}.\underline{f} = P(w_o) \qquad (19)$$

Note that all this formulation firstly reported by Frost 8 could be viewed as the variational approach for spectral estimation in the time domain.

In order to get the final spectral estimate, a criteria to estimate the bandwidth of the analysis filter $\underline{w}$ has to be selected. It is important to note that because $\underline{\underline{c}}$ and/or $\underline{f}$ depend on the central frequency the filter is steered $P(w_o)$ and the spectral estimate will be no longer the same as they are, in an equal bandwidth analysis scheme. One interesting and easy to compute bandwidth is the so-called equal area constraint |7|. In other words, select $B_w(w_o)$ for filter $\underline{w}$ (frequency response $W(w)$) as

$$W(w_o)\cdot B_w(w_o)=\int |W(w)|^2 dw=\underline{w}^T\cdot\underline{w} \qquad (20)$$

This choice will produce the following estimate

$$S(w_o)=\frac{\underline{f}^T(\underline{\underline{c}}\cdot\underline{\underline{R}}^{-1}\cdot\underline{\underline{c}}^T)^{-1}\cdot\underline{f}}{\underline{f}^T(\underline{\underline{c}}\underline{\underline{R}}^{-1}\underline{\underline{c}}^T)^{-1}(\underline{\underline{c}}\underline{\underline{R}}^{-2}\underline{\underline{c}}^T)(\underline{\underline{c}}\underline{\underline{R}}^{-1}\underline{\underline{c}})^{-1}\underline{f}} \qquad (21)$$

Last formula for the scalar case of (16) will produce the already reported normalized maximum likelihood spectral density estimate |9|.

$$P(w_o) = 1/\underline{S}^T\underline{\underline{R}}^{-1}\underline{S} \qquad (22\text{-a})$$

$$S(w_o) = (\underline{S}^T\underline{\underline{R}}^{-1}\underline{S})/(\underline{S}^T\underline{\underline{R}}^{-2}\underline{S}) \qquad (22\text{-b})$$

It is very important to remark that the spectral estimate in (22-b) has very good resolution at the same time that yields the low side-lobe feature of classical maximum likelihood estimate reported by Capon. Also it is worth to remain that the spectral density estimate could be applied in angle of arrival estimation or beamforming in non-equally spaced arrays as well as in two-dimensional problems like radar or sonar applications |9|.

Two extensions are important around the estimate which appears in (22-b). First one is the so-called power estimates which is a family of spectral estimates which steems from generalizing the mentional estimate as it is shown in (23) |10|.

$$S^q(w) = \underline{S}^T\ \underline{\underline{R}}^{-n+1}\ \underline{S}/\underline{S}^T\ \underline{\underline{R}}^{-n}\ \underline{S} \qquad (23)$$

It is easy to check out that n=0 provides the B-T spectral estimate, n=1 the classical MLM estimate and n=2 the reported estimate. Going on in increasing the power of the estimates will perform reduction of $\underline{\underline{R}}$ to its noise subspace, and, as a result, the spectral estimate for high q will be quite similar to SVD methods. All the estimates have convergence to the actual power density estimate in the distributional sense, and also for any q they are homogeneous. Also note that these estimates can be obtained from an objective based on R^n (i.e. minimize $\underline{W}^T\cdot\underline{\underline{R}}^n\cdot\underline{w}$ with the 0dB constraint).

The second application of interest of the background provided in the presentation arising to the spectral estimate of (22-b) is its use in the cross-spectrum problem. Being $\underline{w}_x$ and $\underline{w}_y$ the two filters for chanel x and chanel y, steered at the same frequency and, designed under the same constraints set, the cross-power level could be measured as:

$$P_{xy}(w_o) = E(e_x{}^*(n)\ e_y(n))$$

and using the reduced expressions of e_x and e_y for a single or scalar constraint, the following estimate results |11|.

$$P_{xy}(w_o) = \frac{\underline{S}^T\underline{\underline{R}}_x{}^{-1}\underline{\underline{R}}_{xy}\underline{R}_y{}^{-1}\underline{S}}{(\underline{S}^T\underline{\underline{R}}_x{}^{-1}\underline{S})(\underline{S}^T\underline{\underline{R}}_y{}^{-1}\underline{S})} \qquad (24)$$

Using now the cross-bandwidth estimate as $\underline{w}_x{}^T\cdot\underline{w}_y=B_{xy}(w_o)$, the cross-spectrum estimate is obtained.

$$S_{xy}(w_o) = \frac{\underline{S}^T\underline{R}_x{}^{-1}\underline{R}_{xy}\underline{R}_y{}^{-1}\underline{S}}{\underline{S}^T\underline{R}_x{}^{-1}\underline{R}_y{}^{-1}\underline{S}} \qquad (25)$$

This estimate is a high performance cross-spectrum estimate which performs much better that the well known cross-spectrum estimate $\underline{S}^T\underline{\underline{R}}_{xy}\underline{S}$. Its use in broad-band beamforming provides excellent results and supports the potential of data-dependent filter analysis all the fields of spectral estimation (1D, 2D and multichanel problems).

V. CONCLUSIONS

The interesting alternative provided by the STUSE method in order to obtain good estimates of the data periodogram P(w), extending the very well known WOSA procedure. This approach opens new posibilities to obtain reliable prior estimate P(w) for variational procedures of spectral estimation. This work reported by Mathews and Youn is a valuable tool to enhance the best of the herein described frameworks for spectral estimation.

Using the variational formulation, the importance of objectives in controling the structural model associated to the resulting estimate together with the constraints is reported. It was pointed out that the most important role of the constraints we use in the procedure, is just to be sure enougth that all the information of interest is included in the mathematical formulation of the procedure. That is very important mainly because may authors use to pay more attention and efforts in finding new objectives. If the information provided is the same (i.e. correlation constraints) to different objectives, we will expect very small differences between the estimates, most of the times I like to name them as cosmetic modifications, if the structural model is not the adequate for the signal under processing. In summary, new trends in spectral estimation will devote more efforts in the constraints than in the objectives.

In the time domain variational problem, data-dependent filters are depicted as a general spectral estimation problem. This framework is non-parametric in nature ; and we will expect an excellent performance for any signal or problem, at least from my experience using these methods in practical environments. In other words, these methods perform quite well no matter the signal model and regardless if the problem is 1-D, 2-D, narrowband beamforming for non-equally spaced arrays, broadband beamforming and cross-spectrum or spectral matrix estimation.

REFERENCES

|1| V.J. Mathews and D.H. Youn, "Analysis of the short-time unbiased spectrum estimation algorithm". IEEE Trans. ASSP-33, no. 1, pp. 136-142, Feb. 1985.

|2| M.A. Lagunas, "The use of the most significant autocorrelation lags in iterative ME spectral estimation". IEEE Trans. ASSP-32, no. 2, pp. 445-448, April 1984.

|3| M.A. Lagunas et al., "A linear transform for spectral estimation". IEEE Trans. ASSP-29, no. 5, pp. 989-994, October 1981.

|4| M.A. Lagunas et al., "Arma model maximum entropy power spectral estimation", IEEE Trans. ASSP-32, no. 5, pp. 984-990, October, 1984.

|5| B.R. Musicus, A.M. Kabel, "Maximum Entropy pole-zero estimation". ICASSP-86, Tokyo, paper 27.12.

|6| Richard C. Heyser, "Acoustical Measurements by time delay spectrometry". Journal of Audio Eng. Soc., Vol. 15, no. 4, pp. 370-382, October 1967.

|7| M.A. Lagunas et al., "Maximum likelihood filters in Spectral Estimation problems". Signal Processing, Vol. 10, no. 1, pp. 19-34, January 1986.

|8| O.L. Frost, "An algorithm for linearly constrained adaptive processing". Proc. IEEE, Vol. 60, no. 8, Aug. 1972.

|9| M.A. Lagunas, A. Gasull, An improved ML method for power spectral density estimation". IEEE Trans. ASSP-32, no. 1, Feb. 1984.

|10| V.F. Pisarenko, "On the estimation of spectra by means of nonlinear functions of the variance matrix". Geophys. J. Roy. Astronom. Soc. Vol. 28, pp. 511-531, 1972.

|11| M.E. Santamaría et al., "Improved ML Cross-Spectral estimation". In these proceedings.