ESTIMATION WITH MAXIMUM ERROR REQUIREMENTS

Zvika Ben-Haim *Yonina C. Eldar*

Department of Electrical Engineering
Technion—Israel Institute of Technology
Haifa 32000, Israel

ABSTRACT

We consider the problem of estimating a deterministic parameter vector $\mathbf{x}$ from observations $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$, where $\mathbf{H}$ is known and $\mathbf{w}$ is additive noise. We seek an estimator whose estimation error is within given limits, for as wide a range of conditions as possible. The error limit is a design choice, and is generally lower than the error provided by the well-known least-squares (LS) estimator. We develop estimators guaranteeing the required error for as large a parameter set as possible, and for as large a noise level as possible. We discuss methods for finding these estimators, and demonstrate that in many cases, the proposed estimators outperform the LS estimator.

1. INTRODUCTION

Consider the problem of estimating an unknown deterministic parameter vector $\mathbf{x}$ based on measurements $\mathbf{y} = \mathbf{H}\mathbf{x}+\mathbf{w}$, where $\mathbf{H}$ is a known linear transformation and $\mathbf{w}$ is zero-mean random noise. We would like to design a linear estimator $\hat{\mathbf{x}}$ close to $\mathbf{x}$ in terms of the mean-squared error (MSE). However, since the MSE is a function of the unknown vector $\mathbf{x}$, it cannot be directly minimized. A common approach is to choose the minimum MSE estimator among all linear *unbiased* estimators, in which case the MSE does not depend on the value of $\mathbf{x}$. This approach yields the (weighted) least-squares (LS) estimator [9]. Yet an unbiased estimator does not necessarily guarantee low MSE. Indeed, for any bounded set $\mathcal{U}$, a linear biased estimator exists whose MSE is lower than the MSE of the LS estimator, for *all* $\mathbf{x}$ in $\mathcal{U}$ [3]. Several techniques are aimed at improving the MSE by introducing a bias [10, 13].

Estimator design is based on various *system properties*, such as the noise covariance. If the values of these system properties are not known exactly, one may seek the estimator minimizing the worst-case error among all possible values. This *minimax* approach was first introduced in the context of uncertain noise statistics [7], and has since been applied in a variety of estimation problems [8]. For example, we may seek the *bounded parameter set* estimator, which minimizes the worst-case error for any parameter $\mathbf{x}$ in a given parameter set $\mathcal{U}$. Various error measures may be used, including the worst-case MSE [11, 5] and the worst-case regret [6].

The minimax approach requires specification of system property *bounds*; like the system properties themselves, the bounds may not be known in advance. In some cases, bounds can be estimated from the measurements $\mathbf{y}$, but the resulting estimator is generally nonlinear, and thus requires greater computational complexity. We propose an alternative approach, which is suitable when requirements on the maximum estimation error are more readily available than system property bounds. For example, in communication systems, a minimum SNR may be required for data transmission to be possible. For such cases, following the philosophy of information-gap decision theory [1, 2], we propose a *maximum estimation error* approach, in which an estimator is designed to guarantee the required error for the widest range of system properties possible. This approach can be applied to different system properties. In Section 2, we study the case of uncertain parameter sets, while Section 3 deals with uncertain noise levels. We conclude with a discussion in Section 4.

2. MAXIMUM PARAMETER SET ESTIMATION

In this section, we define the maximum parameter set (MPS) estimator, provide methods for its calculation, and compare its performance with the LS estimator.

2.1. Definition

Let $\mathbf{x}$ be an unknown deterministic vector in $\mathbb{C}^n$ and let $\mathbf{w}$ be a zero-mean random vector in $\mathbb{C}^m$ whose covariance $\mathbf{C_w}$ is known. Suppose $\mathbf{H}$ is a known full-rank $m \times n$ matrix, and let $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$. An *estimator* $\hat{\mathbf{x}}$ is a function of $\mathbf{y}$ which returns an n-vector close to $\mathbf{x}$ in some sense. For example, the least-squares (LS) estimator [9] is given by

$$\hat{\mathbf{x}}_{\mathrm{LS}} = (\mathbf{H}^*\mathbf{C}_\mathbf{w}^{-1}\mathbf{H})^{-1}\mathbf{H}^*\mathbf{C}_\mathbf{w}^{-1}\mathbf{y}, \tag{1}$$

where $\mathbf{P}^*$ denotes the Hermitian conjugate of $\mathbf{P}$.

The following system properties, selected by the designer, are used to construct the MPS estimator:
1) An *error function* $\epsilon(\hat{\mathbf{x}}, \mathbf{x})$, such as the MSE $E\|\mathbf{x} - \hat{\mathbf{x}}\|^2$, which quantifies the degree to which $\hat{\mathbf{x}}$ misrepresents a specific value $\mathbf{x}$.
2) A *maximum error* ϵ_m which defines the error value required for successful operation of the system.
3) A *class of parameter sets* $\{\mathcal{U}_L \subseteq \mathbb{C}^n : \ L \geq 0\}$ which define feasible values of $\mathbf{x}$ under varying parameter set bounds L. We assume that the sets are nested, i.e., $\mathcal{U}_{L_1} \subseteq \mathcal{U}_{L_2}$ for $L_1 \leq L_2$. For regularity, we also require that the parameter sets grow linearly with L, i.e., $\mathcal{U}_L = \left\{\mathbf{x} : \ \frac{1}{L}\mathbf{x} \in \mathcal{U}_1\right\}$; this implies that the parameter sets are centered on the origin, an assumption which we adopt without loss of generality.

We define the *parameter robustness* $\hat{L}(\hat{\mathbf{x}})$ of an estimator $\hat{\mathbf{x}}$ as the maximum bound L for which the maximum error is assured,

$$\hat{L}(\hat{\mathbf{x}}) = \max\{L : \ \epsilon(\hat{\mathbf{x}}, \mathbf{x}) \leq \epsilon_m, \ \forall \mathbf{x} \in \mathcal{U}_L\}. \tag{2}$$

A *maximum parameter set (MPS) estimator* $\hat{\mathbf{x}}_{\mathrm{UP}}$ (among estimators of class $\mathcal{E}$) maximizes the parameter robustness $\hat{L}$,

$$\hat{\mathbf{x}}_{\mathrm{UP}} = \arg\max_{\hat{\mathbf{x}} \in \mathcal{E}} \hat{L}(\hat{\mathbf{x}}). \tag{3}$$

By comparison, a *minimax* (or *bounded parameter set*) *estimator* $\hat{\mathbf{x}}_{\mathrm{M}}$ minimizes the worst-case error in a given parameter set $\mathcal{U}$ (among estimators of class $\mathcal{E}$):

$$\hat{\mathbf{x}}_{\mathrm{M}} = \arg\min_{\hat{\mathbf{x}}\in\mathcal{E}} \max_{\mathbf{x}\in\mathcal{U}} \epsilon(\hat{\mathbf{x}}, \mathbf{x}). \tag{4}$$

2.2. Minimax and MPS Estimators

The MPS estimator maximizes the parameter robustness L defined by the known value of ϵ, while the minimax estimator minimizes the worst-case error ϵ within a range defined by the known value of L. Thus, for an appropriate choice of system properties, minimax estimators may be MPS estimators, and vice versa. Indeed, this occurs for many (though not all) cases of interest.

This similarity notwithstanding, minimax and MPS estimators differ qualitatively in the type of information on which their design is based. The minimax estimator requires that a bound on the uncertain parameter $\mathbf{x}$ be stated, while the MPS estimator requires knowledge of the maximum error under which the system still operates correctly. Thus, proper choice of an estimator depends on the nature of information available to the designer.

The relation between minimax and MPS estimators is formalized by means of the *worst-case error function* $e(L)$, defined as

$$e(L) = \max_{\mathbf{x}\in\mathcal{U}_L} \epsilon(\hat{\mathbf{x}}_{\mathrm{M}}(L), \mathbf{x}), \tag{5}$$

where $\hat{\mathbf{x}}_{\mathrm{M}}(L)$ is a minimax estimator for the parameter set $\mathcal{U}_L$; see Figure 1. For MPS estimators, $e(L)$ represents a trade-off between performance and parameter robustness: modest performance requirements (large ϵ_m) can be guaranteed for a large parameter set $\mathcal{U}_L$, while stringent requirements are only guaranteed for a small parameter set.

We now describe several cases in which MPS estimators are minimax estimators. In such cases, known results regarding minimax estimators are used to efficiently find MPS estimators. We first show that when the error function of interest is the MSE and linear estimators are considered, minimax and MPS estimators are equivalent. Proofs of this and other propositions appear in [3].

Proposition 1. *Let $\{\mathcal{U}_L : L \geq 0\}$ be a class of parameter sets, let $\epsilon(\hat{\mathbf{x}}, \mathbf{x}) = E\|\mathbf{x} - \hat{\mathbf{x}}\|^2$, and let $\mathcal{E}$ be the class of linear estimators. For any L, an estimator $\hat{\mathbf{x}} \in \mathcal{E}$ is a linear minimax estimator over $\mathcal{U}_L$ if, and only if, it is a linear MPS estimator with maximum error ϵ_m equal to the worst-case error $e(L)$ of* (5).

Thus, finding an MPS estimator for a given maximum error ϵ_m is equivalent to finding a minimax estimator whose worst-case error is ϵ_m. This can be accomplished using a line search, in which minimax estimators for different sets $\mathcal{U}_L$ are calculated until an estimator with the required worst-case error is found. Alternatively, when a closed form is known for the minimax estimator of $\mathcal{U}_L$, one can find a closed form for the MPS estimator as well. This is demonstrated by the following proposition, in which the parameter sets are spherical.

Proposition 2. *Let $\epsilon(\hat{\mathbf{x}}, \mathbf{x}) = E\|\mathbf{x} - \hat{\mathbf{x}}\|^2$ and define the spherical parameter sets $\mathcal{U}_L = \{\mathbf{x} : \|\mathbf{x}\| \leq L\}$. Let $\hat{\mathbf{x}}_{\mathrm{LS}}$ be the LS estimator* (1)*, and let $\epsilon_0 = \mathrm{Tr}((\mathbf{H}^*\mathbf{C}_{\mathbf{w}}^{-1}\mathbf{H})^{-1})$ be the MSE of $\hat{\mathbf{x}}_{\mathrm{LS}}$. Among the class of linear estimators, an MPS estimator is*

$$\hat{\mathbf{x}}_{\mathrm{UP}} = \begin{cases} (\epsilon_m/\epsilon_0)\hat{\mathbf{x}}_{\mathrm{LS}}, & \epsilon_m < \epsilon_0 \\ \hat{\mathbf{x}}_{\mathrm{LS}}, & \epsilon_m \geq \epsilon_0. \end{cases} \tag{6}$$

The more general case of ellipsoidal parameter sets is considered in the following proposition. While a closed form is not known for this case, efficient optimization algorithms may be used to calculate the MPS estimator.

Proposition 3. *Let $\epsilon(\hat{\mathbf{x}}, \mathbf{x}) = E\|\mathbf{x} - \hat{\mathbf{x}}\|^2$ and define the ellipsoidal parameter sets $\mathcal{U}_L = \{\mathbf{x} : \mathbf{x}^*\mathbf{T}\mathbf{x} \leq L^2\}$, for some positive definite matrix $\mathbf{T}$. Among linear estimators, an MPS estimator $\hat{\mathbf{x}}_{\mathrm{UP}} = \mathbf{G}\mathbf{y}$ may be found by solving the quasiconvex problem*

$$\min_{\mathbf{G},\lambda,z} z/\lambda \tag{7}$$
$$\text{s.t.} \begin{cases} \begin{bmatrix} z+\epsilon_m & \mathbf{g}^* \\ \mathbf{g} & \mathbf{I} \end{bmatrix} \succeq 0 \\ \begin{bmatrix} \lambda\mathbf{I} & \mathbf{T}^{-1/2}(\mathbf{I}-\mathbf{G}\mathbf{H})^* \\ (\mathbf{I}-\mathbf{G}\mathbf{H})\mathbf{T}^{-1/2} & \mathbf{I} \end{bmatrix} \succeq 0 \end{cases}$$

where $\mathbf{g}$ is the vector obtained by stacking the columns of $\mathbf{G}\mathbf{C}_{\mathbf{w}}^{1/2}$, and $\mathbf{P} \succeq 0$ indicates that $\mathbf{P}$ is positive semidefinite.

We now consider an MPS estimator which guarantees a required *regret*. The regret is defined as the difference between the estimator's MSE and the best MSE obtainable using a linear estimator $\hat{\mathbf{x}}_o = \mathbf{G}(\mathbf{x})\mathbf{y}$ which is a function of $\mathbf{x}$. Because $\hat{\mathbf{x}}_o$ is linear, it does not achieve zero MSE. This approach is intuitively appealing as it attempts to disregard errors resulting from limitations of linear estimators. It has been shown [6] that the regret is given by

$$\epsilon(\mathbf{G}\mathbf{y}, \mathbf{x}) = \mathrm{Tr}(\mathbf{G}\mathbf{C}_{\mathbf{w}}\mathbf{G}^*) + \mathbf{x}^*(\mathbf{I}-\mathbf{G}\mathbf{H})^*(\mathbf{I}-\mathbf{G}\mathbf{H})\mathbf{x} - \frac{\mathbf{x}^*\mathbf{x}}{1+\mathbf{x}^*\mathbf{H}^*\mathbf{C}_{\mathbf{w}}^{-1}\mathbf{H}\mathbf{x}}. \tag{8}$$

For spherical parameter sets, the linear MPS regret estimator is equivalent to the linear minimax regret estimator, as stated in the following proposition. It follows that the MPS regret estimator can be found as easily as the minimax regret estimator, for which an efficient algorithm is known [6].

Proposition 4. *Let $\mathcal{U}_L = \{\mathbf{x} : \|\mathbf{x}\| \leq L\}$ be a class of parameter sets, let $\epsilon(\hat{\mathbf{x}}, \mathbf{x})$ be the regret* (8)*, and consider the class $\mathcal{E}$ of linear estimators. For any L, an estimator $\hat{\mathbf{x}} \in \mathcal{E}$ is a minimax regret estimator over the set $\mathcal{U}_L$ if, and only if, it is an MPS regret estimator with maximum error ϵ_m equal to $e(L)$ of* (5).

As we have seen, MPS and minimax estimators are based on different design requirements; however, results regarding minimax estimators can often be used to find MPS estimators.

2.3. Application: Channel Estimation

As an application of the MPS estimator, we now consider the problem of preamble-based channel estimation. Specifically, we seek to estimate the impulse response of an unknown channel using a training sequence (preamble) transmitted along with payload data, by comparing received symbols to the known preamble sequence. The channel response is used in many detection algorithms, such as maximum likelihood sequence estimation (MLSE) [12].

Let $\mathbf{c} = (c_0, \ldots c_{N_c-1})^T$ denote an unknown channel response of known length N_c, and let $\mathbf{p} = (p_{-N_c+1}, \ldots p_{N_p-N_c})^T$ denote the known vector of preamble symbols of length N_p. The corresponding received symbols are given by

$$r_k = \sum_{l=0}^{N_c-1} c_l p_{k-l} + w_k, \quad k = 0, 1, \ldots N_p - N_c, \tag{9}$$

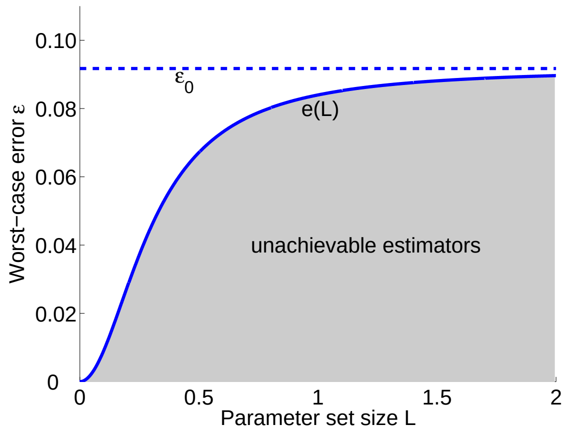

Fig. 1. Worst-case error and parameter set size.

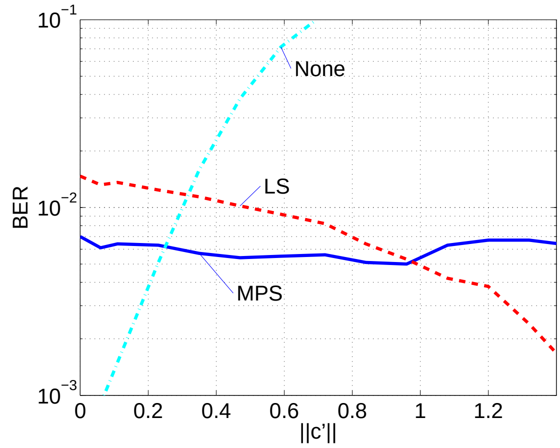

Fig. 2. Bit error rate for various channels.

where w_k is additive white noise with variance σ_w^2. We assume that the channel consists of a direct transmission element $c_0 = 1$ and multipath echoes $\mathbf{c}' = (c_1, \ldots c_{N_c-1})^T$. Thus $\mathbf{r} = \mathbf{p}' + \mathbf{H}\mathbf{c}' + \mathbf{w}$, where $\mathbf{p}' = (p_0, \ldots p_{N_p - N_c})^T$ and $\mathbf{H}$ is a known matrix containing preamble coefficients.

The classical approach to preamble-based channel estimation uses the LS estimator $\hat{\mathbf{c}}' = (\mathbf{H}^*\mathbf{H})^{-1}\mathbf{H}^*(\mathbf{r} - \mathbf{p}')$ [4, 12]. This estimator minimizes the measurement error $\|\mathbf{r} - \mathbf{p}' - \mathbf{H}\hat{\mathbf{c}}'\|^2$. However, we are interested in minimizing the estimation error $\epsilon = E\|\mathbf{c} - \hat{\mathbf{c}}\|^2$, which is correlated with bit error rates (BER) [4]. Since the estimation error is a function of the unknown channel parameter $\mathbf{c}'$, it cannot be directly minimized. The minimax approach allows us to minimize the worst-case error among all possible channels within a known set $\mathcal{U}$. Yet, although we may believe that $\|\mathbf{c}'\|$ is generally small compared with c_0, we cannot explicitly determine a bound on $\|\mathbf{c}'\|$.

The channel estimation error ϵ_m may be viewed as an additional noise source [4], and is therefore a parameter with known implications for the system designer. Thus ϵ_m should be a design parameter, to be chosen together with other system properties, such as SNR requirements and error correction capabilities. The MPS estimator can then be used to maximize the set of channels for which ϵ_m is achieved, in order to guarantee operation for as wide a range of channels as possible.

We consider the problem of estimating a 7-tap channel using the 14-symbol BPSK preamble suggested in [4]. We assume the noise variance is $\sigma_w^2 = 0.1$. The worst-case errors of the LS estimator and of various minimax MSE estimators for spherical parameter sets $\mathcal{U}_L = \{\mathbf{c}' : \|\mathbf{c}'\| \le L\}$ are plotted in Figure 1. By Proposition 1, all of these estimators are also MPS estimators. An engineer constructing a channel estimation system should use such a plot as a design tool, as it demonstrates the tradeoff between channel estimation error and the range of channels for which the error is achieved. For instance, if a maximum error level of 0.1 is acceptable, then clearly the optimal choice is the LS estimator, which guarantees an MSE of $\epsilon_0 < 0.1$ for all channels. However, in some cases the detector degradation provided by such an estimation error may be prohibitive. An engineer may choose a lower maximum channel estimation error while taking into consideration the reduced set of channels for which estimation would be successful. We choose a maximum error of $\epsilon_m = 0.75\epsilon_0$, as this value covers most reasonable channels while substantially reducing the worst-case error.

To compare the performance of LS and MPS estimators, the transmitter, channel and receiver were simulated. Channels with multipath power $\|\mathbf{c}'\|^2$ between 0 and 2 were used. The channel was estimated using both the LS estimator (1) and the MPS estimator (6), and each of the channel estimates was used for MLSE detection of payload data. The bit error rate obtained by the two estimators is plotted in Figure 2. For comparison, a null estimator is also plotted; this "estimator" assumes that $\mathbf{c}' = \mathbf{0}$.

These results demonstrate that in terms of BER, MPS estimation outperforms standard LS estimation for a range of channels. The MPS estimator maintains a BER level around 0.6% in the measured channel range, while the LS estimator results in BER levels above 1% for many common channels.

The LS estimator has modest estimation error requirements, but achieves them for all values of $\mathbf{c}$, while the null estimator can be viewed as an estimator requiring zero estimation error, and achieves this requirement only for $\mathbf{c}' = \mathbf{0}$. With various values of ϵ_m, MPS estimators provide a continuum of choices between these two extremes. An intermediate choice of ϵ_m generally outperforms both extremes.

3. MAXIMUM NOISE LEVEL ESTIMATION

In the previous section, we assumed that the noise covariance $E(\mathbf{w}\mathbf{w}^*)$ is known. In practice, this is rarely the case, and the covariance itself must often be estimated from measurements. In this section we consider the case where $E(\mathbf{w}\mathbf{w}^*) = \sigma^2\mathbf{C}_\mathbf{w}$, for some unknown deterministic *noise level* σ^2, and some known covariance matrix $\mathbf{C}_\mathbf{w}$. This is appropriate, for example, when the noise variables are independent and identically distributed, in which case $\mathbf{C}_\mathbf{w} = \mathbf{I}$ and σ^2 is the noise variance. The estimation techniques used so far require complete knowledge of the noise covariance; thus they cannot be applied to this problem, unless noise parameters are estimated from measurements, which increases computational complexity and may be unreliable.

As an alternative approach, we propose to estimate $\mathbf{x}$ for as large a range of noise levels as possible, while maintaining error requirements. To this end, we define an error function $\epsilon_{\sigma^2}(\hat{\mathbf{x}}, \mathbf{x})$, such as the MSE or the regret, and require some level of performance $\epsilon_{\sigma^2}(\hat{\mathbf{x}}, \mathbf{x}) \leq \epsilon_m$ to be satisfied over a known range $\mathbf{x} \in \mathcal{U}$. The *noise robustness* $\hat{\sigma}^2$ of an estimator $\hat{\mathbf{x}}$ is defined as the maximum σ^2 for which the performance requirement is satisfied,

$$\hat{\sigma}^2(\hat{\mathbf{x}}) = \max\{\sigma^2 : \ \epsilon_{\sigma^2}(\hat{\mathbf{x}}, \mathbf{x}) \leq \epsilon_m, \ \forall \mathbf{x} \in \mathcal{U}\}. \tag{10}$$

The *maximum noise level (MNL) estimator* $\hat{\mathbf{x}}_{\mathrm{UN}}$ (among a class of estimators $\mathcal{E}$) is the estimator maximizing the noise robustness among all estimators in $\mathcal{E}$, for given $\mathcal{U}$, $\epsilon_{\sigma^2}(\hat{\mathbf{x}}, \mathbf{x})$ and ϵ_m,

$$\hat{\mathbf{x}}_{\mathrm{UN}} = \arg \max_{\hat{\mathbf{x}} \in \mathcal{E}} \hat{\sigma}^2(\hat{\mathbf{x}}). \tag{11}$$

We then have the following proposition, which states that MNL estimators are also minimax estimators.

Proposition 5. *Suppose the error function ϵ of interest is continuous in σ^2. Then, an MNL estimator $\hat{\mathbf{x}}_{\mathrm{UN}}$ within a class of estimators $\mathcal{E}$ is a minimax estimator for the parameter set $\mathcal{U}$, with noise level $\sigma_1^2 = \hat{\sigma}^2(\hat{\mathbf{x}}_{\mathrm{UN}})$.*

It follows that the MNL estimator can be found if an algorithm for finding the minimax estimator is known. This is done by searching among all noise levels for the unique minimax estimator with the appropriate worst-case error, which leads to the following proposition.

Proposition 6. *Let $\mathcal{U} = \{\mathbf{x} : \ \|\mathbf{x}\| \leq L\}$ and let $\epsilon_{\sigma^2}(\hat{\mathbf{x}}, \mathbf{x})$ be the MSE. For a given maximum error ϵ_m, a linear MNL estimator is*

$$\hat{\mathbf{x}}_{\mathrm{UN}} = \begin{cases} (1 - \epsilon_m / L^2)\hat{\mathbf{x}}_{\mathrm{LS}}, & L^2 > \epsilon_m \\ \mathbf{0}, & L^2 \leq \epsilon_m, \end{cases} \tag{12}$$

where $\hat{\mathbf{x}}_{\mathrm{LS}}$ is the LS estimator (1).

It is instructive to compare the closed forms obtained for the MPS estimator (6) and the MNL estimator (12), when spherical parameter sets are used. Both are shrunken least-squares estimators [10], and can thus be viewed as a compromise between the least-squares estimator and the zero estimator. When ϵ_m is increased, an increase in either the parameter set or noise level is allowed. However, a larger parameter set is achieved by an estimator closer to the LS estimator (which provides constant error for all $\mathbf{x}$); while a larger noise level is achieved by an estimator closer to the zero estimator (which provides zero error for the nominal value $\mathbf{x} = \mathbf{0}$). Thus, increasing the maximum allowed error has opposite effects, depending on whether the goal is to increase the parameter robustness or the noise robustness.

4. DISCUSSION

In many estimation problems, it is possible to define the maximum allowed estimation error ϵ_m. In some applications, the MSE provided by the LS estimator is acceptable; in these cases, the LS estimator is optimal, as it guarantees this MSE for any value of the parameters. However, when the required error is smaller than the error obtained by the LS estimator, the requirements may still be obtained for some parameter values. In these cases we may seek to maximize the parameter set for which requirements are satisfied, resulting in the MPS estimator. Alternatively, we may seek to maintain the required error for as large a noise level as possible, using the MNL estimator. As we have seen, in many cases, the MPS and MNL estimators equal the minimax estimator whose worst-case error is ϵ_m. This allows us to efficiently calculate many such estimators.

The maximum allowed error is often a function of system design parameters, and can be influenced by design decisions. In such cases, a plot of the worst-case error as a function of the size of the parameter set (as in Figure 1) can be used as a design tool. Such a plot can be interpreted in two complementary ways. It describes the worst-case error obtained if a minimax estimator is used with a given parameter set bound; it also defines the size of the parameter set obtained if an MPS estimator is used with a given maximum error. Thus, such a plot can be used to select a meaningful value for the maximum error, based on the tradeoff between estimation error and parameter set bound.

The choice of an appropriate estimator for a given problem depends on the data available to the designer. The maximum allowed estimation error is an example of added information which may be known to the designer; as we have shown, incorporating this information can considerably improve estimation performance.

5. ACKNOWLEDGEMENT

The authors are deeply grateful to Prof. Y. Ben-Haim for many helpful comments provided throughout the work on this paper.

6. REFERENCES

[1] Y. Ben-Haim, *Information-Gap Decision Theory: Decisions under Severe Uncertainty.* Academic Press, 2001.

[2] Y. Ben-Haim, “Set-models of information-gap uncertainty: axioms and an inference scheme,” *J. Franklin Inst.*, 336: 1093–1117, 1999.

[3] Z. Ben-Haim and Y. C. Eldar, “Estimation with maximum error requirements,” submitted to *IEEE Trans. Sig. Proc.*, 2004; Tech. Rep. #1434, Dept. of Elec. Eng., Technion, May 2004.

[4] S. N. Crozier, D. D. Falconer, and S. A. Mahmoud, “Least sum of square errors (LSSE) channel estimation,” *IEE Proc.-F*, 138 (4): 371–378, Aug. 1991.

[5] Y. C. Eldar, A. Ben-Tal and A. Nemirovski, “Robust mean-squared error estimation with bounded data uncertainties,” to appear in *IEEE Trans. Sig. Proc.*, 2004.

[6] Y. C. Eldar, A. Ben-Tal and A. Nemirovski, “Linear minimax regret estimation with bounded data uncertainties,” to appear in *IEEE Trans. Sig. Proc.*, 2004.

[7] P. J. Huber, “Robust estimation of a location parameter,” *Ann. Math. Statist.*, 35 (1): 73–101, 1964.

[8] S. A. Kassam and H. V. Poor, “Robust techniques for signal processing: a survey,” *Proc. IEEE*, 73: 433–481, 1985.

[9] S. M. Kay, *Fundumentals of Statistical Signal Processing: Estimation Theory.* Prentice Hall, 1993.

[10] L. S. Mayer and T. A. Wilke, “On biased estimation in linear models,” *Technometrics*, 15: 497–508, 1973.

[11] M. S. Pinsker, “Optimal filtering of square-integrable signals in Gaussian noise,” *Problems in Inform. Transmission*, 16: 120-133, 1980.

[12] T. Rappaport, *Wireless Communications: Principles and Practice.* Prentice Hall, 1996.

[13] A. N. Tichonov and V. Y. Arsenin, *Solution of Ill-Posed Problems.* V. H. Winston, 1977.