POWER ALLOCATION FOR GAUSSIAN MIXTURE MODEL PRIOR KNOWLEDGE IN WIRLESS SENSOR NETWORKS

Z. Azmat and H. D. Tuan**

Faculty of Engineering and Information Technology,
University of Technology, Sydney,
Australia

ABSTRACT

This paper presents power allocation in nonlinear sensor networks for Gaussian Mixture (GM) information source. The observations of sensors are transmitted through independent Rayleigh flat fading channels to a fusion centre (FC). Transmit Power is optimally allocated to sensor nodes so as to minimize the mean square error (MSE) of estimate at FC. Bayesian linear and optimal nonlinear estimators are deployed at FC to compare the proposed optimal and uniform power allocation among sensors. Extensive simulations validate that the proposed Bayesian linear estimator with optimized power gains effectively works for GM prior distribution.

Index Terms— Wireless Sensor Networks, Gaussian Mixture Models, Unscented Transformations

1. INTRODUCTION

Target localization in Wireless Sensor Networks (WSNs) has been an active research area due to its immense industrial importance and wide range of applications [1], [2]. Target localization with the aid of fusing data from a set of sensors presents many design challenges due to limited energy resources inherent to WSN. The goal is to produce an estimator that has minimum mean square error (MMSE) under constraints on transmission power at the sensors.

Recently [3] considered power allocation in sensor network assuming that information source follows Gaussian distribution. However, Gaussian Mixture distribution is a more judicious choice for prior knowledge because any non-Gaussian distribution can be represented by sum of Gaussian distributions [4]. It can also be verified that posterior distribution for Gaussian Mixture prior knowledge and Gaussian noise is also a Gaussian Mixture. Motivated by these facts and the MSE optimality of the MMSE estimator for GM distribution, we have adopted this model in our paper.

In [5], the source localization problem has been addressed for Gaussian Mixture Model as prior distribution which is a more judicious choice. [6] shows derivation for closed form expressions for GM prior knowledge with only Bayesian linear model. Moreover, no analytical expression for MSE has been derived in [6]. In this paper, we consider Gaussian Mixture prior distribution in both linear and nonlinear sensor networks. An SDP based convex optimization problem has been solved for Gaussian Mixture Model prior knowledge to allocate power to a group of sensors so as to minimize the MSE subject to a total transmit power. Unlike [5], which only considers Bayesian linear models, we consider nonlinear sensor network in our formulation. Moreover, unlike all previous works [7] consisting mainly of linear sensor networks (LSNs) for locating a static target, the gaussian mixture based optimal power Wiener filters and MMSE estimators are shown computationally tractable in our work.

To the best of our knowledge,we have offered the power allocation for MMSE estimator for nonlinear senor network (NSN) with GM prior knowledge for the first time. The major contribution of this paper is that a suboptimal power allocation for Bayesian MMSE estimator has been proposed for first time using the bounds suggested in [6]. Since Gaussian Mixture is a very generic information source, the proposed estimator can be used to evaluate the MSE performance of different systems under various operating conditions when power allocation for sensor nodes is of interest.

The rest of paper is structured as follows. Section II describes the system model and formulation of the problem. Section III illustrates the derivation of Wiener filtering for Gaussian Mixture prior knowledge. Section V provides the simulation results which is followed by concluding remarks in Section VI.

Most of the notations used in the paper are described here. Bold lower-case and upper-case symbols are used to represent vectors and matrices respectively. By $\mathbf{A} \succeq 0$ it means $\mathbf{A}$ is a positive definite matrix. $\mathsf{diag}[a_i]_{i=1,2,...N}$ or $\mathsf{diag}[a_1, a_2, ..., a_N]$ is a diagonal matrix. $\sqrt{\mathbf{q}}$ for a vector $\mathbf{q}$ with nonnegative components is component-wise understood. $\mathbb{E}[.]$ is the expectation operator. For a random variable (RV) $\boldsymbol{x}$, the notation $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{x}; \boldsymbol{m_x}, \mathbf{R_x})$ means $\boldsymbol{x}$ is Gaussian RV with the moments $\boldsymbol{m_x}$ and $\mathbf{C_x}$.

*zeeshan.a.shaikh@student.uts.edu.au, tuan.hoang@uts.edu.au.

2. PROBLEM FORMULATION

Consider a target $\boldsymbol{x} \sim \sum_{i=1}^{L} \lambda_i \mathcal{N}(\boldsymbol{x}; \boldsymbol{m}_{\boldsymbol{x}}^{(i)}, \boldsymbol{C}_{\boldsymbol{x}}^{(i)})$ in N-dimensional space (i.e. rough initial information of $\boldsymbol{x}$ expressed by $\left\{\lambda_i, \boldsymbol{m}_{\boldsymbol{x}}^{(i)}, \boldsymbol{C}_{\boldsymbol{x}}^{(i)}\right\}_{i=1}^{L}$ is given), which is observed by M spatially distributed sensors. A simple uncoded analog amplify and forward scheme is employed to send the measurements of sensors $\mathbf{y}$ to the fusion center (FC). The channels between sensors and FC are assumed to be independent Rayleigh flat fading orthogonal channels which can be justified by adopting a multiple access technique such as TDMA/FDMA at transmitters of sensors [8]. Thus, all these interactions can be compactly modeled by the following behavioral equations

$$\begin{aligned} \mathbf{y} &= g(\boldsymbol{x}) + \mathbf{n}, && (1) \\ \mathbf{z}(\boldsymbol{\alpha}) &= \mathbf{H}(\boldsymbol{\alpha})\mathbf{y} + \mathbf{w}, && (2) \end{aligned}$$

where $g(\boldsymbol{x}) = (g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), ..., g_M(\boldsymbol{x}))^T$ with each component $g_i(\boldsymbol{x})$ a (linear or nonlinear) deterministic function for expression of i-sensor measuring quantity such as range and/or bearing. Accordingly, $\mathbf{y} = (y_1, y_2, ..., y_M)^T$ are the sensors' measurements and $\mathbf{n} \sim \mathcal{N}(0, \mathbf{C_n})$ with diagonal $\mathbf{C_n}$ is AWGN corrupting noise, which is uncorrelated with the source $\boldsymbol{x}$. These measurements are relayed to the fusion center (FC) and so $\mathbf{H}(\boldsymbol{\alpha}) \in \mathbb{R}^{M\times M}$ is called the relay matrix defined by

$$\mathbf{H}(\boldsymbol{\alpha}) = \mathsf{diag}[\sqrt{\alpha_i}\sqrt{h_i}]_{i=1,2...,M} \qquad (3)$$

which includes the channel gains $\sqrt{h_i}$ between i-th sensor node and FC and amplifier coefficients $\sqrt{\alpha_i}$ to control the transmit power of i-th sensor node. $\mathbf{w} \sim \mathcal{N}(0, \mathbf{C_w})$ with diagonal $\mathbf{C_w}$ is the FC measurement noise.It follows that the power consumed by i-th sensor node is

$$P_i = \alpha_i \mathbf{R_y}(i,i).$$

Hence, for $\boldsymbol{D_\alpha} := \mathsf{diag}[\alpha_i]_{i=1,2...,M}$, the sum power consumed by the entire SN is $\sum_{i=1}^{M} P_i = \mathsf{Trace}(\boldsymbol{D_\alpha}\mathbf{R_y})$ which is normally constrained by a fixed power budget $P_T > 0$

$$\mathsf{Trace}(\boldsymbol{D_\alpha}\mathbf{R_y}) \le P_T. \qquad (4)$$

The Bayesian optimal MMSE estimate based on FC output $\mathbf{z}$ is [9],

$$\begin{aligned} \hat{\boldsymbol{x}}_{\boldsymbol{mmse}} &\triangleq \mathbb{E}[\boldsymbol{x}|\mathbf{z}] \\ &= \sum_{i=1}^{L} \lambda_i(\mathbf{z}) \left[\boldsymbol{m}_{\boldsymbol{x}}^{(i)} + (\mathbf{C}_{\mathbf{z}\boldsymbol{x}}^{(i)})^T (\mathbf{C}_{\mathbf{z}}^{(i)})^{-1} (\mathbf{z} - \bar{\mathbf{z}}^{(i)})\right] \\ &= \sum_{i=1}^{L} \lambda_i(\mathbf{z}) (\boldsymbol{m}_{\boldsymbol{x}}^{(i)} + (\mathbf{C}_{\mathbf{y}\boldsymbol{x}}^{(i)})^T \mathbf{H}^T(\boldsymbol{\alpha})(\mathbf{H}(\boldsymbol{\alpha})\mathbf{C}_{\mathbf{y}}^{(i)}\mathbf{H}^T(\boldsymbol{\alpha}) \\ &+ \mathbf{C_w})^{-1}\left(\mathbf{z} - \mathbf{H}(\boldsymbol{\alpha})\bar{\mathbf{y}}^{(i)}\right)). && (5) \end{aligned}$$

where

$$\begin{aligned} \lambda_i(\mathbf{z}) &= \frac{\lambda_i \frac{1}{\sqrt{\det(2\pi \mathbf{C}_{\mathbf{z}}^{(i)})}} \exp\left(-\frac{1}{2}\|\mathbf{z} - \bar{\mathbf{z}}^{(i)}\|^2_{(\mathbf{C}_{\mathbf{z}}^{(i)})^{-1}}\right)}{\sum_{i=1}^{L} \lambda_i \frac{1}{\sqrt{\det(2\pi \mathbf{C}_{\mathbf{z}}^{(i)})}} \exp\left(-\frac{1}{2}\|\mathbf{z} - \bar{\mathbf{z}}^{(i)}\|^2_{(\mathbf{C}_{\mathbf{z}}^{(i)})^{-1}}\right)} \\ \mathbf{C}_{\mathbf{y}}^{(i)} &= \mathbb{E}[(\mathbf{y} - \bar{\mathbf{y}}^{(i)})(\mathbf{y} - \bar{\mathbf{y}}^{(i)})^T] && (6) \\ \mathbf{C}_{\mathbf{y}\boldsymbol{x}}^{(i)} &= \mathbb{E}[(\mathbf{y} - \bar{\mathbf{y}}^{(i)})(\boldsymbol{x} - \boldsymbol{m}_{\boldsymbol{x}}^{(i)})^T] && (7) \\ \bar{\mathbf{y}}^{(i)} &= g(\boldsymbol{m}_{\boldsymbol{x}}^{(i)}) && (8) \end{aligned}$$

The power optimization problem for MMSE estimate in considered scenario can be put as

$$\min_{\alpha_i \ge 0,\ i=1,2,...,M} \mathsf{Trace}(\mathbf{C_{e,mmse}}) \quad \text{subject to} \quad (4), \qquad (9)$$

In (9), the $\mathbf{C_{e,mmse}}$ has no closed form expression [6]. Consequently, tractable optimization for MMSE estimate of Gaussian Mixture information source cannot be formulated in α_i. Therefore, we move on to linear minimum mean square (LMMSE) estimate for which error covariance matrix exists in closed form and then establish its relation with the upper bound of MMSE estimator. By [10, Theorem 12.1] , the LMMSE estimate for $\boldsymbol{x}$ based on FC output $\mathbf{z}$ is

$$\begin{aligned} \hat{\boldsymbol{x}} &\triangleq \bar{\boldsymbol{x}} + \mathbf{C}_{\mathbf{z}\boldsymbol{x}}^T \mathbf{C}_{\mathbf{z}}^{-1}(\mathbf{z} - \bar{\mathbf{z}}) \\ &= \bar{\boldsymbol{x}} + \mathbf{C}_{\mathbf{y}\boldsymbol{x}}^T \mathbf{H}^T(\boldsymbol{\alpha})(\mathbf{H}(\boldsymbol{\alpha})\mathbf{C_y}\mathbf{H}^T(\boldsymbol{\alpha}) + \mathbf{C_w})^{-1} \\ &\times \ (\mathbf{z} - \mathbf{H}(\boldsymbol{\alpha})\bar{\mathbf{y}}). && (10) \end{aligned}$$

The covariance matrix of the estimator error

$$\mathbf{e} := \boldsymbol{x} - \hat{\boldsymbol{x}},$$

is as follows

$$\begin{aligned} \mathbf{C_e} &= \mathbf{C}_{\boldsymbol{x}} - \mathbf{C}_{\mathbf{z}\boldsymbol{x}}^T(\mathbf{H}(\boldsymbol{\alpha})\mathbf{C_y}\mathbf{H}^T(\boldsymbol{\alpha}) + \mathbf{C_w})^{-1}\mathbf{C}_{\mathbf{z}\boldsymbol{x}} \\ &= \mathbf{C}_{\boldsymbol{x}} - \mathbf{C}_{\mathbf{y}\boldsymbol{x}}^T\mathbf{H}^T(\boldsymbol{\alpha})(\mathbf{H}(\boldsymbol{\alpha})\mathbf{C_y}\mathbf{H}^T(\boldsymbol{\alpha}) + \mathbf{C_w})^{-1} \\ &\times \ \mathbf{H}(\boldsymbol{\alpha})\mathbf{C}_{\mathbf{y}\boldsymbol{x}}. \end{aligned}$$

We are now in a position to formulate the problem of minimization of MSE, subject to the power budget constraint (4) as

$$\min_{\alpha_i \ge 0,\ i=1,2,...,M} \mathsf{Trace}(\mathbf{C_e}) \quad \text{subject to} \quad (4), \qquad (11)$$

Using [3], from (11) the power optimization problem for Gaussian Mixture prior knowledge can be formulated as follows:

$$\min_{t,\mathbf{z},\alpha_i \ge 0,\ i=1,2,..,M} t \quad \text{subject to} (4), \quad (12)$$

$$\mathsf{Trace}(\mathbf{z}) \le t, \ \begin{pmatrix} \mathbf{z} & \mathbf{C}_{\mathbf{y}\boldsymbol{x}}^T \\ \mathbf{C}_{\mathbf{y}\boldsymbol{x}} & \mathbf{C_y} + \mathbf{C_y}\boldsymbol{D_\alpha}\mathbf{C_w}^{-1}\boldsymbol{D_h}\mathbf{C_y} \end{pmatrix} \succeq 0.$$

where $\boldsymbol{D_h} = \mathsf{diag}[h_i]_{i=1,2...,M}$.The above SDP is computationally tractable and can be realized if the sensor output covariance matrix $\mathbf{C_y}$ and its cross-covariance matrix $\mathbf{C}_{\boldsymbol{x}\mathbf{y}}$

with source $\boldsymbol{x}$ can be computed. The MMSE for Bayesian optimal estimator ϵ^2 is bounded as follows [6] :

$$\epsilon^2 \leq \mathsf{Trace}(\mathbf{C}_{\boldsymbol{x}} - \mathbf{C}^T_{\mathbf{y}\boldsymbol{x}}\mathbf{H}^T(\boldsymbol{\alpha})(\mathbf{H}(\boldsymbol{\alpha})\mathbf{C}_{\mathbf{y}}\mathbf{H}^T(\boldsymbol{\alpha}) + \mathbf{C}_{\mathbf{w}})^{-1} \\ \mathbf{H}(\boldsymbol{\alpha})\mathbf{C}_{\mathbf{y}\boldsymbol{x}})$$

Since the MMSE ϵ^2 is upper bounded by the MSE of LMMSE estimator, therefore we use this upper bound to perform the power optimization for MMSE estimator. This owes to upper bound optimization which is sub-optimal solution for optimization of an objective function in absence of its analytic form [11].

3. WIENER FILTERING FOR GAUSSIAN MIXTURE PRIOR KNOWLEDGE

This section illustrates the estimation of a static target characterized by Gaussian Mixture prior knowledge. The localization of such a target is achieved by executing the SDP developed in previous section and for this purpose, the moment approximations are required.In LSNs, model (1)-(2) is completely linear, i.e. the input-output system (1) of the sensor measurements is represented by

$$\mathbf{y} = \mathbf{G}\boldsymbol{x} + \mathbf{n}, \tag{13}$$

where $\mathbf{G} \in \mathbb{R}^{M\times N}$ is a matrix representing the different sensing conditions of sensors, which is known to the FC. Therefore the analytical forms of statistical moments are available

$$\begin{aligned} \mathbf{C}_{\boldsymbol{x}\mathbf{y}} &= \mathbb{E}[(\boldsymbol{x}-\bar{\boldsymbol{x}})(\mathbf{y}-\bar{\mathbf{y}})^T] = \mathbf{C}_{\boldsymbol{x}}\mathbf{G}^T \\ \mathbf{C}_{\mathbf{y}} &= \mathbb{E}[(\mathbf{y}-\bar{\mathbf{y}})(\mathbf{y}-\bar{\mathbf{y}})^T] = \mathbf{G}\mathbf{C}_{\boldsymbol{x}}\mathbf{G}^T + \mathbf{C}_{\mathbf{n}} \end{aligned} \tag{14}$$

where

$$\begin{aligned} \mathbf{C}_{\boldsymbol{x}} &= \sum_{i=1}^{L} \lambda_i \left(\boldsymbol{C}^{(i)}_{\boldsymbol{x}} + (\boldsymbol{m}^{(i)}_{\boldsymbol{x}})(\boldsymbol{m}^{(i)}_{\boldsymbol{x}})^T \right) \\ &- \sum_{i=1}^{L}\sum_{j=1}^{L} \lambda_i\lambda_j (\boldsymbol{m}^{(i)}_{\boldsymbol{x}})(\boldsymbol{m}^{(j)}_{\boldsymbol{x}})^T \end{aligned}$$

$$\bar{\mathbf{y}} = \mathbf{G}\bar{\mathbf{x}}, \qquad \bar{\boldsymbol{x}} = \sum_{i=1}^{L} \lambda_i \boldsymbol{m}^{(i)}_{\boldsymbol{x}}$$

Similarly, for Bayesian estimator, the sensors' output moments can be expressed as follows :

$$\begin{aligned} \mathbf{C}^{(i)}_{\boldsymbol{x}\mathbf{y}} &= \mathbb{E}[(\boldsymbol{x}-\boldsymbol{m}^{(i)}_{\boldsymbol{x}})(\mathbf{y}-\bar{\mathbf{y}}^{(i)})^T] = \mathbf{C}^{(i)}_{\boldsymbol{x}}\mathbf{G}^T \\ \mathbf{C}^{(i)}_{\mathbf{y}} &= \mathbb{E}[(\mathbf{y}-\bar{\mathbf{y}}^{(i)})(\mathbf{y}-\bar{\mathbf{y}}^{(i)})^T] = \mathbf{G}\mathbf{C}^{(i)}_{\boldsymbol{x}}\mathbf{G}^T + \mathbf{C}_{\mathbf{n}} \end{aligned} \tag{15}$$

where

$$\bar{\mathbf{y}}^{(i)} = \mathbf{G}\boldsymbol{m}^{(i)}_{\boldsymbol{x}} \tag{16}$$

It is tedious task to find the output moments of a general nonlinear map $g(\boldsymbol{x})$ which are required for implementation of (12). We approximate $\mathbf{C}_{\boldsymbol{x}\mathbf{y}}$ and $\mathbf{C}_{\mathbf{y}}$ by using unscented transformation (UT) and linear fractional transform (LFT). The UT is used for moderately nonlinear maps and LFT is used for higher order or fractional nonlinear maps. Unlike linearizing the deterministic map $\mathbf{y}$ as done in the conventional Extended Kalman Filter (EKF), the unscented transformation [12] approximates the output moments of a nonlinear transformation by finding a set of regression points of input random variable whose sample pdf approximates the true pdf. These sigma or regression points are transformed through exact nonlinear transformation and the sampled pdf of the outcomes are used to approximate the output moments with reasonable accuracy . In case of Bayesian estimator, the moment approximations are performed for L sets $\left\{\lambda_i, \boldsymbol{m}^{(i)}_{\boldsymbol{x}}, \boldsymbol{C}^{(i)}_{\boldsymbol{x}}\right\}_{i=1}^{L}$ independently to evaluate $\left\{\mathbf{C}^{(i)}_{\mathbf{yx}}, \mathbf{C}^{(i)}_{\mathbf{y}}, \bar{\mathbf{y}}^{(i)}\right\}_{i=1}^{L}$. Due to limited space, we are unable to provide the mathematical details of moment approximation through UKF-LFT. For detailed mathematical description, see [3].

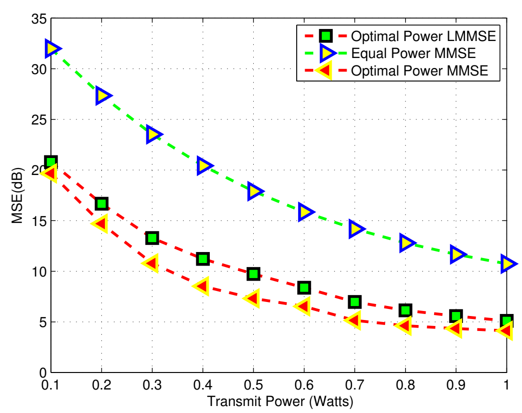

Fig. 1. MSE Performance of LSN for Gaussian Mixture Vector Prior Knowledge

4. SIMULATION RESULTS

The accuracy of the proposed estimation schemes is exhaustively validated via simulative results generated by $10,000$ Monte Carlo source and channel realizations for localization of static target. The number of sensors taking measurements in equal power and optimal power schemes is ten $(M = 10)$. The channels between these sensors and FC are considered to be independent identically distributed (i.i.d.) Rayleigh fading channels with unit power. The noise at FC is assumed to have variance of $\sqrt{0.5}$. We consider a vector valued target whose

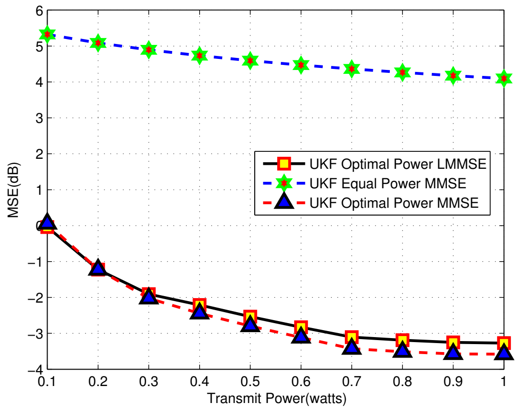

Fig. 2. MSE Performance of NSN for Gaussian Mixture Scalar Prior Knowledge

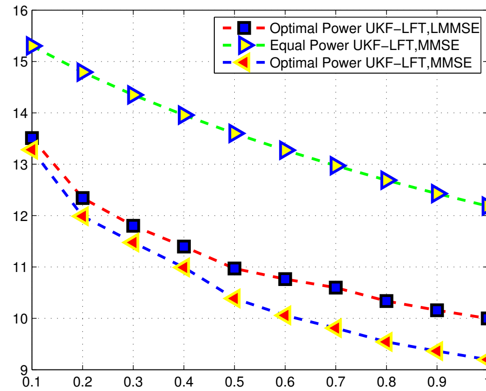

Fig. 3. MSE Performance of NSN for Gaussian Mixture Vector Prior Knowledge

position is characterized by following probability distribution

$$\begin{aligned}\boldsymbol{x} \quad &\sim \quad \frac{1}{3}\mathcal{N}(\boldsymbol{x};\mathbf{0}_{2\times1},\mathbf{I}_2)+\frac{1}{3}\mathcal{N}(\boldsymbol{x};5\times\mathbf{1}_{2\times1},\mathbf{I}_2)\\ &+\quad \frac{1}{3}\mathcal{N}(\boldsymbol{x};-5\times\mathbf{1}_{2\times1},\mathbf{I}_2) \qquad (17)\end{aligned}$$

In Fig. 1, the position of a target lying in $x-y$ coordinate system is observed by ten sensors. The measurements of sensors are linear, i.e., x and y coordinates of target position immersed in noise. The graph plots MSE of position estimate versus total transmit power. The total transmit power is varied from 0.1 to 1 watts. It is obvious from the Fig. 1 that optimal power MMSE based estimator clearly outperforms the equal power based MMSE estimator and optimal power LMMSE based estimator in terms of MSE. Our solution offers atleast 6 dB performance improvement over the equal power MMSE based estimator. This performance improvement comes from the optimal allocation of power among sensors in comparison to uniform distribution of power among sensor nodes.

Fig. 2 delineates the case of a scalar target characterized by $x\sim\frac{1}{2}\mathcal{N}(x;1,1)+\frac{1}{2}\mathcal{N}(x;3,1)$ lying on x-axis. The plot shows three different power allocation schemes for estimation of this target. The range of this scalar target is measured by ten sensors using

$$g_n(x)=\sqrt{(n-1)^2d^2+x^2} \qquad n=1,2,...,10 \qquad (18)$$

The estimation performance of UKF-LFT based optimal power MMSE estimator is comparable to UKF based optimal power LMMSE estimator but the performance improvement of UKF based optimal power estimator is atleast 5.5 dB better than its equal power based counterpart. This performance improvement certainly owes to the optimized power gains which allocates the power in such a fashion so as to minimize the overall distortion.

In Fig. 3, we consider the localization of a static object in $\mathbb{R}^2$ whose position is characterized by (17). A typical sensor's measurements consist of the following ranging and bearing

$$\mathbf{y}_i=\begin{pmatrix}\sqrt{(s_{i,x}-\boldsymbol{x}(1))^2+(s_{i,y}-\boldsymbol{x}(2))^2}\\ \frac{s_{i,y}-\boldsymbol{x}(2)}{s_{i,x}-\boldsymbol{x}(1)}\end{pmatrix}+\boldsymbol{n}_i \qquad (19)$$

The noise at FC has prior knowledge $\sigma_r^2=0.5$ and $\sigma_\phi^2=\pi/180$.It is evident from the graph that optimal power MMSE based estimator beats the performance of equal power MMSE based estimator. In low power region, the distortion gap between optimal power based MMSE and LMMSE estimators is negligible. Moreover for high power region, the performance improvement offered by optimal power based MMSE estimator over equal power based LMMSE estimator is less than 1 dB. However, optimal power based MMSE estimator offers 3 dB performance improvement over equal power MMSE estimator which is significant enhancement. Thus, our optimal power MMSE estimator is better than optimal power LMMSE estimator and equal power MMSE estimator in all the examples.

5. CONCLUSION

The problem of power allocation for localization of a static target characterized by GM information source via linear and nonlinear sensor networks has been dealt with in this paper. The sensors' observations are fused at FC and afterwards estimation is carried out to achieve maximum benefit from multisensor diversity. Both Bayesian optimal and suboptimal estimators are tested to validate the correctness of our proposed power allocation schemes. A suboptimal power allocation scheme has been suggested for Bayesian MMSE estimation of GM static target.

6. REFERENCES

[1] S. Kim, B. Ku, W. Hong, and H. Ko, "Performance comparison of target localization for active sonar systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 4, pp. 1371 –1380, 2008.

[2] L. Liu, X. Zhang, and H. Ma, "Optimal Node Selection for Target Localization in Wireless Camera Sensor Networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 7, pp. 3562 –3576, 2010.

[3] U. Rashid, H.D. Tuan, P. Apkarian, and H.H. Kha, "Globally Optimized Power Allocation in Multiple Sensor Fusion for Linear and Nonlinear Networks," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 903 –915, 2012.

[4] Daniel L. Alspach and Harold W. Sorenson, "Recursive Bayesian Estimation using Gaussian sums," *Automatica*, vol. 7, no. 2, pp. 465–479, 1971.

[5] J.T. Flam, J. Jalden, and S. Chatterjee, "Gaussian mixture modeling for source localization," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process (ICASSP)*, 2011, pp. 2604 –2607.

[6] J.T. Flam, S. Chatterjee, K. Kansanen, and T. Ekman, "On MMSE Estimation: A Linear Model Under Gaussian Mixture Statistics," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3840 –3845, 2012.

[7] I. Bahceci and A. Khandani, "Linear estimation of correlated data in wireless sensor networks with optimum power allocation and analog modulation," *IEEE Trans. Commun.*, vol. 56, no. 7, pp. 1146 –1156, 2008.

[8] S. Cui, J. Xiao, A.J. Goldsmith, Z. Luo, and H.V. Poor, "Estimation Diversity and Energy Efficiency in Distributed Sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 9, pp. 4683 –4695, 2007.

[9] A. Kundu, S. Chatterjee, A. S. Murthy, and T.V. Sreenivas, "GMM based Bayesian approach to speech enhancement in signal / transform domain," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process (ICASSP)*, 2008, pp. 4893 –4896.

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

[11] H.G. Hoang, H.D. Tuan, and T.Q. Nguyen, "Frequency Selective KYP Lemma,IIR Filter, and Filter Bank Design," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 956–965, 2009.

[12] S. Julier, J. Uhlmann, and H.F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Autom. Control*, vol. 45, no. 3, pp. 477 –482, 2000.