Frequency Synchronisation in OFDM
– a Bayesian Analysis

Erik Björnemo
Signals and Systems Group, Uppsala University
Box 528, SE-751 20, Uppsala, Sweden, eb@signal.uu.se

***Abstract*— The Bayesian approach offers three key advantages in the present problem of estimating the frequency offset when synchronising a mobile to a base station in an orthogonal frequency division multiplexing system: a systematic way to include important information via prior probabilities; ability to update our probability distributions consistently using Bayes' theorem; access to reliability measures for the actual estimate at hand. We show that the Bayesian approach in these ways enables fast frequency acquisition without pilots at low SNR, and also provides a good indicator of when the mobile is sufficiently synchronised to transmit for the first time. Our analysis and simulations indicate that oversampling and nulling of subcarriers can speed up the acquisition further. When pilot and (incomplete) channel information is available, it can be successfully included via priors to significantly improve the synchronisation performance.**

I. Introduction

Accurate frequency and time synchronisation of orthogonal frequency division multiplexing (OFDM) systems is required in order to achieve good performance. The very property that these systems rely on – orthogonality of the subcarriers – will be lost if synchronisation is inaccurate. In the uplink of multiple access systems, where several transmitting users must all synchronise to the base station, the need for efficient synchronisation algorithms is especially evident. We focus here on the *frequency* synchronisation of a mobile terminal to the base station in a future OFDM system.

There are several publications on OFDM synchronisation, covering many different scenarios. Most of them rely on the insertion of pilot symbols known to the receiver (e.g. [1]), while some are based only on the redundancy present in the cyclic prefix of each symbol (e.g. [2]), and a few are so-called fully blind and rely solely on the OFDM symbol structure (e.g. [3]). We see from these, and other, examples that the attainable performance is crucially dependent on efficient use of the available information.

In this paper we present a Bayesian analysis of the frequency synchronisation problem based on a model introduced in section II (this model and the analysis relies heavily on the work by Bretthorst [4]). The Bayesian approach is adopted because it focuses on efficient information processing and provides tools for inclusion of cogent prior information, consistent processing of new data, and reliability measures for estimates obtained with the actual data set. We perform a general analysis in section III, which is followed by a study of two special cases in section IV, simulation studies of these cases in section V, and finally the conclusions in section VI.

II. Received signal model

Consider a mobile OFDM terminal whose local oscillator is to be matched to the oscillator in the base station in order to ensure functionality. The adjustment of the oscillator will be based on an estimate of its frequency offset $\Delta\omega$ relative to the base station. This offset is the parameter of interest in this paper. In our model, the mobile is receiving an OFDM symbol consisting of quadrature amplitude modulated (QAM) sinusoids. These complex-valued sinusoids are attenuated and phase-shifted by a dispersive channel, and when received they are also shifted in frequency by the mismatched receiver oscillator. We assume that symbol time synchronisation is sufficiently accurate to avoid inter-symbol interference and that the cyclic prefix is properly removed. The channel is modelled as frequency selective but constant over one symbol period (i.e. as slowly fading). Mathematically, the model of the received data d is

$$d = WB + e. \tag{1}$$

Here, W are the m frequency-shifted sinusoids,

$$W_{ik} = \exp(j[\omega_k + \Delta\omega]t_i) \qquad \begin{cases} k = 1, \dots, m \\ i = 1, \dots, N \end{cases}, \tag{2}$$

corresponding to subcarriers ω_k, sampled uniformly at times t_i. The sinusoids have complex-valued amplitudes B_k, and $B \equiv [B_1 \dots B_m]^T$ therefore models the transmitted QAM symbols, their transmit energy and the channel jointly. The additive thermal measurement noise is represented by e.

III. Bayesian analysis

The product and sum rules of probability theory have been shown to constitute the only internally consistent quantitative rules for reasoning from incomplete information [5]. Probability distributions are from this point of view valid descriptions of our state of knowledge regarding any entities of interest, in this case the frequency offset $\Delta\omega$. When we in this way regard probabilities as carriers of information, it becomes evident that both the data d and all other available cogent prior information – henceforth denoted $\mathcal{I}$ – will determine our ability to infer the value of $\Delta\omega$. The state of knowledge regarding $\Delta\omega$ is summarised in the posterior probability distribution $p(\Delta\omega|d, \mathcal{I})$, the distribution which in turn will determine the best possible estimate based on the available information.

In the following three subsections we study our present problem with respect to the inclusion of prior information, the updating of probability distributions when receiving new data, and the estimation of $\Delta\omega$ and the reliability of this estimate.

A. Including cogent prior information

By the use of Bayes' theorem (i.e. the product rule) we can express the posterior for $\Delta\omega$ as

$$p(\Delta\omega|d,\mathcal{I}) \propto p(d|\Delta\omega,\mathcal{I})p(\Delta\omega|\mathcal{I}). \quad (3)$$

Here, our signal model (1) comes into play when we assign the direct probability, or likelihood, $p(d|\Delta\omega,\mathcal{I})$. In this model, let us first consider the noise term e. If we know that the noise has zero mean – which is reasonable considering the nature of thermally induced noise – and power σ^2, we can apply the maximum entropy principle [5] and thereby assign an uncorrelated Gaussian distribution carrying this information; $p(e|\sigma,\mathcal{I}) = \mathcal{N}(0,\sigma^2 I)$. As a consequence of the additive noise model we then have $p(d|\Delta\omega,B,\sigma,\mathcal{I}) = \mathcal{N}(WB,\sigma^2 I)$. This is essentially our model (1) dressed up as a probability distribution.

Next, we expand the likelihood for $\Delta\omega$ using the sum rule:

$$p(d|\Delta\omega,\mathcal{I}) = \iint p(d|\Delta\omega,B,\sigma,\mathcal{I})p(B,\sigma|\mathcal{I})\mathrm{d}B\mathrm{d}\sigma. \quad (4)$$

In (3) and (4) we now see how our prior information about $\Delta\omega$, B and σ enters the calculation in the shape of prior probability distributions. The prior information can vary considerably between different situations. Consider for example a mobile that is switched on and tries connect to the base station for the first time, and compare with a mobile in full operation with access to pilot symbols and channel predictions. The former will generally have considerably less precise information regarding $\Delta\omega$, and other related parameters, than the latter.

Let us take another, more specific, example in which it is known at the receiver that subcarrier k is not used in the transmission. This information is represented by $p(B_k|\mathcal{I}) = \delta(B_k)$, where $\delta(\cdot)$ is Dirac's delta function. Hereby the amplitude of sinusoid k is fixed to zero and this subcarrier is effectively removed from the model. A receiver to which this information is unavailable will however try to fit sinusoid k in the model to the noise in the data.

In section V we will see examples of the impact of useful prior information: channel predictions and pilot symbols.

B. Updating the state of knowledge using new data

Every new OFDM symbol that is received brings with it new, potentially important, information about the frequency offset. In order to consistently update the information about $\Delta\omega$ when using new data d_2 we apply Bayes' theorem and the sum rule

$$p(\Delta\omega|d_1,d_2,\mathcal{I}) \propto p(d_2|\Delta\omega,d_1,\mathcal{I})p(\Delta\omega|d_1,\mathcal{I}), \quad (5)$$

$$p(d_2|\Delta\omega,d_1,\mathcal{I}) = \iint p(d_2|\Delta\omega,B,\sigma,\mathcal{I})p(B,\sigma|d_1,\mathcal{I})\mathrm{d}B\mathrm{d}\sigma (6)$$

We would like to stress the importance of $p(\Delta\omega|d_1,\mathcal{I})$ as a *complete* description of the state of knowledge regarding $\Delta\omega$ before taking d_2 into account, and that no information is lost when updating this to $p(\Delta\omega|d_1,d_2,\mathcal{I})$ according to (6) and (6). This Bayesian procedure is quite different from the conventional approach in which we obtain only a point estimate. An *ad hoc* way to use several point estimates, calculated for d_1, d_2 and so on, is to average them. By use of the latter approach we may loose crucial information, information that is contained in $p(\Delta\omega|d_1,d_2,\mathcal{I})$. Numeric examples in section V show that repeated updating of the posterior according to (6) eventually gives good estimates even at low signal-to-noise ratios (SNR), despite the fact that every individual $p(\Delta\omega|d_n,\mathcal{I})$ is very broad and multimodal. Averaging of the corresponding point estimates performs, in these examples, considerably worse. (In this paper we do not include any model of time variations in $\Delta\omega$, B and σ between the symbols, but in case such models are used this is the place to introduce them.)

C. Optimal estimates, the posterior and reliability

According to [6], the disturbance (in terms of equivalent noise variance) introduced in an OFDM system by frequency synchronisation errors is roughly proportional to $\Delta\omega^2$. When calculating our estimates it therefore seems reasonable to minimise the mean squared error, i.e. use the mean value of the posterior distribution for $\Delta\omega$ as our estimate.

Although the mean value hereby is the single most important feature of the posterior probability distribution – as it is our point estimate of $\Delta\omega$ which is to be used for adjusting the local oscillator – it is certainly not the only feature of use to us. One important thing that a Bayesian approach provides through the posterior distribution is a measure of the accuracy for the particular estimate at hand. The variance of the posterior tells us something about the reliability of our estimate, not asymptotically in an imagined ensemble of experiments, but for the actual data set we have.

In our present problem we therefore propose to use the variance of the posterior as an indicator for when it is okay for a mobile to transmit. Consider, again, a mobile that is about to connect for the first time. It can, by use of the posterior variance, determine when it is sufficiently well synchronised in frequency in order not to introduce significant interference in the uplink. In addition to the theoretical motivation for using the posterior variance, the simulation studies in section V indicate that there is good correspondence with the actual squared error.

IV. The Ignorant and the Well-Informed

For the model given in (1) and (2) the above analysis is completely general and can in principle include any type of prior information. But there is such a wide range of possible scenarios that we cannot analyse all scenarios in detail here. We choose instead to concentrate on what we believe are two interesting special cases which roughly correspond to acquisition and fine tuning/tracking. These special cases are then simulated as shown in section V.

A. The ignorant receiver – frequency acquisition

While there always is *some* quantitative information available regarding the values of B, σ and $\Delta\omega$, it might be very vague during frequency acquisition. We can acknowledge this large uncertainty by assigning uninformative prior probability distributions. Complete ignorance is in this case represented by $p(B|\mathcal{I}) = constant$, $p(\sigma|\mathcal{I}) \propto 1/\sigma$ and $p(\Delta\omega|\mathcal{I}) = constant$ [5]. From (3) and (4) we obtain

$$p(\Delta\omega|d,\mathcal{I}) \propto \iint \frac{1}{\sigma} p(d|\Delta\omega, B, \sigma, \mathcal{I}) \mathrm{d}B \mathrm{d}\sigma, \tag{7}$$

where

$$p(d|\Delta\omega,B,\sigma,\mathcal{I}) = (\pi\sigma^2)^{-N} e^{-\frac{1}{\sigma^2}\left(d^H d - \frac{1}{N} d^H W W^H d + N X^H X\right)}, \tag{8}$$

$$X \equiv B - \frac{1}{N} W^H d. \tag{9}$$

Here $(\cdot)^H$ denotes the Hermitian transpose. Carrying out the marginalisation in (7) gives

$$p(\Delta\omega|d,\mathcal{I}) \propto \left(d^H d - \frac{1}{N} d^H W W^H d \right)^{m-N}. \tag{10}$$

The term $(1/N) d^H W W^H d$ in (10) is the energy projected onto the m orthogonal subcarriers in W (remember that W is a function of $\Delta\omega$). If the number of samples equals the number of subcarriers, i.e. if $N = m$, all energy in d can be projected onto W regardless of $\Delta\omega$. In this case the posterior distribution will be constant and gives no indication of the value of $\Delta\omega$. There is nothing in the prior information that can help to distinguish between the noise and the sinusoids in the data, and a receiver this ignorant is in this case helpless.

Fortunately, every practical OFDM system uses a couple of subcarriers as guard bands, and this information must be available even to the most ignorant receiver of interest. Therefore, as $m < N$, the posterior distribution will not be flat. It can however be multimodal, which we intuitively can realise as follows. Assume that the true frequency offset is $\Delta\omega_0$. If there are, say, 512 subcarriers, we are likely to get good matches between the model and the data when $\Delta\omega = \Delta\omega_0 + w_k$, for quite a large range of k's since most of the subcarriers in the model then still correspond to a sinusoid in the data. A very good SNR is required if a mismatch of only one subcarrier out of, say, 480 is going to make a decisive difference and suppress all but one mode in the posterior.

We can also see from (10) that the greater the difference $N - m$ is, the sharper the peaks in $p(\Delta\omega|d,\mathcal{I})$ can be. This suggests that nulling of many subcarriers, or oversampling, will improve the estimation accuracy. These strategies have indeed been studied and shown to be quite successful [7], [8]. Our simulations show the same result, but when it comes to nulling of subcarriers we stress that this requires that the receiver has information about the location of the nulled subcarriers. Otherwise, nulling of subcarriers will only make things worse as the receiver then essentially will try to fit the model subcarriers to the noise (see section V).

Finally, we observe that the highest peak in $p(\Delta\omega|d,\mathcal{I})$ corresponds to the maximum likelihood estimate of $\Delta\omega$ (in this case identical to the nonlinear least-squares solution). We use this as a reference in the simulation study in section V.

B. The well-informed receiver – fine tuning and tracking

Let us move on to a more well-informed synchroniser. We now assume that the mobile terminal is sufficiently synchronised to the rest of the system to have access to channel predictions and pilots (inserted for channel prediction purposes). The channel prediction is accompanied by a measure of its accuracy, in this case the covariance matrix computed in the Kalman channel predictor assumed in use [10]. Kalman predictors inherently use Gaussian probability distributions, and we therefore have

$$p(B|bR\mathcal{I}) = \pi^{-m} |R|^{-1} e^{-(B-b)^H R^{-1} (B-b)}, \tag{11}$$

where b is the predicted received signal – pilot and channel – and R is the covariance matrix describing the uncertainty in this prediction. We conservatively consider uncorrelated prediction errors, $R = \alpha^2 I$. Furthermore we assume that the noise power σ^2 is known, at least with an accuracy sufficient to ignore any uncertainty. Marginalising B in (4) then yields

$$p(\Delta\omega|\sigma b\alpha d\mathcal{I}) \propto p(\Delta\omega|\mathcal{I}) \times e^{\frac{\alpha^2}{N\alpha^2+\sigma^2}\left(\frac{1}{\sigma^2} d^H W W^H d + \frac{1}{\alpha^2}(b^H W^H d + d^H W b)\right)}, \tag{12}$$

where the exact shape of $p(\Delta\omega|\mathcal{I})$ depends on the previous synchronisation. Equation (12) shows how the extra channel information results in additional terms $b^H W^H d$ that correlate the predictions b with the projection $W^H d$ of the data onto the subcarriers. One of the main effects is that a multimodal distribution is less likely as these correlations help to distinguish between individual subcarriers. The ignorant receiver can not do this.

While nulling of some pilot subcarriers in an OFDM pilot symbol is recommended for the ignorant receiver by (10), the case is less clear for the well-informed receiver in (12). The term $d^H W W^H d$ will be more discriminating when including fewer subcarriers, but there will on the other hand be less pilot information to exploit through the correlations $b^H W^H d$. This suggests that nulling could be more efficient in cases of low SNR and poor channel predictions, than in cases of high SNR and good channel predictions. Simulations in section V support this suggestion.

V. Simulation studies

We test the synchronisers by simulation of a system centred at 5 GHz, using 512 subcarriers of 50 kHz width. The guard bands are 16 subcarriers at each end. Fading frequency selective channel characteristics are from the Case II Vehicular A model of [9], at a mobile speed of 25 m/s. The channel was in the simulations constant during each symbol interval.

The ignorant receiver was simulated when receiving unknown 4QAM symbols on all subcarriers (excluding guard bands and the central carrier), using both symbol rate sampling and oversampling by a factor of two. It was also simulated when receiving OFDM pilots in which only every eighth subcarrier was used and the rest were nulled. The purpose of these test cases were mainly to study the frequency acquisition performance for an ignorant mobile just switched on.

The well-informed receiver was tried on fully loaded OFDM pilot symbols in which all subcarriers carried known 4QAM symbols, as well as for the previously mentioned sparsely loaded OFDM pilots in which only every eighth subcarrier was used. Its performance for different channel prediction accuracy was assessed by studying the normalised mean squared estimation error.

A. The ignorant receiver needs several symbols for acquisition

We see from Figure 1 that the ignorant synchroniser in (10) is not very successful in estimating the frequency offset by use of only one OFDM symbol, unless the SNR is very high. This is mostly due to the multiple modes in the posterior. But we also see that updating of the posterior according to (6) results in a considerable improvement over the next few symbols (by suppressing all modes but one). Actually, the simulations were carried out using a simplified updating $p(\Delta\omega|d_1, d_2, \mathcal{I}) \propto p(d_2|\Delta\omega, \mathcal{I})p(\Delta\omega|d_1, \mathcal{I})$ where the logical dependence between d_1 and d_2 was ignored.

B. Oversampling improves the ignorant receiver

As we noted in section IV-A, it is suggested by (10) that oversampling might improve the estimation accuracy. This is indeed shown in Figure 1 were oversampling by a factor of two reduces the error noticeably. In order to achieve estimation accuracies approaching 10^{-3} at an SNR as low as 6 dB, oversampling could be an option during acquisition.

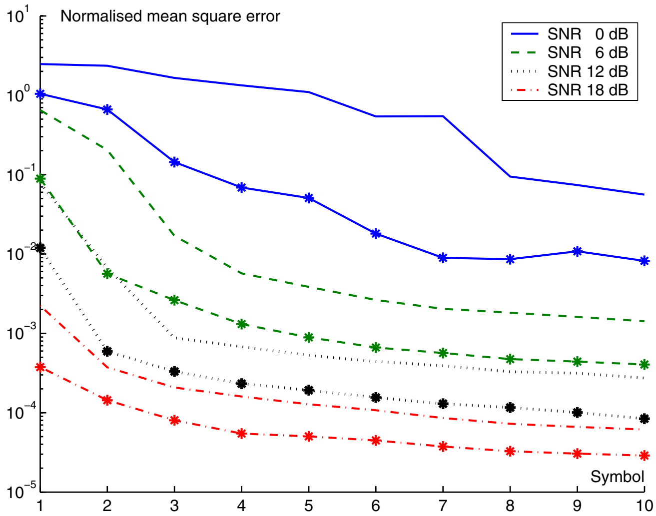

Figure 1. The performance of the ignorant receiver can improve quite significantly when consecutive symbols are used. Two extra symbols can give an improvement of up to 100 times at an SNR of 12 dB, no oversampling (—). Oversampling of the signal (—∗—) yields a noticeable error reduction.

C. Averaging of maximum likelihood estimates not sufficient

In the present problem there are indeed good reasons to retain the whole posterior rather than just the corresponding point estimate. This is shown by the simulations presented in Figure 2 where we compare successive averaging of the point estimates with updating according to Bayes' rule. We clearly see that there must be crucial information present in the posterior $p(\Delta\omega|d_n, \mathcal{I})$ that is not captured by the maximum likelihood reference estimate. It is most likely the ability to eventually rule out all the "false modes" that in this case renders the Bayesian approach so much more powerful.

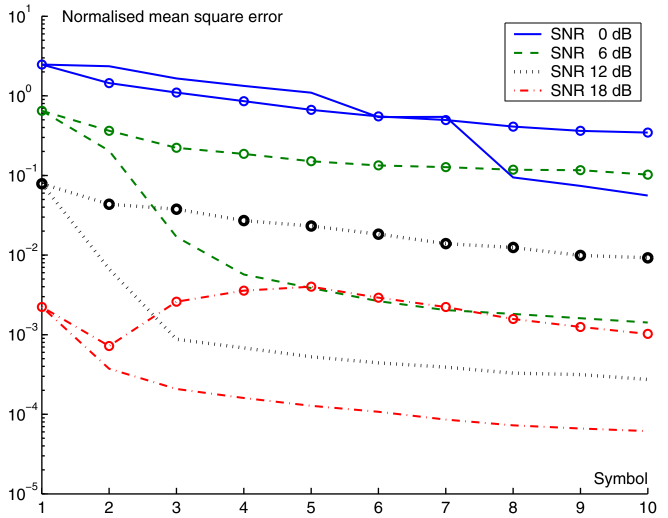

Figure 2. Successive averaging of maximum likelihood estimates (—○—) does not yield the same performance as updating the posterior by the use of Bayes' rule (—). Results pertain to the ignorant receiver using symbol rate sampling.

D. The width of the posterior is a good reliability measure

What we also get from the posterior probability distribution for $\Delta\omega$ is a reliability measure: how good is our present estimate. We propose that a mobile should use the variance of the posterior distribution as an indicator of when it can transmit its first message without causing excessive interference. Figure 3 shows the good correspondence between simulated errors and the uncertainty as given by the posterior variance.

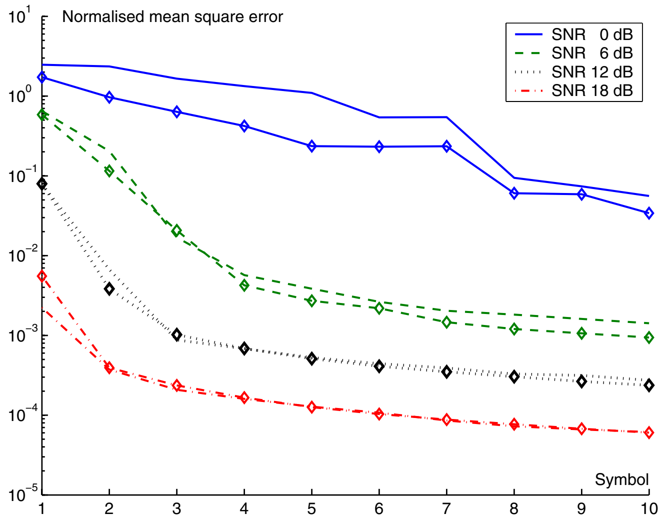

Figure 3. Simulations support the use of the variance of the posterior (—◇—) as an indicator of when the mobile is sufficiently synchronised to the system. It is here compared with the actual error in the simulations (—).

E. The impact of pilot symbols and channel predictions

A mobile that is sufficiently synchronised to the base station to have access to the OFDM pilot symbols can make use of these when fine tuning its frequency synchronisation. In Figure 4 we present the increase in estimation accuracy obtained when using fully loaded OFDM pilots and channel predictions of different quality, as given by the signal-to-estimation error ratio (SER) $B^H B/(m\alpha^2)$ [10]. The results were obtained with a vague prior in (12); $p(\Delta\omega|\mathcal{I}) = \mathcal{N}(0, 4)$. We see from a comparison of Figure 1 and Figure 4 that a SER of 9 dB can give an improvement of almost three orders of magnitude at an SNR of 6 dB (regarding the accuracy obtained by use of one OFDM symbol).

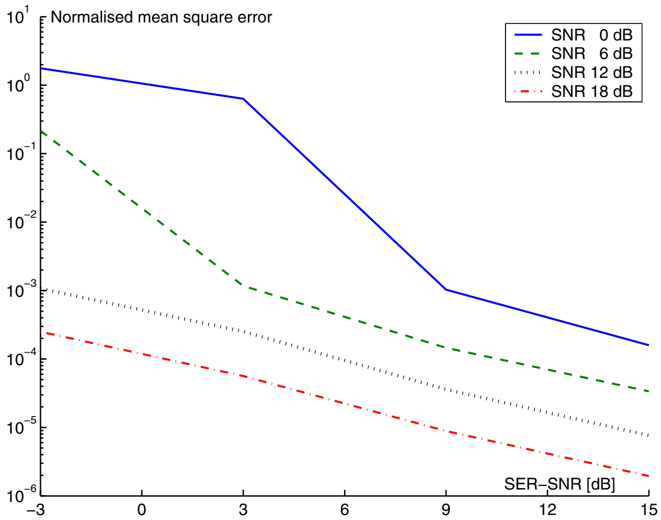

Figure 4. Access to pilots and channel predictions can substantially increase the estimation accuracy. This fact can be exploited when fine-tuning the local oscillator during operation.

F. The effects of subcarrier nulling

If we consider a sparsely loaded OFDM pilot symbol, i.e. one in which some subcarriers are nulled while the others carry QAM pilot symbols, instead of a fully loaded we note some quite dramatic effects. Figure 5 shows the performance of the well-informed receiver when only every eighth subcarrier is used. The channel predictions now seem to have only marginal impact. The estimation accuracy is increased significantly for low SNR and low SER, but seems to decrease for high SNR and high SER (compare with Figure 4).

These results then also suggest that subcarrier nulling can speed up frequency acquisition, i.e. increase the performance of the ignorant receiver, by increasing the discriminating power of the term $d^H W W^H d$ in (10). But this requires that the receiver is a little bit less ignorant than presently: it must know that these sparsely loaded pilots exists (which is reasonable). Otherwise the performance is decreased by subcarrier nulling: simulations show that the normalised mean square error is around 1.5 at an SNR of 18 dB.

VI. Conclusions

Our main conclusions from the preceding analysis and simulations are: 1) Updating of the full posterior probability

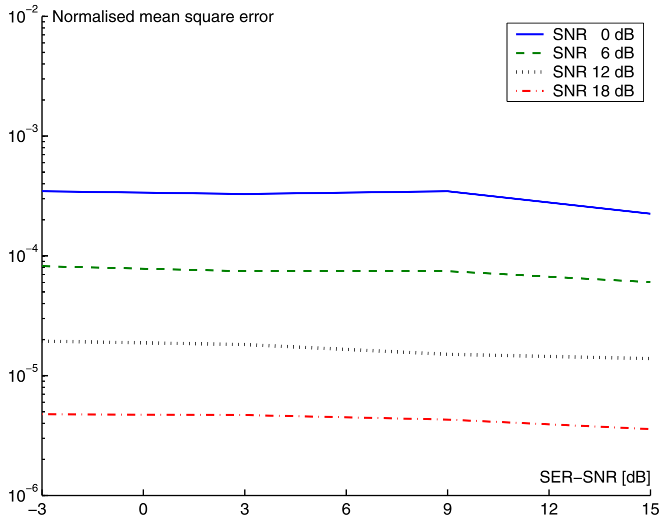

Figure 5. Nulling of subcarriers in an OFDM pilot seems to render the channel predictions almost useless for the well-informed receiver, unless they are of extremely good quality. This also means that the ignorant receiver, if it knew the locations of the nulls, could perform very well with sparse pilots.

distribution for $\Delta\omega$ according to Bayes' rule retains information that can be crucial at low SNR. 2) The variance of the posterior is a good indicator for when a mobile can transmit without incurring other users excessive interference due to a frequency offset. 3) Fast frequency acquisition is possible at low SNR without the use of pilot information. Pilots with nulled subcarriers can on the other hand speed up the acquisition provided that the receiver knows which subcarriers are nulled (it need not know the QAM pilot information, however). Oversampling can also speed up acquisition.

References

[1] C. R. N. Athaudage, K. Sathananthan and R. R. V. Angiras, "Enhanced Frequency Synchronization for OFDM Systems using Timing Error Feedback Compensation", *IEEE PIMRC 2004*, September 2004.

[2] J.-J. van de Beek, M. Sandell and P. O. Börjesson, "ML Estimation of Time and Frequency Offset in OFDM", *IEEE Trans. Signal Processing*, vol. 45, pp. 1800-1805, July 1997.

[3] U. Tureli, H. Liu and M. D. Zoltowski, "OFDM Blind Carrier Offset Estimation: ESPRIT", *IEEE Trans. Commun.*, vol. 48, pp. 1459-1461, September 2000.

[4] G. L. Bretthorst, *Bayesian Spectrum Analysis and Parameter Estimation*, Lecture Notes in Statistics, vol. 48, Springer-Verlag, 1988.

[5] E. T. Jaynes, *Probaility Theory – The Logic of Science*, Cambridge University Press, April 2003.

[6] M. Speth, S. A. Fechtel, G. Fock and H. Meyr, "Optimum Receiver Design for Wireless Broad-Band Systems Using OFDM – Part I", *IEEE Trans. Commun.*, vol. 47, pp. 1668-1677, November 1999.

[7] M. Ghogho, A. Swami and G. B. Giannakis, "Optimized Null-Subcarrier Selection for CFO Estimation in OFDM over Frequency-Selective Channels", *IEEE GLOBECOM '01*, vol. 1, pp. 202-206, November 2001.

[8] B. Chen and H. Wang, "Blind Estimation of OFDM Carrier Frequency Offset via Oversampling", *IEEE Trans. Signal Processing*, vol. 52, pp.2047-2057, July 2004.

[9] 3GPP TR25.996 V6.1.0, "Spatial channel model for Multiple Input Multiple Output (MIMO) simulations", realease 6, September 2003.

[10] M. Sternad and D. Aronsson, "Channel Estimation and Prediction for Adaptive OFDM Downlinks", *IEEE VTC 2003-Fall*, vol. 2, pp. 1283-1287, October 2003.