Chapter 8

Conclusions and Topics for Future Research

In this thesis, optimal *blind* second-order estimators are deduced considering the true distribution of the nuisance parameters. Quadratic estimators are formulated assuming that the nuisance parameters distribution is known and a certain side information on the unknown parameters is available. Adopting the Bayesian formulation, the referred side information is introduced by means of the parameters prior distribution. This approach allows unifying the formulation of the open-loop (large-error) estimators in Chapter 3 and the closed-loop (small-error) estimators in Chapter 4. In the former case, the prior knowledge is rather vague whereas a very informative prior is considered in the latter case.

The first important conclusion is that, in most estimation problems, second-order techniques are severely degraded due to the bias term unless the small-error condition is satisfied. As an illustrative example, it is shown in Chapter 3 that it could be difficult to have unbiased frequency estimates using second-order open-loop schemes. However, the interest of second-order open-loop estimators is motivated by the problematic convergence of closed-loop schemes in noisy scenarios, in which large-error and small-error estimators are shown to yield approximately the same mean square error.

To avoid the bias limitations, the Best Quadratic Unbiased Estimator (BQUE) is deduced in Chapter 4 under the small-error condition. The covariance matrix associated with the BQUE estimator constitutes the tighest lower bound on the variance of any second-order unbiased estimator. Formally, it is claimed in Chapter 4 that

$$E\left\{(\widehat{\boldsymbol{\alpha}}-\mathbf{g}(\boldsymbol{\theta}))^2\right\} \geq \mathbf{B}_{BQUE}(\boldsymbol{\theta}) = \mathbf{D}_g(\boldsymbol{\theta})\left(\mathbf{D}_r^H(\boldsymbol{\theta})\,\mathbf{Q}^{-1}(\boldsymbol{\theta})\,\mathbf{D}_r(\boldsymbol{\theta})\right)^{-1}\mathbf{D}_g^H(\boldsymbol{\theta})$$

for *any quadratic* estimator of $\boldsymbol{\alpha} = \mathbf{g}(\boldsymbol{\theta})$. In the above expression, $\mathbf{D}_g(\boldsymbol{\theta})$ and $\mathbf{D}_r(\boldsymbol{\theta})$ are the Jacobian of $\mathbf{g}(\boldsymbol{\theta})$ and $\mathrm{vec}(\mathbf{R}(\boldsymbol{\theta}))$, respectively, where $\mathbf{R}(\boldsymbol{\theta})$ stands for the covariance matrix of the observed vector $\mathbf{y}$. On the other hand, $\mathbf{Q}(\boldsymbol{\theta})$ contains all the central fourth-order moments of $\mathbf{y}$. The matrix $\mathbf{Q}(\boldsymbol{\theta})$ can be splitted into two terms, as it was pointed out in Chapter 3, obtaining

$$\mathbf{Q}(\boldsymbol{\theta}) \triangleq \mathbf{R}^*(\boldsymbol{\theta}) \otimes \mathbf{R}(\boldsymbol{\theta}) + \mathcal{A}(\boldsymbol{\theta}) \mathbf{K} \mathcal{A}^H(\boldsymbol{\theta}) \tag{8.1}$$

where the second term accounts for all the *non-Gaussian* information about the nuisance parameters ($\mathbf{K} \neq \mathbf{0}$) that is profitable to estimate the vector of parameters $\boldsymbol{\theta}$ by means of quadratic processing. The evaluation of the potential benefits gained when considering this second term has become one of the most important issues in this thesis.

In many problems, the Gaussian assumption (i.e., $\mathbf{K} = \mathbf{0}$) is adopted to design second-order schemes when the actual distribution is unknown or it becomes an obstacle to obtain analytically the ML estimator. The most relevant contribution in this thesis is proving that the Gaussian assumption leads, in some scenarios, to *suboptimal* second-order estimation methods. Conversely, the Gaussian assumption is proved to supply the optimal second-order estimator –independently of the actual parameterization– in all these cases:

- The nuisance parameters are normally distributed.
- The SNR is low.
- All the derivatives of the transfer matrix $\mathbf{A}(\boldsymbol{\theta})$ are orthogonal to the columns of $\mathbf{A}(\boldsymbol{\theta})$ (Section 7.3). Formally,

$$\mathbf{P}_{\mathbf{A}}(\boldsymbol{\theta}) \frac{\partial \mathbf{A}(\boldsymbol{\theta})}{\partial \theta_p} = \mathbf{0}$$

  where $\mathbf{P}_{\mathbf{A}}(\boldsymbol{\theta})$ is the orthogonal projector onto the subspace generated by the columns of $\mathbf{A}(\boldsymbol{\theta})$.[1]

- The SNR is high and the nuisance parameters are drawn from a multilevel alphabet (e.g., the QAM constellation).

Otherwise, in case of dealing with constant-modulus nuisance parameters (e.g., the MPSK or CPM modulations), some improvement can be expected in case of exploiting the second term of $\mathbf{Q}(\boldsymbol{\theta})$ for medium-to-high SNR. The actual improvement is a function of the observation size and depends on the actual parameterization.

All these conclusions have also been evidenced in Chapter 5 where the design of the optimal second-order tracker is presented. In this chapter, the Kalman filter formulation is adopted to optimize both the acquisiton and steady-state performance. In that way, in Chapter 5, the

[1] A more general condition is presented in (7.47) in case of circular nuisance parameters.

Gaussian assumption is validated in the acquisition phase, concluding that the acquistion time can be significantly shortened for practical SNRs if the nuisance parameters are drawn from a constant modulus alphabet. Otherwise, if the nuisance parameters have multiple amplitudes, the Gaussian assumption is also optimal in terms of acquisition performance.

Despite the last statements, in some significant estimation problems, the Gaussian assumption applies asymptotically as long as the observation interval grows to infinity, as proved in Chapter 7. In that case, the importance of the second term of $\mathbf{Q}(\boldsymbol{\theta})$ is relegated to those scenarios in which the observation interval is short. Despite this, most conclusions from the asymptotic study in Chapter 7 are problem-dependent. In the following paragraphs, the main conclusions from every estimation problem addressed in this thesis are summarized:

- **Synchronization.** The Gaussian assumption is asymptotically optimal when the number of observed symbols is infinite (Section 7.4.4). In a continuous mode transmission, the asymptotic condition allows neglecting the so-called "edge effect" related to the partial observation of the border symbols (Section 6.1.2).

  On the other hand, if the observation time is limited, the fourth-order information about the received constellation is crucial to filter out the self-noise at high SNR in case of a partial response CPM modulation (e.g., LREC or GMSK) but it is negligible in case of linear modulations (e.g., MPSK or QAM). The importance of this result is beyond the actual interest of CPM modulations. Actually, it shows that the optimal second-order estimator is able to take advantage of the statistical dependence of the Laurent's expansion pseudo-symbols. In this case, as it happens in *coded transmissions*, the received symbols are not statistically independent in spite of being uncorrelated. Thus, the results for the CPM format could be translated to optimize existing second-order synchronizers in case of coded communication systems.

  Finally, note that the Gaussian assumption always applies in TDMA communication systems whatever the observation length.

- **Channel Estimation**. If the channel amplitude is not estimated, the Gaussian assumption yields minor losses at high SNR *on the average*, that is, if the estimator performance is averaged considering multiple channel realizations (Section 6.4). However, the Gaussian assumption is expected to fail for some particular realizations of the channel impulse response. Indeed, this point is currently being investigated [LS04][LS05a][LS05b].

  Another important conclusion is that the Gaussian assumption yields a severe degradation at high SNR when the channel amplitude is estimated too. If the transmitted symbols belong to a multilevel constellation (e.g., QAM), second-order estimators exhibit the typical variance floor at high SNR. On the other hand, if the transmitted symbols have *constant* *modulus* (e.g., MPSK or CPM), the BQUE estimator is able to avoid the aforementioned variance floor whereas, if the Gaussian assumption is adopted, the estimator performance degrades at high SNR because the channel amplitude estimate is drastically degraded. In fact, the higher is the transmitted pulse bandwidth (roll-off factor) the more important is the incurred loss at high SNR.

The asymptotic study in Section 7.4.4 states that the above conclusions are still valid if the observation time goes to infinity. The loss incurred by the Gaussian assumption becomes a function of the actual channel impulse response.

- **Direction-of-Arrival (DOA).** The Gaussian assumption is asymptotically optimal when the number of antennas is infinite (Section 7.4.5). The Gaussian assumption also applies in the single user case whatever the array size or the working SNR (Section 7.4.5). Likewise, the Gaussian assumption is optimal at high SNR if the transmitted symbols are drawn from a multilevel constellation (e.g., QAM or APK).

  On the other hand, if the transmitted symbols belong to a *constant-modulus alphabet* (e.g., MPSK or CPM), the fourth-order statistics of the transmitted symbols can be exploited to discriminate the DOA of those signals impinging into the array from near directions. When the Gaussian assumption is adopted and this fourth-order information is omitted, a significant loss is manifested at high SNR that is a function of the number of antennas and the users angular separation. Furthermore, the incurred loss cannot be reduced even if the number of received symbols goes to infinity (Section 7.4.5).

8.1 Further Research

In this thesis, the ultimate limits of second-order estimation are studied from both the practical and theoretical point of view. However, some interesting points are still open and should be investigated in the future. From the author's opinion, the most promising topics for further study are outlined in the following paragraphs:

1. **Multiuser estimation problems.** Second-order methods are able to exploit the constant modulus property of the random nuisance parameters. This property appears reflected in the eigendecomposition of the kurtosis matrix $\mathbf{K}$. In some estimation problems, this information is crucial to deal with the intersymbol interference as well as the multiple access interference in *multiuser applications*. Actually, the results of this thesis suggest that the constant modulus property is mainly relevant in multiuser or MIMO scenarios. In these scenarios, the constant-modulus property could be exploited to discriminate the parameters associated to non-orthogonal interfering users.

2. **Asymptotic Gaussian assumption.** The Gaussian assumption does not apply for practical SNRs if the nuisance parameters have constant-modulus and the observation length is rather short in case of low-cost implementations. However, in some important problems, the Gaussian assumption is rapidly satisfied as the number of observations is augmented because of the Central Limit Theorem. This asymptotic study was addressed in Chapter 7 for different estimation problems. In all the studied problems, if the nuisance parameters have constant modulus, the second term in equation (8.1) generates a favourable term that persists at high SNR. In the problem of DOA estimation, this term becomes negligible if the number of antennas goes to infinity. However, this term survives if the number of antennas is finite, even if the number of snapshots goes to infinity. Therefore, the results in Chapter 7 could be useful to identify those estimation problems in which the non-Gaussian information persists as the number of observations goes to infinity.

3. **Noncircular and coded nuisance parameters.** If we deal with *noncircular nuisance parameters* (e.g., CPM signals), the kurtosis matrix $\mathbf{K}$ provides additional information regarding the statistical dependence of the nuisance parameters. In that way, it is possible to remove the self-noise –including the multiple access interference– even if the number of parameters exceeds the number of observations. This feature should be throroughly investigated because it could be exploited to improve second-oder estimators in case of *coded transmissions*. Moreover, other noncircular constellations should be studied in detail as, for example, BPSK, digital PAM and staggered formats as the offset QPSK modulation

Besides these three principal research lines, some other topics for future research are listed next:

1. **Large error bounds with nuisance parameters.** In Section 2.6, the most important lower bounds in the literature were classified and briefly described. Among all the existing bounds, the Cramér-Rao bound (CRB) is without doubt the most widespread one due to its simplicity. However, the true CRB is still unknown in a lot of estimation problems in the presence of nuisance parameters. To fill this gap, the CRB is derived in some particular scenarios: low SNR, high-SNR, Gaussian nuisance parameters (Gaussian UCRB), deterministic and continuous nuisance parameters (CCRB), and known nuisance parameters (MCRB). Moreover, in this thesis we have deduced the CRB under the quadratic constraint.

   In the context of digital communications, it would be useful to apply the last assumptions to the *large-error bounds* in Section 2.6 in order to characterize the large-error region and the SNR threshold in the presence of nuisance parameters. Among all the lower bounds in Section 2.6, the Hammersley-Chapman-Robbins, Weiss-Weinstein and Ziv-Zakai lower bounds are surely the most promising candidates. Then, the obtained large-error bounds should be compared to the *second-order* large-error estimators deduced in Chapter 3. From this comparison, we could determine whether quadratic estimators are optimal or not at low SNR in the large-error regime. Also, we could evaluate the performance loss due to the presence of the random nuisance parameters.

2. **Acquisition optimization.** The QEKF was proposed in Chapter 5 with the aim of improving the acquisition performance of classical closed loop schemes. Initially, the QEKF supplies the large-error MMSE solution derived in Chapter 3 considering the initial non-informative prior. Thereafter, the QEKF converges progressively to the small-error solution in Chapter 4 every time that a new observation is processed. The new datum is used to update the prior distribution so that the prior becomes every time more and more informative.

   Unfortunately, the QEKF convergence is not guaranteed unless the observations and parameters are jointly Gaussian distributed. Moreover, although the QEKF had converged, the acquisition time could have been shortened by optimizing the prior update. Therefore, an important topic for research is to find the optimal prior update for optimizing the acquisition probability and delay. In that sense, the Unscented Kalman Filter proposed in [Jul97][Wan00] should be considered since it is known to guarantee the convergence under mild conditions.

3. **Low-cost implementation**. The optimal second-order estimator is formulated in Chapter 3 and Chapter 4 resorting to the vec(·) transformation. Consequently, it is necessary to compute the inverse of the $M^2 \times M^2$ square matrix $\mathbf{Q}(\boldsymbol{\theta})$ and, thus, the estimator computational cost increases rapidly when the number of observations M is augmented. In some problems, matrix $\mathbf{Q}^{-1}(\boldsymbol{\theta})$ can be computed offline before processing the first sample (e.g., in digital synchronization). However, in other relevant problems such as channel and DOA estimation, the inverse needs to be computed every time (online).

   The closed-loop architecture introduced in Section 2.5.1 as well as the QEKF formulation in Chapter 5 allow reducing the number of observations M that are jointly processed each time. Additionally, suboptimal quadratic estimators could be investigated by considering rank-reduction techniques [Sic92][Sch91b] and transversal filtering implementations. In the latter case, the (scalar) parameter θ could be estimated from the sample covariance matrix at time n applying a *one-rank* matrix $\mathcal{M} = \mathbf{h}\mathbf{h}^H$. Thus, we have

$$\widehat{\theta}_n = b + \operatorname{Tr}\left(\mathcal{M}\widehat{\mathbf{R}}_n\right) = b + \mathbf{y}_n^H \mathcal{M} \mathbf{y}_n = b + \left\|\mathbf{h}^H \mathbf{y}_n\right\|^2$$

   where $\mathbf{y}_n$ is the observation at time n and $\mathbf{h}$ collects the coefficients of the estimator time-invariant impulse response. In that case, we aim at optimizing the coefficients of $\mathbf{h}$ according to the criteria presented in Chapter 3 and Chapter 4. Evidently, we can consider multiple transversal filters:

$$\widehat{\theta}_n = b + \sum_{m=1}^{R} \left\| \mathbf{h}_m^H \mathbf{y}_n \right\|^2$$

where R is the rank of

$$\mathcal{M} = \sum_{m=1}^{R} \mathbf{h}_m \mathbf{h}_m^H.$$

Actually, the original estimators in Chapter 3 and Chapter 4 were the sum of M transversal filters because matrix $\mathcal{M}$ was originally full rank, i.e., $R = M$.