DOA Parameter Estimation with 1-bit Quantization Bounds, Methods and the Exponential Replacement

(Extended Abstract for WSA 2016, For Review Purpose only!)

Manuel Stein, Kurt Barbé* and Josef A. Nossek
Institute for Circuit Theory and Signal Processing (NWS), Technische Universität München, Germany
*Research Team Stochastics (STOX), Dept. Mathematics (DWIS), Vrije Universiteit Brussel, Belgium
Email: manuel.stein@tum.de, kurt.barbe@vub.ac.be, josef.a.nossek@tum.de

***Abstract*—We discuss the problem of direction-of-arrival (DOA) estimation with 1-bit quantization. While 1-bit analog-to-digital converters (ADC) allows to simplify the analog complexity of a wireless receiver significantly, efficient processing in the digital domain becomes non-trivial with such nonlinear measurements. In particular, for parametric covariance estimation problems like DOA estimation, where a multivariate Gaussian variable is passed through an element-wise symmetric hard-limiter, exact representation of the likelihood function is only possible up to four real-valued dimensions. This makes a rigorous DOA performance analysis and the derivation of efficient algorithms difficult when more than two sensors are involved. In the final paper we present a replacement framework, which aims at substituting the multivariate 1-bit output distribution by an equivalent and matched representation with in the class of the exponential family. This allows to approximate the Fisher information measure and therefore the Cramér-Rao lower bound (CRLB) for 1-bit DOA estimation with any number of sensors in a conservative way. Further, by the presented framework an unbiased estimator can be formulated, which asymptotically achieves the pessimistic CRLB.**

***Index Terms*—DOA estimation, 1-bit quantization, nonlinear stochastic system, Fisher information, CRLB, exponential family.**

I. Introduction

Concerning hardware complexity and energy consumption of signal processing systems, the circuit forming the analog-to-digital converter (ADC) at the receiver has been identified as one of the bottlenecks [1]. While a high number of bits b allows accurate representation of analog signals in the digital domain and therefore good processing performance, the power dissipation and production cost of the ADC device scales exponentially $\mathcal{O}(2^b)$ with the number of bits b. An interesting approach is to reduce the resolution of the ADC device and resort to simple ADC concepts without feedback. In the extreme case the continuous analog waveform at each receive sensor output is converted into a binary representation by a symmetric hard-limiter. The limiter sets its output to 1 if the continuous input signal is positive and to -1 otherwise. The circuit for such an ADC can be realized by a single comparator. This allows to perform ADC in the most efficient way. A serious drawback of hard-limiting 1-bit ADC is that such a low-complexity device executes a highly nonlinear and noninvertible operation on the receive signal. This is associated with a performance loss. It is well established, that for certain problems (univariate location parameter estimation) in the low signal-to-noise ratio (SNR) regime the relative performance gap between a symmetric hard-limiting 1-bit system and an ideal receiver with infinite ADC resolution is $2/\pi$ (-1.96 dB) [2] and therefore moderate. Advanced problems with multivariate input to the quantizer are less well understood. Here the main problem is to find an appropriate representation of the parametric likelihood function at the output of the quantizer.

II. Problem Statement

A. Ideal Receiver (∞-bit ADC)

Consider a multivariate receive signal $\boldsymbol{y} \in \mathbb{R}^M$ which is well represented by a multivariate Gaussian probability density function

$$p(\boldsymbol{y};\theta) = \frac{1}{(2\pi)^{\frac{N}{2}}\sqrt{\det \boldsymbol{\Sigma}_y(\theta)}} \exp\left(-\frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Sigma}_y^{-1}(\theta)\boldsymbol{y}\right) \quad (1)$$

with a single parameter $\theta \in \Theta$ and covariance

$$\boldsymbol{\Sigma}_y(\theta) = \mathrm{E}_{\boldsymbol{y};\theta}\left[\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}\right]. \quad (2)$$

The optimum unbiased estimator is easily obtained by maximizing the likelihood [3] given N data snapshots

$$\boldsymbol{Y} = \begin{bmatrix}\boldsymbol{y}_1 & \boldsymbol{y}_2 & \dots & \boldsymbol{y}_N\end{bmatrix}, \quad (3)$$

i.e.,

$$\begin{aligned}\hat{\theta}(\boldsymbol{Y}) &= \arg\max_{\theta\in\Theta} \ln p(\boldsymbol{Y};\theta) \\ &= \arg\max_{\theta\in\Theta} \sum_{n=1}^{N} \ln p(\boldsymbol{y}_n;\theta) \\ &= \arg\min_{\theta\in\Theta} \ln\big(\det \boldsymbol{\Sigma}_y(\theta)\big) + \mathrm{Tr}\left(\bar{\boldsymbol{\Sigma}}_y(\boldsymbol{Y})\boldsymbol{\Sigma}_y^{-1}(\theta)\right), \quad (4)\end{aligned}$$

where

$$\bar{\boldsymbol{\Sigma}}_y(\boldsymbol{Y}) = \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{y}_n\boldsymbol{y}_n^{\mathrm{T}} \quad (5)$$

is the sample mean of the receive covariance. As the maximum-likelihood estimator is unbiased and asymptotically efficient, it is possible to characterize it's performance in an analytical way through the celebrated Cramér-Rao lower bound [4], [5]

$$\mathrm{E}_{\boldsymbol{Y};\theta}\left[\left(\theta-\hat{\theta}(\boldsymbol{Y})\right)^2\right] \geq \left(NF_y(\theta)\right)^{-1}, \tag{6}$$

where the Fisher information is defined by

$$F_y(\theta)=\int_{\mathcal{Y}}\left(\frac{\partial \ln p(\boldsymbol{y};\theta)}{\partial\theta}\right)^2 \mathrm{d}\boldsymbol{y}. \tag{7}$$

For the multivariate Gaussian model (1), we obtain [3, p. 47]

$$F_y(\theta)=\frac{1}{2}\operatorname{Tr}\left(\boldsymbol{\Sigma}_y^{-1}(\theta)\frac{\partial\boldsymbol{\Sigma}_y(\theta)}{\partial\theta}\boldsymbol{\Sigma}_y^{-1}(\theta)\frac{\partial\boldsymbol{\Sigma}_y(\theta)}{\partial\theta}\right). \tag{8}$$

B. Low-Complexity Receiver (1-bit ADC)

The situation changes fundamentally if a nonlinear transformation

$$\boldsymbol{z}=\boldsymbol{f}(\boldsymbol{y}) \tag{9}$$

is involved. Then the derivation of an exact representation of likelihood $p(\boldsymbol{z};\theta)$ can become a challenging problem. If we consider an element-wise hard-limiter with 1-bit output

$$\boldsymbol{z}=\operatorname{sign}(\boldsymbol{y}), \tag{10}$$

where $\operatorname{sign}(\boldsymbol{x})$ is the element-wise signum function defined by

$$[\operatorname{sign}(\boldsymbol{x})]_n=\begin{cases}+1 & \text{if } x_n \geq 0\\ -1 & \text{if } x_n<0,\end{cases} \tag{11}$$

the likelihood function for one output constellation is found by evaluating the intergral

$$p(\boldsymbol{z};\theta)=\int_{\mathcal{Y}(\boldsymbol{z})} p(\boldsymbol{y};\theta)\mathrm{d}\boldsymbol{y}, \tag{12}$$

where $\mathcal{Y}(\boldsymbol{z})$ is the subset in $\mathcal{Y}$ which is mapped to $\boldsymbol{z}$. Such an integral is identified as the orthant probability of a multivariate Gaussian variable (multivariate version of the Q-function). Unfortunately, a general closed-form expression for the orthant probability is an open mathematical problem. Only for the cases $N\leq 4$ solutions are provided in literature [6] [7]. The problem becomes even worse, if one is interested in evaluating the Fisher information measure

$$\begin{aligned}F_z(\theta)&=\int_{\mathcal{Z}}\left(\frac{\partial \ln p(\boldsymbol{z};\theta)}{\partial\theta}\right)^2 \mathrm{d}\boldsymbol{z}\\ &=\sum_{\mathcal{Z}}\left(\frac{\partial \ln p(\boldsymbol{z};\theta)}{\partial\theta}\right)^2\end{aligned} \tag{13}$$

by summing the squared score function over the discrete support of $\boldsymbol{z}$. As $\mathcal{Z}$ contains 2^M possible receive vectors direct computation of $F(\theta)$ is prohibitively complex when M is large. Due to this fact the literature on performance bounds and efficient algorithms for parametric covariance estimation with 1-bit quantization is limited. While [8] [2] are classical references for signal processing with 1-bit quantizer, more recently [9] covers the problem of signal parameter estimation from coarsely quantized data with uncorrelated noise. [10] is concerned with 1-bit DOA estimation, but has to restrict the discussion to $K=2$ sensors due to the outlined problem (13). In contrast [11] studies estimation with a multivariate model and dithering, where the threshold of the hard-limiter is distributed randomly over the spatial or temporal domain.

III. Analysis by an Information Bound

In order to avoid intractable situations like (13), we have recently developed lower bounds for the Fisher information of nonlinear stochastic systems [12] [13], which extend the results of [14] [15] for additive systems to the broader class of non-additive systems. These are extremely useful, when the radio front-end prior to the 1-bit quantizer is to be designed [16]. In a final attempt we have generalized the result to [17]

$$F_z(\theta)\geq\left(\frac{\partial\boldsymbol{\mu}_{\boldsymbol{\phi}}(\theta)}{\partial\theta}\right)^{\mathrm{T}}\boldsymbol{R}_{\boldsymbol{\phi}}^{-1}(\theta)\frac{\partial\boldsymbol{\mu}_{\boldsymbol{\phi}}(\theta)}{\partial\theta}, \tag{14}$$

where

$$\boldsymbol{\mu}_{\boldsymbol{\phi}}(\theta)=\mathrm{E}\left[\boldsymbol{\phi}(\boldsymbol{z})\right] \tag{15}$$

and

$$\boldsymbol{R}_{\boldsymbol{\phi}}(\theta)=\mathrm{E}\left[\boldsymbol{\phi}(\boldsymbol{z})\boldsymbol{\phi}^{\mathrm{T}}(\boldsymbol{z})\right]-\boldsymbol{\mu}_{\boldsymbol{\phi}}(\theta)\boldsymbol{\mu}_{\boldsymbol{\phi}}^{\mathrm{T}}(\theta), \tag{16}$$

with

$$\boldsymbol{\phi}(z)=\begin{bmatrix}\phi_1(\boldsymbol{z}) & \phi_2(\boldsymbol{z}) & \dots & \phi_L(\boldsymbol{z})\end{bmatrix}^{\mathrm{T}} \tag{17}$$

being a vector with L arbitrary transformations $\phi_l(\boldsymbol{z})$ of the output variable $\boldsymbol{z}$. It can be shown that the bound (14) is tight if the output distribution $p(\boldsymbol{z};\theta)$ with parameter θ is an exponential family distribution and $\boldsymbol{\phi}(\boldsymbol{z})$ contains all sufficient statistics of $p(\boldsymbol{z};\theta)$ [18]. In [17] we show how such Fisher information bounds can be applied in order to verify the quality of nonlinear devices based on calibrated measurements and applied this approach for the estimation theoretic analysis of nonlinear amplification devices, a Rician model and a cubic polynomial. In our extended article [18], we also discuss how to design unbiased estimates for nonlinear systems which asymptotically perform according to the inverse of the approximated Fisher information measure (14).

IV. Outline of the Final Paper

In the final paper we will review the bounding framework (14) and apply it to the 1-bit DOA problem in order to analyze the performance gap between an ideal ∞-bit receiver and a 1-bit receiver

$$\chi(\theta)=\frac{F_z(\theta)}{F_y(\theta)}. \tag{18}$$

To this end, we will assume a real-valued model with K sensors [19], where the parametric covariance obtains the form

$$\boldsymbol{\Sigma}_y(\theta)=\gamma\boldsymbol{A}(\theta)\boldsymbol{A}^T(\theta)+\boldsymbol{I}_{2K} \tag{19}$$

with θ being the DOA parameter. The parametric covariance is characterized by the steering matrix

$$\boldsymbol{A}(\theta)=\begin{bmatrix}\boldsymbol{A}_I^T(\theta) & \boldsymbol{A}_Q^T(\theta)\end{bmatrix}^T\in\mathbb{R}^{2M\times 2} \tag{20}$$

containing the two submatrices (in-phase and quadrature channel)

$$\boldsymbol{A}_I(\theta) = \begin{bmatrix} \xi_1(\theta) & \nu_1(\theta) \\ \xi_2(\theta) & \nu_2(\theta) \\ \vdots & \vdots \\ \xi_K(\theta) & \nu_K(\theta) \end{bmatrix} \in \mathbb{R}^{K \times 2} \tag{21}$$

$$\boldsymbol{A}_Q(\theta) = \begin{bmatrix} -\nu_1(\theta) & \xi_1(\theta) \\ -\nu_2(\theta) & \xi_2(\theta) \\ \vdots & \vdots \\ -\nu_K(\theta) & \alpha_K(\theta) \end{bmatrix} \in \mathbb{R}^{K \times 2}, \tag{22}$$

with entries

$$\begin{aligned} \xi_k(\theta) &= \cos\big((k-1)\pi \sin(\theta)\big) \\ \nu_k(\theta) &= \sin\big((k-1)\pi \sin(\theta)\big). \end{aligned} \tag{23}$$

Making use of the results of [7], we will analytically derive a specific version of (14) for the DOA problem and discuss the obtained approximation quality. The obtained performance analysis will cover a discussion of the behavior of $\chi(\theta)$ for different numbers of sensors K and different signal-to-noise ratios γ^2.

REFERENCES

[1] R. H. Walden, “Analog-to-digital converter survey and analysis,” *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 4, pp. 539–550, 1999.

[2] J. H. Van Vleck, D. Middleton, “The spectrum of clipped noise,” *Proceedings of the IEEE*, vol. 54, no. 1, pp. 2–19, 1966.

[3] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddler River, NJ: Prentice Hall, 1993.

[4] C. R. Rao, “Information and accuracy attainable in the estimation of statistical parameters,” *Bull. Calcutta Math. Soc.*, vol. 37, no. 3, pp. 81–91, 1945.

[5] H. Cramér, *Mathematical Methods of Statistics*. Princeton, NJ: Princeton Univ. Press, 1946.

[6] B. Kedem, *Binary Time Series (Lecture Notes in Pure and Applied Mathematics)*. New York: Marcel Dekker, 1980.

[7] M. Sinn, K. Keller, ”Covariances of Zero Crossings in Gaussian Processes,” Theory Probab. Appl., vol. 55, no. 3, pp. 485–504, 2011.

[8] W. R. Bennett, “Spectra of quantized signals”, *Bell System Technical Journal*, vol. 27, pp. 446–472, 1948.

[9] A. Host-Madsen, P. Handel, “Effects of sampling and quantization on single-tone frequency estimation,” *IEEE Trans. Signal Processing*, vol. 48, no. 3, pp. 650–662, 2000.

[10] O. Bar-Shalom, A. J. Weiss, “DOA estimation using one-bit quantized measurements,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 3, pp. 868–884, 2002.

[11] O. Dabeer, E. Masry, “Multivariate Signal Parameter Estimation Under Dependent Noise From 1-Bit Dithered Quantized Data,” *IEEE Transactions on Information Theory*, vol. 54, no. 4, pp.1637–1654, 2008.

[12] M. Stein, A. Mezghani, J. A. Nossek, “A lower bound for the Fisher information measure,” *IEEE Signal Process. Letters*, vol. 21, no. 7, pp. 796–799, July 2014.

[13] M. Stein, J. A. Nossek, ”A pessimistic approximation for the Fisher information measure,” submitted to *IEEE Trans. Signal Process.*, preprint: http://arxiv.org/abs/1508.03878, 2015.

[14] P. Stoica, P. Babu, ”The Gaussian data assumption leads to the largest Cramèr-Rao bound,” *IEEE Signal Process. Mag.*, vol. 28, pp. 132–133, May 2011.

[15] S. Park, E. Serpedin, K. Qaraqe, “Gaussian assumption: The least favorable but the most useful,” *IEEE Signal Process. Mag.*, vol. 30, pp. 183–186, 2013.

[16] M. Stein, S. Theiler, J. A. Nossek, “Overdemodulation for high-performance receivers with low-resolution ADC,” *IEEE Wireless Commun. Letters*, vol. 4, no. 2, pp.169–172, 2015.

[17] M. Stein, J. A. Nossek and K. Barbé, “Measurement-driven Quality Assessment of Nonlinear Systems by Exponential Replacement,” submitted for presentation at the IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Taipei, Taiwan, 2016

[18] M. Stein, J. A. Nossek and K. Barbé, “Fisher Information Bounds with Applications in Nonlinear Learning, Compression and Inference,” in preparation, 2015.

[19] M. Stein, M. Castañeda, J. A. Nossek, ”Information-Preserving Spatial Filtering for Direction-of-Arrival Estimation,” *18th International ITG Workshop on Smart Antennas (WSA)*, 2014.