Influence of Oversampling on Channel Parameter Estimation for Joint Communication and Positioning

Kathrin Schmeink, Rebecca Adam, and Peter Adam Hoeher
Information and Coding Theory
Faculty of Engineering, University of Kiel
Kaiserstrasse 2, 24143 Kiel, Germany
Email: {kas, rbl, ph}@tf.uni-kiel.de

***Abstract*—Recently, a maximum-likelihood channel parameter estimator was proposed by the authors for joint communication and positioning. The parameters of the physical channel (e.g. amplitudes and propagation delays of the propagation paths) are estimated jointly with the channel coefficients of the equivalent discrete-time channel model by exploiting a priori information about pulse shaping and receive filtering. In this paper, the influence of oversampling on the proposed estimator is investigated. On the one hand, more samples can improve the performance of the estimator. On the other hand, the complexity of the estimator increases with oversampling and the noise gets more colored. The Cramer-Rao lower bounds (CRLBs) of the time of arrival for different channel models with different oversampling factors are compared. Taking into account the tradeoff between performance and complexity, it can be concluded that only an oversampling factor of two is reasonable. This design issue is different from satellite navigation receiver designs, where usually extensive oversampling is applied. Furthermore, the performance of the estimator is compared to the CRLB by means of Monte Carlo simulations.**

I. Introduction

Interest in joint communication and positioning is steadily increasing [1]. The combination of communication and positioning offers attractive features including synergetic effects like improved resource allocation and new applications such as location-based services or a precise localization of emergency calls. However, it is a challenging task to combine communication and positioning since their system requirements are quite different: For positioning the parameters of the physical channel like the time of arrival (TOA) or the angle of arrival (AOA) have to be estimated as accurate as possible [2], [3]. Typically, much training is spent for that purpose. In contrast, for communication purposes high data rates with little training overhead are desirable. It is sufficient to estimate the channel coefficients of the equivalent discrete-time channel model, that comprises pulse shaping and receive filtering in addition to the physical channel. The parameters of the physical channel are not relevant for estimating the equivalent discrete-time channel model.

This work has partly been funded by the German Research Foundation (DFG project number HO 226/11-1).

Recently, the authors proposed a maximum-likelihood channel parameter estimator in order to combine communication and positioning [4]. The proposed estimator determines *both* the parameters of the physical channel (including the TOA) and the channel coefficients of the equivalent discrete-time channel model by exploiting a priori information about pulse shaping and receive filtering. On the one hand, the usage of a priori information about pulse shaping and receive filtering has already been suggested in [5]–[7] for improved channel estimation in communication systems. But the information that is obtained concerning the physical channel is discarded in these publications. On the other hand, channel parameter estimation is well known for channel sounding [8]–[10]. But, to the authors best knowledge, the proposed parameter estimation methods are not applied for estimation of the equivalent discrete-time channel model. The proposed estimator combines both approaches mentioned above. Similar estimators have been investigated for example in the context of RAKE receivers in DS-CDMA [11] or for pure navigation purposes [12]. In [4] the performance of the proposed estimator was examined for symbol-rate sampling. This paper expands the concept to oversampling.

The remainder of this paper is organized as follows: The system and channel model is described in Sec. II. In Sec. III two channel parameter estimators are presented. The performance of these estimators is investigated in Sec. IV. Numerical results obtained by Monte Carlo simulations are shown and compared to the Cramer-Rao lower bound (CRLB). Finally, conclusions are drawn in Sec. V.

II. System and Channel Model

Throughout this paper, the discrete-time complex baseband notation is used. Let $a[n]$ denote the nth symbol of a coded and modulated transmission block of length N. Some symbols $a[n]$ are known at the receiver side ("training symbols"), whereas others are not known ("data symbols"). The oversampled sequence $x[k]$ is modeled by inserting $J-1$ zeros between

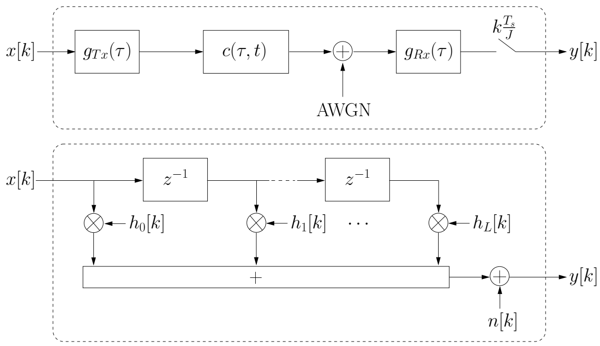


Fig. 1. Continuous-time channel and equivalent discrete-time channel model.

neighboring symbols, where J is the oversampling factor:

$$x[k] = \begin{cases} a[k/J], & \text{if } k \bmod J = 0 \\ 0, & \text{else.} \end{cases} \quad (1)$$

The oversampled sequence has a length of $K = J \cdot N$. The received sample $y[k]$ at time index k can be modeled as

$$y[k] = \sum_{l=0}^{L} h_l[k] \cdot x[k-l] + n[k], \qquad 0 \le k \le K+L-1, \quad (2)$$

where $h_l[k]$ is the lth channel coefficient of the equivalent discrete-time channel model with effective channel memory length L and $n[k]$ is a Gaussian noise sample.

The input/output behavior of the equivalent discrete-time channel model corresponds *exactly* to the input/output behavior of the continuous-time channel as shown in Fig. 1. The channel coefficients $h_l[k]$ are samples of the overall (continuous-time) channel impulse response, that includes pulse shaping $g_{\text{Tx}}(\tau)$ and receive filtering $g_{\text{Rx}}(\tau)$ in addition to the physical channel $c(\tau, t)$. The delay elements z^{-1} of the equivalent discrete-time channel model correspond to the sampling rate J/T_s, where T_s is the symbol duration. The noise process $n[k]$ is generally colored because the additive white Gaussian noise is filtered by $g_{\text{Rx}}(\tau)$ in the continuous-time channel. If a square-root Nyquist pulse is applied at the receiver the noise remains white for symbol-rate sampling ($J = 1$). In the following, it is assumed that the channel is quasi time-invariant over the data burst (block fading). Thus, the time index k can be omitted concerning the channel.

The physical channel can be modeled by a weighted sum of delayed Dirac impulses, i.e., each propagation path m is described by a complex amplitude f_m and a delay τ_m. If M denotes the number of propagation paths, the physical channel is fully characterized by the parameter vector $\boldsymbol{\theta} = [\text{Re}\{f_1\}, \text{Im}\{f_1\}, \tau_1, ..., \text{Re}\{f_M\}, \text{Im}\{f_M\}, \tau_M]$. If the pulse shaping and the receive filter are combined to a filter $g(\tau) = g_{\text{Tx}}(\tau) * g_{\text{Rx}}(\tau)$, the channel coefficients can be expressed as a function of the parameter vector as [4]

$$h_l(\boldsymbol{\theta}) = \sum_{\substack{\mu=1 \\ \nu=3(\mu-1)}}^{M} (\theta_{\nu+1} + j\theta_{\nu+2})\, g\left(l\frac{T_s}{J} - \theta_{\nu+3}\right). \quad (3)$$

If the parameter vector $\boldsymbol{\theta}$ and the shape of the filter $g(\tau)$ are known, the channel coefficients are also known.

III. Channel Parameter Estimation

In the following, two channel parameter estimators are presented. Both estimators determine at first the parameter estimate $\hat{\boldsymbol{\theta}}$ according to the maximum-likelihood principle. Based on this parameter estimate $\hat{\boldsymbol{\theta}}$ channel estimates $\hat{\boldsymbol{h}} = \boldsymbol{h}(\hat{\boldsymbol{\theta}})$ are calculated according to (3).

A training preamble of length $N_T < N$ ($K_T = JN_T$) is assumed. For the interval $L \le k \le K_T - 1$ the received samples of (2) can also be expressed in vector/matrix notation as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{h}(\boldsymbol{\theta}) + \boldsymbol{n}, \quad (4)$$

where $\boldsymbol{X}$ is the training matrix with Toeplitz structure, $\boldsymbol{y} = [y[L], y[L+1], ..., y[K_T - 1]]^T$ is the observation vector, $\boldsymbol{h}(\boldsymbol{\theta}) = [h_0(\boldsymbol{\theta}), h_1(\boldsymbol{\theta}), ..., h_L(\boldsymbol{\theta})]^T$ is the channel coefficient vector, and $\boldsymbol{n}$ is a zero mean Gaussian noise vector with covariance matrix $\boldsymbol{C}_n$. The covariance matrix describes the color of the noise and depends on the oversampling factor J and the shape of the receive filter $g_{\text{Rx}}(\tau)$:

$$[\boldsymbol{C}_n]_{i,j} = \sigma_n^2 \cdot g_{\text{Rx}}^2\left((i-j)\frac{T_s}{J}\right). \quad (5)$$

With a maximum-likelihood approach the first estimator is given by [13]

$$\hat{\boldsymbol{\theta}}_y = \arg\min_{\tilde{\boldsymbol{\theta}}} \underbrace{\left\{\left(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{h}(\tilde{\boldsymbol{\theta}})\right)^H \boldsymbol{C}_n^{-1}\left(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{h}(\tilde{\boldsymbol{\theta}})\right)\right\}}_{\Omega_y(\tilde{\boldsymbol{\theta}})}. \quad (6)$$

The parameter estimate $\hat{\boldsymbol{\theta}}_y$ is obtained directly from the received samples $\boldsymbol{y}$. The second estimator performs two steps. At first the channel coefficients are pre-estimated by a standard least-squares channel estimator:

$$\check{\boldsymbol{h}} = \left(\boldsymbol{X}^H\boldsymbol{C}_n^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^H\boldsymbol{C}_n^{-1}\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{\theta}) + \boldsymbol{\epsilon}. \quad (7)$$

Afterwards the parameter estimate $\hat{\boldsymbol{\theta}}_h$ is obtained by fitting the model function (3) to these pre-stage channel estimates:

$$\hat{\boldsymbol{\theta}}_h = \arg\min_{\tilde{\boldsymbol{\theta}}} \underbrace{\left\{\left(\check{\boldsymbol{h}} - \boldsymbol{h}(\tilde{\boldsymbol{\theta}})\right)^H \boldsymbol{C}_\epsilon^{-1}\left(\check{\boldsymbol{h}} - \boldsymbol{h}(\tilde{\boldsymbol{\theta}})\right)\right\}}_{\Omega_h(\tilde{\boldsymbol{\theta}})}, \quad (8)$$

where $\boldsymbol{C}_\epsilon = (\boldsymbol{X}^H\boldsymbol{C}_n^{-1}\boldsymbol{X})^{-1}$ is the covariance matrix of the estimation error $\boldsymbol{\epsilon}$.

Both metrics $\Omega_y(\tilde{\boldsymbol{\theta}})$ and $\Omega_h(\tilde{\boldsymbol{\theta}})$ are nonlinear and can not be solved in closed form. Exhaustive search is prohibitive because the search space is continuous and of high dimension ($3M$). Hence, an optimization method has to be applied. Most optimization methods work iteratively: In each iteration the metric has to be recalculated given the recent hypothesis of the parameter vector $\tilde{\boldsymbol{\theta}}$. The vectors and matrices involved in the first metric $\Omega_y(\tilde{\boldsymbol{\theta}})$ are larger than the vectors and matrices involved in the second metric $\Omega_h(\tilde{\boldsymbol{\theta}})$. Hence, the second metric is advantageous from a complexity point of view.

For multipath channels the metrics $\Omega_y(\tilde{\boldsymbol{\theta}})$ and $\Omega_h(\tilde{\boldsymbol{\theta}})$ have many local minima. Either a global optimizer or a local optimizer with a good initial guess that is close to the global minimum are required. It is suggested to divide the problem into acquisition and tracking: If no a priori information about the physical channel is available a global optimizer is applied (acquisition). Afterwards a local optimizer can be employed with the parameter estimate of the proceeding burst as initial guess (tracking). This procedure is suitable for slowly varying channels. Only acquisition is considered in the following. In [4] particle swarm optimization (PSO) has been proposed as global optimizer. PSO is a heuristic optimization method that does not need any gradient information about the metric (which is also called *fitness function* in this context) [14]–[16]. Only function evaluations are performed. So-called particles explore the search space moving randomly through it and being attracted by good fitness values in their past and of their neighbors. Many different variants of PSO are available in the literature. In this paper, PSO as described in [15] is applied.

IV. Numerical Results

In the following, the performance of both estimators is investigated by means of Monte Carlo simulations. Positioning based on the TOA is assumed. Hence, only the delay of the line-of-sight (LOS) path $\hat{\theta}_3 = \hat{\tau}_1$ is of interest concerning the parameter estimates $\hat{\boldsymbol{\theta}}$. The mean squared error (MSE) of the delay estimate $\hat{\theta}_3$ is compared to the CRLB which is given by

$$\mathrm{CRLB}(\theta_3) = \left[\boldsymbol{I}^{-1}(\boldsymbol{\theta})\right]_{3,3}, \tag{9}$$

where $\boldsymbol{I}(\boldsymbol{\theta})$ is the Fisher information matrix

$$\boldsymbol{I}(\boldsymbol{\theta}) = 2\mathrm{Re}\left\{\boldsymbol{J}(\boldsymbol{\theta})^H \boldsymbol{X}^H \boldsymbol{C}_n^{-1} \boldsymbol{X} \boldsymbol{J}(\boldsymbol{\theta})\right\} \tag{10}$$

and $\boldsymbol{J}(\boldsymbol{\theta})$ is the Jacobian matrix of the channel coefficients (3) with entries

$$[\boldsymbol{J}(\boldsymbol{\theta})]_{l,m} = \frac{\delta h_l(\boldsymbol{\theta})}{\delta \theta_m}. \tag{11}$$

Furthermore, the MSE of the channel estimates $\hat{\boldsymbol{h}}$ is compared to the MSE of the pre-stage channel estimates $\check{\boldsymbol{h}}$ obtained by standard least-squares channel estimation (LSCE).

The simulation setup is as follows: A training preamble of length $N_T = 125$ is assumed, that covers 10% of the total transmission block. A pseudo-random sequence of BPSK symbols is used as training. The convolution of pulse shaping and receive filtering has a Gaussian shape $g(\tau) \sim \exp\left(-(\tau/T_s)^2\right)$. As mentioned before PSO as proposed in [15] is applied with 30 particles for minimizing the metrics $\Omega_y(\tilde{\boldsymbol{\theta}})$ (YPSO) and $\Omega_h(\tilde{\boldsymbol{\theta}})$ (HPSO). Three different channel models with memory length $L = J \cdot 10$ are assumed:

A) a single path channel ($M = 1$),
B) a two path channel with minimum phase and large excess delay ($M = 2$, $|f_1|^2/|f_2|^2 = 7$ dB, $\tau_2 - \tau_1 = 0.31T_s$),
C) a two path channel with maximum phase and small excess delay ($M = 2$, $|f_1|^2/|f_2|^2 = -7$ dB, $\tau_2 - \tau_1 = 0.12T_s$).

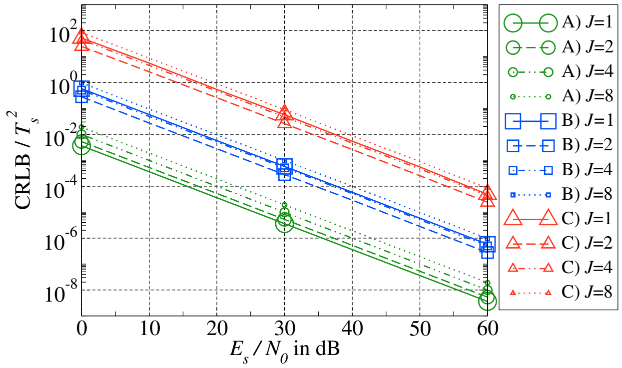


Fig. 2. Mean CRLB of the delay estimate $\hat{\theta}_3$ for different oversampling factors ($J = 1, 2, 4, 8$) in A) a single path channel, B) a two path channel with minimum phase and large excess delay, C) a two path channel with maximum phase and small excess delay.

The phases for each path are generated randomly with uniform distribution between 0 and 2π. Furthermore, a LOS propagation delay is generated randomly with uniform distribution between 0 and $5T_s$. For each run of the Monte Carlo simulation a different parameter vector $\boldsymbol{\theta}$ is generated. Hence, a mean CRLB is introduced,

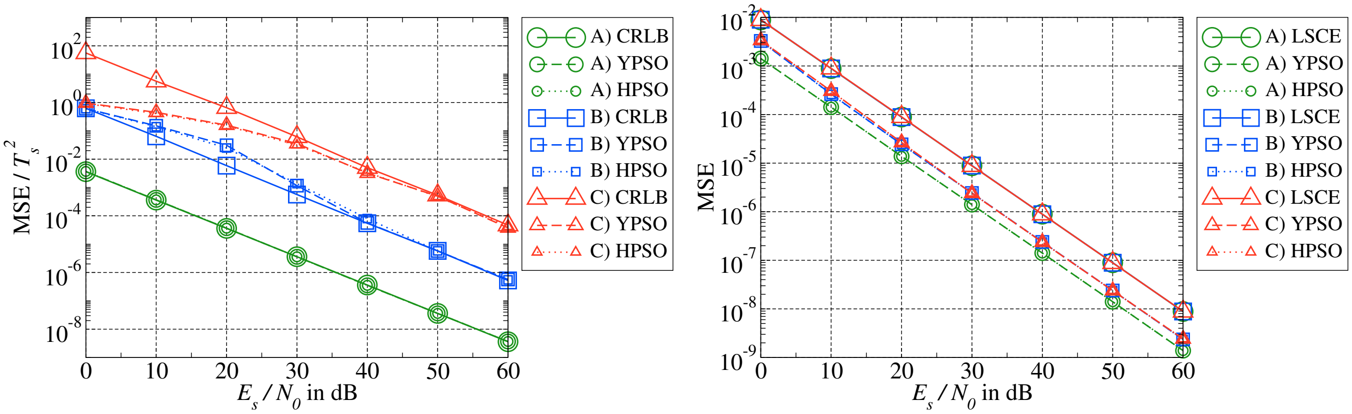
$$\mathrm{CRLB}(\overline{\theta_3}) = E\left\{\left[\boldsymbol{I}^{-1}(\boldsymbol{\theta})\right]_{3,3}\right\}, \tag{12}$$

where the expectation is taken with respect to $\boldsymbol{\theta}$. The CRLB and the MSE of the delay estimate $\hat{\theta}_3$ are normalized with respect to the symbol duration T_s.

In Fig. 2 the mean CRLB for the delay estimate $\hat{\theta}_3$ is shown for all channel models with different oversampling factors ($J = 1, 2, 4, 8$) over the signal-to-noise ratio (SNR, E_s/N_0). The single path channel A) is the best case and, hence, a lower bound for all other channel models. The smaller the excess delay in the two path channels, the more difficult it is to separate the propagation paths and the worse is the CRLB. For oversampling factors $J \geq 2$ the CRLBs show the same behavior for all channels: With increasing oversampling factor the CRLBs degrade by an equal amount since the noise gets more colored. The difference between symbol rate sampling ($J = 1$) and oversampling ($J \geq 2$) depends on the channel model. In contrast to the single path channel, oversampling can provide a slight gain for the two path channels. The achievable gain decreases with increasing excess delay. This gain is probably due to finite filter lengths. For both two path channels the CRLB for $J = 1$ and $J = 4$ are almost the same. Taking into account the tradeoff between performance and complexity, it can be concluded that only an oversampling factor of $J = 2$ is reasonable for multipath channels.

In Fig. 3 the performance of the proposed estimators is compared for all channel models and $J = 1$. In Fig. 3(i) the MSE[1] of the delay estimate $\hat{\theta}_3$ is shown in comparison to the mean CRLB. Both estimators (YPSO and HPSO)

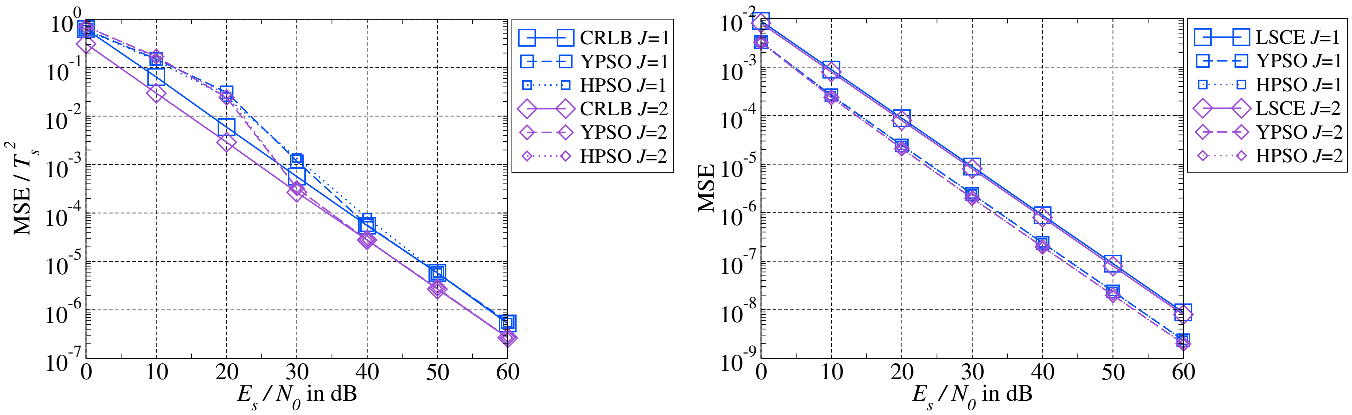
[1]PSO does not assure global convergence as already discussed in [4]. The outliers due to premature convergence were not considered for the calculation of the MSE.



(i) MSE of the delay estimate $\hat{\theta}_3$ in comparison to the mean CRLB.

(ii) MSE of the channel estimates in comparison to LSCE.

Fig. 3. Performance of the proposed estimators (YPSO and HPSO) for $J = 1$ in A) a single path channel, B) a two path channel with minimum phase and large excess delay, C) a two path channel with maximum phase and small excess delay.



(i) MSE of the delay estimate $\hat{\theta}_3$ in comparison to the mean CRLB.

(ii) MSE of the channel estimates in comparison to LSCE.

Fig. 4. Performance of the proposed estimators (YPSO and HPSO) for $J = 1$ and $J = 2$ in a two path channel with minimum phase and large excess delay.

provide approximately the same performance. In case of the single path channel, the estimators correspond to a minimum variance unbiased (MVU) estimator since their MSE coincides with the CRLB over the whole SNR range. For the two path channels, the estimators are biased at low SNR since maximum-likelihood estimators are only asymptotically (for a high number of observations or at high SNR) unbiased [13]. Hence, the CRLB is not valid at low SNR: The MSE can be below the CRLB as in the case of channel C). For decreasing SNR the MSE for both two path channels converges to 1, which corresponds to the symbol duration T_s, i.e., even in the worst case a resolution of a symbol duration can be achieved in a one shot measurement. For high SNR the MSE coincides with the CRLB for both two path channels. Thus, the estimators are asymptotically optimal and efficient. In Fig. 3(ii) the MSE of the channel estimates is plotted in comparison to LSCE. Since the parameters of the physical channel are irrelevant for the channel estimates, the MSE for LSCE is the same for all three channel models. The channel estimates of the proposed estimators are more accurate in comparison to LSCE. For the two path channels a gain of approximately 6 dB is achieved. In case of the single path channel the gain is even larger (approximately 8.5 dB). The MSE of the channel estimates is reduced in comparison to LSCE because the number of parameters to estimate, $3M$, is less than the number of channel coefficients $L + 1$. This also explains why the gain is larger for the single path channel.

In Fig. 4 the performance of the proposed estimators is compared for $J = 1$ and $J = 2$ given channel B). Again, Fig. 4(i) shows the MSE of the delay estimate $\hat{\theta}_3$ in comparison to the mean CRLB. At low SNR the same performance is achieved for both oversampling factors, but for $J = 2$ the MSE converges faster to the CRLB. The estimator is unbiased earlier since more samples are available. Concerning channel estimation the influence of oversampling is negligible as shown in Fig. 4(ii).

The number of iterations for the PSO was limited to 4000. On average much less iterations are performed. In the beginning, the particles move fast through the search space, but their velocity decreases from iteration to iteration. The algorithm stops if the velocity of all particles is below a given threshold. The number of required iterations depends on the channel model and is independent of the oversampling factor J: For the single path channel approximately 200 iterations are performed. For channel B) and C) roughly 1000 and 1500 iterations are necessary, respectively. The metrics $\Omega_y(\tilde{\boldsymbol{\theta}})$ and $\Omega_h(\tilde{\boldsymbol{\theta}})$ are evaluated for each particle in each iteration. As mentioned before, the vectors and matrices involved in the calculation of $\Omega_h(\tilde{\boldsymbol{\theta}})$ are much smaller than those involved in the calculation of $\Omega_y(\tilde{\boldsymbol{\theta}})$. Since both estimators perform equally well, it is suggested to apply the estimator based on $\Omega_h(\tilde{\boldsymbol{\theta}})$ (HPSO).

V. Conclusions

In this paper, the influence of oversampling on the channel parameter estimators proposed by the authors in [4] is investigated. The color of the noise has to be taken into account. With increasing oversampling factor the noise gets more colored and the complexity of the estimator increases. Both estimators (YPSO and HPSO) perform equally well for all channel models and all oversampling factors. It is suggested to apply the estimator based on the metric $\Omega_h(\tilde{\boldsymbol{\theta}})$ (HPSO) due to complexity reasons. Comparing the CRLBs of the delay estimate for different oversampling factors it can be concluded that only an oversampling factor of $J = 2$ is reasonable for multipath channels. The achievable gain depends on the channel model and decreases with increasing excess delay. Concerning the channel estimates the influence of oversampling is negligible.

References

[1] R. Raulefs, S. Plass, and C. Mensing, "The where project - combining wireless communications and navigation," in *Proc. 20th Wireless World Research Forum (WWRF)*, Apr. 2008, Ottawa, Canada.

[2] K. Pahlavan and A. H. Levesque, *Wireless information networks*. Hoboken, New Jersey, USA: John Wiley & Sons, 2005, ch. 13: RF location sensing.

[3] K. Cheung, H. So, W.-K. Ma, and Y. Chan, "A constrained least squares approach to mobile positioning: Algorithms and optimality," *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp. 1–23, 2006, Article ID 20858.

[4] K. Schmeink, R. Block, and P. A. Hoeher, "Joint channel and parameter estimation for combined communication and navigation using particle swarm optimization," in *Proc. Workshop on Positioning, Navigation and Communication (WPNC)*, Mar. 2010, Dresden, Germany.

[5] A. Khayrallah, R. Ramesh, G. Bottomley, and D. Koilpillai, "Improved channel estimation with side information," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, vol. 2, May 1997, pp. 1049–1053, Phoenix, Arizona, USA.

[6] J.-W. Liang, B. Ng, J.-T. Chen, and A. Paulraj, "GMSK linearization and structured channel estimate for GSM signals," in *Proc. IEEE Military Communications Conference (MILCOM)*, vol. 2, Nov. 1997, pp. 817–821, Monterey, California, USA.

[7] H.-N. Lee and G. J. Pottie, "Fast adaptive equalization/diversity combining for time-varying dispersive channels," *IEEE Transactions on Communications*, vol. 46, no. 9, pp. 1146–1162, Sep. 1998.

[8] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 477–489, Apr. 1988.

[9] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dalhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 434–450, Mar. 1999.

[10] A. Richter, M. Landmann, and R. S. Thomae, "Maximum likelihood channel parameter estimation from multidimensional channel sounding measurements," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, Apr. 2003, pp. 1056–1060, Jeju, Korea.

[11] G. Fock, P. Schulz-Rittich, A. Schenke, and H. Meyr, "Low complexity high resolution subspace-based delay estimation for DS-CDMA," in *Proc. IEEE International Conference on Communications (ICC)*, Apr. 2002, New York, USA.

[12] F. Antreich, J. Nossek, and W. Utschick, "Maximum likelihood delay estimation in a navigation receiver for aeronautical applications," *Aerospace Science and Technology*, vol. 12, no. 3, pp. 256 – 267, 2008.

[13] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Upper Saddle River, New Jersey, USA: Prentice Hall, 1993.

[14] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE International Conference on Neural Networks*, vol. 4, Nov./Dec. 1995, pp. 1942–1948, Perth, Australia.

[15] D. Bratton and J. Kennedy, "Defining a standard for particle swarm optimization," in *Proc. IEEE Swarm Intelligence Symposium (SIS)*, Apr. 2007, pp. 120–127, Honolulu, Hawaii, USA.

[16] C. Blum and X. Li, "Swarm intelligence in optimization," in *Swarm Intelligence: Introduction and Applications*, ser. Natural Computing Series, C. Blum and D. Merkle, Eds. Springer, 2008, pp. 43–85.