

Copyright ©IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

PERFORMANCE OF THE STOCHASTIC MV-PURE ESTIMATOR WITH EXPLICIT MODELING OF UNCERTAINTY

Tomasz Piotrowski

Dept. of Informatics
Nicolaus Copernicus University
Grudziądzka 5, 87-100 Toruń, Poland
E-mail address: tpotrowski@is.umk.pl

Isao Yamada

Dept. of Communications and Integrated Systems
Tokyo Institute of Technology
Tokyo 152-8552, Japan
E-mail address: isao@sp.ss.titech.ac.jp

ABSTRACT

The stochastic MV-PURE estimator is a robust to mismatches in model knowledge linear estimator for highly noisy and ill-conditioned cases. To date, its properties were analyzed in the theoretical settings of perfect model knowledge and thus could not explain clearly the reason behind its superior performance compared to the Wiener filter observed in simulations in practical cases of imperfect model knowledge. In this paper we derive closed form expressions of the mean-square-error (MSE) of both Wiener filter and the stochastic MV-PURE estimator for the case of perturbed singular values of a model matrix in the linear model considered. These expressions provide in particular conditions under which the stochastic MV-PURE estimator achieves smaller MSE not only than Wiener filter, but also than its full-rank version, the minimum-variance distortionless (MVDR) estimator in such settings. We provide numerical simulations confirming the main theoretical results presented.

Index Terms— Stochastic MV-PURE estimator, parameter estimation, reduced-rank estimation, uncertainty modeling

1. INTRODUCTION

The stochastic MV-PURE estimator [1], based on the previously introduced deterministic *minimum-variance pseudo-unbiased reduced-rank estimator* (MV-PURE) [2, 3], is a reduced-rank linear estimator designed for the stochastic linear model $\mathbf{y} = H\mathbf{x} + \sqrt{\epsilon}\mathbf{n}$. Its reduced-rank approach serves to combat ill-conditioning of the model by introducing small amount of bias for huge savings in variance [1–4] (see also [5–7] for an in-depth discussion of the benefits of reduced-rank approach in estimation and filtering). Moreover, in the stochastic case, where the mean vector and the covariance matrix of the input random vector \mathbf{x} to be estimated are assumed available, the stochastic MV-PURE estimator exhibited in simulations in [1] a significantly improved robustness to mismatches in model knowledge compared to the theoretically optimal [in the mean-square-error (MSE) sense] minimum-mean-square-error (MMSE) estimator (Wiener filter) [8, 9]. However, while the improved performance of the stochastic MV-PURE estimator over the widely used minimum-variance distortionless (MVDR) estimator [10, 11] in highly noisy and ill-conditioned situations has been established in [4], the improved performance of the stochastic MV-PURE estimator over the MMSE estimator has been only demonstrated via numerical simulations in [1]. What is known, however, is that the MVDR estimator which is recognized in [1] as the full-rank (a special case of) stochastic MV-PURE estimator, provides improved performance

over the MMSE estimator in many cases of imperfect model knowledge [10, 11]. This fact encouraged us to begin work on comparing the performance of the stochastic MV-PURE, MVDR, and MMSE estimators with explicit modeling of uncertainty in the stochastic linear model considered, and this paper describes first results in this direction.

In this paper we consider the case of white additive noise \mathbf{n} and white input vector \mathbf{x} , which occurs in many scenarios in signal processing case. We take uncertainty in the knowledge of a model matrix H into account by perturbing its singular values, and then observing its effect on the performance of the stochastic MV-PURE, MVDR, and MMSE estimators.

Namely, in the settings described above, we derive in Section 3 explicit expressions of the MSE for the stochastic MV-PURE, MVDR, and MMSE estimators. Based on them, we provide the main result of this paper: we define a function measuring the gap in performance between the MMSE and stochastic MV-PURE estimators, and find analytically its argument of the minimum ($\arg \min$) in the most important case from the practical applications standpoint, where the perturbed singular values are not far removed from the singular values of H . Additionally, we prove that the larger the power of the additive noise, the lower the optimal rank of the stochastic MV-PURE estimator. This result extends theoretical results of [4] to the case of imperfect model knowledge, and in particular, confirms that in highly noisy settings the (full-rank) MVDR estimator is inherently inadequate.

We close with a numerical example in Section 4 showing the strength of the main result of this paper, where in particular we observe that the stochastic MV-PURE estimator achieves lower MSE than the theoretically MSE-optimal MMSE estimator in our settings in vicinity of the analytically found $\arg \min$ point.

2. PRELIMINARIES

Consider the stochastic linear model of the form:

$$\mathbf{y} = H\mathbf{x} + \sqrt{\epsilon}\mathbf{n}, \quad (1)$$

where \mathbf{y} , \mathbf{x} , \mathbf{n} are random vectors representing observed signal, signal to be estimated, and additive noise, respectively, and $H \in \mathbb{R}^{n \times m}$ is a known matrix of rank m , and $\epsilon > 0$ is a known constant representing noise power. We assume that \mathbf{x} and \mathbf{n} have zero mean, are uncorrelated: $R_{\mathbf{x}\mathbf{n}} = 0$, and white: $R_{\mathbf{x}} = I_m$ and $R_{\mathbf{n}} = I_n$. From the previous assumptions, $R_{\mathbf{y}} = HH^t + \epsilon I_n$ is positive definite and $R_{\mathbf{y}\mathbf{x}} = H$. We denote the singular value decomposition (SVD) of H by

$$SVD(H) = U\Sigma V^t, \quad (2)$$

with singular values $\sigma_i, i = 1, 2, \dots, m$ organized in non-increasing order, and by $V_r = (v_1, \dots, v_r) \in \mathbb{R}^{m \times r}$ we denote the first r columns of V .

We consider the problem of linear estimation of \mathbf{x} given \mathbf{y} , with MSE as the performance criterion. Thus, we seek to find a fixed matrix $W \in \mathbb{R}^{m \times n}$, called here an estimator, for which the estimate of \mathbf{x} of the form $\hat{\mathbf{x}} = W\mathbf{y}$ is optimal with respect to a measure related to the mean-square-error of $\hat{\mathbf{x}}$ given by:

$$J(W) = \text{tr}[WR_y W^t] - 2\text{tr}[WR_{y\mathbf{x}}] + \text{tr}[R_{\mathbf{x}}]. \quad (3)$$

In this paper we study the performance of the following three estimators: the minimum-mean-square-error (MMSE) estimator (Wiener filter) [8, 9], the minimum-variance distortionless (MVDR) estimator [8, 10, 11], and the stochastic MV-PURE estimator [1].

The MMSE estimator is defined as an estimator achieving lowest MSE (3) among linear estimators, minimizing directly the MSE:

$$\text{minimize } J(W). \quad (4)$$

Under our assumptions, this estimator can be expressed as:

$$W_{MMSE} = R_{\mathbf{xy}} R_{\mathbf{y}}^{-1} = H^t (H H^t + \epsilon I_n)^{-1} = V \Sigma^t (\Sigma \Sigma^t + \epsilon I_n)^{-1} U^t, \quad (5)$$

with:

$$J(W_{MMSE}) = m - \sum_{i=1}^m \frac{\sigma_i^2}{\sigma_i^2 + \epsilon}. \quad (6)$$

The MVDR estimator introduces the distortionless constraint to the MSE optimization problem as follows:

$$\begin{cases} \text{minimize } J(W) \\ \text{subject to } WH = I_m, \end{cases} \quad (7)$$

which under our assumptions produces as the solution

$$W_{MVDR} = (H^t H)^{-1} H^t = H^\dagger = V \Sigma^\dagger U^t, \quad (8)$$

where H^\dagger is the Moore-Penrose pseudoinverse of H [12] and Σ^\dagger is a diagonal matrix with diagonal entries $\sigma_1^{-1}, \dots, \sigma_m^{-1}$. We have:

$$J(W_{MVDR}) = \epsilon \sum_{i=1}^m \frac{1}{\sigma_i^2}. \quad (9)$$

Finally, the stochastic MV-PURE estimator is defined as an optimal reduced-rank generalization of the MVDR estimator, and is a solution to the following problem for a given rank constraint $r \leq m$:

$$\begin{cases} \text{minimize } J(W_r) \\ \text{subject to } W_r \in \bigcap_{\ell \in \mathcal{J}} \mathcal{P}_r^\ell, \end{cases} \quad (10)$$

where

$$\mathcal{P}_r^\ell = \arg \min_{W_r \in \mathcal{X}_r^{m \times n}} \|W_r H - I_m\|_\ell^2, \quad \ell \in \mathcal{J}, \quad (11)$$

where $\mathcal{X}_r^{m \times n} := \{W_r \in \mathbb{R}^{m \times n} : \text{rk}(W_r) \leq r \leq m\}$ where $\text{rk}(X)$ stands for the rank of $X \in \mathbb{R}^{m \times n}$ and where \mathcal{J} is the

index set of all unitarily invariant norms.¹ As shown in [1], under our assumptions of white input vector \mathbf{x} and noise \mathbf{n} , the stochastic MV-PURE estimator is of specially simple form:

$$W_{MV-PURE}^t = V \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} V^t W_{MVD R} = V_r V_r^t H^\dagger, \quad (12)$$

with:

$$J(W_{MV-PURE}^t) = m - r + \epsilon \sum_{i=1}^r \frac{1}{\sigma_i^2}. \quad (13)$$

In the theoretical case of model (1), the MMSE estimator (5) is by definition MSE-optimal. However, this need not be the case if we introduce an uncertainty in the knowledge of model (1) as discussed below.

3. RESULTS FOR PERTURBED SINGULAR VALUES OF THE MODEL MATRIX

We introduce uncertainty into our considerations by replacing model (1) with

$$\mathbf{y} = (H + \Delta H)\mathbf{x} + \sqrt{\epsilon}\mathbf{n}, \quad (14)$$

where

$$SVD(H + \Delta H) = U \Gamma V^t, \quad (15)$$

i.e., we replace the singular values $\sigma_i, i = 1, 2, \dots, m$ of H [cf. (2)] with perturbed singular values $\gamma_i, i = 1, 2, \dots, m$.

We add superscript ⁽¹⁴⁾ to denote the quantities such as covariance matrices and MSE which are related to model (14). In particular, $R_{\mathbf{y}}^{(14)}$ is the covariance matrix of \mathbf{y} in (14), and $J^{(14)}(W)$ is the MSE of estimator W in model (14).

We have:

$$R_{\mathbf{y}}^{(14)} = (H + \Delta H)(H + \Delta H)^t + \epsilon I_n, \quad (16)$$

and

$$R_{\mathbf{yx}}^{(14)} = H + \Delta H, \quad (17)$$

and the MSE is expressed in terms of model (14) as

$$J^{(14)}(W) = \text{tr}[WR_{\mathbf{y}}^{(14)}W^t] - 2\text{tr}[WR_{\mathbf{yx}}^{(14)}] + m. \quad (18)$$

Using expressions (2), (5) and (15), it can be readily verified that the mean-square-error (18) of the MMSE estimator (5) for the perturbed model (14) can be expressed as:

$$J^{(14)}(W_{MMSE}) = \sum_{i=1}^m \frac{\sigma_i^2(\gamma_i^2 + \epsilon)}{(\sigma_i^2 + \epsilon)^2} - 2 \sum_{i=1}^m \frac{\gamma_i \sigma_i}{\sigma_i^2 + \epsilon} + m. \quad (19)$$

Similarly, using (2), (8), and (15) we obtain that the mean-square-error (18) of the MVDR estimator (8) for the perturbed model (14) is of the form:

$$J^{(14)}(W_{MVDR}) = \epsilon \sum_{i=1}^m \frac{1}{\sigma_i^2} + \sum_{i=1}^m \left(\frac{\gamma_i}{\sigma_i} - 1 \right)^2, \quad (20)$$

¹Matrix norm ℓ is unitarily invariant if it satisfies $\|UXV\|_\ell = \|X\|_\ell$ for all orthogonal $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ and all $X \in \mathbb{R}^{m \times n}$ [13, p. 203]. The Frobenius, spectral, and trace (nuclear) norms are examples of unitarily invariant norms.

and from (2), (12), and (15) we obtain analogously for the stochastic MV-PURE estimator (12) that:

$$J^{(14)}(W_{MV-PURE}^r) = m - r + \epsilon \sum_{i=1}^r \frac{1}{\sigma_i^2} + \sum_{i=1}^r \left(\frac{\gamma_i}{\sigma_i} - 1 \right)^2. \quad (21)$$

Naturally, expressions (19), (20) and (21) reduce to (6), (9) and (13), respectively, for $\Delta H = 0$ (and hence $\gamma_i = \sigma_i$), as they should.

Although the above expressions for the MSE in the perturbed model (14) of the estimators under consideration look complicated at first, they are in fact standard quadratic functions of singular values γ_i , $i = 1, 2, \dots, m$ of the perturbed matrix $H + \Delta H$, parametrized by singular values σ_i , $i = 1, 2, \dots, m$ of H , and the noise power ϵ .

We are now in a position to compare directly the performance of the MMSE and stochastic MV-PURE estimators for the perturbed model (14). To this end, consider the following function $f_r : \mathbb{R}_+^m \rightarrow \mathbb{R}$ of $\gamma = (\gamma_1, \dots, \gamma_m)$ with $\gamma_1 \geq \dots \geq \gamma_m > 0$, for $r \in (1, 2, \dots, m)$:

$$f_r(\gamma) := J^{(14)}(W_{MV-PURE}^r) - J^{(14)}(W_{MMSE}) = A - B - r, \quad (22)$$

where

$$A = \left[\epsilon \sum_{i=1}^r \frac{1}{\sigma_i^2} + \sum_{i=1}^r \left(\frac{\gamma_i}{\sigma_i} - 1 \right)^2 \right] - \left[\sum_{i=1}^r \frac{\sigma_i^2(\gamma_i^2 + \epsilon)}{(\sigma_i^2 + \epsilon)^2} - 2 \sum_{i=1}^r \frac{\gamma_i \sigma_i}{\sigma_i^2 + \epsilon} \right], \quad (23)$$

and

$$B = \sum_{i=r+1}^m \frac{\sigma_i^2(\gamma_i^2 + \epsilon)}{(\sigma_i^2 + \epsilon)^2} - 2 \sum_{i=r+1}^m \frac{\gamma_i \sigma_i}{\sigma_i^2 + \epsilon}. \quad (24)$$

For $r = m$, we have $B = 0$ and the quadratic function $f_r = f_m$ is strictly convex, since it has a convex domain, and its Hessian: a diagonal matrix with diagonal entries (which are positive for all $\sigma_i > 0$ and $\epsilon > 0$) of the following form:

$$H(f_r)_{i,i} = 2 \left(\frac{1}{\sigma_i^2} - \frac{\sigma_i^2}{(\sigma_i^2 + \epsilon)^2} \right), \quad i = 1, \dots, m, \quad (25)$$

is positive definite on the domain, which conditions ensure strict convexity of f_r [14]. Calculating derivatives show that the global minimum is achieved at

$$\gamma_i^{\min} = \frac{\epsilon \sigma_i (\sigma_i^2 + \epsilon)}{(\sigma_i^2 + \epsilon)^2 - \sigma_i^4}, \quad i = 1, \dots, m. \quad (26)$$

By taking the limits $\epsilon \rightarrow 0$ and $\epsilon \rightarrow \infty$ in (26), it is seen that $\gamma_i^{\min} \in (\sigma_i/2, \sigma_i)$, and the derivative of (26) with respect to ϵ reveals that the value of γ_i grows monotonically from $\sigma_i/2$ to σ_i with increasing noise level ϵ . Moreover, the optimal solution is feasible, as it can be verified by calculating the derivative of (26) with respect to σ_i , that if $\sigma_i > \sigma_{i+1}$, then also $\gamma_i^{\min} > \gamma_{i+1}^{\min}$ for a given noise level ϵ .

When $r < m$, f_r is no longer convex, as it can be observed that its Hessian ceases to be positive semidefinite on the domain of f_r , since the term $2\sigma_i^{-2}$ vanishes in (25) for $i = r+1, \dots, m$ in such a case. However, for $r < m$ we

can exploit the fact that A (23) and B (24) can be optimized independently as follows.

Term A is a function only of $\gamma_1, \dots, \gamma_r$, and is a strictly convex function of $\gamma_1, \dots, \gamma_r$ even for $r < m$ since its Hessian is the $r \times r$ principal submatrix of (25), with the global minimum being achieved at the point $(\gamma_1^{\min}, \dots, \gamma_r^{\min})$, where γ_i^{\min} are given by (26) for $i = 1, \dots, r$. Analogously, calculating Hessian of $-B$ reveals that it is a strictly concave function of $\gamma_{r+1}, \dots, \gamma_m$ with the global maximum at $\gamma_j^{\max} = \sigma_j + \epsilon \sigma_j^{-1}$, $j = r+1, \dots, m$.² We note further that for all $j = r+1, \dots, m$ we have $\gamma_j^{\max} > \sigma_j$.

From the practical perspective, we are primarily interested in the case where distances $|\sigma_i - \gamma_i|$ do not exceed some reasonable threshold. The above discussion shed light on such settings, as it is seen that the global minimum of term A (23) lies in the segment $[\sigma_i/2, \sigma_i]$ for $i = 1, \dots, r$, and the minimum of $-B$ [with B given in (24)] is always found at the boundary of the feasible set, due to its strict concavity and the global maximum of $-B$ at $\gamma_j^{\max} = \sigma_j + \epsilon \sigma_j^{-1}$, $j = r+1, \dots, m$. Thus, as an example, the following optimization problem, for a given $r \in (1, 2, \dots, m)$ and $0 < c < \sigma_m$:

$$\begin{cases} \text{find} & \arg \min f_r \\ \text{subject to} & c \leq \gamma_i \leq \sigma_i + \epsilon \sigma_i^{-1}, \quad i = 1, \dots, m, \end{cases} \quad (27)$$

produces as the solution

$$\begin{cases} \gamma_i^{\min} = \epsilon \sigma_i (\sigma_i^2 + \epsilon) / ((\sigma_i^2 + \epsilon)^2 - \sigma_i^4) & i = 1, \dots, r, \\ \gamma_i^{\min} = c & i = r+1, \dots, m, \end{cases} \quad (28)$$

such that $\gamma_1^{\min} \geq \gamma_2^{\min} \geq \dots \geq \gamma_m^{\min} > 0$ with $\gamma_i^{\min} \in (\sigma_i/2, \sigma_i)$ for $i = 1, \dots, r$. We will demonstrate this result in a numerical example in Section 4, where we will obtain in particular that $f_r(\gamma^{\min}) < 0$. This fact is of major importance as it shows that the reduced-rank stochastic MV-PURE estimator is capable of achieving lower MSE under explicit modeling of uncertainty in model (14) than the (MSE-optimal under perfect model knowledge) MMSE estimator.³

An in-depth discussion of conditions under which one has $f_r(\gamma^{\min}) < 0$ for a given rank constraint r is beyond the scope of this paper, and is the subject of an ongoing research. Clearly, one may simply insert expressions (28) into (22) to find it out (possibly aided with a symbolic solver), but a more general and elegant solution is possible.

The above problem is clearly related to the problem of finding rank constraint r under which the MSE of the stochastic MV-PURE estimator is minimized. We present below a Proposition showing a simple but insightful condition which shed light onto this problem.

Proposition 1 *Let us choose natural r_0, r_1 such that $1 \leq r_0 < r_1 \leq m$. If the power of additive noise ϵ is such that*

$$\sqrt{\epsilon} > \sigma_{r_0+1} \quad (29)$$

then

$$J^{(14)}(W_{MV-PURE}^{r_0}) < J^{(14)}(W_{MV-PURE}^{r_1}). \quad (30)$$

²Note that depending on the value of ϵ , this optimum point may be infeasible, as $\sigma_j > \sigma_{j+1}$ does not necessarily imply $\gamma_j^{\min} > \gamma_{j+1}^{\min}$, $j = r+1, \dots, m-1$. However, this does not affect the discussion below.

³In our case perfect model knowledge corresponds to $\gamma_i = \sigma_i$ for $i = 1, \dots, m$.

Proof: From (20) and (21) we note that to prove the Proposition it suffices to show that

$$\epsilon \sum_{i=r_0+1}^{r_1} \frac{1}{\sigma_i^2} + \sum_{i=r_0+1}^{r_1} \left(\frac{\gamma_i}{\sigma_i} - 1 \right)^2 > r_1 - r_0. \quad (31)$$

Assume now that $\sqrt{\epsilon} > \sigma_{r_0+1}$, which implies that $\sqrt{\epsilon} > \sigma_i$ for all $i = r_0 + 1, \dots, m$ due to singular values organized in nonincreasing order. By this assumption, we have that

$$\epsilon \sum_{i=r_0+1}^{r_1} \frac{1}{\sigma_i^2} > r_1 - r_0, \quad (32)$$

which completes the proof. \square

We note that in the above Proposition condition (29) guarantees that (30) holds uniformly for all feasible γ_i , $i = 1, \dots, m$. Indeed, the second term on the left hand side of (31) shows that the larger the discrepancy between γ_i and σ_i , the larger the gain in performance in (30) under condition (29).

Moreover, the above result mirrors the results of [4] derived in theoretical settings of perfect model knowledge. It is encouraging therefore that results similar to those of [4] hold under explicit modeling of uncertainty in the model knowledge.

4. NUMERICAL EXAMPLE

We close with a numerical example illustrating the relation between the MSE of the MMSE (19) and the stochastic MV-PURE (21) estimators for the perturbed model (14). For clarity, we limit ourselves to a very small example, where $n = 4$, $m = 2$ and the rank constraint $r = 1$, which can be thus easily pictured on a 2-dimensional figure. The matrix H has Gaussian entries of zero mean and unit variance, with singular values $\sigma_1 = 1.8777$, $\sigma_2 = 0.9498$. The noise level is such that $\sqrt{\epsilon} = \sigma_2 + 0.05$ (thus, in view of Proposition 1, we must have $J^{(14)}(W_{MV-PURE}^1) < J^{(14)}(W_{MV-PURE}^2) = J^{(14)}(W_{MVD R})$ for all $\gamma_1 \geq \gamma_2 > 0$), and we set $c = 0.01$ for (28). We draw below function f_r (22) for $r = 1$ in vicinity of its minimum $f_r(\gamma^{\min})$, where γ^{\min} is given by (28). In our case the argument of the minimum is $\gamma_1^{\min} = 1.0554$, $\gamma_2^{\min} = c = 0.01$ [cf. (28)], and this minimum is marked by the big red dot on the figure. The value at the minimum is $f_r(1.0554, 0.01) = -0.2521$.

We would like also to note that the above numerical example is perhaps simplest possible, as the only non-trivial rank constraint is $r = 1$, which essentially means that the stochastic MV-PURE estimator (12) is a matrix of scaled columns (rows). Nevertheless, even in such simple settings we could present the benefits of the reduced-rank approach of the stochastic MV-PURE estimator in estimation under model uncertainty.

5. REFERENCES

- [1] T. Piotrowski, R. L. G. Cavalcante, and I. Yamada, "Stochastic MV-PURE estimator: Robust reduced-rank estimator for stochastic linear model," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1293–1303, Apr. 2009.
- [2] I. Yamada and J. Elbadraoui, "Minimum-variance pseudo-unbiased low-rank estimator for ill-conditioned inverse problems," in *Proc. ICASSP*, Toulouse, France, May 2006, pp. 325–328.
- [3] T. Piotrowski and I. Yamada, "MV-PURE estimator: Minimum-variance pseudo-unbiased reduced-rank estimator for linearly constrained ill-conditioned inverse problems," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3408–3423, Aug. 2008.
- [4] T. Piotrowski and I. Yamada, "Why the stochastic MV-PURE estimator excels in highly noisy situations?," in *Proc. ICASSP*, Taipei, Taiwan, Apr. 2009, pp. 3081–3084.
- [5] L. L. Scharf, "The SVD and reduced rank signal processing," *Signal Process.*, vol. 25, pp. 113–133, 1991.
- [6] Y. Yamashita and H. Ogawa, "Relative Karhunen-Loeve transform," *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 371–378, Feb. 1996.
- [7] Y. Hua, M. Nikpour, and P. Stoica, "Optimal reduced-rank estimation and filtering," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 457–469, Mar. 2001.
- [8] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, New Jersey, 2000.
- [9] D. G. Luenberger, *Optimization by Vector Space Methods*, John Wiley & Sons, New York, 1969.
- [10] X. Wang and H. V. Poor, *Wireless Communication Systems*, Prentice Hall, Upper Saddle River, 2004.
- [11] A. B. Gershman and N. D. Sidiropoulos, Eds., *Space-Time Processing for MIMO Communications*, John Wiley & Sons, Chichester, 2005.
- [12] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications, Second Edition*, Springer Verlag, New York, 2003.
- [13] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.
- [14] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, 2004.

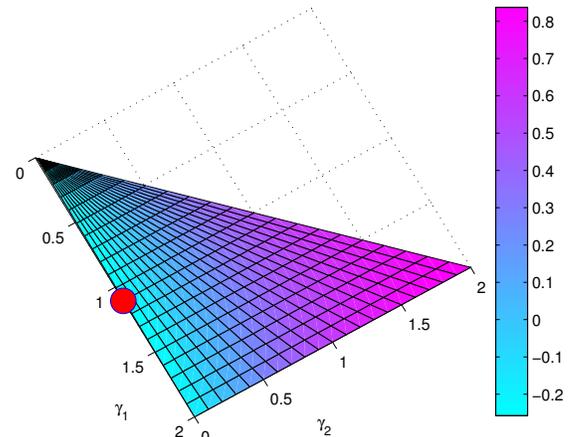


Fig. 1. We obtain the improved performance of the stochastic MV-PURE estimator over the MMSE estimator around minimum of f_r , as in this region we have $f_r < 0$ for $r = 1$.