

GNSS Data Demodulation over Fading Environments: Antipodal and M-ary CSK Modulations

 ISSN 1751-8644
 doi: 0000000000
 www.ietdl.org

Lorenzo Ortega^{1,*}, Jordi Vilà-Valls², Charly Poulliat³, and Pau Closas⁴

¹ Telecommunications for Space and Aeronautics Lab (TéSA), Toulouse, 31500, France

² ISAE-SUPAERO, University of Toulouse, 31055, Toulouse, France.

³ INP/ENSEEIH, University of Toulouse, Toulouse, 31000, France.

⁴ Dept. of Electrical and Computer Engineering, Northeastern University, Boston, MA, 02115, USA.

* E-mail: lorenzo.ortega@tesa.prd.fr

Abstract: This article investigates new strategies to compute accurate low-complexity Log Likelihood Ratio (LLR) values based on the Bayesian formulation under uncorrelated fading channels for both antipodal and CSK modulations when no Channel State Information (CSI) is available at the receiver. These LLR values are then used as input to modern error correcting schemes used in the data decoding process of last generation GNSS signals. Theoretical analysis based on the maximum achievable rate is presented for the different methods in order to evaluate the performance degradation with respect to the optimal CSI channel. Finally, Frame Error Rate (FER) simulation results are shown, validating the appropriate performance of the proposed LLR approximation methods.

1 Introduction

Reliable and precise position, navigation and timing (PNT) information is fundamental in safety-critical applications such as Intelligent Transportation Systems (ITS), automated aircraft landing or autonomous unmanned ground/air vehicles (robots/drones), to name a few. The main source of positioning information is provided by Global Navigation Satellite Systems (GNSS) [1], a technology which has attracted a lot of interest in the last decade. Even if most of the research has been in the signal processing aspects [2], in order to overcome the system limitations under non-nominal conditions [3, 4], and data fusion strategies with alternative ranging technologies [5–8], a key part of GNSS receivers is the data demodulation stage which allow to recover essential information. The latter has been long disregarded but may be a critical point in harsh environments, being the main object of this contribution.

In the last generation of GNSS signals, modern error correcting codes (i.e., such as low-density parity-check (LDPC) or convolutional codes) were considered in the GNSS signal design in order to enhance the data demodulation performance, especially over harsh scenarios [9, 10]. The inputs to the corresponding soft decoding algorithms are the so-called Log-Likelihood Ratio (LLR) values [11, 12], which represent a statistical test to compare the goodness of fit between probabilities of receiving a positive or negative logic bit. These LLR values can be shown to be sufficient statistics for the decoding and detection process [13]. Typically, in order to compute the LLRs, the entire knowledge of the propagation channel behaviour referred as perfect channel state information (CSI) is considered. However, this assumption does not hold necessary true in real-life applications since the CSI might not be fully available at the receiver [14], yielding a possible decoding loss due to the incorrect information at the decoding input. This situation can be further aggravated under urban environments where effects such as shadowing or multipath reduce the channel capacity at the receiver.

In this work, we focus on the uncorrelated fading channel [11] which is commonly used to model phenomena such as shadowing or multipath. This channel is modeled by a fading gain h_n and an additive Gaussian noise with variance σ^2 . Note that if this parameters are perfectly known at the receiver (i.e., perfect CSI) the LLR has a well known closed-form expression [11]. Otherwise, the LLR expression

is unknown and LLR approximations are required. Indeed, when no CSI is available at the GNSS receiver only one method to compute such LLR approximation is available in the literature [11], which has several limitations: 1) this method can only be used with antipodal modulations and it is not valid for M -ary modulations, such as the code shift keying (CSK) modulation [15], and 2) it requires a high decoder complexity in order to compute the LLR approximation, which may not be useful in practice.

Bearing in mind the lack of practical solutions in the literature, the goal of this contribution is to provide new strategies to compute low-complexity closed-form LLR approximation expressions for both antipodal and CSK modulations, considering the uncorrelated fading channel and with no CSI available at the receiver. Thus, the article focus on the following cases:

- *Antipodal modulation and GNSS pilot signal (Section 3):* most of the new generation GNSS signals are composed by a data and a pilot component, therefore two LLR approximation methods are presented first considering the GNSS pilot signal case, which allows to estimate some channel parameters at the receiver. The first method was proposed in [11] based on [16], which seeks to maximize the mutual information between the transmitted symbol and the LLR, and is given for completeness. This method provides a good LLR approximation at expenses of a high complexity burden. To reduce the computational complexity, and resorting to a Bayesian formulation, a second method based on our previous work [17] is presented. This method not only reduces the complexity but also provides similar LLR values. In order to evaluate the performance w.r.t. the perfect and statistical CSI solutions a study of the capacity is also provided.
- *Antipodal modulation and GNSS data signal (Section 4):* legacy GNSS signals may not have a pilot component and/or simple receiver implementations may only track the data component. In that context, and w.r.t. the pilot signal case in Section 3, new alternatives must be studied. We propose two Bayesian approximations considering an uncorrelated fading channel, no CSI and a data signal component. Again, we also provide the channel capacity performance analysis of the proposed methods, w.r.t. the perfect CSI case and the Bayesian approximation with a pilot signal component.
- *M-ary CSK modulation and GNSS pilot signal (Section 6):* the CSK is a M -ary modulation that can increase the data rate without losing synchronization performance, thus being a suitable signal

candidate for futures GNSS applications [15]. We derive a new low-complexity LLR approximation expression that can be used over uncorrelated fading channels when no CSI is available at the receiver. Moreover, since this modern signal is expected to be transmitted as a data component [12, 18] along with a GNSS pilot component, the latter can be exploited to infer some of the channel parameters. Finally, we also provide the channel capacity performance analysis of the proposed methods.

The article is organized as follows: the system model for the antipodal modulation and background on LLR expressions considering both perfect CSI and statistical CSI [17, 19] are provided in Section 2; LLR approximations for an antipodal modulation without CSI considering a GNSS pilot signal are provided in Section 3, and without a pilot signal in Section 4; the system model for the CSK modulation and background on LLR expressions considering perfect CSI are provided in Section 5; the LLR approximations for the M -ary CSK modulation without CSI considering a GNSS pilot signal are provided in Section 6; Results are analyzed for two uncorrelated fading channels in Section 7 and conclusions are drawn in Section 8.

2 System Model for the Antipodal Modulation and LLR with Perfect/Statistical CSI

2.1 System Model and LLR with Perfect CSI

Current GNSS signals transmit binary data information, we assume the transmitted message as a binary vector $\mathbf{u} = [u_1, \dots, u_K]$ of K bits. This message is encoded into a codeword $\mathbf{c} = [c_1, \dots, c_N]$ of length $N > K$ and mapped to Antipodal symbols (e.g. Binary Phase-Shift Keying) $x_n = \mu(c_n) \in \{-1, 1\}$, where we impose $\mu(0) = 1$ and $\mu(1) = -1$. Each symbol x_n is then spread using a pseudo-random noise (PRN) sequence that can be expressed in vector form as $\mathbf{p}_n \in \mathbb{R}^L$, where L corresponds to the number of chips of the PRN sequence. Then, the transmitted symbol per coded bit is given by

$$\mathbf{x}_n = x_n \cdot \mathbf{p}_n \in \mathbb{R}^L, \quad n = \{1, \dots, N\}, \quad (1)$$

where by convention vectors are defined as column vectors. Then, chip-level rectangular pulse shaping is used before transmission. Considering the data demodulation stage of a GNSS receiver, a key task is to obtain the posterior probability of a transmitted code symbol c_n given the observed sample y_n . Received signal models for two relevant (open sky and fading) scenarios are discussed in this subsection, after relevant LLR concepts are reviewed.

The information in y_n is used to compute the LLR value, defined for the n -th symbol as,

$$\mathcal{L}_n = \ln \left(\frac{p(c_n = 0|y_n)}{p(c_n = 1|y_n)} \right) = \ln \left(\frac{p(x_n = 1|y_n)}{p(x_n = -1|y_n)} \right) \quad (2)$$

This LLR can be used to feed the input of a Soft Input Soft Output (SISO) error correction decoder [13]. Assuming that c_n are identically and uniformly distributed (i.u.d.) [13] $\forall n = 1, \dots, N$, (2) can also be written as

$$\mathcal{L}_n = \ln \left(\frac{p(y_n|x_n = 1)}{p(y_n|x_n = -1)} \right), \quad (3)$$

where equiprobable symbols are assumed. Note that $p(y_n|x_n)$ represents the likelihood distribution given a transmitted symbol, which directly depends on the transmission channel.

• Open Sky Environment

Standard techniques typically assume an additive white Gaussian noise (AWGN) propagation model. Assuming perfect time and frequency synchronisation, the received baseband symbol sequence at

the chip-level can be written as

$$y_n = \mathbf{x}_n + \mathbf{w}_n \in \mathbb{R}^L, \quad n = \{1, \dots, N\}, \quad (4)$$

where $\mathbf{w}_n \sim \mathcal{N}(\mathbf{0}, L^2 \cdot \sigma^2 \mathbf{I}_L)$, with \mathbf{I}_L being the identity matrix of size L . Thus, we denote the normalized output of the matched filter as $y_n = \frac{\mathbf{y}_n^\top \mathbf{p}_n}{L} \in \mathbb{R}$. Then, the normalized received symbol sequence is

$$y_n = x_n + w_n \in \mathbb{R}, \quad n = \{1, \dots, N\}, \quad (5)$$

where $w_n \sim \mathcal{N}(0, \sigma^2)$ and σ^2 is known. In a standard GNSS receiver this is the symbol model at the output of the prompt correlator, for which the channel is considered to be static over a symbol period. Note that all symbols y_n are affected by the same noise statistics then the Gaussian likelihood $p(y_n|x_n)$ is

$$p(y_n|x_n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_n-x_n)^2}{2\sigma^2}}, \quad (6)$$

and considering equiprobable symbols the LLR can be computed as

$$\mathcal{L}_n = \frac{2y_n}{\sigma^2}. \quad (7)$$

• Fading Environment

If we consider now a GNSS environments characterized by effects such as shadowing or multipath, the detection function typically used in this context corresponds to an uncorrelated fading channel with additive real-valued AWGN. Again, we assume perfect time and frequency synchronisation, the received baseband symbol sequence at the chip-level can be written as

$$y_n = \mathbf{h}_n \cdot \mathbf{x}_n + \mathbf{w}_n \in \mathbb{R}^L, \quad n = \{1, \dots, N\}, \quad (8)$$

where \mathbf{h}_n denotes the fading gain per chip and $\mathbf{w}_n \sim \mathcal{N}(\mathbf{0}, L^2 \cdot \sigma^2 \mathbf{I}_L)$, with \mathbf{I}_L being the identity matrix of size L . Thus, we denote the normalized output of the matched filter as $y_n = \frac{\mathbf{y}_n^\top \mathbf{p}_n}{L} \in \mathbb{R}$. Then, the normalized received symbol sequence is

$$y_n = h_n \cdot x_n + w_n \in \mathbb{R}, \quad n = \{1, \dots, N\}, \quad (9)$$

where both w_n and h_n are independent random processes. w_n are i.i.d. centered Gaussian random variables with variance σ^2 , i.e. $w_n \sim \mathcal{N}(0, \sigma^2)$. h_n are also i.i.d. random variables with an associated probability density function (pdf) given by $p(h)$, i.e. $h_n \sim p(h)$. It is assumed that $h_n \geq 0$ and change independently from one sample to another. All the symbols are again affected by the same noise statistics and the LLR simplifies to [16]

$$\mathcal{L}_n = \frac{2}{\sigma^2} h_n \cdot y_n, \quad (10)$$

which explicitly implies perfect CSI, i.e. h_n and the variance σ^2 are assumed known. In practice, this assumption does not hold and even if σ^2 can be precisely estimated, the fading gain remains unknown in most of the situations.

2.2 LLR with Statistical CSI

A relaxation of the perfect CSI situation is to consider that full statistical CSI is available at the receiver, i.e. σ^2 is assumed known or accurately estimated and h_n is an unknown random quantity whose pdf and parameters are well characterized. Additionally, we consider a binary input memoryless channel where symbols x_n are unknown. From a Bayesian perspective [20], since both x_n and h_n are unknown quantities it is sound to consider them as random

variable. All the statistically relevant information about these variables is contained in their joint posterior distribution $p(x_n, h_n|y_n)$. Assuming that x_n and h_n are independent

$$\begin{aligned} p(x_n, h_n|y_n) &\propto p(y_n|x_n, h_n)p(x_n, h_n) \\ &= p(y_n|x_n, h_n)p(x_n)p(h_n), \end{aligned} \quad (11)$$

where the first term corresponds to the likelihood of observations given the unknowns and the second term represents the a priori knowledge on x_n and h_n . From (9) the likelihood distribution is

$$p(y_n|x_n, h_n) \sim \mathcal{N}(\mu, \sigma^2), \quad (12)$$

with mean $\mu = h_n \cdot x_n$ and known variance σ^2 . According to the LLR definition (2) we are interested in the marginal distribution,

$$p(x_n|y_n) = \int_{-\infty}^{\infty} p(x_n, h|y_n)dh, \quad (13)$$

which leads to

$$\mathcal{L}_n = \ln \left(\frac{\int_{-\infty}^{\infty} p(x_n = 1, h|y_n) dh}{\int_{-\infty}^{\infty} p(x_n = -1, h|y_n) dh} \right) \quad (14)$$

$$= \ln \left(\frac{\int_{-\infty}^{\infty} p(y_n|x_n = 1, h)p(h) dh}{\int_{-\infty}^{\infty} p(y_n|x_n = -1, h)p(h) dh} \right). \quad (15)$$

since x_n are equiprobable. In order to model a GNSS urban environment it is common practice to use the 2-state Prieto model [21]. Nevertheless, this model does not have a closed-form expression for the channel gain pdf $p(h)$. **An alternative proposed in [12] is to consider a Rice distribution. However, as noted in [12], the LLR has no closed-form and in practice its use is computationally too complex. In addition, the statistical CSI required to compute the LLR may not be available, therefore different alternatives must be accounted for, being the object of the rest of the article.**

3 Antipodal GNSS Data Demodulation without CSI for a Pilot Signal Component

In this section we present two methods to compute the LLR approximation when no CSI is available at the receiver, but some channel parameters can be inferred by means of a GNSS pilot component, which is typically available in new generation GNSS signals. For instance, the GPS L1C signal is composed of two different components: a data component used to transmit the C/NAV-2 message [22] and a pilot component which transmits a secondary known code [22]. This secondary code can be used as a learning/training sequence as described in this section.

3.1 Best LLR Linear Approximation (BLA)

For completeness we introduce the method proposed in [16] and used for a GNSS data component in [12]. Motivated by (10), this method computes the coefficient α that provides the best linear approximation of the LLR as

$$\mathcal{L}_n = \alpha y_n. \quad (16)$$

The scaling factor α is obtained by maximizing the mutual information $I(\mathcal{L}; X)$ between the transmitted symbol X and the detector input \mathcal{L} , both being random variables whose realizations x_n and \mathcal{L}_n are observed at the receiver,

$$\alpha = \arg \max_{\alpha} I(\mathcal{L}; X), \quad (17)$$

where the mutual information is defined as,

$$I(\mathcal{L}; X) = H(X) - H(X|\mathcal{L}), \quad (18)$$

with $H(X)$ and $H(X|\mathcal{L})$ the entropy of X and the conditional entropy of X given \mathcal{L} , respectively. When considering a memoryless

binary input symmetric output channel and consistent LLR values, this expression can be expressed as a function of the LLR pdf at the receiver input [23], considering $\{X = +1\}$:

$$I(\mathcal{L}; X) = 1 - \int_{-\infty}^{\infty} \log_2(1 + e^{-\mathcal{L}}) p(\mathcal{L}|X = +1) d\mathcal{L}. \quad (19)$$

This approximate criterion is derived from the capacity associated with a binary input memoryless channel, $C = I(\mathcal{L}; X) = 1 - \mathbb{E}_{\mathcal{L}|X=+1}[\log_2(1 + e^{-\mathcal{L}})]$, for which the conditional pdf of the true LLRs has been replaced by the conditional pdf of the approximated ones. When considering $C = I(\mathcal{L}; X)$ it can be shown that the conditional pdfs given the true LLRs are both symmetric and consistent (see [16]). The latter condition is not necessary fulfilled for $p(\hat{\mathcal{L}}|X)$ and the quantity $\hat{I}(\hat{\mathcal{L}}; X)$ cannot be interpreted as a true mutual information quantity. However, this quantity can be used as a good approximate measure, as proved in [16] where this quantity is maximized for the best linear LLR approximation. Notice that the proposed optimization method [16] assumes the knowledge of the linearly approximated LLRs conditional pdf. However, in real scenarios this is unknown at the receiver and one has to resort to a numerical resolution (which it is often computationally demanding) in order to estimate the approximated LLRs conditional pdf. To overcome this limitation, one can resort to the corresponding empirical mean estimator as done in [11, 12] through the time average estimation proposed in [24],

$$\hat{I}(\hat{\mathcal{L}}; X) \approx 1 - \frac{1}{K} \sum_{k=1}^K \log_2(1 + e^{-x_k \hat{\mathcal{L}}_k}), \quad (20)$$

where K is the number of samples used to estimate $\hat{I}(\hat{\mathcal{L}}; X)$. Notice that this method needs a learning sequence (i.e., known values x_k). This information can be directly obtained through the symbols of the GNSS pilot component. Finally, in order to compute α , one can apply a one-dimensional search method [25] based on the objective function (20). We underline that that this method does not require knowledge of σ^2 , therefore no CSI is required.

3.2 Bayesian LLR Linear Approximation

Recall from Section 2.2 that the problem of computing the LLR values involve solving the integrals in (14), for which we have to select a prior distribution for h_n . In [11] the pdf $p(h)$ was selected to be a Rice distribution, leading to a complex LLR expression for practical applications. In Bayesian inference, a common approach is to select a prior distribution to be conjugate of the likelihood distribution, which results in a posterior distribution that is of the same family as the a priori, where only the parameters need to be updated [26]. This idea was exploited in [17] to provide a simple low-complexity closed-form LLR approximation for M-ary modulations over uncorrelated fading channels. Given that the likelihood (12) is Gaussian, the conjugate prior distribution for h_n in (11) is also Gaussian [20],

$$h_n \sim \mathcal{N}(\mu_h, \sigma_h^2), \quad (21)$$

where the pdf parameters (i.e., μ_h and σ_h^2) need to be adjusted according to the unknown parameters' uncertainty. The marginal distribution in (14) is obtained by solving the integral

$$p(x_n|y_n) \propto \int_{-\infty}^{\infty} e^{-\frac{(y_n - hx_n)^2}{2\sigma^2}} e^{-\frac{(h - \mu_h)^2}{2\sigma_h^2}} dh, \quad (22)$$

which can be shown to be another Gaussian (refer to [17])

$$p(x_n|y_n) \propto e^{-\frac{(x_n - y_n / \mu_h)^2}{2\left(\frac{\sigma^2 + \sigma_h^2}{\mu_h^2}\right)}} \propto \mathcal{N}\left(\frac{y_n}{\mu_h}, \left(\frac{\sigma^2 + \sigma_h^2}{\mu_h^2}\right)\right), \quad (23)$$

and the corresponding LLR are

$$\mathcal{L}_n = \frac{2y_n\mu_h}{(\sigma^2 + \sigma_{h_n}^2)}, \quad (24)$$

where similar to (16) the resulting LLR approximation is a linear function of y_n . Additionally, we underline that under the AWGN channel (i.e., $\mu_h = 1$ and $\sigma_h^2 = 0$) the LLR in (24) reduces to the Gaussian LLR solution in (7). Notice that (24) requires the knowledge of σ^2 , as well as μ_h and σ_h^2 . The latter values can be obtained by resorting to the maximum likelihood (ML) estimates [17]:

$$\hat{\mu}_h = \frac{1}{K} \sum_{k=1}^K y_k x_k, \quad \hat{\sigma}_h^2 = \frac{1}{K} \sum_{k=1}^K (y_k x_k - \hat{\mu}_h)^2 - \sigma^2. \quad (25)$$

where K is the number of samples used to estimate μ_h and σ_h^2 , x_k is the k -th symbol of a binary learning sequence (i.e., the known secondary code from the GNSS pilot component) and y_k is the k -th received pilot symbol. Finally, the Bayesian LLR linear approximation is given by

$$\mathcal{L}_n = \frac{2y_n\hat{\mu}_h}{\left(\frac{1}{K} \sum_{k=1}^K (y_k x_k - \hat{\mu}_h)^2\right)}. \quad (26)$$

Figure 1 summarizes the different approaches in Section 3.

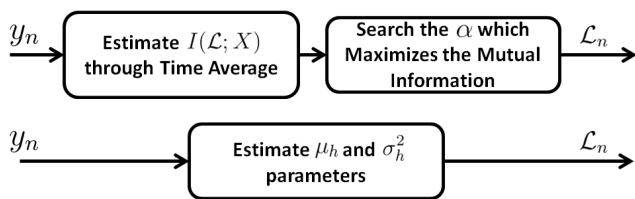


Fig. 1: Linear LLR approximations: (top) estimating $I(\mathcal{L}; X)$ through time average, and (bottom) the Bayesian approach.

3.3 Performance of the LLR Approximation Methods for the Antipodal Modulation and a Pilot Signal

In this section we evaluate the LLR approximations over well-known uncorrelated fading channels in order to compare the channel capacity performance with the perfect CSI and statistical CSI cases.

• Normalized Rayleigh Fading Channel

The normalized Rayleigh fading channel [27] is typically used to describe phenomena such as shadowing or multipath, and it is well known to provide a closed-form solution of the LLR (14), i.e., when full statistical CSI is considered. As a consequence, we can obtain an upper bound for the LLR approximation performances. Following [27], the LLR expression (14) when considering a normalized Rayleigh fading channel is

$$\mathcal{L}_n = \log \left(\frac{\Phi \left(y / \sqrt{2\sigma^2(1 + 2\sigma^2)} \right)}{\Phi \left(-y / \sqrt{2\sigma^2(1 + 2\sigma^2)} \right)} \right), \quad (27)$$

where $\Phi(z) = 1 + \sqrt{\pi} z e^{z^2} \operatorname{erfc}(-z)$, and $\operatorname{erfc}(\cdot)$ represents the complementary error function. We evaluate the upper bound of the maximum achievable rate R_0 [28] of the LLR linear approximations as well as the LLR expression with full statistical CSI (27) and perfect CSI (7), (10). To compute such maximum achievable rate we follow the methodology proposed in [29] based on the extrinsic information transfer (EXIT) chart analysis.

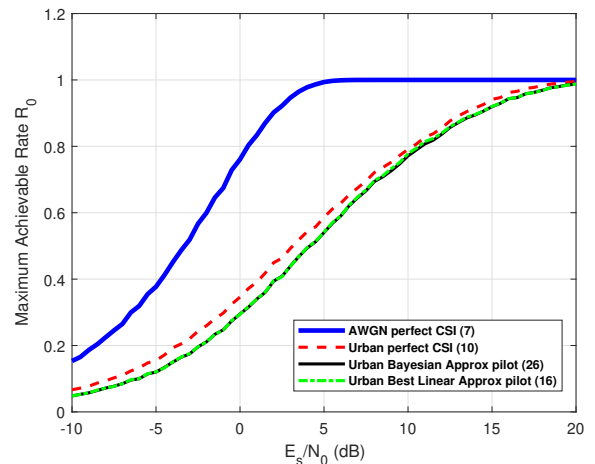
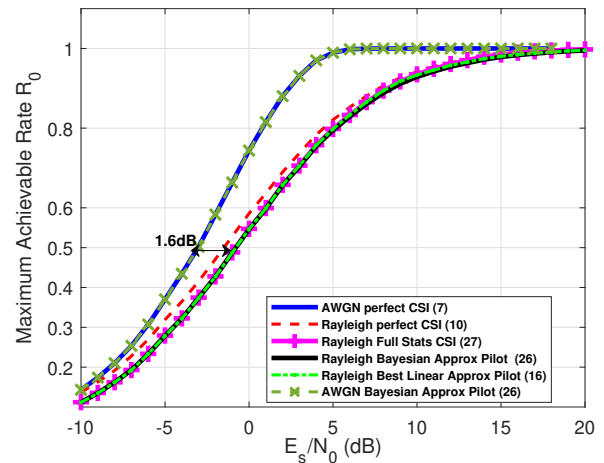


Fig. 2: Antipodal modulation, no CSI and pilot signal: upper bounds of maximum achievable rate R_0 for: (top) a normalized Rayleigh channel, and (bottom) a 2-state Prieto channel model.

Figure 2 (top plot) illustrates the open sky R_0 upper bound, i.e., AWGN channel, for i) the perfect CSI case (7), and ii) the Bayesian LLR linear approximation (26) when 12 seconds of pilot symbols are retrieved. Moreover, considering the normalized Rayleigh channel the R_0 upper bound is shown for iii) perfect CSI (10), iv) full statistical CSI (27), v) Bayesian LLR linear approximation (26), and vi) the LLR approximation (16), considering that the mutual information $I(\mathcal{L}; X)$ in (17) is computed from (19) when 12 seconds of pilot symbols are retrieved. We can observe that:

- i) *Channel capacity loss caused by the fading effect:* considering an ideal coding scheme of rate $R = 1/2$, the channel capacity loss between the AWGN channel and the normalized Rayleigh channel (when perfect CSI solutions are assumed) is around 1.6 dB. Notice that this loss can be reduced when applying lower rate channel coding schemes. As an example, the channel capacity difference with a channel coding scheme of rate $R = 1/3$ is around 0.9 dB.
- ii) *Channel capacity loss due to channel uncertainty:* considering an ideal coding scheme of rate $R = 1/2$, the channel capacity loss is around 0.8 dB between the full statistical CSI solution and the perfect CSI solution. Moreover, the best and Bayesian LLR linear approximations provide the same channel capacity than the full statistical CSI solution, proving that when no perfect CSI is available, only the first and second order moments of the fading distribution are required to achieve an optimal solution. Finally, when the transmission channel is characterized by an AWGN, the Bayesian solution (26) converges to the perfect CSI solution (7).

• 2-State Prieto Channel

In a second scenario, we propose to evaluate the performance in a more realistic GNSS urban scenario. We consider a 2-state Prieto channel model [21] for a vehicle speed of 50 km/h and a satellite elevation angle of 40 degrees. In this example, we consider the data component of the signal GPS L1C which is characterized by a symbol rate of 100 symbols/s and a PRN code of length 10230 chips. The results are shown in Figure 2 (bottom plot), where we illustrate again the upper bound of the maximum achievable rate.

In this case, considering perfect CSI and a coding scheme of rate 1/2, the channel capacity loss is around 6 dB w.r.t. the AWGN scenario. Notice that for this type of scenarios, error correcting schemes with lower rates are highly recommended. On the other hand, we verify that the best and Bayesian linear approximations, ((16) and (24)), almost converge to the perfect CSI solution (loss around 0.8 dB), proving the validity of such approximations. Finally, we underline that no full statistical CSI expression is available since $p(\mathcal{L}|X = +1)$ has no closed-form and is unknown at the receiver.

To conclude, we provide a brief comment on the complexity of the LLR linear approximation methods (refer to Figure 1): i) the first method requires an online estimation technique which needs to resort to an iterative one-dimensional search method, based on a cost function involving log/exp function evaluations. Note that the complexity of this method directly depends on the number of samples K to estimate $\hat{I}(\hat{\mathcal{L}}; X)$; ii) the second method avoids to compute the one-dimensional search, but instead the first and second order moments of the fading distribution have to be estimated. Note from (25) that only simple arithmetical operations are required. Again, the complexity of the method depends on the number of samples K used to estimate μ_h and σ_h^2 , but for equal number of samples the Bayesian solution is computationally less expensive than the best linear one.

4 Antipodal GNSS Data Demodulation without CSI for a Data Signal Component

The previous section focused on LLR approximations when a pilot component (i.e., training sequence) is available. However, legacy GNSS signals may not have a pilot component, therefore different alternatives must be accounted for. In the sequel we introduce data demodulation strategies considering an uncorrelated fading channel, no CSI and a data signal component.

4.1 Bayesian LLR Approximation without Training Data: MLE of μ_h and σ_h^2 through Expectation-Maximization

In contrast with Section 3.2, where the underlying pilot signal assumption allowed to compute the ML estimates in (25) which are in turn used to compute the LLR approximation (26), in this case we do not have access to such training sequence. Therefore we propose a method to derive the μ_h and σ_h^2 ML estimates when no learning sequence is available. The marginal likelihood $p(y_n) = p(y_n|x_n = 1)p(x_n = 1) + p(y_n|x_n = -1)p(x_n = -1)$ is a mixture of two Gaussian distribution:

$$\begin{aligned} y_n &\sim \mathcal{N}(\mu_h x_n, \sigma_h^2 + \sigma^2)p(x_n = +1) \\ &+ \mathcal{N}(\mu_h x_n, \sigma_h^2 + \sigma^2)p(x_n = -1) \\ &= \frac{1}{2}\mathcal{N}(\mu_h, \sigma_h^2 + \sigma^2) + \frac{1}{2}\mathcal{N}(-\mu_h, \sigma_h^2 + \sigma^2) \\ &= \frac{1}{2}\mathcal{N}(\mu_h, \sigma_a^2) + \frac{1}{2}\mathcal{N}(-\mu_h, \sigma_a^2), \end{aligned} \quad (28)$$

with $\sigma_a^2 = \sigma_h^2 + \sigma^2$. Thus, we can compute $\Lambda(y_n; \mu_h, \sigma_a^2) = \log p(y_1, \dots, y_N)$ as

$$\log(\Lambda) = \sum_{n=1}^N \log \left(\frac{1}{2\sqrt{2\pi\sigma_a^2}} \left(e^{-\frac{(y_n - \mu_h)^2}{2\sigma_a^2}} + e^{-\frac{(y_n + \mu_h)^2}{2\sigma_a^2}} \right) \right), \quad (29)$$

and obtain the maximum likelihood estimates of μ_h and σ_a^2 as the roots of the partial derivatives with respect to μ_h and σ_a^2 . The partial derivative with respect to μ_h is

$$\frac{d \log(\Lambda)}{d\mu_h} = \sum_{n=1}^N \left(\frac{\left(\frac{y_n + \mu_h}{\sigma_a^2} \right) e^{-\frac{y_n \mu_h}{\sigma_a^2}} - \left(\frac{y_n - \mu_h}{\sigma_a^2} \right) e^{-\frac{y_n \mu_h}{\sigma_a^2}}}{e^{-\frac{y_n \mu_h}{\sigma_a^2}} + e^{-\frac{y_n \mu_h}{\sigma_a^2}}} \right), \quad (30)$$

and we have to solve for $\frac{d \log(\Lambda)}{d\mu_h} = 0$, which has no analytical solution. However, conditional on a specific realization of the latent variables x_n we could use the μ_h estimate from the previous Section 3.2. We first obtain the (discrete) posterior distribution of x_n given the observations:

$$\begin{aligned} \gamma_{x_1} &= p(x_n = 1|y_n) = \frac{p(y_n|x_n = 1)p(x_n = 1)}{p(y_n)} \\ &= \frac{\frac{1}{2}\mathcal{N}(\mu_h, \sigma_a^2)}{\frac{1}{2}\mathcal{N}(\mu_h, \sigma_a^2) + \frac{1}{2}\mathcal{N}(-\mu_h, \sigma_a^2)} = \frac{e^{-\frac{y_n \mu_h}{\sigma_a^2}}}{e^{-\frac{y_n \mu_h}{\sigma_a^2}} + e^{-\frac{y_n \mu_h}{\sigma_a^2}}}, \end{aligned} \quad (31)$$

and

$$\begin{aligned} \gamma_{x_{-1}} &= p(x_n = -1|y_n) = \frac{p(y_n|x_n = -1)p(x_n = -1)}{p(y_n)} \\ &= \frac{\frac{1}{2}\mathcal{N}(-\mu_h, \sigma_a^2)}{\frac{1}{2}\mathcal{N}(\mu_h, \sigma_a^2) + \frac{1}{2}\mathcal{N}(-\mu_h, \sigma_a^2)} = \frac{e^{-\frac{y_n \mu_h}{\sigma_a^2}}}{e^{-\frac{y_n \mu_h}{\sigma_a^2}} + e^{-\frac{y_n \mu_h}{\sigma_a^2}}}. \end{aligned} \quad (32)$$

Then, (30) can be expressed as

$$\frac{d \log(\Lambda)}{d\mu_h} = \sum_{n=1}^N \left(\gamma_{x_1} \frac{y_n + \mu_h}{\sigma_a^2} - \gamma_{x_{-1}} \frac{y_n - \mu_h}{\sigma_a^2} \right), \quad (33)$$

and we can estimate μ_h as,

$$\hat{\mu}_h = \frac{1}{N} \left(\sum_{n=1}^N \gamma_{x_1} y_n - \sum_{n=1}^N \gamma_{x_{-1}} y_n \right). \quad (34)$$

We can proceed similarly to estimate σ_a^2 ,

$$\begin{aligned} \frac{d \log(\Lambda)}{d\sigma_a^2} &= \sum_{n=1}^N \left(\frac{\left(\frac{(y_n + \mu_h)^2}{2(\sigma_a^2)^2} \right) e^{-\frac{y_n \mu_h}{\sigma_a^2}} + \left(\frac{(y_n - \mu_h)^2}{2(\sigma_a^2)^2} \right) e^{-\frac{y_n \mu_h}{\sigma_a^2}}}{e^{-\frac{y_n \mu_h}{\sigma_a^2}} + e^{-\frac{y_n \mu_h}{\sigma_a^2}}} - \frac{1}{2\sigma_a^2} \right) \\ &= \sum_{n=1}^N \left(\frac{(y_n + \mu_h)^2}{2(\sigma_a^2)^2} \gamma_{x_1} + \frac{(y_n - \mu_h)^2}{2(\sigma_a^2)^2} \gamma_{x_{-1}} - \frac{1}{2\sigma_a^2} \right), \end{aligned} \quad (35)$$

and obtain σ_a^2 as (i.e., $\frac{d \log(\Lambda)}{d\sigma_a^2} = 0$),

$$\hat{\sigma}_a^2 = \frac{1}{N} \left(\sum_{n=1}^N \gamma_{x_1} (y_n - \hat{\mu}_h)^2 + \sum_{n=1}^N \gamma_{x_{-1}} (y_n + \hat{\mu}_h)^2 \right). \quad (36)$$

Notice that in the previous equations it was not taken into account that γ_{x_1} and $\gamma_{x_{-1}}$ depend on the unknown parameters. An iterative Expectation-Maximization (EM) algorithm can be used in order to estimate $\hat{\mu}_h$ and $\hat{\sigma}_a^2$. The resulting EM algorithm is as follows:

1. Initialize μ_h and σ_a^2 and evaluate the log-likelihoods.
2. E-step: Evaluate the posterior probabilities γ_{x_1} and $\gamma_{x_{-1}}$ using the current values of μ_h and σ_a^2 with (31) and (32).
3. M-step: Estimate the new parameters $\hat{\mu}_h$ and $\hat{\sigma}_a^2$ with the updated values of γ_{x_1} and $\gamma_{x_{-1}}$ using (34) and (36).

4. Evaluate the log-likelihoods with the updated parameter estimates. If the log-likelihood change is below a given *small* threshold ϵ , stop. Otherwise, go back to the E-step.

Finally, the resulting expression for the LLR values with no CSI and a data channel signal is

$$\mathcal{L}_n = \frac{2y_n \hat{\mu}_h}{\hat{\sigma}_a^2}. \quad (37)$$

4.2 Bayesian LLR Approximation without Training Data: Rough Estimation of μ_h and σ_h^2

To avoid the iterations in the previous EM-based solution, a simpler low complexity approach to derive μ_h and σ_h^2 ML estimates is proposed. The ML estimates in (25) can be approximated as,

$$\hat{\mu}_h = \frac{1}{N} \sum_{n=1}^N |y_n|, \quad \hat{\sigma}_h^2 = \frac{1}{N} \sum_{n=1}^N (|y_n| - \hat{\mu}_h)^2 - \sigma^2. \quad (38)$$

where y_n is the n -th received symbols, and N is the number of received symbols used to estimate μ_h and σ_h^2 . The corresponding LLR expression without CSI and a data signal is,

$$\mathcal{L}_n = \frac{2y_n \hat{\mu}_h}{\left(\frac{1}{N} \sum_{n=1}^N (|y_n| - \hat{\mu}_h)^2\right)}. \quad (39)$$

Notice that the approximated ML estimates in (38) are biased, in contrast to the true ones in (25). As expected, this bias is made apparent only at low SNR values. As it will be shown in next Section 4.3 this has a minor impact in terms of maximum achievable rate R_0 with respect to the perfect CSI case, but a clear impact is present at low SNR for the fading distribution first and second order estimation performance (refer to Fig.4). The fact that such bias has almost no impact on the maximum achievable rate translates into a marginal impact in the frame error rate analysis shown in Section 7.1.

4.3 Performance of the LLR Approximation Methods for the Antipodal Modulation and a Data Signal

As done in Section 3.3 for the pilot signal scenario, we can assess the upper bound of the maximum achievable rate R_0 for approximations (37) and (39), results shown in Figure 3. The two approximations are compared to the perfect CSI case for both AWGN and Rayleigh channel, as well as to the Bayesian LLR linear approximation (26) considering a pilot signal.

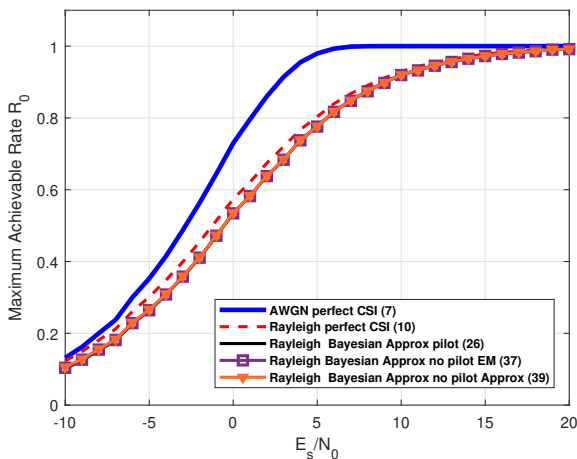


Fig. 3: Antipodal modulation, no CSI and data signal: upper bounds of maximum achievable rate R_0 for a normalized Rayleigh channel.

Notice that all the LLR approximation methods converge to the same solution. However, it is well known that finite length codes used in real coding schemes are not optimal, thus it is likely that in real scenarios these methods will not be equivalent. Then, in order to evaluate this issue we illustrate in Figure 4 the estimation accuracy of the first and second order moments of the fading distribution when 12 seconds of signal are retrieved (i.e., obtained from 2000 Monte Carlo runs). Note from these results that when a pilot is available at the receiver, an accurate estimation of the first and second order moments is achieved independently of the signal-to-noise ratios (SNR). On the other hand, when no pilot is available, the EM-based method does not correctly estimate the parameters in low SNR regimes. In addition, the estimation accuracy using the rough approximation (38) is significantly degraded and only performs well for high SNR. Thus, it is expected that the error correcting performance of real channel coding schemes when considering both LLR approximations (37) and (39) will be degraded w.r.t the Bayesian approximation exploiting a pilot signal (26). Because the EM-based solution performs better than (38), we expect the error correction performance using the former (37) be better than with the latter (39).

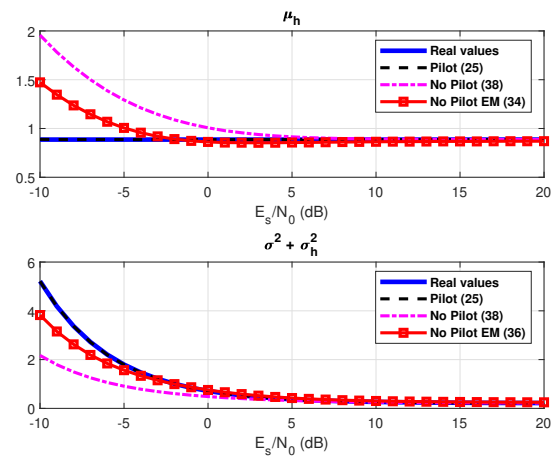


Fig. 4: Fading distribution first (top) and second (bottom) order moments estimation, using the Bayesian LLR approximation methods with pilot and data signal components.

5 System Model for the M-ary CSK and LLR with Perfect CSI

The code shift keying (CSK) modulation [30] is a M -ary orthogonal modulation which was first proposed as a GNSS signal candidate in [15]. Each symbol CSK x_ℓ corresponds to a different circular shift of a unique PRN sequence c . Let $S_\ell = \{\ell, 1 \leq \ell \leq 2^Q = M\}$ be the set of data symbols, with Q the number of bits to be transmitted, then the PRN sequence c_ℓ associated to the symbol $x_\ell, \ell \in S_\ell$, satisfies the following rule:

$$c_\ell(i) = c(\text{mod}(i - m_\ell, L)), \forall \ell \in [1, 2^Q], \forall i \in [1, L], \quad (40)$$

where i represents the PRN chip, m_ℓ is the integer number corresponding to the ℓ -th symbol shift, L is the number of chips in the PRN sequence and $\text{mod}(x, y)$ is the modulus operation. As an example, in Fig. 5, it is illustrated the PRN sequences associated to the 4-ary CSK modulation with a number of chips $L = 10230$.

At the transmitter the information bits are usually encoded by an error correction code, generating a codeword of length N bits. Then, the N codeword bits are grouped in N/Q CSK symbols of Q bits. Finally, the M -ary CSK associates each CSK symbol x_ℓ to a PRN sequence c_ℓ by right shifting the fundamental PRN sequence c .

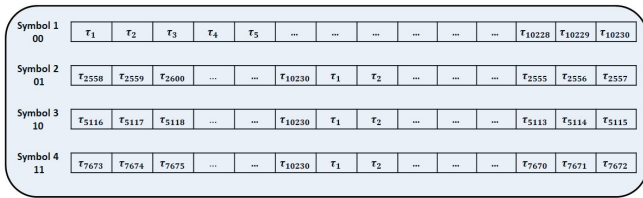


Fig. 5: CSK Symbol Waveform Example

5.1 CSK Data Demodulation in Open Sky Environments

As presented in Section 1, an open sky environment can be modelled by an AWGN channel. Then, assuming an AWGN channel and perfect time and frequency synchronisation, the n -th received sequence y_n corresponding to the transmitted CSK PRN sequence $c_{n,\ell}$ associated to the vector sequence $[b_1, b_2, \dots, b_Q]$ and the CSK symbol x_ℓ can be expressed as:

$$y_{n,i} = c_{n,i,\ell} + n_{n,i}, \quad (41)$$

where i represent the PRN chip, $n_{n,i} \sim \mathcal{N}(0, \sigma^2)$ are zero-mean *i.i.d.* Gaussian noise samples with variance $\sigma^2 = N_0/2$. Note that coherent reception is a valid assumption since a GNSS receiver capable to demodulate the CSK signal is also tracking in parallel the pilot signal component, which may provide a precise phase estimation. Let us now define $x_j, 1 \leq j \leq 2^{Q-1}$ the transmitted symbol if $b_q = 1$, and $x_t, 1 \leq t \leq 2^{Q-1}$ the transmitted symbol if $b_q = 0$. Therefore, considering perfect synchronization and following the LLR derivation in [18], the LLR expression for the bit b_q is given by

$$\mathcal{L}_{n,b_q} = \log \left(\frac{\sum_{\forall j} \left(e^{\frac{1}{\sigma^2} \frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_j}} \prod_{z \neq q} P(b_{j,z}) \right)}{\sum_{\forall t} \left(e^{\frac{1}{\sigma^2} \frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_t}} \prod_{z \neq q} P(b_{t,z}) \right)} \right), \quad (42)$$

where $P(b_{j,z})$ denotes the probability of $b_{j,z}$ which is z -th bit of the transmitted symbol x_j and $P(b_{t,z})$ denotes the probability of $b_{t,z}$ which is the z -th bit of the transmitted symbol x_t . Moreover, the term $\frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_l}$ corresponds to the normalized correlation between the n -th transmitted and n -th received PRN sequences. When a Bit Interleaver Coded Modulation (BICM) scheme [18, 28] is implemented at the receiver and equiprobable transmission bits are assumed, the LLR can be simplified to,

$$\mathcal{L}_{n,b_q} = \log \left(\frac{\sum_{\forall j} \left(e^{\frac{1}{\sigma^2} \frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_j}} \right)}{\sum_{\forall t} \left(e^{\frac{1}{\sigma^2} \frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_t}} \right)} \right). \quad (43)$$

5.2 CSK Data Demodulation in Fading Environments with Perfect CSI

Assuming the uncorrelated fading channel defined in Section 2, the received sequence can be written as

$$y_{n,i} = h_n \cdot c_{n,i,\ell} + n_{n,i}, \quad (44)$$

where $n_{n,i} \sim \mathcal{N}(0, \sigma^2)$ are zero-mean *i.i.d.* Gaussian noise samples with variance $\sigma^2 = N_0/2$ and h_n is the fading gain, which is assumed to be invariant within the symbol, and is also defined as an *i.i.d.* random variable with associated pdf give $p(h)$ and $h \geq 0$. Again, coherent reception and perfect synchronization are assumed. The LLR expression for the bit b_q over the uncorrelated fading

channel can be derived from (42) as

$$\mathcal{L}_{n,b_q} = \log \left(\frac{\sum_{\forall j} \left(e^{\frac{1}{\sigma^2} \frac{1}{L} \sum_{i=1}^L y_{n,i} h_n c_{n,i,x_j}} \prod_{z \neq q} P(b_{j,z}) \right)}{\sum_{\forall t} \left(e^{\frac{1}{\sigma^2} \frac{1}{L} \sum_{i=1}^L y_{n,i} h_n c_{n,i,x_t}} \prod_{z \neq q} P(b_{t,z}) \right)} \right), \quad (45)$$

where h_n is considered known at the receiver. Note that the previous equation can be simplified considering a BICM scheme,

$$\mathcal{L}_{n,b_q} = \log \left(\frac{\sum_{\forall j} \left(e^{\frac{1}{\sigma^2} \frac{1}{L} \sum_{i=1}^L y_{n,i} h_n c_{n,i,x_j}} \right)}{\sum_{\forall t} \left(e^{\frac{1}{\sigma^2} \frac{1}{L} \sum_{i=1}^L y_{n,i} h_n c_{n,i,x_t}} \right)} \right), \quad (46)$$

where perfect CSI is assumed in order to compute the LLR.

6 A Bayesian Approach to CSK Demodulation in Fading Environments without CSI

In this sequel we derive the LLR values for a CSK modulation considering that no CSI is available at the receiver. In that perspective, we adapt the Bayesian method in Section 3.2 to compute a closed-form LLR expression. From the definition of the LLR in (3), we redefine the LLR expression for the bit b_q over the uncorrelated fading channel as

$$\mathcal{L}_{n,b_q} = \log \left(\frac{p(y_n | b_q = 1, h_n)}{p(y_n | b_q = 0, h_n)} \right), \quad (47)$$

where

$$p(y_n | b_q = 1, h_n) = \sum_{\forall j} p(x_j) \int_{-\infty}^{\infty} p(y_n | x_j, h) p(h) dh \quad (48)$$

$$p(y_n | b_q = 0, h_n) = \sum_{\forall t} p(x_t) \int_{-\infty}^{\infty} p(y_n | x_t, h) p(h) dh \quad (49)$$

We follow the approach in [17] and consider a conjugate prior distribution for h_n ,

$$h_n \sim \mathcal{N}(\mu_h, \sigma_h^2), \quad (50)$$

where as in (21) μ_h and σ_h^2 need to be adjusted according to the channel uncertainty. From (48) and (49) we are interested in

$$\int_{-\infty}^{\infty} p(y_n | x_j, h) p(h) dh \propto \int_{-\infty}^{\infty} \prod_{i=1}^L e^{-\frac{(y_{n,i} - h c_{n,i,x_j})^2}{2\sigma^2}} e^{-\frac{(h - \mu_h)^2}{2\sigma_h^2}} dh \quad (51)$$

$$\int_{-\infty}^{\infty} p(y_n | x_t, h) p(h) dh \propto \int_{-\infty}^{\infty} \prod_{i=1}^L e^{-\frac{(y_{n,i} - h c_{n,i,x_t})^2}{2\sigma^2}} e^{-\frac{(h - \mu_h)^2}{2\sigma_h^2}} dh \quad (52)$$

In order to compute the LLR expression, we only need to compute those terms which depend on c_{n,i,x_j} and c_{n,i,x_t} (refer to Appendix 9),

$$\kappa_{n,1} = -\frac{2\mu_h \frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_j} + \frac{\sigma_h^2}{\sigma^2} \left(\frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_j} \right)^2}{2(\sigma^2 + \sigma_h^2)}, \quad (53)$$

$$\kappa_{n,2} = -\frac{2\mu_h \frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_t} + \frac{\sigma_h^2}{\sigma^2} \left(\frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_t} \right)^2}{2(\sigma^2 + \sigma_h^2)}. \quad (54)$$

Finally, (47) is given by

$$\mathcal{L}_{n,b_q} = \log \left(\frac{\sum_{\forall j} e^{-\kappa_{n,1}} P(x_j)}{\sum_{\forall t} e^{-\kappa_{n,2}} P(x_t)} \right). \quad (55)$$

Considering that the q -th bit of the symbol x_j is always equal to 1 and the q -th bit of the symbol x_t is always equal to 0, then (55) is

$$\mathcal{L}_{n,b_q} = \log \left(\frac{P(b_q = 1) \sum_{\forall j} e^{-\kappa_{n,1}} \prod_{z \neq q} P(b_{j,z})}{P(b_q = 0) \sum_{\forall t} e^{-\kappa_{n,2}} \prod_{z \neq q} P(b_{t,z})} \right). \quad (56)$$

When considering that a BICM scheme is implemented at the receiver, the previous expression can be written as,

$$\mathcal{L}_{n,b_q} = \log \left(\frac{\sum_{\forall j} e^{-\frac{2\mu_h \frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_j} + \frac{\sigma_h^2}{\sigma^2} \left(\frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_j} \right)^2}{2(\sigma^2 + \sigma_h^2)}}}{\sum_{\forall t} e^{-\frac{2\mu_h \frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_t} + \frac{\sigma_h^2}{\sigma^2} \left(\frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_t} \right)^2}{2(\sigma^2 + \sigma_h^2)}}} \right). \quad (57)$$

Notice that to compute (57) several exponential operations are required. A useful metric to reduce the computational complexity is the log-sum approximation [31], and the LLR approximation is

$$\mathcal{L}_{n,b_q} = \log \left(\frac{\max_{x_j} e^{-\frac{2\mu_h \frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_j} + \frac{\sigma_h^2}{\sigma^2} \left(\frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_j} \right)^2}{2(\sigma^2 + \sigma_h^2)}}}{\max_{x_t} e^{-\frac{2\mu_h \frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_t} + \frac{\sigma_h^2}{\sigma^2} \left(\frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_t} \right)^2}{2(\sigma^2 + \sigma_h^2)}}} \right). \quad (58)$$

The previous expression avoids to use \log/exp function evaluations, reducing the computational burden at the receiver. Notice that the first μ_h and second σ_h^2 order moments of $p(h)$ are assumed to be known, i.e. partial statistical CSI is assumed. However, these parameters might not be available at the receiver and therefore they must be estimated online. Assuming a binary learning sequence (e.g., symbols of a pilot component), we can infer μ_h and σ_h^2 as in (25). Moreover, this LLR approximation considers σ^2 known at the receiver, as typically done in the literature [17, 27], a result which also holds true in those scenarios where σ^2 was precisely estimated before the fading effect. In any case, if one wants to avoid the knowledge of σ^2 and provide a LLR without CSI close-form expression, (58) can be replaced by

$$\mathcal{L}_{n,b_q} = \log \left(\frac{\max_{x_j} e^{-\frac{2\hat{\mu}_h \frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_j} + \beta \cdot \left(\frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_j} \right)^2}{2 \left(\frac{1}{K} \sum_{m=1}^K (y_k x_k - \hat{\mu}_h)^2 \right)}}}{\max_{x_t} e^{-\frac{2\hat{\mu}_h \frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_t} + \beta \cdot \left(\frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_t} \right)^2}{2 \left(\frac{1}{K} \sum_{m=1}^K (y_k x_k - \hat{\mu}_h)^2 \right)}}} \right). \quad (59)$$

where $\hat{\mu}_h$ is estimated as in (25), K is the number of pilot symbols used to estimate the channel fading parameters and β is a coefficient which weighs the second term in (53)-(54). Notice that $\beta = 0$ involves neglecting the second term in (53)-(54) and provides a LLR approximation based on the metric obtained in (24), where the observed symbol is the output of the matched filter of the CSK demodulator $\frac{1}{L} \sum_{i=1}^L y_{n,i} c_{n,i,x_l}$. On the other hand, high values of β can mask the information provided by the first term in (53)-(54), generating an inaccurate LLR approximation. Based on simulations, we have found that $\beta = 1$ is the optimal value, yielding (59) finally to

$$\mathcal{L}_{b_q} = \alpha \left[\max_{x_j} \left(\sum_{i=1}^L y_{n,i} c_{i,x_j} + \frac{1}{2L} \left(\sum_{i=1}^L y_{n,i} c_{i,x_j} \right)^2 \right) \right] - \alpha \left[\max_{x_t} \left(\sum_{i=1}^L y_{n,i} c_{i,x_t} + \frac{1}{2L} \left(\sum_{i=1}^L y_{n,i} c_{i,x_t} \right)^2 \right) \right], \quad (60)$$

where $\alpha = \frac{\hat{\mu}_h}{\frac{1}{K} \sum_{m=1}^K (y_k x_k - \hat{\mu}_h)^2}$.

6.1 Performance of the LLR Approximation Methods for the CSK Modulation

As done in Section 3.3, we compute the maximum achievable rate of the Bayesian LLR approximation method (60) for different CSK modulation orders $Q = \{2, 4, 6\}$, considering a BICM CSK demodulator, and for both uncorrelated normalized Rayleigh fading channel and 2-state Prieto channel. These results are summarized in Figure 6. The new Bayesian CSK demodulation is compared to the perfect CSI cases given by (43) and (46). Both fading scenarios are also compared to the LLR approximation case (58) where partial statistical CSI is available at the receiver (i.e., where μ_h , σ_h^2 and σ^2 are assumed to be known).

• Normalized Rayleigh Fading Channel

As previously seen in Section 3.3, two different effects can cause a channel capacity loss. In Figure 6 (top) we can first see the impact of the fading channel, which induces for an ideal coding scheme of rate $R = 1/2$ a loss of 1.2 dB, 1.3 dB and 1.4 dB for modulation orders of $Q = 2$, $Q = 4$ and $Q = 6$, respectively, w.r.t. the AWGN case. Moreover, an additional 0.5 dB are lost due to channel uncertainty. Note that the channel capacity loss can be reduced by using coding scheme of lower rates. The latter is highly recommended for modulation orders greater than $Q = 2$, because transmitting more bits in one symbol increases the demodulation threshold.

• 2-State Prieto Channel

Results for the maximum achievable rate considering a 2-state Prieto channel model, for a vehicle speed of 50 km/h and a satellite elevation angle of 40 degrees, are shown in Figure 6 (bottom). As an example, we consider the data component of the GPS L1C signal which is characterized by a symbol rate of 100 symbols/s and a PRN code of length 10230 chips. W.r.t. the Rayleigh fading channel, we can see a channel capacity loss which is further degraded. For an ideal channel coding scheme of rate $R = 1/2$, a loss of around 6 dB is found for the CSK modulation scheme of order $Q = 2$, $Q = 4$ and $Q = 6$. Moreover, an additional 1 dB is lost due to the channel uncertainty. Note that in this limit fading channel scenario, high-order CSK modulations are not recommended without low rate channel coding schemes. Furthermore, due to the bad channel quality, specific channel coding structure such as rate compatible Root-LDPC codes [9], which allows to retrieve the entire diversity of the channel, are also highly recommended.

7 Results: FER Performance for the CED with the GPS L1C signal and LDPC Codes

In this section we compare soft decoding performance corresponding to the different LLR approximations introduced in the previous sections. Particularly, as an example, we provide the Frame Error Rate (FER) (i.e., w.r.t. the carrier-to-noise density (C/N_0)) performance for the Clock and Ephemerides Data (CED) considering the GPS L1C signal [32] with irregular LDPC codes, decoded with a sum-product algorithm [23].

7.1 Results for Antipodal GNSS Modulations

First, we consider a normalized Rayleigh fading channel and a channel coding scheme based on the standard irregular LDPC code of rate 1/2 used to encode the GPS L1C subframe 2 [32]. The FER results are summarized in Figure 7, where we show the performance of the LLR approximation in Section 3, (16) (BLA) and (26) (Bayesian), when 12 seconds of pilot symbols are retrieved. For comparison, we also show the FER results corresponding to the perfect CSI LLR (10) (Perfect CSI) and the full statistical CSI LLR (27) (Stats CSI). From the FER results we can see that both BLA and Bayesian approximation methods achieve a similar data demodulation performance w.r.t the statistical CSI case, which proves that

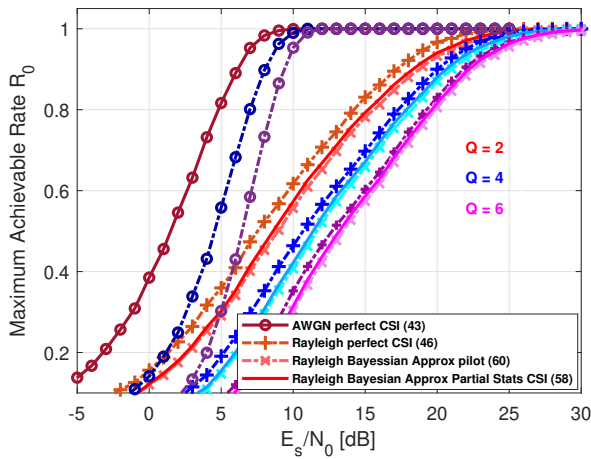
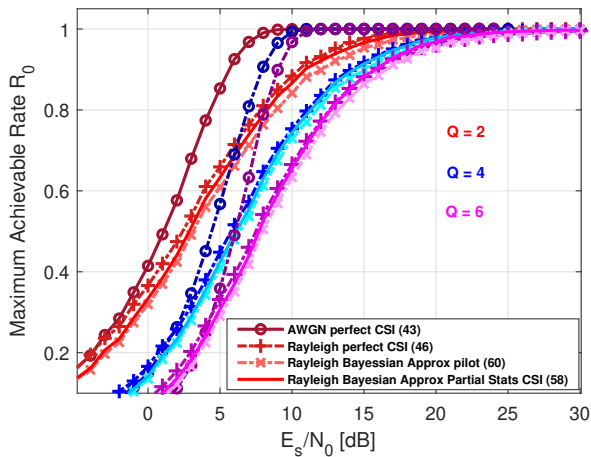


Fig. 6: Upper bounds of maximum achievable rate R_0 for a CSK modulation with $Q = 2, 4, 6$: (top) normalized Rayleigh channel, and (bottom) 2-state Prieto channel.

when no CSI is available only a good estimation of the first and second order moments of the fading distribution are required. Moreover, when the channel transmission is characterized by an AWGN channel, the LLR approximation in (26) converges to the perfect CSI LLR solution (7). Figure 8 shows the same comparison but for a data signal component. For comparison, we show the FER results corresponding to the LLR approximation with a pilot component (i.e., with a learning sequence and 12 seconds of pilot symbols retrieved) (26). Considering that no pilot sequence is available, we show the FER performance for the approximations introduced in Section 4, (37) (No Pilot EM) and (39) (No Pilot Approx). Notice first that the method which estimates the first and second order moments of the fading distribution through the EM algorithm provides a FER very similar to the one obtained with the pilot case (26), with only a 0.1 dB performances loss. On the other hand, the simpler method which estimate the corresponding first and second order moments of the fading distribution based on the absolute value of the received symbol reduces the data demodulation performance around 0.4 dB. Note from Figure 4 that LLR values from the methods in Section 4 were expected to provide lower error correction performance than the method obtained with the pilot case (26).

As a second scenario, we consider a 2-state Prieto channel model for a vehicle speed of 50 km/h and a satellite elevation angle of 40 degrees. We show the FER results again for the GPS LIC data within subframe 2 but in this case with a regular rate compatible Root LDPC codes of rate $R = 1/3$ (proposed in [9] to reduce the demodulation threshold) channel coding scheme. The comparison of the different methods with and without a pilot signal is illustrated in Figure 9. Notice that the EM-based method provides a FER very similar to

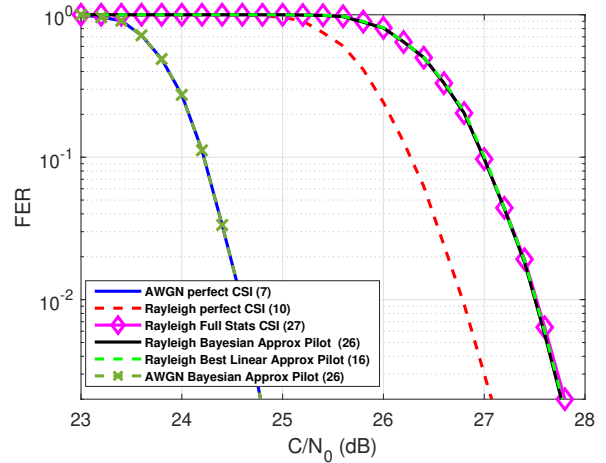


Fig. 7: FER of standard GPS LIC CED over a normalized Rayleigh channel for antipodal modulations and a pilot signal component.

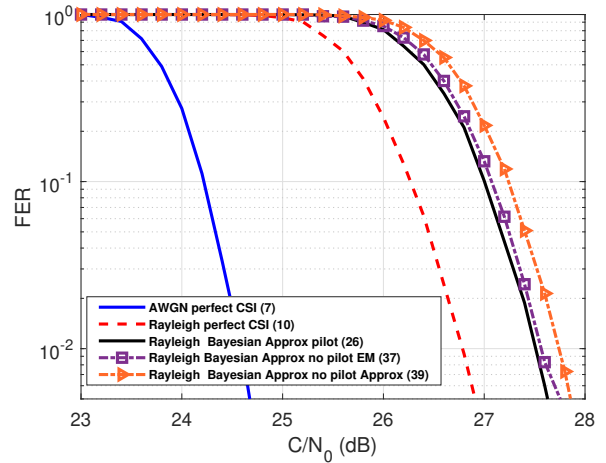


Fig. 8: FER of standard GPS LIC CED over a normalized Rayleigh channel for antipodal modulations and a data signal component.

the one obtained with the pilot case (26). In addition, a very small performance loss is observed w.r.t. the perfect CSI case. On the other hand, the simpler approximation (39) reduces the data demodulation performance around 1 dB w.r.t. the EM-based solution.

7.2 Results for M -ary CSK Modulations

As in the previous subsection we consider first a normalized Rayleigh fading channel, with a channel coding scheme based on the standard irregular LDPC used to encode the GPS LIC subframe 2. The FER results for the CSK modulation with $Q = \{2, 4, 6, 8, 10\}$ are summarized in Figure 10, where we show the performance of the LLR approximations in Section 6: i) with partial statistical CSI (58), and ii) when 12 seconds of pilot symbols are retrieved (60). For comparison, we also show the FER results corresponding to the perfect CSI LLR (46) and the perfect CSI LLR values under AWGN channel (43). From the FER results we can see that both partial statistics CSI (58) or no CSI (60) achieve a similar data demodulation performance, independently of the CSK modulation order. On the other hand, we can appreciate a FER performances loss in the order of 0.6-0.8 dB due to the channel uncertainty. The channel uncertainty impact seems to increase along with the CSK modulation order. Finally, we can see a significant FER performance loss due to the fading effect, i.e., in the order of 2-4 dB. Again, this loss of performances is related with the modulation order. From these

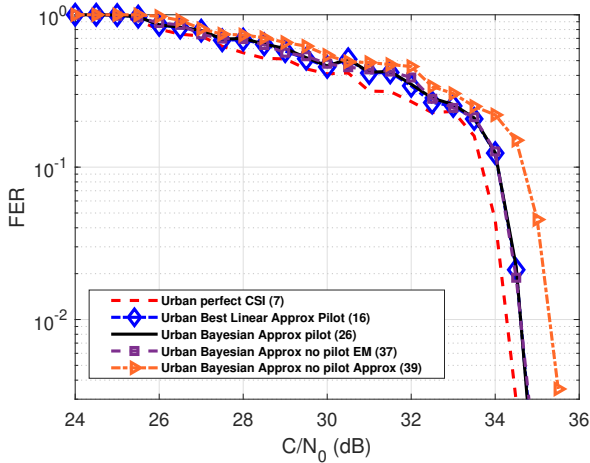


Fig. 9: FER GPS L1C CED with a regular rate compatible Root-LDPC code of rate $R = 1/3$ over a 2-state Prieto channel, for antipodal modulations.

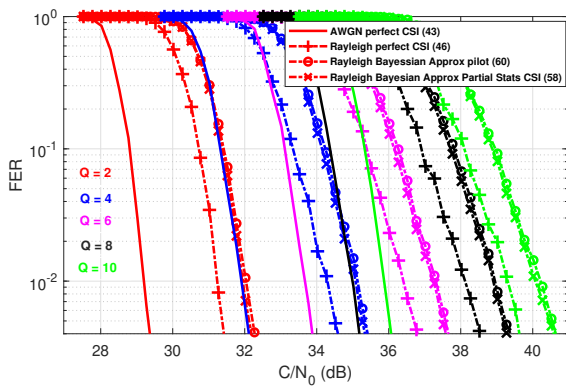


Fig. 10: FER of the CED for the GPS L1C data component over a normalized Rayleigh channel, for the CSK modulation.

results we can suggest to use channel coding schemes of lower rate for higher-order CSK modulations.

Results for a second scenario, considering a 2-state Prieto channel model for a vehicle speed of 50 km/h and a satellite elevation angle of 40 degrees, and a channel coding scheme based on rate compatible Root LDPC codes of rate $R = 1/4$, are summarized in Figure 11. We show FER performance results for CSK modulations with $Q = \{2, 4, 6\}^*$ and the different approximations in Section 6: i) partial statistical CSI (58), and ii) when 12 seconds of pilot symbols are retrieved (60). For comparison, we also show the FER results corresponding to the perfect CSI LLR case (46). From the FER results we can see that both (58) and (60) achieve almost a similar data demodulation performance independently of the CSK modulation order, as for the previous normalized Rayleigh fading channel scenario. On the other hand, we can appreciate a FER performance loss in the order of 1.5-2 dB due to the channel uncertainty. Again, these results suggest to consider low rate error correcting schemes in order to reduce the demodulation threshold.

8 Conclusion

This article addressed the problem of GNSS data demodulation over fading environments, for which no practical solutions existed in the

*Notice that results for higher order CSK modulations have not been included due to the lack of practical use (i.e., C/N_0 around 50-55 dB-Hz).

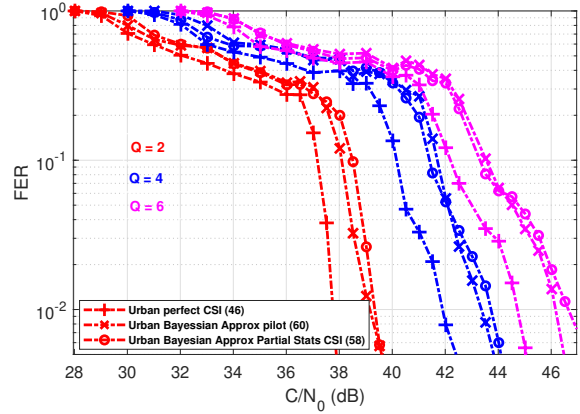


Fig. 11: FER of the CED for the GPS L1C data component encoded with a Root Code of Rate $R = 1/4$ over the 2-state prieto channel, for the CSK modulation.

literature when no CSI is available at the receiver. In that perspective, we derived several closed-form LLR approximations for both state-of-the-art antipodal GNSS modulations and new GNSS candidate CSK modulations. Regarding the former, both pilot and data signals were considered. Since modern GNSS signals always have a pilot component only this case was considered for the CSK modulation.

If a pilot component is available at the receiver some channel parameters can be inferred. For antipodal modulations, two LLR approximation methods were introduced, the first one being the state-of-the-art for GNSS data demodulation without CSI but computationally too expensive. Therefore, a second low complexity Bayesian LLR approximation was also presented. Results over two uncorrelated fading channels showed that both methods converge to the full statistical CSI method thus being optimal. For the data signal case (i.e., no training sequence available) two Bayesian LLR approximations were derived, the first one using an EM-based algorithm to estimate the first and the second order moments of the fading distribution, and the second one being a low complexity alternative. Results over two uncorrelated fading channels showed that the EM-based method data demodulation performance is very close to the pilot signal case, which validated its good behaviour. The low complexity alternative was shown to perform well only at high SNR.

Regarding the CSK modulation, we derived a Bayesian LLR approximation without CSI and a pilot component being tracked. The maximum achievable rate over two uncorrelated fading channels was presented to provide the performance degradation w.r.t. the AWGN channel. Simulation results showed the good performance of the proposed method, being a promising data demodulation solution.

Acknowledgment

This research was supported by the TéSA Lab Postdoctoral Research Fellowship, the DGA/AID project 2019.65.0068.00.470.75.01, and the National Science Foundation under Awards CNS-1815349 and ECCS-1845833.

9 Derivation of $\int_{-\infty}^{\infty} P(Y|x_j, h)p(h)dh$ to Compute the Closed-form CSK LLR Approximation

In this Appendix we are interested in the computation of the integrals $\int_{-\infty}^{\infty} P(Y|x_j, h)p(h)dh$ in (51)-(52), therefore we want to obtain,

$$\int_{-\infty}^{\infty} \prod_{i=1}^N e^{-\frac{(y_{n,i} - hc_{n,i,x_j})^2}{2\sigma^2}} e^{-\frac{(h-\mu_h)^2}{2\sigma_h^2}} dh \propto \quad (61)$$

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2} \frac{1}{N} \sum_{i=1}^N (y_{n,i}^2 - 2hy_{n,i}c_{n,i,x_j} + h^2c_{n,i,x_j}^2)} e^{-\frac{(h-\mu_h)^2}{2\sigma_h^2}} dh =$$

$$\int_{-\infty}^{\infty} e^{-\beta_1 \frac{1}{N} \sum_{i=1}^N (y_{n,i}^2 - 2hy_{n,i}c_{n,i,x_j} + h^2)} e^{-\beta_2 (h-\mu_h)^2} dh$$

where $\beta_1 = \frac{\sigma_h^2}{2\sigma^2\sigma_h^2}$ and $\beta_2 = \frac{\sigma^2}{2\sigma^2\sigma_h^2}$. Since the product of two Gaussian distributions is in turn a Gaussian distribution we proceed by finding the resulting mean (μ_a) and variance (σ_a^2) as,

$$\frac{(h - \mu_a)^2}{\sigma_a^2} + \kappa_n = \frac{h^2}{\sigma_a^2} + \frac{2h\mu_a}{\sigma_a^2} + \frac{\mu_a^2}{\sigma_a^2} + \kappa_n = \frac{\beta_1}{N} \sum_{i=1}^N y_{n,i}^2 \quad (62)$$

$$+ \beta_2 \mu_h^2 - 2h \left(\frac{\beta_1}{N} \sum_{i=1}^N y_{n,i}c_{n,i,x_j} + \beta_2 \mu_h \right) + h^2 (\beta_1 + \beta_2)$$

where κ_n is an auxiliary constant, and after identifying terms on both sides of (62),

$$\frac{1}{\sigma_a^2} = (\beta_1 + \beta_2), \quad (63)$$

$$\frac{\mu_a}{\sigma_a^2} = \left(\frac{\beta_1}{N} \sum_{i=1}^N y_{n,i}c_{n,i,x_j} + \beta_2 \mu_h \right), \quad (64)$$

$$\frac{\mu_a^2}{\sigma_a^2} = \frac{\left(\frac{\beta_1}{N} \sum_{i=1}^N y_{n,i}c_{n,i,x_j} + \beta_2 \mu_h \right)^2}{(\beta_1 + \beta_2)}. \quad (65)$$

Then, the constant κ_n can be computed as,

$$\kappa_n = \frac{\beta_1}{N} \sum_{i=1}^N y_{n,i}^2 + \beta_2 \mu_h^2 - \frac{\left(\frac{\beta_1}{N} \sum_{i=1}^N y_{n,i}c_{n,i,x_j} + \beta_2 \mu_h \right)^2}{(\beta_1 + \beta_2)}$$

$$= \frac{\beta_1 \beta_2 \left(\frac{1}{N} \sum_{i=1}^N y_i^2 + \mu_h^2 - 2\mu_h \frac{1}{N} \sum_{i=1}^N y_{n,i}c_{n,i,x_j} \right)}{\beta_1 + \beta_2}$$

$$+ \frac{\beta_1^2 \left(\frac{1}{N} \sum_{i=1}^N y_i^2 - \left(\frac{1}{N^2} \sum_{i=1}^N y_{n,i}c_{n,i,x_j} \right)^2 \right)}{\beta_1 + \beta_2} \quad (66)$$

where $\frac{\beta_1 \beta_2}{\beta_1 + \beta_2} = \frac{1}{2(\sigma^2 + \sigma_h^2)}$ and $\frac{\beta_1^2}{\beta_1 + \beta_2} = \frac{\sigma_h^2}{2\sigma^2(\sigma^2 + \sigma_h^2)}$. Finally, (61) yields to

$$\int_{-\infty}^{\infty} e^{-\frac{(h-\mu_a)^2}{\sigma_a^2}} e^{-\kappa_n} dh = e^{-\kappa_n} \quad (67)$$

where the pdf definition is applied [17]. Note that in order to compute the LLR expression, we are only interested in those terms which depend on c_{n,i,x_j} . Then, we define the constant $\kappa_{n,1}$ as those values of κ_n which depends on c_{n,i,x_j} ,

$$\kappa_{n,1} = - \frac{2\mu_h \frac{1}{N} \sum_{i=1}^N y_{n,i}c_{n,i,x_j} + \frac{\sigma_h^2}{\sigma^2} \left(\frac{1}{N} \sum_{i=1}^N y_{n,i}c_{n,i,x_j} \right)^2}{2(\sigma^2 + \sigma_h^2)}, \quad (68)$$

and the corresponding $\kappa_{n,2}$ for c_{n,i,x_t} (see (53)-(54)).

10 References

- P. J. G. Teunissen and O. Montenbruck, editors. *Handbook of Global Navigation Satellite Systems*. Springer, Switzerland, 2017.
- Daniel Medina, Lorenzo Ortega, Jordi Vilà-Valls, Pau Closas, Francois Vincent, and Eric Chaumette. Compact crb for delay, doppler, and phase estimation – application to gnss spp and rtk performance characterisation. *IET Radar, Sonar & Navigation*, 14(10):1537–1549, 2020.
- M. G. Amin, P. Closas, A. Broumandan, and J. L. Volakis. Vulnerabilities, Threats, and Authentication in Satellite-Based Navigation Systems [scanning the issue]. *Proceedings of the IEEE*, 104(6):1169–1173, 2016.
- D. Medina, H. Li, J. Vilà-Valls, and Pau Closas. Robust Statistics for GNSS Positioning under Harsh Conditions: a Useful Tool? *Sensors*, 19(24), Dec. 2019.
- D. Dardari, P. Closas, and P. Djuric. Indoor tracking: Theory, Methods, and Technologies. *IEEE Trans. on Vehicular Tech.*, 64(4):1263–1278, April 2015.
- L. Liu, B. G. Cai, and J. Wang. Cooperative Localization of Connected Vehicles: Integrating GNSS With DSRC Using a Robust Cubature Kalman Filter. *IEEE Trans. on Intel. Transp. Syst.*, 18(8):2111–2125, Aug. 2017.
- J. Wang, J. Liu, and N. Kato. Networking and Communications in Autonomous Driving: A Survey. *IEEE Commun. Surveys & Tutorials*, 21(2):1243–1274, 2nd Quart. 2019.
- Z. Kassas, M. Maaref, J. Morales, J. Khalife, , and K. Shamaei. Robust Vehicular Navigation and Map-matching in Urban Environments with IMU, GNSS, and Cellular Signals. *IEEE Intelligent Transportation Magazine*, accepted for publication.
- L. Ortega, C. Poulliat, M.L. Boucheret, M.Aubault-Roudier, and H. Al-Bitar. Optimal Channel Coding Structures for Fast Acquisition Signals in Harsh Environment Conditions. ION GNSS+, Miami, Florida, USA, 2019.
- L. Ortega, C. Poulliat, M.L. Boucheret, M.Aubault-Roudier, and H. Al-Bitar. Optimizing the co-design of message structure and channel coding to reduce the ttd for a galileo 2nd generation signal. *NAVIGATION*, 67(3):471–492, 2020.
- M. Roudier. *Analysis and Improvement of GNSS Navigation Message Demodulation Performance in Urban Environments*. Theses, INP Toulouse, January 2015.
- M. Roudier, T. Grelier, L. Ries, A. Garcia-Pena, O. Julien, C. Poulliat, M-L. Boucheret, and D. Kubrak. Optimizing GNSS Navigation Data Message Decoding in Urban Environment. In *IEEE/ION PLANS*, pages 581–588, 2014.
- W. Ryan and S. Lin. *Channel Codes: Classical and Modern*. Cambridge university press, 2009.
- J. Curran, M. Navarro, M. Anghileri, P. Closas, and S. Pfletschinger. Coding Aspects of Secure GNSS Receivers. *Proceedings of the IEEE*, 104(6):1271–1287, 2016.
- A. Garcia Peña, D. Salós, O. Julien, L. Ries, and T. Grelier. Analysis of the Use of CSK for Future GNSS Signals. In *ION GNSS 2013, 26th International Technical Meeting of The Satellite Division of the Institute of Navigation*, pages pp 1461–1479, Nashville, United States, September 2013. ION.
- R. Yazdani and M. Ardakani. Linear LLR Approximation for Iterative Decoding on Wireless Channels. *IEEE Trans. Communications*, 57(11):3278–3287, Nov 2009.
- L. Ortega, M. Aubault-Roudier, C. Poulliat, M. L. Boucheret, H. Al-Bitar, and P. Closas. LLR Approximation for Fading Channels using a Bayesian Approach. *IEEE Communications Letters*, 24(6):1244–1248, 2020.
- L. Ortega, M. Aubault-Roudier, C. Poulliat, M. L. Boucheret, and H. Al-Bitar. Binary Root Protograph LDPC Codes for CSK Modulation to Increase the Data Rate and Reduce the TTD. ION GNSS+, Miami, Florida, USA, 16/09/19-20/09/19, 2019. Institute of Navigation (ION).
- L. Ortega, C. Poulliat, M.L. Boucheret, M.Aubault-Roudier, H. Al-Bitar, and P. Closas. Data Decoding Analysis of Next Generation GNSS Signals. ION GNSS+, Miami, Florida, USA, 2019.
- C.M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, Berlin, Heidelberg, 2006.
- R. Prieto-Cerdeira, F. Perez-Fontan, P. Burzigotti, A. Bolea-Alamañac, and I. Sanchez-Lago. Versatile Two-state Land Mobile Satellite Channel Model with First Application to DVB-SH analysis. *International Journal of Satellite Communications and Networking*, 28(5-6):291–315, 2010.
- U.S Government, Interface Specification IS-GPS-8000 Navstar GPS Space/Segment/Navigation User Interface. Technical report, 2019.
- S.J. Johnson. *Iterative Error Correction: Turbo, Low-Density Parity-Check and Repeat-Accumulate Codes*. Cambridge University Press, 2009.
- J. Hagenauer. The EXIT chart - introduction to extrinsic information transfer in iterative processing. In *2004 12th European Signal Processing Conference*, pages 1541–1548, Sept 2004.
- E.K.P. Chong and S.H. Zak. *An Introduction to Optimization*. Wiley Series in Discrete Mathematics and Optimization. Wiley, 2013.
- J.M. Bernardo and A.F.M. Smith. *Bayesian theory*, volume 405. John Wiley & Sons, 2009.
- R. Asvadi, A.H. Banihashemi, M. Ahmadian-Attari, and Hamid H. Saeedi. LLR Approximation for Wireless Channels Based on Taylor Series and its Application to BICM with LDPC Codes. *IEEE Transactions on Communications*, 60(5):1226–1236, 2012.
- G. Caire, G. Taricco, and E. Biglieri. Bit interleaved coded modulation. *IEEE Transactions on Information Theory*, 44(3):927–946, 1998.
- J. Hagenauer. The exit chart - introduction to extrinsic information transfer in iterative processing. In *2004 12th European Signal Processing Conference*, pages 1541–1548, 2004.
- AY-C. Wong and V.C.M. Leung. Code phase shift keying: a power and bandwidth efficient spread spectrum signaling technique for wireless local area network applications. In *CCECE'97. Canadian Conference on Electrical and Computer Engineering. Engineering Innovation: Voyage of Discovery. Conference Proceedings*, volume 2, pages 478–481. IEEE, 1997.

- 31 F. Tosato and P. Bisaglia. Simplified Soft-output Demapper for Binary Interleaved COFDM with Application to HIPERLAN/2. In *IEEE International Conference on Communications*, volume 2, pages 664–668, 2002.
- 32 Interface Specification IS-GPS-800 NavStar GPS Space Segment/ UserSegment L1C Interface. Technical report.