



On the performance of digital adaptive spur cancellation for multi-standard radio frequency transceivers

Robin Gerzaguët, Laurent Ros, Fabrice Belvèze, Jean-Marc Brossier

► To cite this version:

Robin Gerzaguët, Laurent Ros, Fabrice Belvèze, Jean-Marc Brossier. On the performance of digital adaptive spur cancellation for multi-standard radio frequency transceivers. Digital Signal Processing, 2014, 33 (oct.), pp.83-97. 10.1016/j.dsp.2014.07.006 . hal-01057201

HAL Id: hal-01057201

<https://hal.science/hal-01057201>

Submitted on 21 Aug 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the Performance of Digital Adaptive Spur Cancellation for Multi-Standard Radio Frequency Transceivers[☆]

Robin Gerzaguét^{a,b,*}, Laurent Ros^a, Fabrice Belvèze^b, Jean-Marc Brossier^a

^a*GIPSA-Lab, Image and Signal Department, BP 46, 38402 Saint Martin d'Hères, France*

^b*ST-Microelectronics, Grenoble, France*

Abstract

This study deals with the asymptotic performance of a multiple-spur cancellation scheme. Radio frequency transceivers are now multi-standard and specific impairment can occur. The clock harmonics, called spurs, can leak into the signal band of the reception stage, and thus degrade the performance. The performance of a fully digital approach is presented here. A one-spur cancellation scheme is first described, for which we exploit the *a priori* knowledge of the spur frequency to create a reference of the polluting tone with the same frequency. A least-mean-square (LMS) algorithm block that uses this reference to mitigate the polluter is designed. However, due to imperfections in the physical components, there is a shift between the *a priori* frequency and the actual frequency of the spur, and the spur is affected by

[☆]Part of this work was presented at the french Symposium GRETSI (Groupe d'études du traitement du signal et des images) 2013

*Corresponding author. Tel.: +33 4 76 82 62 68

Email addresses: `robin.gerzaguete@gipsa-lab.fr` (Robin Gerzaguete), `laurent.ros@gipsa-lab.fr` (Laurent Ros), `fabrice.belveze@st.com` (Fabrice Belvèze), `jean-marc.brossier@gipsa-lab.fr` (Jean-Marc Brossier)

Brownian phase noise. Under these circumstances, we study the asymptotic and transient performance of the algorithm. We next improve the transient performance by adding a previously proposed adaptive-step-size process. In a second part of this paper, we present a multiple-spur parallel approach that is based on the one-spur cancellation scheme, for which we provide a closed-form expression of the asymptotic signal-plus-noise interference ratio in the presence of frequency shifts and phase noise.

Keywords: Multi-standard transceiver, radio frequency impairment, spurs, digital cancellation, least-mean-square (LMS) algorithm, adaptive-step-size algorithm.

1. Introduction

The development of multiple standards for wireless communication, from Global System for Mobile (GSM) to Long-Term Evolution (LTE), has been motivated by the increasing demand for mobility and new telecommunication services (e.g., data, audio, video) [1]. The coexistence of different wireless systems has resulted in multi-standard solutions for the mobile architecture loading, for instance, GSM, 3G, LTE, and WiFi devices, each of which comes with their specific impairments [2, 3, 4, 5].

Due to a multiplicity of the operating frequencies and the physical proximity between the components, harmonics from clocks, called spurs, can appear at detrimental frequencies. These spurs can have at least two harmful consequences. First, due to substrate coupling between components, they can

leak into the phase-locked loop. If they fall near the transmission frequency, they can demodulate the received signal into the receiver path, which leads to degradation of the receiver signal-to-noise ratio (SNR) [6]. The second case is when spurs couple directly into the receiver path, which leads to a mixture of desired noisy signal and additive spurs (see Figure 1).

Spurs can be mitigated using different passive and active techniques. From frequency planning [7] to clock spreading [8, 9], all of these passive methods aim at mitigating the level of the spur, to reduce the polluting interference that is synthesized.

Nevertheless, the multiplicity of standards that are supported has lead to the constant growth of spur sources (which is related to the growing number of operating frequencies), and passive cancellation has become more difficult. At the same time, radio frequency (RF) transceivers contain more and more digital parts, and signal processing techniques are becoming an area of interest for problematic RF impairment [10, 11, 12, 6]. Thus, active methods for spur cancellation that are based on notch filter techniques or signal processing algorithms have been developed over the last few years [13, 14, 15, 16, 17].

In this study, we focus on direct pollution from spurs. The observed noisy signal is polluted by a fixed number of spurs that need to be mitigated. The aim of this study is to propose a fully digital approach for spur attenuation and to derive its analytic performance. We first focus on a one-spur model, for easy comprehension, and then extend the problem and results to a multiple-spur compensation model.

Our method is based on (more or less accurate) knowledge of the spur frequency. The algorithms studied here are based on synthesized references of the spurs, which are tones with the same frequency. Due to imperfections in the physical components, the reference cannot be considered as a perfect replica of the polluting spur. We assume here that there is a shift between the *a priori* or expected frequency (assumed to be known or estimated) and the actual frequency of the spur. However, in real cases, the spur is not a pure tone, as it is modulated by phase noise (PN). This PN should follow a Brownian model, which is a classical model used in the literature for free-running oscillators [18, 19, 20, 21].

For the one-spur model, we introduce a mono-spur cancellation block. For this real-case model, we derive closed-form formulae of the signal-plus-noise-to-interference ratio (SNIR) at the output of the algorithm. We also give an equivalent scheme, which is a notch filter, and propose a frequency interpretation of the system design. Then, we improve the transient response of the algorithm by adding a previously proposed adaptive-step-size process to the least-mean-square (LMS) algorithm, which leads to a more rapid solution that has the same asymptotic performance.

In the second part of this paper, we consider a problematic P -spur mixture, and propose a multiple-spur cancellation scheme based on a parallel structure that is composed of P mono-spur cancellation blocks. In this model, we still consider imperfections, e.g., PN and frequency shift (FS), for all of the spurs and references, and compute the theoretical values of the asymp-

otic SNIR in the most general case. Our main contributions, in addition to being a comprehensive study, can be summarized below:

- Proposition and analytical analysis of a LMS-based algorithm for multiple-spur pollution with specific impairments (frequency shift and phase noise), with a frequency interpretation of the compensation system (equivalence to a first-order infinite impulse response (IIR) notch filter),
- Derivation of closed-form expressions usable to tune the step-size parameter of the LMS-based spur cancellation algorithms, as well as to predict the SIR performance with respect to the polluter properties (frequency shift values, phase noise variances, reference amplitudes). To the best of our knowledge, no such analytical results have been found to date.
- Proposition of an adaptive step overlay to the LMS algorithm to improve the transient response while keeping the same asymptotic performance. The improved algorithm has almost the same convergence speed as the recursive least squares (RLS) algorithm, but can better cope with a sudden change in the configuration of the spurs (if the parameters of the reference remain unchanged).

This paper is organized as follows. We give the one-spur model in Section 2 and present the cancellation scheme in Section 3, for which we express

asymptotic and transient studies. Then, we develop a multiple-spur cancellation scheme and provide closed-form formulae of the overall performance in Section 4. Section 5 validates our method and theoretical results through simulations.

2. One-spur model

In this section, we first consider the cancellation of a single spur. The discrete-time observation model (at sample time nT_s) is

$$d(n) = x(n) + b(n) + s(n) , \quad (1)$$

where $d(n)$ is the observed signal, $x(n)$ is the zero-mean data signal of variance σ_x^2 , $b(n)$ is circular complex additive white Gaussian noise with zero mean and variance σ_b^2 , and $s(n)$ is the additive spur. The spur is expressed as

$$s(n) = Ae^{j(\omega - \delta_\omega)n + j\phi(n) + j\phi_0} , \quad (2)$$

where A is the unknown amplitude, ϕ_0 is the unknown phase, $\phi(n)$ is the PN that affects the spur, and ω is the *a priori* spur (natural) frequency, which is normalized by the sample time. Nevertheless, due to imperfections in the oscillators, the spur is affected by an unknown (natural) frequency shift (FS) δ_ω , which corresponds to the difference between the normalized *a priori* frequency and the normalized effective frequency of the spur.

As mentioned in the introduction, we consider here that the PN is under

a Brownian model [19, 20, 21]. The evolution of the phase is

$$\phi_{n+1} = \phi_n + \xi(n) , \quad (3)$$

with $\phi(0) = 0$ and $\xi(n)$ is an additive white noise of variance σ_ξ^2 .

As the *a priori* frequency of the spur is known, a reference signal is synthesized as an image of the polluting spur. The evolution of this reference is written as

$$u(n) = B e^{j\omega n + j\phi_R} , \quad (4)$$

where B is the amplitude of the reference and ϕ_R is its phase. In a real case, the synthesized reference can also have a FS and PN. Nevertheless, as it is only the global FS (spur + reference) and the global PN that matter, we assume for simplicity that all of the imperfections are attributed to the spur. The aim of the cancellation is to use the synthesized reference $u(n)$ to cancel the polluting spur and to evaluate the influence of the PN and FS on the performance. The influence of the polluting spur, in terms of the signal-plus-noise-to-interference-ratio (SNIR) expressed in dB, is

$$\text{SNIR}_{\text{Mixture}} = -10 \log_{10} \left(\frac{A^2}{\sigma_x^2 + \sigma_b^2} \right) , \quad (5)$$

We denote $e(n)$ as the compensated signal after application of the algorithm,

and introduce the estimation error $\epsilon(n)$ as

$$\epsilon(n) = x(n) + b(n) - e(n) \quad . \quad (6)$$

The asymptotic variance of the estimation error σ_ϵ^2 is expressed as

$$\sigma_\epsilon^2 = \lim_{n \rightarrow \infty} E[|\epsilon(n)|^2] \quad . \quad (7)$$

The spur canceller is not intended to cancel the additive noise term. As we focus here on the performance of the spur canceller and not on the impact of the additive white noise, we consider for the performance measurement that the noise is part of the desired signal. In other words, regarding the spur cancellation algorithm, the desired signal is the observed signal without the spur. The performance of the algorithm will be described by the SNIR after compensation, defined by:

$$\text{SNIR}_{\text{comp}} = -10 \log_{10} \left(\frac{\sigma_\epsilon^2}{\sigma_x^2 + \sigma_b^2} \right) \quad . \quad (8)$$

For the theoretical results, several hypotheses are made:

- The desired signal $x(n)$ is a white complex zero-mean circular signal;
- The normalized FS $\delta_\omega \ll \omega$ (normalized frequency);
- The PN variance $\sigma_\xi^2 \ll 1$.

We will see in Section 5.1 that the last two assumptions are completely

satisfied in practical cases. The link between our model parameters (δ_ω and σ_ξ^2) and physical parameters of oscillator used in practice will also be explained later.

3. One-spur cancellation scheme and performance

3.1. Filter scheme

The compensated signal $e(n)$ is obtained by subtracting the estimated spur value $\hat{s}(n)$ from the input sample $d(n)$. This value is obtained by multiplying the reference signal $u(n)$ by a complex adaptive coefficient $w(n)$, which is updated using the output of the algorithm $e(n)$, as shown in Figure 2.

3.2. Ideal coefficient

From an analysis perspective, let us introduce the ideal coefficient for this cancellation scheme, which is defined by a complex coefficient $w_{\text{opt}}(n)$ that provides an output $\hat{s}(n)$ that is strictly equal to the polluting spur $s(n)$ for all n . Thus, $w_{\text{opt}}(n)$ verifies:

$$w_{\text{opt}}(n)u(n) = s(n) \quad \forall n \in \mathbb{N} . \quad (9)$$

The iterative expression of the ideal coefficient using (2) and (4) is:

$$w_{\text{opt}}(n) = \frac{A}{B} e^{j(\phi_0 - \phi_R)} e^{-j\delta_\omega n} e^{j\phi(n)} . \quad (10)$$

Due to the hypotheses made at the end of Section 2 ($\delta_\omega \ll \omega, \sigma_\xi^2 \ll 1$), the ideal coefficient can be expressed recursively as

$$w_{\text{opt}}(n+1) \approx w_{\text{opt}}(n) + w_{\text{opt}}(n)[j\xi(n) - j\delta_\omega] . \quad (11)$$

Using $w(n) = w_{\text{opt}}(n)$ specified in (10) or (11) in Figure 2 gives an ideally recovered signal, which means that in this case $e(n) = x(n) + b(n)$, $\forall n \in \mathbb{N}$. As most of the parameters (e.g., amplitudes, phases, FS, PN) are not known, the ideal coefficient cannot be used, and a recursive estimation of $w_{\text{opt}}(n)$ has to be performed.

3.3. The LMS algorithm

We propose to use the LMS algorithm [22, 23], which is based on the error signal $e(n) = d(n) - \hat{s}(n)$, to perform the updating of the coefficient $w(n)$:

$$\hat{s}(n) = w(n)u(n) \quad (12)$$

$$e(n) = d(n) - \hat{s}(n) \quad (13)$$

$$w(n+1) = w(n) + \mu u^*(n)e(n) \quad (14)$$

where μ is the constant step size of the algorithm and $()^*$ denotes complex conjugate. The LMS algorithm is a stochastic-gradient descent method that computes at each iteration n the gradient of the instantaneous square error as $1/2 \partial|e(n)|^2/\partial w = -u^*(n)e(n)$, in order to use the steepest descend to

update the coefficient $w(n)$ in (14). In Appendix A, we briefly review how the LMS algorithm approach can also be legitimized thanks to a weighted-least-squares error cost function to be minimized. In the case of a perfect situation (i.e., no PN and no FS), it can be shown that (10) is the solution of the Yule-Walker equation, to which the LMS algorithm converges [24].

3.4. Asymptotic performance

We now introduce the difference between the ideal coefficient and the iterative coefficient $w(n)$, called the misalignment $v(n)$:

$$v(n) = w(n) - w_{\text{opt}}(n) . \quad (15)$$

From (13), with the model defined in (1) and using the property of the ideal coefficient of (9), the output of the algorithm then becomes

$$e(n) = x(n) + b(n) - v(n)u(n) . \quad (16)$$

We see from (16) and (6), that $\epsilon(n) = v(n)u(n)$, and with (8), we can link the SNIR after compensation to the variance of the misalignment as

$$\text{SNIR}_{\text{comp}} = -10 \log_{10} \left(\frac{B^2 \sigma_v^2}{\sigma_x^2 + \sigma_b^2} \right) . \quad (17)$$

The variance of the misalignment can be expressed recursively using (11), (14), (16), and (17) (see Appendix B). We use the same hypotheses as mentioned at the end of Section 2, and since the PN follows a Brownian

model, only the scale of this model $\phi(n+1) - \phi(n)$ [which is the white noise $\xi(n)$] has any influence. Therefore,

$$\begin{aligned} E[|v(n+1)|^2] &= (1 - \mu B^2)^2 E[|v(n)|^2] + \mu^2 B^2 (\sigma_x^2 + \sigma_b^2) \\ &\quad + Q(\delta_\omega, \sigma_\xi^2) , \end{aligned} \quad (18)$$

where

$$Q(\delta_\omega, \sigma_\xi^2) = \frac{A^2}{B^2} (\sigma_\xi^2 + \delta_\omega^2) + \frac{2A^2(1 - \mu B^2)\delta_\omega^2(\mu B^2\delta_\omega^2 - \delta_\omega^2 + \mu B^2)}{B^2(\mu^2 B^4\delta_\omega^2 - 2\mu B^2\delta_\omega^2 + \mu^2 B^4)} \quad (19)$$

Equation (18) is composed of three terms: the first is the recursive term; the second is the contribution of the noisy signal; and the last term $Q(\delta_\omega, \sigma_\xi^2)$ represents the impact of both the PN and FS on the overall performance.

Asymptotically, we assume that $E[|v(n+1)|^2] = E[|v(n)|^2]$, and substituting (18) into (17) and using the definition of (19), we obtain the asymptotic SNIR after convergence as

$$\text{SNIR}_{\text{comp}} = -10 \log_{10} \left[\frac{\mu B^2}{2 - \mu B^2} + \frac{Q(\sigma_\xi^2, \delta_\omega)}{\mu(\sigma_x^2 + \sigma_b^2)(2 - \mu B^2)} \right] . \quad (20)$$

A maximal value of the step size μ , to ensure the convergence of the algorithm (see [24, Section 9-4]), is given by

$$\mu < \mu_{\text{max}} = \frac{2}{B^2} . \quad (21)$$

The asymptotic SNIR is composed of a static term (misalignment due to the presence of the input noise) and the pollution term (due to the PN and FS). It appears that the step size μ defined in (14) leads to a trade-off between the accuracy of the estimate and resistance against the contribution of the polluter.

We consider here that $\mu \ll 1$. The optimal step size μ_{opt} is then obtained by the maximisation of (20) (see Appendix C), and is described on

$$\begin{aligned} \mu_{\text{opt}} = & \left(\frac{A^2 \sqrt{\frac{108 (\sigma_x^2 + \sigma_b^2) \delta_\omega^4 - \sigma_\xi^2 A^2}{\sigma_x^2 + \sigma_b^2}}}{3^{\frac{3}{2}} (\sigma_x^2 + \sigma_b^2) B^6} + \frac{2 A^2 \delta_\omega^2}{(\sigma_x^2 + \sigma_b^2) B^6} \right)^{\frac{1}{3}} \\ & + \frac{\sigma_\xi^2 A^2}{3 (\sigma_x^2 + \sigma_b^2) B^4 \left(\frac{A^2 \sqrt{\frac{108 (\sigma_x^2 + \sigma_b^2) \delta_\omega^4 - \sigma_\xi^2 A^2}{\sigma_x^2 + \sigma_b^2}}}{3^{\frac{3}{2}} (\sigma_x^2 + \sigma_b^2) B^6} + \frac{2 A^2 \delta_\omega^2}{(\sigma_x^2 + \sigma_b^2) B^6} \right)^{\frac{1}{3}}} . \end{aligned} \quad (22)$$

3.5. Equivalent scheme in z domain

From (12), (13), (14), the update of the filter coefficient can be expressed as

$$w(n+1) = (1 - \mu B^2)w(n) + \mu d(n)u^*(n) . \quad (23)$$

This means that the algorithm can be considered as a low-pass filter (LPF) noise canceller, based on the filtering of $d(n)u^*(n)$. The equivalent scheme is presented in Figure 3.

The observation $d(n)$ is first frequency shifted downward (by $-\omega$) and

then filtered using a LPF. As a consequence, the translated observation $d(n)u^*(n)$ has its spur at the frequency $-\delta_\omega$ and is selected by the LPF. The result is then frequency shifted back upward by the reference frequency ω and then subtracted from the observed data.

From (23), the transfer function of the LPF in the z domain is

$$L(z) = \frac{W(z)}{U^*(z) * D(z)} = \frac{\mu}{z - (1 - \mu B^2)} \quad (24)$$

where $*$ denotes convolution in the z domain. This filter has a cut-off frequency of μB^2 . As a consequence, there is a direct link between the value of the LMS step size and the bandwidth of the LPF.

From this filter scheme, a global notch structure can be obtained as the result of the direct filtering of the observation $d(n)$ by a filter $H(z)$ (cf Figure 3). In the z domain, as $L(z) * U(z) = BL(ze^{-j\omega})$, the output of the global filter H is [24]

$$\begin{aligned} E(z) &= D(z) - W(z) * U(z) \\ &= D(z) - \{L(z)[D(z) * U^*(z)]\} * U(z) \\ &= D(z) - [BL(z)D(ze^{j\omega})] * U(z) . \end{aligned}$$

We denote

$$G(z) = BL(z)D(ze^{j\omega})$$

so that

$$\begin{aligned} E(z) &= D(z) - BG(ze^{-j\omega}) = D(z) - B^2L(ze^{-j\omega})D(z) \\ &= D(z) - [B^2L(z) * U(z)]D(z) . \end{aligned}$$

We can define $H(z)$ as the global notch filter response of the overall structure in the z domain, expressed as

$$H(z) = \frac{E(z)}{D(z)} = 1 - B^2L(z) * U(z) . \quad (25)$$

The filter thus obtained is a band stop filter, as $L(z) * U(z)$ is a LPF that is up-shifted in frequency by ω . Furthermore, from (25), the z transform of the filter becomes

$$\begin{aligned} H(z) &= 1 - \frac{\mu B^2}{ze^{j\omega} - (1 - \mu B^2)} \\ &= \frac{1 - z^{-1}e^{j\omega}}{1 - (1 - \mu B^2)e^{j\omega}z^{-1}} . \end{aligned} \quad (26)$$

It is of note that the equivalent notch filter $H(z)$ coincides with the first-order IIR notch filter implemented in [14].

3.6. Special cases and approximate performance formulae

In this subsection, we consider two special cases (frequency shift only and phase noise only) to obtain simplified and more interpretable asymptotic performance formulae.

3.6.1. Frequency shift only

We consider here that there is no PN, i.e., $\sigma_\xi^2 = 0$. The pollution term from (19) thus becomes

$$Q^{\text{FS}}(\delta_\omega) = \frac{2A^2(1 - \mu B^2)\delta_\omega^2(\mu B^2\delta_\omega^2 - \delta_\omega^2 + \mu B^2)}{B^2(\mu^2 B^4\delta_\omega^2 - 2\mu B^2\delta_\omega^2 + \mu^2 B^4)} . \quad (27)$$

We assume that the FS δ_ω is smaller than the step μ . The asymptotic SNIR with only FS can therefore be approximated as:

$$\text{SNIR}^{\text{FS}} \simeq -10 \log_{10} \left[\frac{\mu B^2}{2} + \frac{A^2\delta_\omega^2}{(\sigma_x^2 + \sigma_b^2)\mu^2 B^4} \right] . \quad (28)$$

It should be noted that due to the simplifications of (19) and (20), (28) can be obtained equivalently with a frequency-domain representation based on the LPF $L(z)$ and a power spectral density computation (see Appendix D). The hypothesis $\delta_\omega < \mu$ consists of having a LPF cut-off frequency that is greater than the FS. Indeed, the input of the LPF $d(n)u^*(n)$ has a spur at frequency $-\delta_\omega$, which can only be mitigated if the LPF has a cut-off frequency that is greater than the FS.

From the maximisation of (28), a simplified expression of the optimal step size in this case can be obtained as

$$\mu_{\text{opt}}^{\text{FS}} = \sqrt[3]{\frac{4A^2\delta_\omega^2}{B^6(\sigma_x^2 + \sigma_b^2)}} . \quad (29)$$

From (28) and (29), the maximal value of the SNIR can be analytically

expressed as

$$\text{SNIR}_{\text{Max}}^{\text{FS}} = -\frac{1}{3}10 \log_{10} \left[\frac{27A^2\delta_\omega^2}{16(\sigma_x^2 + \sigma_b^2)} \right] . \quad (30)$$

Equation (30) means that the SNIR performance will decrease by 6.66 dB per decade of the increase in the FS. This gives a direct link between the expected performance and the acceptable FS.

3.6.2. Phase noise only

We now consider that the reference frequency perfectly matches the spur frequency, i.e., $\delta_\omega = 0$. With only PN, the pollution term from (19) is

$$Q(\delta_\omega, \sigma_\xi^2) = \frac{A^2}{B^2} \sigma_\xi^2 . \quad (31)$$

The associated SNIR is

$$\text{SNIR}^{\text{PN}} = -10 \log_{10} \left[\frac{\mu B^2}{2 - \mu B^2} + \frac{A^2 \sigma_\xi^2}{(\sigma_x^2 + \sigma_b^2) \mu B^2 (2 - \mu B^2)} \right] . \quad (32)$$

In the frequency domain, the PN will widen the spur spectrum, so if the cut-off frequency of the LPF $L(z)$ is too small, the LPF will not mitigate all of the spur band, which introduces performance loss .

As in the case of FS only, a simplified optimal step size can be obtained from (32) in the case of PN only:

$$\mu_{\text{opt}}^{\text{PN}} = \frac{1}{B^2} \sqrt{\sigma_\xi^2 \frac{A^2}{\sigma_x^2}} \quad (33)$$

and from (33), the maximal SNIR for a given PN can be analytically expressed as

$$\text{SNIR}_{\max}^{\text{PN}} = 10 \log_{10} \left(\sqrt{\frac{\sigma_x^2 + \sigma_b^2}{\sigma_\xi^2 A^2}} - \frac{1}{2} \right). \quad (34)$$

3.7. Transient mode and adaptive step process

Equations (19) - (28) describe the performance in the asymptotic mode, after convergence of the system. In the transient mode, where the filter $w(n)$ converges to the ideal coefficient, the average SNIR is a function of the sample index n , and is denoted as

$$\text{SNIR}(n) = -10 \log_{10} \left[\frac{B^2 E[|v(n)|^2]}{\sigma_x^2 + \sigma_b^2} \right]. \quad (35)$$

In the asymptotic mode, where $n \rightarrow \infty$, the transient SNIR defined in (35) converges to the asymptotic SNIR defined in (17). From (18), and assuming $w(1) = 0$, the average SNIR can be obtained recursively, which leads to the closed-form formula

$$\begin{aligned} \text{SNIR}(n) = -10 \log_{10} \left\{ \frac{B^2}{\sigma_x^2 + \sigma_b^2} \left[(1 - \mu B^2)^{2n} \frac{A^2}{B^2} \right. \right. \\ \left. \left. + [\mu^2 B^2 (\sigma_x^2 + \sigma_b^2) + Q(\delta_\omega, \sigma_\xi^2)] \frac{1 - (1 - \mu B^2)^{2n}}{\mu B^2 (2 - \mu B^2)} \right] \right\}. \end{aligned} \quad (36)$$

First, it is important to note that $\lim_{n \rightarrow \infty} \text{SNIR}(n) = \text{SNIR}_{\text{comp}}$, as defined in (20). At the same time, (36) demonstrates that the asymptotic convergence speed is inversely proportional to the step size μ . To have a faster algorithm in the transient mode, we propose to add an adaptive-step-

size process to the LMS algorithm. Thus, the step size will have a large value if the convergence is not settled, and will have to converge to a deterministic value. The purpose is to have the same asymptotic performance as theoretically expressed in (20), but also to limit the transient mode of the algorithm.

The step-size update is based on the instantaneous square error $|e(n)|^2$, to which a gradient descent is applied. We use the step-size update proposed in [25], which is based on a geometric update [26] with a forgetting factor $\gamma < 1$ [27]:

$$\mu_{n+1} = \{\mu_n [\gamma + \eta |u^*(n)e(n)G(n)|]\}_{\mu_{\min}}^{\mu_{\max}} \quad (37)$$

where $G(n) = \partial w(n)/\partial \mu$ can be expressed recursively as

$$G(n) = (1 - \mu_n B^2)G(n) + u^*(n)e(n) \quad (38)$$

η is a constant factor, and $\{\}_{\mu_{\min}}^{\mu_{\max}}$ means that the step size is bounded between a maximal and a minimal value. The maximal value is set for convergence reasons from (21), and the minimal value is set to have the desired asymptotic performance [for instance, by setting the value of $\mu_{\min}$ to the optimal value (22) in a given scenario, typically, the worst case, where $Q(\delta_\omega, \sigma_\xi^2)$ is maximum].

The term $|u^*(n)e(n)G(n)|$ represents the recursive expression of the cost function derivative $\partial |e(n)|^2 / \partial \mu|_n$. By construction, this has an important

value if the convergence is not settled, and in the asymptotic mode, it converges to 0.

With a geometric update combined with a forgetting factor, the step size will necessarily increase during transient mode, where $\gamma + \eta|u^*(n)e(n)G(n)| > 1$, and then decrease to finally converge to $\mu_{\min}$ in the asymptotic mode (where $\gamma + \eta|u^*(n)e(n)G(n)| < 1$).

The LMS algorithm combined with the adaptive-step-size process designed is summarized as

$$\begin{aligned}
\hat{s}(n) &= w(n)u(n) \\
e(n) &= d(n) - \hat{s}(n) \\
w(n+1) &= w(n) + \mu_n u^*(n)e(n) \\
G(n+1) &= (1 - \mu_n B^2)G(n) + u^*(n)e(n) \\
\mu_{n+1} &= \{\mu_n [\gamma + \eta|u^*(n)e(n)G(n)|]\}_{\mu_{\min}}^{\mu_{\max}}.
\end{aligned} \tag{39}$$

Here, the first three equations correspond to the LMS algorithm, and the last two correspond to the adaptive-step-size process overlay. In the asymptotic mode, (39) is equivalent to (12), (13), and (14) with $\mu_n = \mu_{\min}$.

4. Multi-spur cancellation scheme

Several spurs can be present in the baseband signal at different frequencies. We consider here the case where P spurs pollute the noisy signal. We assume that all of the spur frequencies are different from each other. The

mixture is given by

$$d(n) = x(n) + b(n) + \sum_{i=1}^P s_i(n) \quad (40)$$

with

$$s_i(n) = A_i e^{j(\omega_i - \delta_{\omega_i})n + j\phi_{0i} + j\phi_i(n)} \quad (41)$$

where A_i is the unknown amplitude, ϕ_{0i} is the unknown phase, $\phi_i(n)$ is the PN that affects Spur i , ω_i is the *a priori* frequency, and δ_{ω_i} is the FS of Spur i . We consider that all of the PNs are under a Brownian model of variance $\sigma_{\xi_i}^2$, as in Section 2. For each spur, a dedicated reference is synthesized as

$$u_i(n) = B_i e^{j\omega_i n + j\phi_{Ri}} . \quad (42)$$

As all of the spur frequencies are assumed to be different, the spurs are temporally uncorrelated. From the one-spur cancellation scheme defined in Section 3, a parallel structure can be designed, where all of the spurs are simultaneously filtered (the P references are weighted by P complex coefficients). The parallel cancellation scheme is presented in Figure 4. In this structure, all of the estimates $\hat{s}_i(n)$ coming from the LMS blocks are summed and then subtracted from the observation $d(n)$. The resulting output $e(n)$ is then used as feedback to the P blocks.

The output of the algorithm at time index n is written as

$$e(n) = d(n) - \hat{\Sigma}(n) \quad , \quad (43)$$

where

$$\hat{\Sigma}(n) = \sum_{i=1}^P w_i(n) u_i(n) \quad (44)$$

is the global estimate. The iteration process of the P filter coefficients is then set from (14) as

$$w_i(n+1) = w_i(n) + \mu u_i^*(n) e(n) \quad ; \forall i \quad (45)$$

where μ is the global constant step size used in the structure. For all of the blocks in Figure 4, the adaptive-step-size process described in (39) can also be used.

At the same time, for each block, we introduce the related pollution factor from (19) as

$$Q_i(\delta_{\omega_i}, \sigma_{\xi_i}^2) = \frac{A_i^2}{B_i^2} (\sigma_{\xi_i}^2 + \delta_{\omega_i}^2) + \frac{2A_i^2(1 - \mu B_i^2) \delta_{\omega_i}^2 (\mu B_i^2 \delta_{\omega_i}^2 - \delta_{\omega_i}^2 + \mu B_i^2)}{B_i^2 (\mu^2 B_i^4 \delta_{\omega_i}^2 - 2\mu B_i^2 \delta_{\omega_i}^2 + \mu^2 B_i^4)} \quad . \quad (46)$$

Assuming there is no correlation between the spurs and the spur estimates, the asymptotic SNIR can be expressed (see Appendix E) as

$$\text{SNIR}_{\text{PaLMS}_{[P]}} = -10 \log_{10} \left[\frac{\mu \sum_{i=1}^P B_i^2}{2 - \mu \sum_{i=1}^P B_i^2} + \frac{\sum_{i=1}^P Q_i(\delta\omega_i, \sigma_{\xi_i}^2)}{\mu(\sigma_x^2 + \sigma_b^2)(2 - \mu \sum_{i=1}^P B_i^2)} \right]. \quad (47)$$

The SNIR is composed of two terms: the first is due to the desired noisy signal, and the second is due to the polluting terms. In the case of one spur, the expression in (47) is equal to the asymptotic SNIR expressed in (20).

5. Simulation

In this section, analytic formulae for the asymptotic and transient performance of the digital spur cancellation algorithms are compared to simulations.

5.1. Link between model parameters and physical values

First, we have supposed that the *a priori* spur frequency is perfectly known, as in the case of multi-standard transceivers. The clock harmonics are exact multiples of the operating clock frequency, and a spur will fall into the reception stage if one of the harmonics is within the signal bandwidth at the receiver carrier frequency f_{Rx} (see Figure 1). In such a case, the *a priori* spur natural frequency ω (normalized between 0 and 2π), defined in (2) and (4), is obtained from the clock frequency harmonic, which is a multiple of the operating clock frequency F_{clock} , the received carrier frequency f_{Rx} and

the sampling frequency F , where

$$\omega = 2\pi \frac{pF_{\text{clock}} - f_{\text{Rx}}}{F}, \quad (48)$$

with p the index of the polluting harmonic.

As an example, if we consider a clock operating at $F_{\text{clock}} = 26$ MHz, and a carrier frequency $f_{\text{Rx}} = 930$ MHz for a LTE 10 MHz signal (at sampling frequency $F = 15.36$ MHz), the 36th clock harmonic, (at 936 MHz) will pollute the received signal. The spur will be at $f_s = 6$ MHz in baseband, leading to

$$\omega = 2\pi \frac{36 \cdot 26e6 - 930e6}{15.36e6} = 2\pi \cdot 0.39.$$

Our models and performance equations are based on values of the frequency shift δ_ω and the phase noise variance σ_ξ^2 . First, the parameter δ_ω is related to the practical imprecision Δ of a synthesized frequency, which is expressed in parts-per-million (ppm). For example, in LTE systems, a typical frequency drift is about 10 ppm [28, 29]. The link between δ_ω and the physical parameters (spur frequency f_s , sampling frequency F , expressed in Hertz) is $\delta_\omega = \Delta f_s / F$. For example, in LTE, for a signal bandwidth of 10 MHz sampled at $F = 15.36$ MHz, with a polluting spur at 6 MHz, the resulting δ_ω is $4 \cdot 10^{-6}$.

Secondly, we have to link our parameter model σ_ξ^2 to a practical phase noise specification. A state-of-the-art discretized model for Brownian phase

noise is $\phi(n) = 2\pi f_s \sqrt{c} B(n)$, where f_s is the spur frequency, $B(n)$ represents the Brownian motion process of variance $\frac{1}{F}$ (where F is the sampling frequency), and c is a parameter of the oscillator defined in [18]. Thus, our state noise variance σ_ξ^2 is $4\pi^2 f_s^2 c / F$. The oscillator constant c cannot be obtained directly, but is a function of the oscillator power spectral density L expressed in dBc/Hz at a specific frequency f (> 0), with the relation $L(f) = 10 \log_{10} (f_s^2 c / f^2)$ [18, 30]. The link between a phase noise specification and the state noise variance of our model is finally

$$\sigma_\xi^2 = 4\pi^2 f^2 10^{-\frac{L(f)}{10}} \frac{1}{F} .$$

For example, in LTE, for a signal bandwidth of 10 MHz ($F = 15.36$ MHz) and for a power spectral density of -100 dBc/Hz at $f = 10$ kHz, the equivalent variance of the Brownian model is $\sigma_\xi^2 = 2.5 \cdot 10^{-8}$.

5.2. One-spur case

We focus first on the transient performance for one spur. We consider here that the observation is composed of a white complex signal of unit variance, an additive white complex Gaussian noise with a SNR of 10 dB and a spur of known frequency, where the SNIR before compensation is 0 dB. The spur follows the model presented in (2), which is affected by a frequency shift $\delta_\omega = 10^{-7}$, and the corresponding Brownian PN has a variance σ_ξ^2 of 10^{-7} . The reference is constructed from (4) and is used in the filter scheme described in Figure 2. At the middle of the simulation, a change

in the configuration is applied: the frequency of the spur and the reference are not changed, but the level of the spur is raised from 0 dB to 5 dB to force the algorithm to re-converge after having reached the asymptotic mode (assuming unchanged parameters for the reference). The two algorithms presented in Section 3 are compared: the constant-step solution of (12)-(14), and the adaptive-step-size solution of (39). For these algorithms to have the same asymptotic performance, the value of the constant LMS step size is the same as the minimal step of the adaptive-step-size solution, and this is set to 2^{-11} (which is close to optimal). The maximal value of the adaptive step size is bounded by (21). The simulated mean transient SNIRs for the two algorithms, the theoretical SNIR evolution for the constant step size of (36), the theoretical asymptotic value of (20), and the adaptive-step-size algorithm of (39) are depicted in Figure 5. Also, two other methods in the literature are computed for comparizon: the recursive least squares (RLS) algorithm [24] and the equivalent IIR as described in (26) and presented in [14]. It is first shown that the theoretical formulae are validated by the simulations: the mean transient performance of the constant-step-size LMS follows (36) and constant and adaptive-step-sized algorithms have the same asymptotic performance, as defined in (20). Furthermore, the adaptive-step-size solution is more rapid than the constant-step-size solution for the same asymptotic performance. The constant-step-size LMS solution has the same transient and asymptotic performance as the IIR filter, which is in agreement with the equivalent notch filter of (26). Finally, the adaptive solution has

a convergence rate that is similar to the RLS solution, but can better cope with a change in the configuration of the spurs, and offers a more suitable hardware implementation (as there is no division in our algorithm).

Figure 6 compares the asymptotic SNIR to the simulated results for one spur with input SNIR of 10 dB, in the case of FS only (no PN). The SNIR is computed versus the step size (which is equivalent to the minimal step size in the case of the adaptive-step-size solution) for different values of FS. The simulated data is white and has an additive white noise of variance 10 dB. The figure validates the asymptotic formula (28). The presence of the optimal step size is also validated, and the loss of performance, due to the second term of (28), is proportional to the square of the step size, as expected by the formulae.

At the same time, the optimal step size gives the maximal value of the SNIR that can be obtained for a given FS. Figure 7 shows the evolution of this maximal SNIR (30) versus FS. This shows that the theory is in agreement, and that the FS leads to a loss of performance of 6.66 dB per decade.

Figure 8 compares the asymptotic SNIR to the simulated results in the case of PN (no FS). The asymptotic SNIR is computed versus the step size, for different variances of the Brownian model of (3), again, showing agreement with the theory described in (32).

The maximal SNIR is given by (34), and is validated in Figure 9 for an initial SNIR of 10 dB.

In Figure 10, we consider both the PN and FS for one spur, and the

theory elaborated in (20) is validated by the simulation for several pairs of σ_ξ^2 and δ_ω .

5.3. Multi-spur case

In order to show the impact of the number of spurs on the asymptotic performance, Figure 11 shows the asymptotic SNIR of (47), as compared to the simulations. The noisy signal is still white, and is polluted by 1 to P spurs of different frequencies. All of the P spurs are affected by a Brownian PN of variance $\sigma_{\xi_i}^2$ and their reference has a FS of δ_{ω_i} , following the model described in (41) and (42). The parameters used are described in Table 1. It should be noted that the frequency values are given for information only, and have no impact on the overall performance, as (47) does not depend on w_i .

5.4. Bit error rate

5.4.1. QPSK Case

We compare the influence of the spurs and the compensated spurs in terms of the bit error rate (BER). We first consider a data signal modulated by quadrature phase shift keying (QPSK) modulation through an additive white Gaussian noise channel with various SNRs. We suppose that the noisy signal is polluted by two spurs, affected by a FS of 10^{-7} and a PN of variance 10^{-7} . We show in Figure 12 the results of the BER without the compensation system and with the parallel LMS compensation scheme (with step size of 2^{-11}) for different values of the channel SNR and for different SIRs. (Here,

the amplitudes of the spurs are computed from only the power of the desired signal (SIR) neglecting the additive white noise in order to have constant spur amplitude versus the channel SNR: the amplitudes of the spurs are obtained directly by $A_1 = A_2 = \sqrt{\sigma_x^2/10^{SIR/10}}$, where σ_x^2 is the variance of the data signal without noise.) We show that our methods greatly reduce the BER and that the compensated BER comes close to the ideal channel influence $Q\left(\sqrt{2E_b/N_0}\right)$.

5.4.2. LTE Case

Finally, we compare the BER for the more realistic case of a simplified LTE modulation that is based on the Third Generation Partnership Project (3GPP) model [31]. The data signal follows a LTE 1.4 MHz bandwidth specifications based on orthogonal frequency division multiplexing (OFDM), with 72 QPSK subcarriers, a fast Fourier transform (FFT) of size of 128, and a cyclic prefix of length nine samples. The BER is computed on the uncoded bits of the data and through an additive white Gaussian noise channel with various SNRs. We consider that the noisy signal is polluted by two spurs. These spurs are affected by a FS of 10^{-7} and are modulated with a Brownian PN of variance 10^{-7} .

Figure 13 shows the results of the BER without compensation and with the parallel LMS compensation scheme for different values of the channel SNR and for different SIRs. The theoretical performance without spurs (i.e.,

only the influence of the channel) is shown in black ¹. Without compensation, it is shown that the presence of the spurs reduces the performance of the overall system, which leads to a lower limit on the BER. The decrease in performance depends on input SIR, although it impacts the signal as soon as the SNR is greater than about 4 dB. After compensation, the BER is greatly reduced compared to the initial BER, and reaches a floor. Thus, the designed system has good performance if the spurs have an initial SIR between 0 dB and 20 dB.

5.5. Practical Phase Noise case

Up to now, we have verified the theoretical performance of the proposed algorithm in the presence of Brownian phase noise, which is the usual model chosen in practise for free running oscillators [19, 20, 21]. However, if a more practical PLL-type oscillator is used, the polluting model for the phase noise will not be Brownian. In such a case, the theoretical asymptotic performance expressions (20) and (32) are no longer valid, and the performance will depend on the scale and shape of chosen phase noise.

The profile of a practical phase noise is depicted in Figure 14. Such a phase noise spectrum can result from a sum of $1/f^2$ noise modeled as a

¹For the simulation, the theoretical curve is given by the link between the SNR and E_b/N_0 . In the case of LTE with 1.4-MHz bandwidth and QPSK modulation, the link is given by $E_b/N_0 = \text{SNR}(n_{\text{FFT}} + n_{\text{CP}})/(2n_{\text{Re}})$, where $n_{\text{FFT}} = 128$ is the size of the fast Fourier transform, $n_{\text{CP}} = 9$ is the length of the cyclic prefix, and $n_{\text{Re}} = 72$ is the number of received carriers. This leads to the expression of $\text{BER} = Q \left[\sqrt{\text{SNR}(n_{\text{FFT}} + n_{\text{CP}})/n_{\text{Re}}} \right]$

Lorenz spectrum, $1/f$ noise (flicker noise) filtered by a second-order PLL, and a noise floor as described in [32, Section 2.3]. An example of a phase noise specification from [32, Section 2.3] is proposed in Table 2.

In Figure 15, we represent in red the asymptotic performance of our algorithm versus the step size in the presence of one spur, impaired by the practical phase noise described in Table 2, obtained by Monte-Carlo simulation. For comparison, we also plot the theoretical asymptotic performance with Brownian phase noise with specific variance obtained in (32). We note that, for large values of the step size, the performance is similar for the two models [since the first term of (32), $\mu B^2/(2 - \mu B^2)$, is predominant and does not depend on the model], whereas for small values of the step size, the theoretical curve under the Brownian assumption does not match with the practical performance [since the second term of (32) is predominant and is model dependent].

6. Conclusions

This study has focused on the performance of a multiple-spur cancellation scheme. We first proposed a one-spur canceller based on the LMS algorithm, which uses *a priori* knowledge of spur frequency to create a reference. We proposed a model with imperfections, due to the limited accuracy of frequency knowledge and Brownian phase noise. In this case, we derived closed-form formulae of the asymptotic SNR and show that the step size of the algorithm can be optimally tuned. We showed that the LMS algorithm is

equivalent to a notch filter, which leads to simplified expressions in the case of PN only or FS only. We added an adaptive-step-size process to improve the convergence speed without lowering the asymptotic performance. This solution offers performance similar to the RLS algorithm but can better cope with configuration changes. In a second part of the paper, we extended the problem to the multiple-spur case, and developed a parallel scheme that uses the one-spur block previously described. We developed closed-form formulae of the asymptotic SNIR for P spurs, in a general case of P Brownian noises and P FSs.

Appendix A. Establishment of the LMS algorithm

The goal of this appendix is to recall how the LMS approach can be legitimized, from simplification of the recursive-least-squares approach. The weighted least-squares cost function used is [24, 33]

$$J_n(\hat{w}) = \sum_{p=1}^n \lambda^{n-p} |e(p)|^2$$

with $\lambda < 1$, because of the nonstationarity of the parameters to be estimated. Due to the structure of the filter defined in Figure 2, we have:

$$\begin{aligned} J_n(\hat{w}) &= \sum_{p=1}^n \lambda^{n-p} |d(p) - \hat{w}u(p)|^2 \\ &= \sum_{p=1}^n \lambda^{n-p} [d(p)d^*(p) - \hat{w}^*d(p)u^*(p) - \hat{w}d^*(p)u(p) + \hat{w}\hat{w}^*u(p)u^*(p)] . \end{aligned}$$

At each iteration, the weight estimate update $w(n+1)$ is based on the minimisation of $J_n(\hat{w})$. As the cost function is a real function of complex variables, using complex derivative theory [34], the minimisation is obtained by setting $\partial J_n(\hat{w})/\partial \hat{w}^* = 0$ and assuming $\partial \hat{w}/\partial \hat{w}^* = 0$. As a consequence, the derivative component is written as

$$\frac{\partial J_n(\hat{w})}{\partial \hat{w}^*} = \nabla_{\hat{w}^*} J_n(\hat{w}) = \sum_{p=1}^n \lambda^{n-p} \nabla_{\hat{w}^*} |e(p)|^2 = \sum_{p=1}^n \lambda^{n-p} [-d(p)u^*(p) + \hat{w}u(p)u^*(p)] .$$

Setting $\partial J_n(\hat{w})/\partial \hat{w}^* = 0$ then leads to

$$w(n+1) = \frac{\sum_{p=1}^n \lambda^{n-p} [d(p)u^*(p)]}{\sum_{p=1}^n \lambda^{n-p} [u(p)u^*(p)]} . \quad (49)$$

We denote the temporal autocorrelation of the reference by $R_\lambda(n)$ and the temporal crosscorrelation between the reference and the observation by $p_\lambda(n)$, which are estimated as

$$R_\lambda(n) = \sum_{p=1}^n \lambda^{n-p} u(p)u^*(p) \quad (50)$$

$$p_\lambda(n) = \sum_{p=1}^n \lambda^{n-p} d(p)u^*(p) , \quad (51)$$

and finally have

$$w(n+1) = \frac{p_\lambda(n)}{R_\lambda(n)} . \quad (52)$$

The temporal crosscorrelation and autocorrelation can be recursively expressed as

$$p_\lambda(n) = \lambda p_\lambda(n-1) + d(n)u^*(n) \quad (53)$$

$$R_\lambda(n) = \lambda R_\lambda(n-1) + u(n)u^*(n) . \quad (54)$$

Substituting (53) in (52), we have

$$w(n+1) = \frac{1}{R_\lambda(n)} [\lambda p_\lambda(n-1) + d(n)u^*(n)] . \quad (55)$$

From (52) and due to the recursive update of $R_\lambda(n)$ expressed in (54)

$$\begin{aligned} p_\lambda(n-1) &= w(n)R_\lambda(n-1) \\ &= w(n) \frac{R_\lambda(n) - u(n)u^*(n)}{\lambda} , \end{aligned}$$

and substitution into (55) gives

$$w(n+1) = w(n) + \frac{1}{R_\lambda(n)} [-w(n)u(n)u^*(n) + d(n)u^*(n)] . \quad (56)$$

As by construction $e(n) = d(n) - w(n)u(n)$, the update becomes

$$w(n+1) = w(n) + \frac{1}{R_\lambda(n)} [e(n)u^*(n)] . \quad (57)$$

The standard update of the LMS algorithm [22] is finally obtained by substituting the inverse recursive update of the temporal autocorrelation of the reference $[1/R_\lambda(n)]$ with a constant step size μ , to have a less complex algorithm, and thus the recursive equation of $w(n)$ becomes

$$\begin{aligned} w(n+1) &= w(n) + \mu u^*(n)e(n) , \\ &= w(n) - \frac{\mu}{2} \nabla_w |e(n)|^2 \end{aligned} \quad (58)$$

Appendix B. Proof of asymptotic misalignment variance for one spur

The expression of the misalignment can be expressed recursively using (14) and (11), as

$$\begin{aligned} v(n+1) &= w(n+1) - w_{\text{opt}}(n+1) \\ &= w(n) + \mu u^*(n)[x(n) + b(n)] - \mu B^2 v(n) - w_{\text{opt}}(n) \\ &\quad - w_{\text{opt}}(n)[j\xi(n) - j\delta_\omega] \\ &= [w(n) - w_{\text{opt}}(n)] + \mu u^*(n)[x(n) + b(n)] \\ &\quad - \mu B^2 v(n) - w_{\text{opt}}(n)[j\xi(n) - j\delta_\omega] \\ &= (1 - \mu B^2)v(n) + \mu u^*(n)[x(n) + b(n)] - w_{\text{opt}}(n)[j\xi(n) - j\delta_\omega] . \end{aligned} \quad (59)$$

The LMS algorithm is based on minimization of the gradient of $J(n)$, so the complex sequence of the misalignment $v(n)$ is white. Assuming temporal uncorrelation between the white sequences [the misalignment $v(n)$, the signal $x(n)$, the additive noise $b(n)$, and the phase noise $\xi(n)$], we only have dependence between the misalignment $v(n)$ and the recursive expression of the optimal filter $w_{\text{opt}}(n)$. Note that there is no single dependency between $v(n)$ and $[x(n) + b(n)]$ due to causality and the whiteness of the desired noisy signal. The variance of the misalignment can be expressed as

$$\begin{aligned} E[|v(n+1)|^2] &= (1 - \mu B^2)^2 E[|v(n)|^2] + \mu^2 B^2 (\sigma_x^2 + \sigma_b^2) \\ &\quad + \frac{A^2}{B^2} (\sigma_\xi^2 + \delta_\omega^2) + \delta_\omega (1 - \mu B^2) E[|v(n)w_{\text{opt}}(n)|] . \end{aligned} \quad (60)$$

The term $E[|v(n)w_{\text{opt}}(n)|]$ can be expressed through the separation between the real part and the imaginary part of the two components:

$$E[|v(n)w_{\text{opt}}(n)|] = \frac{2A^2(\mu B^2 \delta_\omega^2 - \delta_\omega^2 + \mu B^2)}{B^2(\mu^2 B^4 \delta_\omega^2 - 2\mu B^2 \delta_\omega^2 + \mu^2 B^4)} . \quad (61)$$

Then, replacing (61) in (60) leads to (18).

Appendix C. Proof of optimal step value

We still assume the aforementioned hypotheses

- The normalized FS $\delta_\omega \ll \omega$ (normalized frequency),

- The PN variance $\sigma_\xi^2 \ll 1$,
- The step size $\mu \ll 1$.

With these hypotheses, the linear asymptotic SNIR obtained in (20) can be expressed as

$$\text{SNIR}_{\text{comp}}^{\text{lin}} \approx \left[\frac{\mu B^2}{2} + \frac{A^2 \delta_\omega^2}{\mu^2 B^4 (\sigma_x^2 + \sigma_b^2)} + \frac{A^2 \sigma_\xi^2}{2\mu B^2 (\sigma_x^2 + \sigma_b^2)} \right] . \quad (62)$$

We denote

$$\begin{aligned} S(\mu) &= \frac{\partial \text{SNIR}_{\text{comp}}^{\text{lin}}}{\partial \mu} \\ &= \frac{B^2}{2} - \frac{2A^2 \delta_\omega^2}{B^4 (\sigma_x^2 + \sigma_b^2) \mu^3} - \frac{A^2 \sigma_\xi^2}{2(\sigma_x^2 + \sigma_b^2) B^2 \mu^2} . \end{aligned} \quad (63)$$

The optimal step value is obtained with $S(\mu_{\text{opt}}) = 0$, leading to

$$-B^6 (\sigma_x^2 + \sigma_b^2) \mu_{\text{opt}}^3 + B^2 A^2 \sigma_\xi^2 \mu_{\text{opt}} + 4A^2 \delta_\omega^2 = 0 . \quad (64)$$

Then, solving (64) leads to the optimal step expression derived in (22).

Appendix D. Proof of the simplified SNIR of a frequency shift in the frequency domain

In the asymptotic mode, assuming that there is no PN, the algorithm is equivalent to a notch filter after baseband frequency shifting of the spur. The

estimation error at the output of the structure is expressed after shifting as

$$\epsilon_{\text{BB}}(n) = [x(n) + b(n) - e(n)] u^*(n) . \quad (65)$$

Expressing the z transform of the filter obtained, using the z transform of the equivalent LPF (24), the estimation error can be expressed as:

$$\begin{aligned} E_{\text{BB}}(z) &= S(z) * U^*(z) - B^2 W(z) \\ &= S(z) * U^*(z) - B^2 L(z) [D(z) * U^*(z)] . \end{aligned}$$

With (1) for the observation equation

$$E_{\text{BB}}(z) = -B^2 L(z) [X(z) * U^*(z)] + [1 - B^2 L(z)] [S(z) * U^*(z)] . \quad (66)$$

The variance of the estimation error is from (66)

$$\begin{aligned} \sigma_{\epsilon_{\text{BB}}}^2 &= \sigma_x^2 B^4 T \int_{-\frac{1}{2T}}^{\frac{1}{2T}} |L(e^{2j\pi fT})|^2 df \\ &+ \int_{-\frac{1}{2T}}^{\frac{1}{2T}} |1 - B^2 L(e^{2j\pi fT})|^2 \Gamma_{su^*}(f) df . \end{aligned}$$

The first term corresponds to the static error due to the presence of the noisy data (assumed to be white and with a power spectral density $\sigma_x^2 T$). The second term corresponds to the dynamic error related to the frequency shift, where $\Gamma_{su^*}(f)$ is the power spectral density of su^* (which corresponds to the downshifted spur frequency). After computation and using the same

hypotheses ($\delta_\omega \ll 1, \delta_\omega < \mu; \mu \ll 1$), the variance of the estimation error is

$$\sigma_{\epsilon_{\text{BB}}}^2 \approx \sigma_x^2 \frac{\mu B^2}{2 - \mu B^2} + \left(\frac{A \delta_\omega}{\mu B^2} \right)^2, \quad (67)$$

which leads to the same value of the asymptotic SNIR as described in the simplified form of (28).

Appendix E. Proof of SNIR formula in the parallel LMS architecture

In the parallel structure, the global output of the algorithm is composed of the signal $x(n)$, the noise $b(n)$, the spurs $s_i(n)$, and the spur estimates $w_i(n)u_i(n)$, such that

$$e(n) = x(n) + b(n) + \sum_{i=1}^P s_i(n) - \sum_{i=1}^P w_i(n)u_i(n). \quad (68)$$

The misalignments for each block are introduced with the same definition as in (15), as $s_i(n) = w_{\text{opt}_i}(n)u_i(n)$, so that

$$e(n) = x(n) + b(n) - \sum_{i=1}^P v_i(n)u_i(n). \quad (69)$$

The estimation error of the structure, which is defined in (6), leads to the expression of the asymptotic SNIR of the P stages:

$$\text{SNIR}_P = -10 \log_{10} \left(\frac{\sum_{i=1}^P B_i^2 E[v_i^2]}{\sigma_x^2 + \sigma_b^2} \right). \quad (70)$$

For each block, the misalignment is recursively described using (11), (15), (45), (59), and (69) as

$$v_i(n+1) = v_i(n) + \mu u_i^*(n)[x(n) + b(n)] - w_{\text{opt}_i}(n)[j\xi_i(n) - j\delta_{\omega_i}] - \sum_{j=1}^P \mu B_i B_j v_j(n) \quad , \forall i \in [1; P] . \quad (71)$$

Assuming the same hypothesis between uncorrelated misalignment, the variance of Block i is expressed $\forall i \in [1; P]$ as

$$E[v_i^2(n)] = \frac{\mu(\sigma_x^2 + \sigma_b^2)}{2 - \mu B_i^2} + \frac{Q_i(\sigma_{\xi_i}^2, \delta_{\omega_i})}{\mu B_i^2(2 - \mu B_i^2)} + \sum_{\substack{j=1 \\ j \neq i}}^P \left[\frac{\mu B_j^2 E[v_j^2(n)]}{2 - \mu B_i^2} \right] . \quad (72)$$

From (72), the problem can be set in matrix form as

$$\begin{pmatrix} 2 - \mu B_1^2 & \dots & -\mu B_P^2 \\ -\mu B_1^2 & \dots & -\mu B_P^2 \\ \vdots & \vdots & \vdots \\ -\mu B_1^2 & \dots & 2 - \mu B_P^2 \end{pmatrix} \begin{pmatrix} E[v_1^2(n)] \\ E[v_2^2(n)] \\ \vdots \\ E[v_P^2(n)] \end{pmatrix} = \begin{pmatrix} \mu(\sigma_x^2 + \sigma_b^2) \\ \mu(\sigma_x^2 + \sigma_b^2) \\ \vdots \\ \mu(\sigma_x^2 + \sigma_b^2) \end{pmatrix} + \begin{pmatrix} \frac{Q_1(\sigma_{\xi_1}^2, \delta_{\omega_1})}{\mu B_1^2} \\ \frac{Q_2(\sigma_{\xi_2}^2, \delta_{\omega_2})}{\mu B_2^2} \\ \vdots \\ \frac{Q_P(\sigma_{\xi_P}^2, \delta_{\omega_P})}{\mu B_P^2} \end{pmatrix} .$$

The system can be solved for the P blocks, which leads to the same expression of the variance, i.e., $\forall i$:

$$E[v_i^2] = \frac{\mu(\sigma_x^2 + \sigma_b^2)}{2 - \mu \sum_{j=1}^P B_j^2} + \frac{\frac{1}{2} \sum_{j=1}^P Q_j}{2 - \mu \sum_{j=1}^P B_j^2} + \frac{Q_i(\sigma_{\xi_i}^2, \delta_{\omega_i})}{2\mu B_i^2} \quad (73)$$

Finally, substituting (73) into (70) gives

$$\begin{aligned} \sum_{i=1}^P B_i^2 E[v_i^2] &= \frac{\mu(\sigma_x^2 + \sigma_b^2) \sum_{i=1}^P B_i^2}{2 - \mu \sum_{j=1}^P B_j^2} + \frac{\frac{1}{2} \sum_{i=1}^P B_i^2 \sum_{j=1}^P Q_j(\sigma_{\xi_j}^2, \delta_{\omega_j})}{2 - \mu \sum_{j=1}^P B_j^2} \\ &\quad + \frac{\sum_{i=1}^P Q_i(\sigma_{\xi_i}^2, \delta_{\omega_i})}{2\mu}, \forall i. \end{aligned}$$

The above equation can be simplified to

$$\sum_{i=1}^P B_i^2 E[v_i^2] = \frac{\mu(\sigma_x^2 + \sigma_b^2) \sum_{i=1}^P B_i^2}{2 - \mu \sum_{j=1}^P B_j^2} + \frac{\sum_{i=1}^P Q_i(\sigma_{\xi_i}^2, \delta_{\omega_i})}{\mu(2 - \mu \sum_{j=1}^P B_j^2)}. \quad (74)$$

Then, substituting (74) into (70) leads to the expression of the asymptotic SNIR expressed in (47).

References

- [1] J. Atallah, M. Ismail, Future 4G front-ends enabling smooth vertical handovers, IEEE Circuits and Devices Magazine, 22 (1) (2006) 6–15. doi:10.1109/MCD.2006.1598074.
- [2] S. Maruyama, S. Ogawa, K. Chiba, Mobile terminals toward LTE and requirements on device technologies, in: Proc. IEEE Symposium on VLSI Circuits, 2007, pp. 2–5. doi:10.1109/VLSIC.2007.4342673.
- [3] A. Rusu, M. Ismail, Design challenges of wireless mobile radios, in: Proc. International Symposium on Signals, Circuits and Systems. (ISSCS), 2007, pp. 1–4. doi:10.1109/ISSCS.2007.4292669.

- [4] P. Vainikainen, J. Holopainen, C. Icheln, O. Kivekas, M. Kyro, M. Mustonen, S. Ranvier, R. Valkonen, J. Villanen, More than 20 antenna elements in future mobile phones, threat or opportunity?, in: Proc. 3rd European Conference on Antennas and Propagation. (EuCAP), 2009, pp. 2940–2943.
- [5] E. Keehr, A. Hajimiri, Successive regeneration and adaptive cancellation of higher order intermodulation products in RF receivers, *IEEE Transactions on Microwave Theory and Techniques* 59 (5) (2011) 1379–1396. doi:10.1109/TMTT.2011.2123107.
- [6] S.-K. Ting, A. Sayed, Digital suppression of spurious PLL tones in A/D converters, *IEEE Transactions on Signal Processing*, 59 (11) (2011) 5275–5288. doi:10.1109/TSP.2011.2162833.
- [7] C. Mishra, A. Valdes-Garcia, F. Bahmani, A. Batra, E. Sanchez-Sinencio, J. Silva-Martinez, Frequency planning and synthesizer architectures for multiband OFDM UWB radios, *IEEE Transactions on Microwave Theory and Techniques*, 53 (12) (2005) 3744–3756. doi:10.1109/TMTT.2005.856087.
- [8] K. Hardin, J. Fessler, D. Bush, Spread spectrum clock generation for the reduction of radiated emissions, in: Proc. IEEE International Symposium on Electromagnetic Compatibility, 1994, pp. 227–231. doi:10.1109/ISEMC.1994.385656.

- [9] S.-F. Ho, H.-Y. Huang, A wideband programmable spread-spectrum clock generator, in: Proc. Asian Solid-State Circuits Conference, 2005, pp. 521–524. doi:10.1109/ASSCC.2005.251792.
- [10] P. Baudin, F. Belveze, Impact of RF impairments on a DS-CDMA receiver, IEEE Transactions on Communications, 52 (1) (2004) 31–36. doi:10.1109/TCOMM.2003.822159.
- [11] A. Raghavan, E. Gebara, E. Tentzeris, J. Laskar, Analysis and design of an interference canceller for collocated radios, IEEE Transactions on Microwave Theory and Techniques 53 (11) (2005) 3498–3508. doi:10.1109/TMTT.2005.859042.
- [12] M. Valkama, A. Springer, G. Hueber, Digital signal processing for reducing the effects of RF imperfections in radio devices; an overview, in: Proc. IEEE International Symposium on Circuits and Systems (ISCAS), 2010, pp. 813–816. doi:10.1109/ISCAS.2010.5537444.
- [13] P. Zahradnik, M. Vlcek, B. Simak, DC-notch FIR filters for zero-IF receivers, in: Proc. Sixth International Conference on Networking (ICN), 2007, pp. 68–68. doi:10.1109/ICN.2007.33.
- [14] S. Gunturi, J. Balakrishnan, Mitigation of narrowband interference in differentially modulated communication systems, in: Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2009, pp. 2549–2552. doi:10.1109/ICASSP.2009.4960142.

- [15] C. Samori, M. Zanuso, S. Levantino, A. Lacaita, Multipath adaptive cancellation of divider non-linearity in fractional-N PLLs, in: Proc. IEEE International Symposium on Circuits and Systems (ISCAS), 2011, pp. 418–421. doi:10.1109/ISCAS.2011.5937591.
- [16] R. Gerzaguet, L. Ros, J. Brossier, F. Belveze, Performances d’un supprimeur numérique adaptatif de *spurs* dans un terminal radio-fréquence multistandard en présence d’imperfections sur la référence, in: Proc. 24eme colloque Grets, 2013, pp. 1–4.
- [17] R. Gerzaguet, L. Ros, J. Brossier, F. Belveze, Architectures numériques parallèles et successives pour la suppression de *spurs* multiples dans un terminal radio fréquence multi-standard et performances théoriques, in: Proc. 24eme colloque Grets, 2013, pp. 1–4.
- [18] A. Demir, A. Mehrotra, J. Roychowdhury, Phase noise in oscillators: a unifying theory and numerical methods for characterization, Proc. IEEE Transactions on Circuits and Systems 47 (5) (2000) 655–674. doi:10.1109/81.847872.
- [19] D. Petrovic, W. Rave, G. Fettweis, Common phase error due to phase noise in OFDM-estimation and suppression, in: Proc. 15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2004, pp. 1901–1905. doi:10.1109/PIMRC.2004.1368329.

- [20] V. Syrjala, M. Valkama, N. Tchamov, J. Rinne, Phase noise modelling and mitigation techniques in OFDM communications systems, in: Proc Wireless Telecommunications Symposium (WTS), 2009, pp. 1–7. doi:10.1109/WTS.2009.5068965.
- [21] J. Vila-Valls, J.-M. Brossier, L. Ros, Oversampled phase tracking in digital communications with large excess bandwidth, *Signal Processing* 90 (3) (2010) 821–833.
- [22] B. Widrow, J. McCool, M. Ball, The complex LMS algorithm, *Proceedings of the IEEE* 63 (4) (1975) 719 – 720. doi:10.1109/PROC.1975.9807.
- [23] B. Widrow, J. Glover, J.R., J. McCool, J. Kaunitz, C. Williams, R. Hearn, J. Zeidler, J. Eugene Dong, R. Goodlin, Adaptive noise cancelling: Principles and applications, *Proceedings of the IEEE* 63 (12) (1975) 1692–1716. doi:10.1109/PROC.1975.10036.
- [24] S. Haykin, *Adaptive filter theory*, Prentice Hall, 1996.
- [25] R. Gerzaguet, L. Ros, J. Brossier, F. Belveze, S. Ghandour-Haidar, Self-adaptive stochastic rayleigh flat fading channel estimation, in: Proc. 18th International Conference on Digital Signal Processing (DSP 2013), 2013, pp. 1–6.
- [26] S.-L. Goh, D.-P. Mandic, Stochastic Gradient-Adaptive Complex-Valued nonlinear neural adaptive filters with a Gradient-Adaptive step

- size, *IEEE Transactions on Neural Networks* 18 (5) (2007) 1511–1516.
doi:10.1109/TNN.2007.895828.
- [27] A. Wee-Peng, B. Farhang-Boroujeny, A new class of gradient adaptive step-size LMS algorithms, *IEEE Transactions on Signal Processing* 49 (4) (2001) 805–810. doi:10.1109/78.912925.
- [28] Q. Wang, C. Mehlhruer, M. Rupp, Carrier frequency synchronization in the downlink of 3GPP LTE, in: *Proc. IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2010, pp. 939–944. doi:10.1109/PIMRC.2010.5671968.
- [29] G. Jia, Y. Xiong, X. Yang, W. Fang, Frequency synchronization for TD-LTE-A downlink receiver, *The Journal of China Universities of Posts and Telecommunications* 19 (1) (2012) 31–37.
- [30] D. Petrovic, W. Rave, G. Fettweis, Effects of phase noise on OFDM systems with and without PLL: Characterization and compensation, *Proc. IEEE Transactions on Communications*, 55 (8) (2007) 1607–1616. doi:10.1109/TCOMM.2007.902593.
- [31] 3rd Generation Partnership Project, Evolved universal terrestrial radio access (e-utra) ; physical channels and modulation, Tech. rep., 3GPP TS (2011).
- [32] L. Smaini, RF Analog Impairments Modeling for Communication Sys-

tems Simulation: Application to OFDM-based Transceivers, Wiley, 2012.

- [33] M. Bellanger, Adaptive digital filters, Vol. 11, CRC Press, 2001.
- [34] D. Brandwood, A complex gradient operator and its application in adaptive array theory, IEE Proceedings H (Microwaves, Optics and Antennas) 130 (1) (1983) 11–16. doi:10.1049/ip-f-1:19830003.

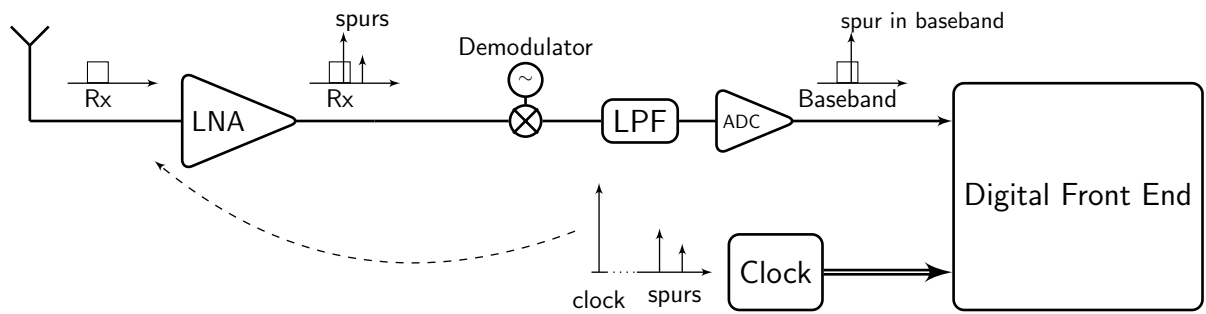


Figure 1: Block diagram of a receiver stage with spur pollution due to clock harmonics. LNA denotes the low noise amplifier, LPF denotes the low pass filter, and ADC denotes the analog to digital converter.

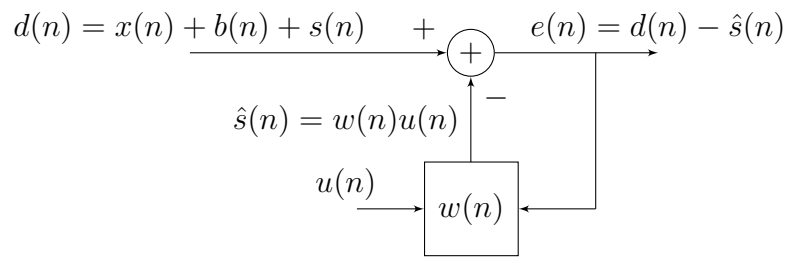


Figure 2: Filtering scheme used for spur cancellation.

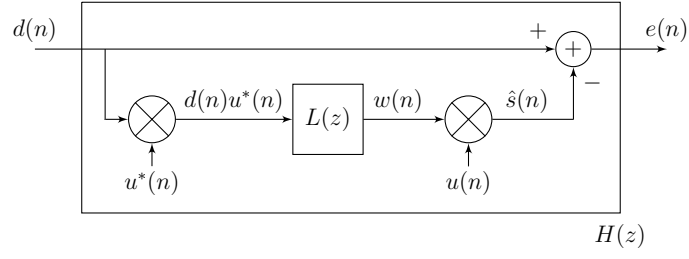


Figure 3: Equivalent scheme of the LMS algorithm where $L(z)$ and $H(z)$ denote, respectively, a LPF in the z domain and the response of the global notch structure in the z domain.

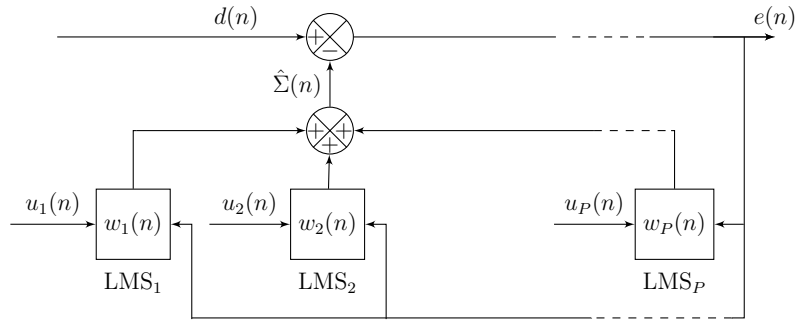


Figure 4: Parallel LMS cancellation scheme

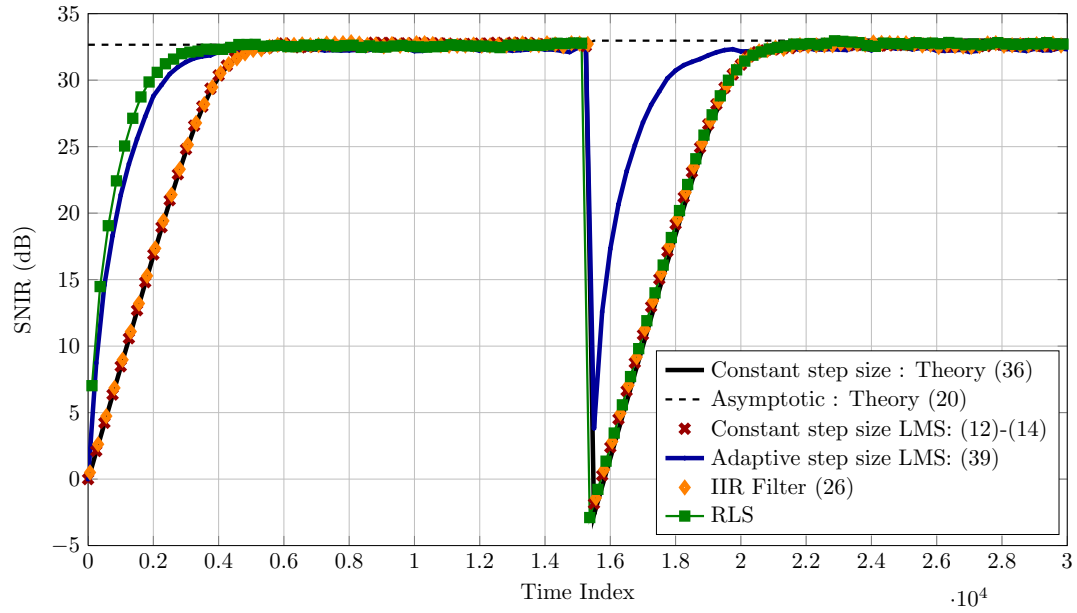


Figure 5: Comparison of the mean transient SNIR for the two proposed solutions and for other methods in the literature.

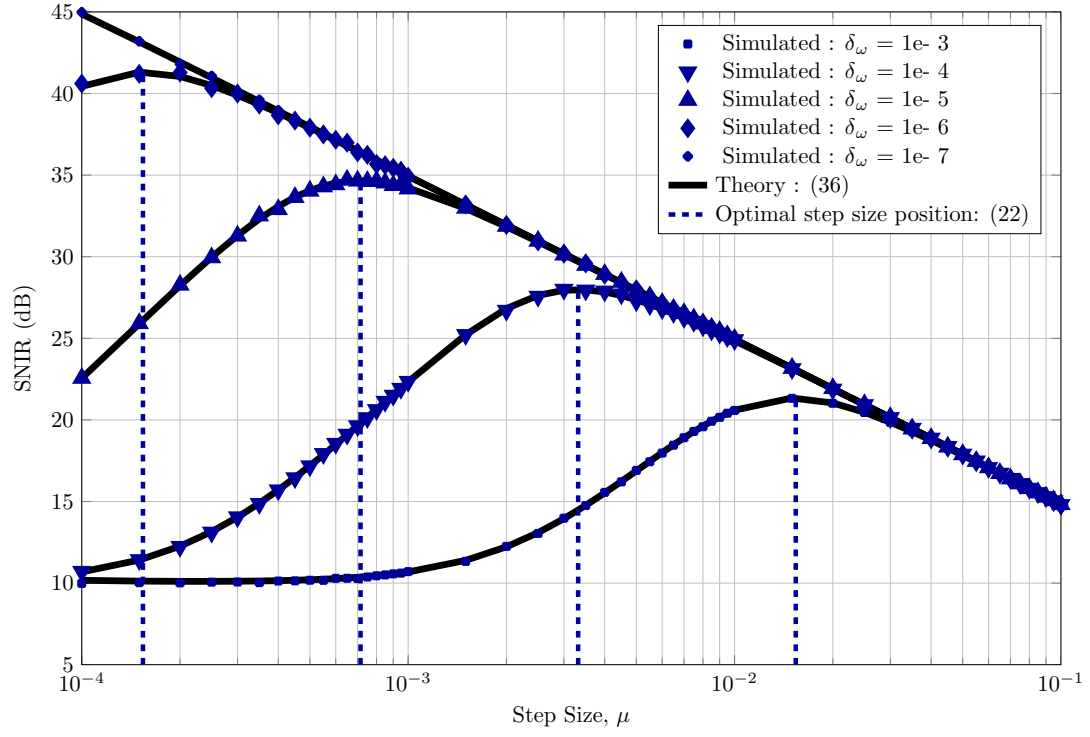


Figure 6: Comparison of the theoretical SNIR obtained in (28) with the simulated results, as a function of the step size, for one spur and with a FS of δ_ω (PN=0).

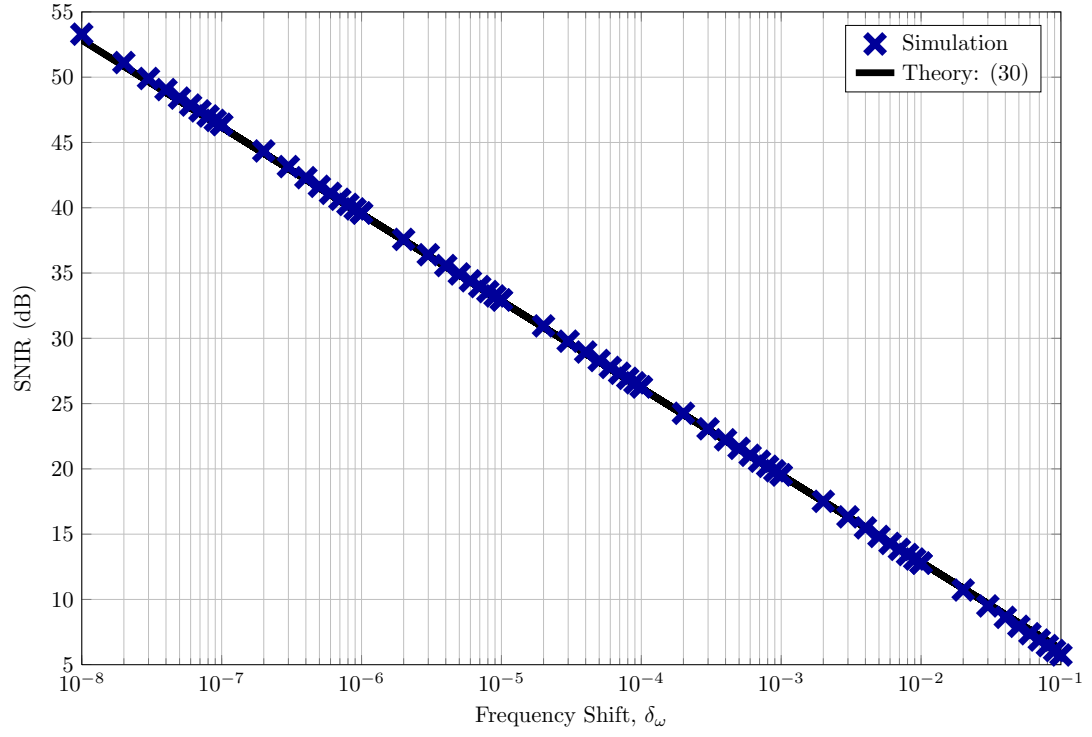


Figure 7: Comparison of the maximal SNIR obtained by simulation and the theoretical SNIR (30), versus the FS of δ_ω (PN=0).

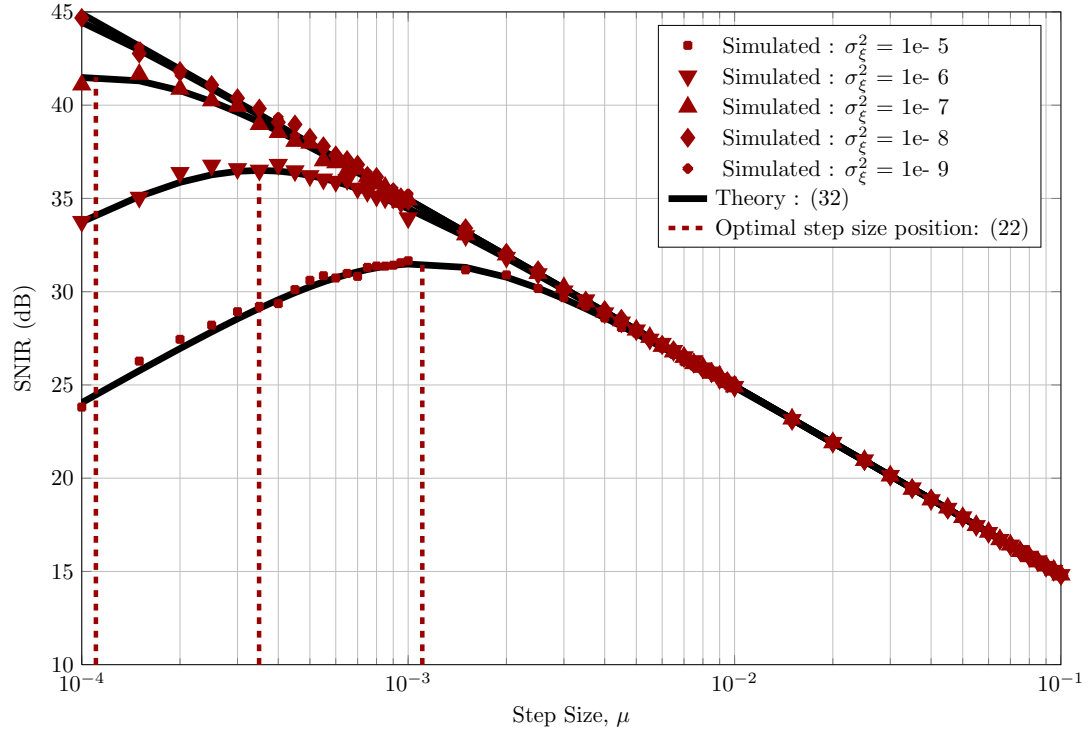


Figure 8: Comparison of the theoretical SNIR obtained in (32) with the simulated results, versus the step size for one spur, when the spur is affected by a Brownian PN of variance σ_{ξ}^2 (FS=0).

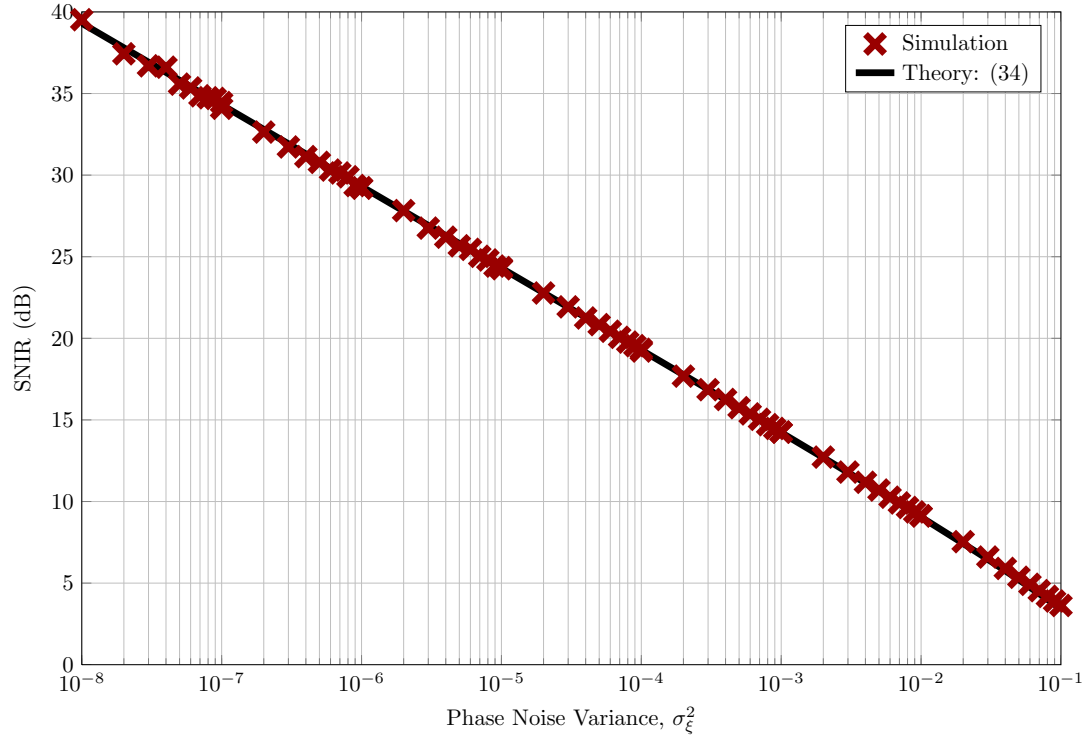


Figure 9: Comparison of the maximal SNIR obtained by simulation and the value of the SNIR with (34), versus phase noise variance σ_{ξ}^2 (FS=0).

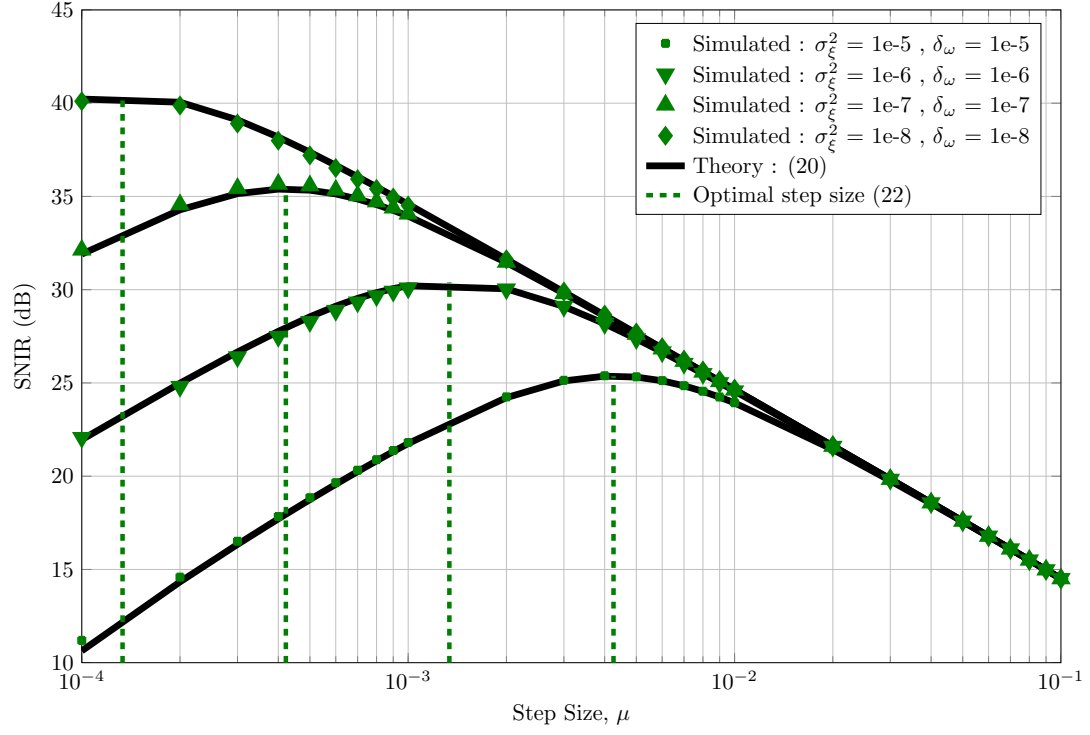


Figure 10: Comparison of the theoretical SNIR obtained in (20) with the simulated results, versus the step size for one spur, when the spur is affected by a Brownian PN of variance σ_{ξ}^2 and a FS δ_{ω} .

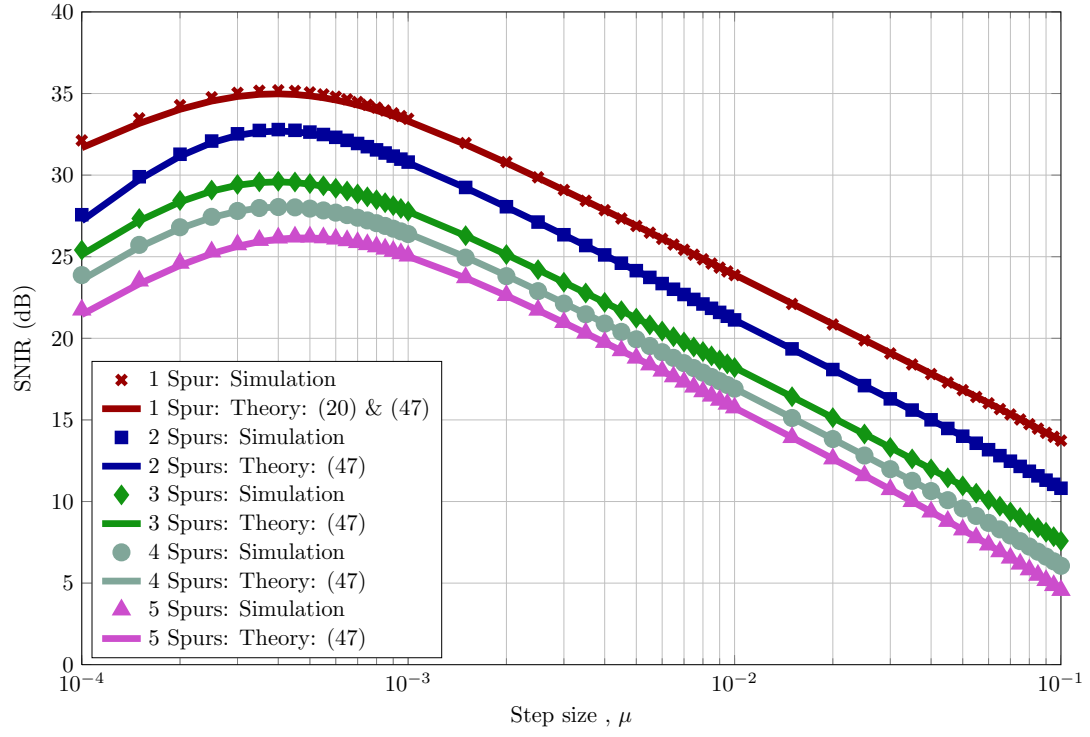


Figure 11: Asymptotic performance of the parallel LMS structure in the case of one to five spurs, with the parameters given in Table 1, versus step size of the LMS algorithm, or the minimal step of the adaptive-step-size solution.

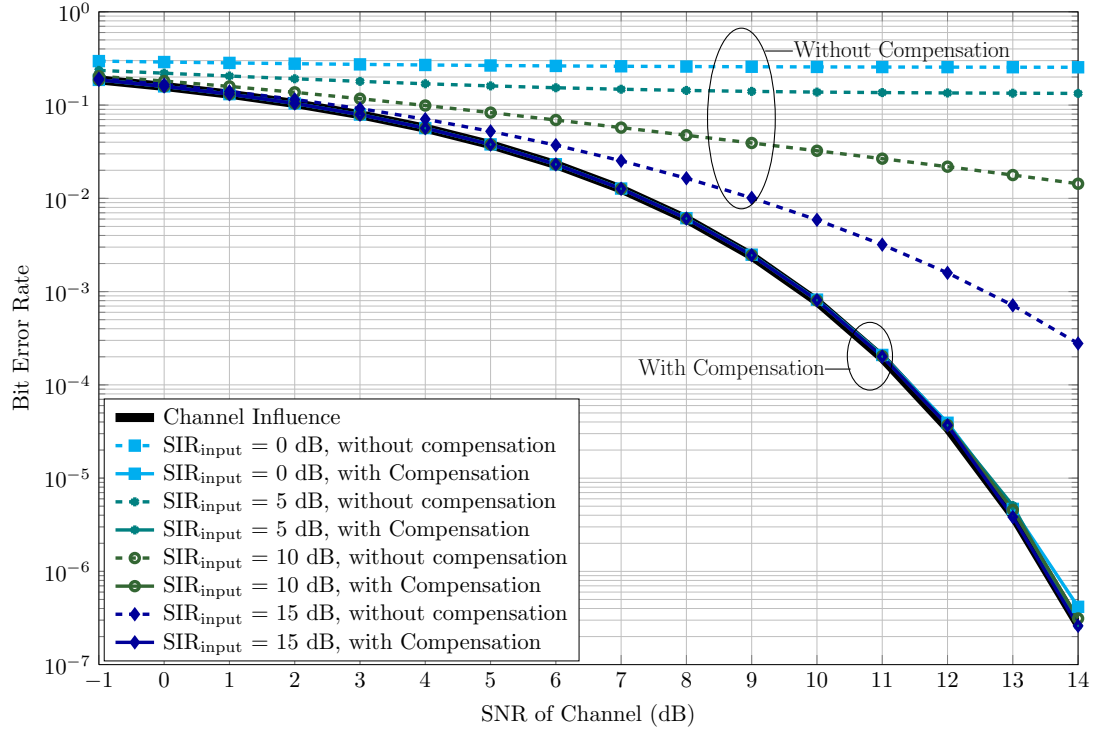


Figure 12: Bit error rate of simple QPSK modulation versus SNR for two spurs of varying SIR, with Brownian PN and affected by a FS. The BER is computed with and without compensation.

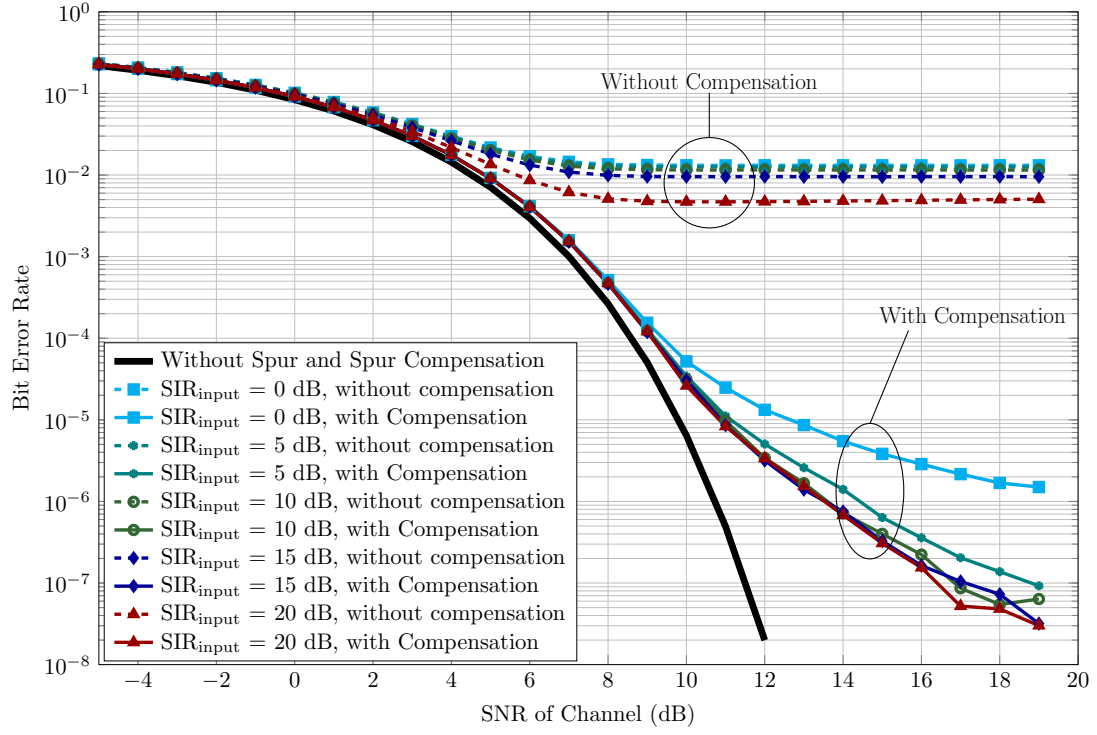


Figure 13: Bit error rate of LTE with signal bandwidth of 1.4 MHz, versus the SNR for two spurs of varying SIR, with Brownian PN and affected by a FS. The BER is computed with and without compensation.

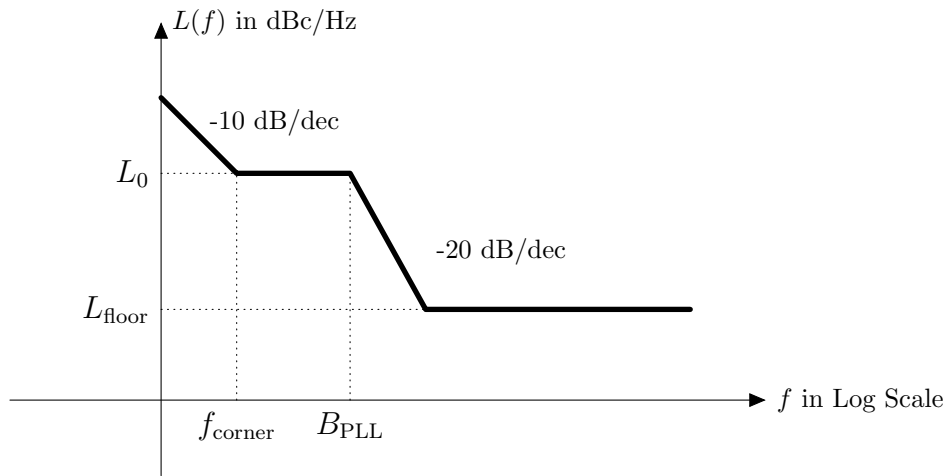


Figure 14: Double sideband phase noise in dBc/Hz

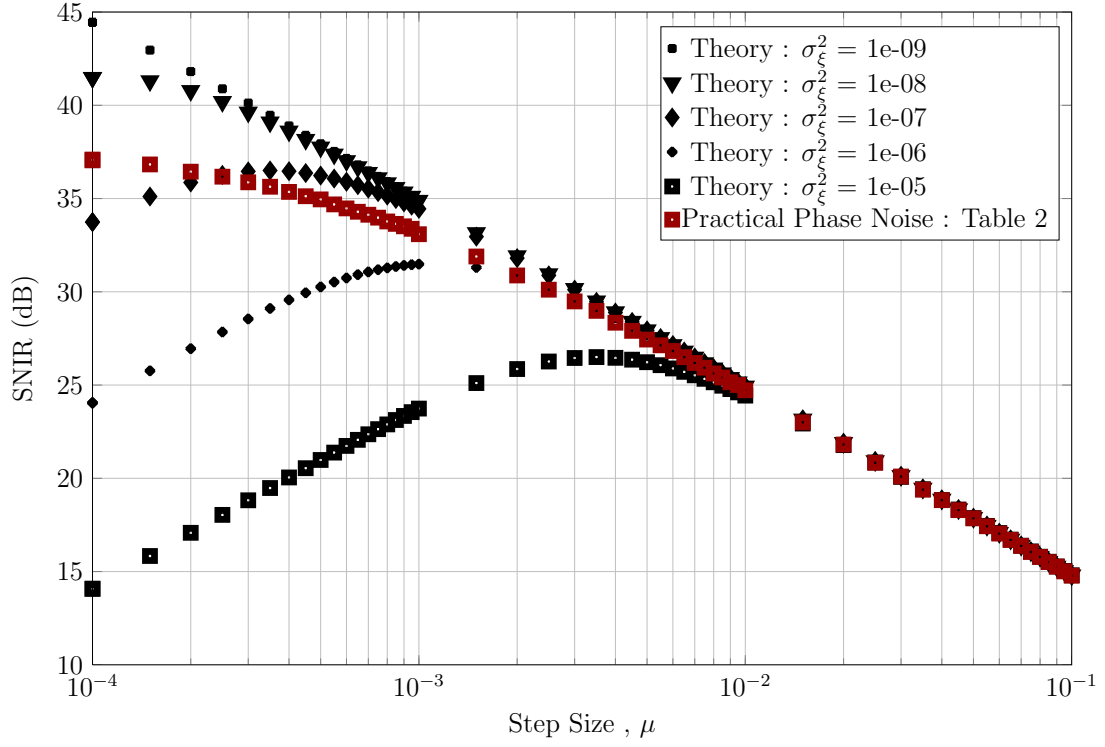


Figure 15: Comparizon between the theoretical performance for a Brownian model (32) and simulated performance with phase noise described in Table 2

Index_{<i>i</i>}	ω_i	Input SNIR (dB)	B	PN : $\sigma_{\xi_i}^2$	FS : δ_{ω_i}
1	$0.19 \cdot 2\pi$	10	0.9	10^{-6}	10^{-6}
2	$0.22 \cdot 2\pi$	15	0.85	10^{-8}	10^{-5}
3	$0.4 \cdot 2\pi$	5	1.2	10^{-6}	10^{-7}
4	$0.6 \cdot 2\pi$	7	1	10^{-6}	10^{-6}
5	$0.8 \cdot 2\pi$	12	1.1	10^{-5}	10^{-6}

Table 1: Parameters for the multiple-spur simulation in asymptotic mode for various (normalized) frequencies $\omega_i = 2\pi f_{s_i}/F$ (f_{s_i} is the frequency of the i th spur in baseband, and F is the sampling frequency)..

Phase Noise Parameter	Value	Unit
L_0	-95	dBc/Hz
f_{corner}	1	kHz
B_{PLL}	100	kHz
L_{Floor}	-150	dBc/ Hz

Table 2: Specification example for phase noise (see [32, section 2.3])