

Hierarchical Data Decision Aided 2-Source BPSK H-MAC Channel Phase Estimator with Feed-Back Gradient Solver for WPNC Networks

Jan Sykora

Abstract—This paper focuses on the *channel state estimation* (CSE) problem in *parametrized Hierarchical MAC (H-MAC)* stage in Wireless Physical Layer Network Coding (WPNC) networks with Hierarchical Decode and Forward (HDF) relay strategy. We derive a non-pilot based H-MAC channel phase estimator for 2 BPSK alphabet sources. The CSE is aided only by the knowledge of H-data decisions. At HDF relay, there is no information on individual source symbols available. The estimator is obtained by a marginalization over the hierarchical dispersion. The estimator uses a gradient additive update solver and the indicator function (gradient) is derived in exact closed form and in approximations for low and high SNR. We analyze the properties of the equivalent solver model, particularly the equivalent gradient detector characteristics and its main stable domain properties under a variety of channel parameterization scenarios.

I. INTRODUCTION

A. Background and Context

Wireless Physical Layer Network Coding (WPNC) is a PHY layer concept for communications in dense radio networks with highly interacting signals. We have reached a relatively solid understanding of some fundamental limits and the system design and performance analysis in basic topologies and scenarios [1], [2], [3], [4], [5]. This includes knowledge of achievable and converse rates for specific strategies and scenarios (e.g. Compute & Forward, Noisy Network Coding, Hierarchical Decode and Forward (HDF)), design of Network Coded Modulation (NCM) and related hierarchical demodulation and decoding strategies including their performance analysis. Most of the results assume relatively idealized assumptions related to *relative channel parametrization* and *NCM codebooks* used at source nodes.

This paper focuses on the *channel state estimation* (CSE) problem in *parametrized Hierarchical MAC (H-MAC)* stage¹. The H-decoding and H-demodulation performance in H-MAC depends strongly on the knowledge of the channel parametrization. This holds generally for all communications systems. However in the case of WPNC, particularly with HDF strategy, the *relative* (among all source nodes) channel parametrization becomes of paramount importance, since (apart of significant performance consequences) it affects the overall structure

of the hierarchical (soft-output) demodulator (H-SODEM). Namely, it affects the shape of H-constellation which is a highly complicated nonlinear function of the relative fading. This problem becomes even more significant in the *randomly time-variant fading channels*, where traditional pilot based approaches have limited applicability. We develop a non-pilot based H-MAC channel state estimator which provides good tracking performance in time-variant environments. Due to the paper length restriction, we constrain the exposition of background WPNC theory and terminology/notation at the minimum necessary level needed for the contribution of the paper. Thorough additional details can be found in [1].

B. Goals and Contributions

For a classical single user (orthogonally separated) communications, the pilot based CSE is a widely adopted and relatively easily implementable concept. In the WPNC networks, it has however a limited applicability. In order to obtain acceptable accuracy of the estimates (see e.g. Cramer-Rao lower bound analysis in [1]), the pilots from individual sources must be orthogonal. This is however in a strong contradiction of the whole WPNC concept which uses mutually interacting *non-orthogonal* signals. Orthogonal resources can be created only by time-division — letting only one source active at given time. In a randomly time-varying channel environments, this would dramatically decrease the duty rate of data payload and thus the overall efficiency of the system. Moreover the H-decoding in H-MAC is quite critically dependent on the CSE quality, particularly for the *relative phase*. Under these conditions, the CSE performed directly on the random data payload stream is a viable solution. Apart of avoiding, or minimizing the need for pilots (possibly only used for the ambiguity resolution), it also opens a natural way for CSE algorithms to continuously track the time-varying H-MAC channel.

The H-MAC CSE in the WPNC networks is significantly different from the classical CSE cases. Most importantly, we estimate the channel parameters of several nonorthogonal signals received in superposition. This makes the problem a *multidimensional composite* estimation problem of several channel parameters over a random payload data stream. There are several options how to approach this problem, ranging from non-data aided (where we marginalize the observation likelihood over a priori data statistic), through various data

WiMob-2018, A1.1 (3.9.2018).

jan.sykora@fel.cvut.cz, Czech Technical University in Prague. This work was supported by the Ministry of education, youth and sports of the Czech Republic, grant LTC17042.

¹A prefix ‘H-’ denotes *hierarchical* entities, i.e. many-to-one functions of source node entities (see [1] for details).

soft-information aided forms, to decision aided forms (where the decisions are provided from the decoder).

The data decision aided strategies have one specific problem in WPNC networks. The core principle underlying the whole treatment (otherwise, it would be quite straightforward) is the fact that we do *not* have available *individual* component source symbols s_A, s_B (nor messages b_A, b_B) decisions available at the relay. This is a fundamental principle of HDF strategy. The whole system is operated from the information-theoretic point of view at the operational point where we do not have a chance to have reliable decoded individual component messages b_A, b_B (and consequently all c_A, c_B , and s_A, s_B) but *only* their *many-to-one* HNC map $c = \chi_c(c_A, c_B)$ which is called *H-data* (or H-symbols).

The CSE may be thus aided only by H-data information. Due to the many-to-one HNC map property, there is still a randomness in the payload data even for fixed given H-data knowledge. This is called H-dispersion and CSE will have to marginalize it from its utility function.

II. SYSTEM MODEL

We consider a 2-component (2 source nodes SA, SB, one relay) H-MAC topology (Fig. 1) with *isomorphic layered NCM* and symbol-timing synchronized nodes. Component messages are b_A, b_B and the target on the relay is to decode a hierarchical message $b = \chi(b_A, b_B)$ which is a many-to-one function (Hierarchical Network Code (HNC) map) of the source messages. We consider only the processing on one relay and assume that the end-to-end solvability is fulfilled by a proper choice of all remaining components of WPNC network (see [1] for end-to-end solvability conditions). The component messages are encoded by a *common* codebook \mathcal{C} , $\mathbf{c}_A = \mathcal{C}(b_A)$, $\mathbf{c}_B = \mathcal{C}(b_B)$. The coded symbols are symbol-wise mapped on the channel constellation points $s_{A,n} = \mathcal{A}_s(c_{A,n})$, $s_{B,n} = \mathcal{A}_s(c_{B,n})$. \mathcal{A}_s is the source 1-dimensional constellation alphabet with size M_s .

The isomorphic property of NCM guarantees that H-code symbols $\mathbf{c} = \mathcal{C}(b)$ are HNC maps of component coded symbols $c_n = \chi_c(c_{A,n}, c_{B,n})$. In this paper, we consider *minimal* HNC map χ_c , i.e. such that given the knowledge of H-symbol c_n and any of component symbols $c_{A,n}$ (or $c_{B,n}$), we can uniquely determine the other source symbol $c_{B,n}$ (or $c_{A,n}$).

The H-MAC memoryless AWGN channel is modeled as

$$\mathbf{x} = h_A \mathbf{s}_A(\mathbf{c}_A) + h_B \mathbf{s}_B(\mathbf{c}_B) + \mathbf{w} \quad (1)$$

where $h_A = \eta_A e^{j\varphi_A}$, $h_B = \eta_B e^{j\varphi_B}$ are fading coefficients and \mathbf{w} is AWGN vector with σ_w^2 variance per dimension. The observation vector has length N . Signal to noise ratio (SNR) is defined w.r.t. source A signal and it is $\gamma_x = E[|s_A|^2]/\sigma_w^2$.

Since the channel is memoryless, the noise is white, and the constellation mappers are symbol-wise, the complete frame observation likelihood is $p(\mathbf{x}|\varphi_A, \varphi_B, \mathbf{c}_A, \mathbf{c}_B) = \prod_n p(x|\varphi_A, \varphi_B, c_{A,n}, c_{B,n})$. The symbol-wise likelihood is²

$$p(x|\varphi_A, \varphi_B, c_A, c_B) \equiv e^{-\frac{1}{\sigma_w^2} |x - \eta_A e^{j\varphi_A} s_A(c_A) - \eta_B e^{j\varphi_B} s_B(c_B)|^2} \quad (2)$$

²For a notational simplicity (unless it is explicitly needed), we drop the explicit sequential index n notation, e.g. $x \triangleq x_n$, and similarly others.

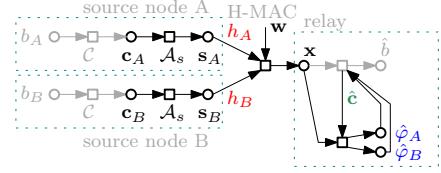


Figure 1. 2-component H-MAC system.

where we dropped unimportant scaling in the Gaussian density. Channel coefficient magnitudes $\eta_A, \eta_B \in \mathbb{R}^+$ are assumed to be known at the relay.

In this paper, we focus on a *special* case of both sources having BPSK alphabets $\mathcal{A}_s = \{\pm 1\}$, i.e. $s_A(c_A) = 2c_A - 1$, $c_A \in \{0, 1\}$, and similarly for s_B . In the case of binary symbol-wise HNC map, the only possible choice fulfilling the minimal cardinality condition is the XOR map $c = c_A \oplus c_B$, $c \in \{0, 1\}$.

III. H-DATA DECISION AIDED H-MAC CSE

We derive the *H-data decision aided* ML (Maximum Likelihood) CSE for 2-source BPSK H-MAC. The estimator does *not* have available the individual messages b_A, b_B (nor c_A, c_B , and s_A, s_B) but *only* their many-to-one HNC map $c = \chi_c(c_A, c_B)$. This is provided by the H-decoder (Fig. 1), and we assume that the estimates are reliable $\hat{\mathbf{c}} = \mathbf{c}$. The random hierarchical dispersion [1] remains unknown. It must be marginalized from the observation likelihood. Under special conditions (namely the minimal HNC map), we will be able to model its stochastic properties needed for the marginalization.

A. H-Dispersion Marginalization for 2-Source BPSK

First, we perform a simple manipulation with symbol-wise observation likelihood. We use the fact that the source constellation mappers $s_A(c_A), s_B(c_B)$ are one-to-one functions, and both of them are from the PSK class of alphabets $|s_A| = |s_B| = 1$

$$p(x|\varphi_A, \varphi_B, s_A, s_B) \equiv \exp \left(-\frac{2}{\sigma_w^2} \eta_A \eta_B \Re \left[e^{j(\varphi_A - \varphi_B)} s_A s_B^* \right] \right) \\ \times \exp \left(\frac{2}{\sigma_w^2} \left(\eta_A \Re \left[x e^{-j\varphi_A} s_A^* \right] + \eta_B \Re \left[x e^{-j\varphi_B} s_B^* \right] \right) \right). \quad (3)$$

We dropped all scaling which do not depend on $\varphi_A, \varphi_B, s_A, s_B$.

The marginalization of the H-dispersion means a marginalization over all source symbols consistent with H-symbol, i.e. $\sum_{(s_A, s_B):c} f(s_A, s_B)$ where $f(\cdot)$ is the marginalized entity. The minimal HNC map defines a unique value of s_B for a given pair of s_A, c . For the 2-source BPSK NCM with XOR map, it is

$$s_B = s_B(c_B(c, c_A)) = s_B(c \oplus c_A) = (1 - 2c)s_A. \quad (4)$$

Then the H-dispersion marginalization can be evaluated as

$$p(x|\varphi_A, \varphi_B, c) = \frac{1}{p(c)} \sum_{(s_A, s_B):c} p(x|\varphi_A, \varphi_B, s_A, s_B) p(s_A) p(s_B) \\ = \frac{1}{p(c)} \sum_{s_A} p(x|\varphi_A, \varphi_B, s_A, s_B(s_A, c)) p(s_A) p(s_B). \quad (5)$$

The minimal map together with IID source symbols guarantee [1] that $p(s_A) = p(s_B) = p(c) = 1/M_s$.

We define *cross inner product symbol* $s = s_A s_B^*$ and a class of NCM called *H-symbol inner product unique* ones, that fulfill the property that the inner product symbol s is uniquely defined by the H-symbol c , i.e. $s(c)$ is one-to-one function. The 2-source BPSK NCM with XOR map fulfills that and it holds

$$s(c) = s_A s_B^* = (1 - 2c)s_A^2 = 1 - 2c. \quad (6)$$

We also see, that for any PSK alphabet $s_B^* = s(c)s_A^*$.

The marginalized H-symbol conditioned channel phases likelihood is then (dropping all unimportant scaling)

$$\begin{aligned} p(x|\varphi_A, \varphi_B, c) &\equiv \\ &\equiv \sum_{s_A \in \{\pm 1\}} \exp \left(-\frac{2}{\sigma_w^2} \eta_A \eta_B \Re \left[e^{j(\varphi_A - \varphi_B)} s(c) \right] \right) \\ &\quad \times \exp \left(\frac{2}{\sigma_w^2} \left(\eta_A \Re \left[x e^{-j\varphi_A} s_A^* \right] + \eta_B \Re \left[x e^{-j\varphi_B} s(c) s_A^* \right] \right) \right) \\ &= \exp \left(-\frac{2}{\sigma_w^2} \eta_A \eta_B s(c) \cos(\varphi_A - \varphi_B) \right) \\ &\quad \times 2 \cosh \left(\frac{2}{\sigma_w^2} \left(\eta_A \Re \left[x e^{-j\varphi_A} \right] + \eta_B s(c) \Re \left[x e^{-j\varphi_B} \right] \right) \right). \end{aligned} \quad (7)$$

The phase estimator metric ρ is defined as a logarithm of scaled (dropping all unimportant scaling) $p(x|\varphi_A, \varphi_B, c)$, i.e.

$$\begin{aligned} \rho &= -\frac{2}{\sigma_w^2} \eta_A \eta_B s(c) \cos(\varphi_A - \varphi_B) \\ &\quad + \ln \cosh \left(\frac{2}{\sigma_w^2} \left(\eta_A \Re \left[x e^{-j\varphi_A} \right] + \eta_B s(c) \Re \left[x e^{-j\varphi_B} \right] \right) \right). \end{aligned} \quad (8)$$

For the whole observation frame, we have $\rho_N(\varphi_A, \varphi_B, \mathbf{c}) = \sum_{n=1}^N \rho(\varphi_A, \varphi_B, c_n)$.

The ML CSE is then

$$\hat{\boldsymbol{\varphi}} = [\hat{\varphi}_A, \hat{\varphi}_B] = \arg \min_{\varphi_A, \varphi_B} \rho_N(\varphi_A, \varphi_B, \mathbf{c}) \quad (9)$$

for H-symbols \mathbf{c} provided by the H-decoder.

B. Feed-Back Gradient Solver

Clearly, a direct solution of ML CSE is unfeasible. We need to resort to the feed-back iterative solver based on the gradient based additive updates. At i th iteration, the estimate update is

$$\hat{\boldsymbol{\varphi}}(i+1) = \hat{\boldsymbol{\varphi}}(i) + K \nabla_{\boldsymbol{\varphi}} \rho_N(\hat{\boldsymbol{\varphi}}(i), \mathbf{c}) \quad (10)$$

where K is the loop gain. The gradient indicator has symbol-wise components $\mu_A = \partial \rho / \partial \varphi_A$ and $\mu_B = \partial \rho / \partial \varphi_B$, $\boldsymbol{\mu} = [\mu_A, \mu_B]$. The complete frame observation gradient is $\boldsymbol{\mu}_N = \nabla_{\boldsymbol{\varphi}} \rho_N(\hat{\boldsymbol{\varphi}}(i), \mathbf{c}) = \sum_n \mu(\hat{\boldsymbol{\varphi}}(i), c_n)$.

Performing the derivatives, we get

$$\begin{aligned} \mu_A &= \frac{2}{\sigma_w^2} \eta_A \eta_B s(c) \sin(\varphi_A - \varphi_B) + \frac{2}{\sigma_w^2} \eta_A \Im \left[x e^{-j\varphi_A} \right] \\ &\quad \times \tanh \left(\frac{2}{\sigma_w^2} \left(\eta_A \Re \left[x e^{-j\varphi_A} \right] + \eta_B s(c) \Re \left[x e^{-j\varphi_B} \right] \right) \right), \end{aligned} \quad (11)$$

$$\begin{aligned} \mu_B &= \frac{2}{\sigma_w^2} \eta_A \eta_B s(c) \sin(\varphi_B - \varphi_A) + \frac{2}{\sigma_w^2} \eta_B \Im \left[x e^{-j\varphi_B} \right] \\ &\quad \times \tanh \left(\frac{2}{\sigma_w^2} \left(\eta_B \Re \left[x e^{-j\varphi_B} \right] + \eta_A s(c) \Re \left[x e^{-j\varphi_A} \right] \right) \right). \end{aligned} \quad (12)$$

The expressions are, as expected, symmetric w.r.t. φ_A and φ_B , and are the exact expressions including a proper scaling (needed e.g. for Cramer-Rao lower bound analysis). A practical implementation of these expressions calls for approximations avoiding complicated nonlinear functions. In the next, we determine approximation for low and high SNR.

For low SNR scenarios, we can assume that η_A^2/σ_w^2 and η_B^2/σ_w^2 are very small, and thus $\eta_A x / \sigma_w^2 \ll 1$ and $\eta_B x / \sigma_w^2 \ll 1$. Then we can use the approximation $\tanh(A) \approx A$. The *low SNR gradient indicator approximation* is then (after some manipulations)

$$\begin{aligned} \mu_{A0} &= \frac{2}{\sigma_w^4} \eta_A \eta_B s(c) \sin(\varphi_A - \varphi_B) (\sigma_w^2 - |x|^2) \\ &\quad + \frac{2}{\sigma_w^4} \eta_A^2 \Im \left[x^2 e^{-j2\varphi_A} \right] + \frac{2}{\sigma_w^4} \eta_A \eta_B s(c) \Im \left[x^2 e^{-j(\varphi_A + \varphi_B)} \right], \end{aligned} \quad (13)$$

$$\begin{aligned} \mu_{B0} &= \frac{2}{\sigma_w^4} \eta_A \eta_B s(c) \sin(\varphi_B - \varphi_A) (\sigma_w^2 - |x|^2) \\ &\quad + \frac{2}{\sigma_w^4} \eta_B^2 \Im \left[x^2 e^{-j2\varphi_B} \right] + \frac{2}{\sigma_w^4} \eta_A \eta_B s(c) \Im \left[x^2 e^{-j(\varphi_A + \varphi_B)} \right]. \end{aligned} \quad (14)$$

For high SNR scenarios, we can assume that η_A^2/σ_w^2 and η_B^2/σ_w^2 are large, and thus $\eta_A x / \sigma_w^2 \gg 1$ and $\eta_B x / \sigma_w^2 \gg 1$. Then we can use the approximation $\tanh(A) \approx \text{sign}(A)$. The *high SNR gradient indicator approximation* is then

$$\begin{aligned} \mu_A &= \frac{2}{\sigma_w^2} \eta_A \eta_B s(c) \sin(\varphi_A - \varphi_B) + \frac{2}{\sigma_w^2} \eta_A \Im \left[x e^{-j\varphi_A} \right] \\ &\quad \times \text{sign} \left(\eta_A \Re \left[x e^{-j\varphi_A} \right] + \eta_B s(c) \Re \left[x e^{-j\varphi_B} \right] \right), \end{aligned} \quad (15)$$

$$\begin{aligned} \mu_B &= \frac{2}{\sigma_w^2} \eta_A \eta_B s(c) \sin(\varphi_B - \varphi_A) + \frac{2}{\sigma_w^2} \eta_B \Im \left[x e^{-j\varphi_B} \right] \\ &\quad \times \text{sign} \left(\eta_B \Re \left[x e^{-j\varphi_B} \right] + \eta_A s(c) \Re \left[x e^{-j\varphi_A} \right] \right). \end{aligned} \quad (16)$$

C. Equivalent Model of Feed-Back Solver

The performance analysis of the feed-back iterative solver with additive updates operating under time-variant channel parameter conditions and under random noise excitation is a complicated task, namely due to a combination of two phenomena — the nonlinearity and the random excitation. In our case, it is even more complex due to the multidimensional tracked entity (φ_A, φ_B) . The analysis of the dynamic nonlinear state-space systems with random excitation is, on its own, an extremely wide field and exposing all details is completely out of the scope of this paper (see e.g. [6], [7], [8]).

Usually, the analysis is approached by forming an *equivalent model*. The equivalent model is formed in such a way that the variables of the interest, typically the tracking error $\Delta\boldsymbol{\varphi} = \hat{\boldsymbol{\varphi}} - \boldsymbol{\varphi}$, are *directly* processed as an *open* state variables. This is in contrast with the true system, where these variables may be buried deeply inside the complex nonlinear functions. Also, the equivalent model tries to separate the nonlinearity and the memory of the system and also considers only additive random excitation. Of course, with exception of some very specific scenarios, this all makes the equivalent model just an approximation, however the approximation that is mathematically tractable by solving stochastic differential equations using Fokker-Planck equation.

For the purpose of this paper, we concentrate only on the most important property of the equivalent model, and this is the *equivalent characteristic of the gradient indicator function* (frequently also simply called *equivalent detector characteristics*). It is defined as a mean value of the indicator function (gradient) for a fixed local estimate parameter $\hat{\boldsymbol{\varphi}}$ and for fixed value of the true channel $\boldsymbol{\varphi}$ observed in the received signal $x(\boldsymbol{\varphi})$

$$g(\boldsymbol{\varphi}, \hat{\boldsymbol{\varphi}}) = E_{w,s} [\boldsymbol{\mu}(\hat{\boldsymbol{\varphi}})|_{x=x(\boldsymbol{\varphi})}]. \quad (17)$$

The expectation can be evaluated in a closed-form expression or numerically using Monte Carlo simulation.

The most important properties that we want to investigate are (1) the position of the main stable node, which should ideally give $g(\boldsymbol{\varphi}, \boldsymbol{\varphi}) = \mathbf{0}$ for any $\boldsymbol{\varphi}$, and (a) the gradient direction $g(\boldsymbol{\varphi}, \boldsymbol{\varphi} + \Delta\boldsymbol{\varphi})$ in the plane of all $\Delta\boldsymbol{\varphi}$, which should ideally direct to the main stable node (the origin) $\Delta\boldsymbol{\varphi} = \mathbf{0}$ for any $\boldsymbol{\varphi}$. Since the feed-back estimators have typically many stable nodes, this requirement is usually relaxed to just a reasonable neighborhood of the main stable node. The resulting ambiguity needs to be resolved by external means (e.g. pilots).

The equivalent detector characteristic and gradients under various SNRs, parameter values, and for exact and low/high SNR approximations are shown in Figs. 2,3,4,5,6,7,8,9. We also show an example of received constellation points for random values of source symbols and AWGN.

IV. DISCUSSION AND CONCLUSIONS

A. H-decision Aided H-MAC 2-source Phase Estimator

We have derived the H-MAC channel phase estimator for 2 BPSK source nodes, which is aided only by the knowledge of H-data decision, i.e. no information on individual source symbols is neither needed nor available at the relay. The estimator is obtained by a marginalization over the hierarchical dispersion. The estimator uses gradient additive update solver and the indicator function (gradient) is derived in exact closed form and in an approximation for low and high SNR.

B. Good Tracking Capabilities of Feed-back Gradient Solver

The (C1) parts of Figs. 2,3,4,5,6,7,8,9 show the exact solution gradients. We see that the attraction domain of the

main stable node (i.e. the region in $\Delta\boldsymbol{\varphi}$ plane, where gradient aims in a correct direction towards the main stable node $\Delta\boldsymbol{\varphi} = \mathbf{0}$) is relatively very wide. As expected there are other stable nodes (not the main one) since the H-constellation and source constellations are rotationally invariant. This is unavoidable in the random data case. It can be resolved only by using rotationally unique pilots or codebook. However, it is instructive to see that the pattern of other stable nodes is (unlike for classical single user PSK case) irregular and depends on true H-MAC channel phases.

C. Exact and Low/High SNR Approximation

The exact gradient indicator function solution exhibits desired behavior, mainly the presence of the main stable node at $\Delta\boldsymbol{\varphi} = \mathbf{0}$. However the expression contains complex nonlinear functions which makes its practical implementation difficult. We have investigated *low* SNR approximation, which contains only quadratic functions of received signal. However this approximation exhibits a problematic behavior — its main stable node is *not* located at the origin (see parts (B) and row#2 of parts (C) in Figs. 2,3,4,5,6,7,8,9). This appears to be caused by an asymmetry in the approximation in two gradient contribution parts that are added together. The first one is left intact but the other one is linearized which clearly makes an imbalance between those two contributions dependent on actual H-MAC phases. On the other hand, the *high* SNR approximation performs surprisingly well even for moderate-low SNRs. It still contains a nonlinear function, but only easy-to-be-handled sign.

D. Invariance to True Channel Phase

Invariance of the estimator solver to the actual channel phase is a property that greatly simplifies the performance analysis. It means that the equivalent detector characteristic is a function only of the estimation error $g(\boldsymbol{\varphi}, \hat{\boldsymbol{\varphi}}) = g(\Delta\boldsymbol{\varphi})$. Although all forms (exact and approximations) of gradient expression mix component of the received signal (containing both individual true phases) and the individual phase estimate values, the numerical graphical evaluation of the gradient demonstrates only minor differences (compare Figs. 2,3 and 4,5). However this phenomenon will require further and deeper investigation.

REFERENCES

- [1] J. Sykora and A. Burr, *Wireless Physical Layer Network Coding*. Cambridge University Press, 2018.
- [2] S. Zhang, S. C. Liew, and P. P. Lam, "Hot topic: Physical-layer network coding," in *MobiCom'06*, pp. 358–365, Sept. 2006.
- [3] T. Koike-Akino, P. Popovski, and V. Tarokh, "Optimized constellations for two-way wireless relaying with physical network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, pp. 773–787, June 2009.
- [4] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, pp. 6463–6486, oct 2011.
- [5] J. Sykora and A. Burr, "Layered design of hierarchical exclusive codebook and its capacity regions for HDF strategy in parametric wireless 2-WRC," *IEEE Trans. Veh. Technol.*, vol. 60, pp. 3241–3252, Sept. 2011.
- [6] S. Haykin, *Adaptive Filter Theory*. Prentice-Hall, third ed., 1996.
- [7] U. Spagnolini, *Statistical Signal Processing in Engineering*. John Wiley & Sons, 2017.
- [8] H. Meyr and G. Ascheid, *Synchronization in Digital Communications: Volume 1*. John Wiley & Sons, 1990.

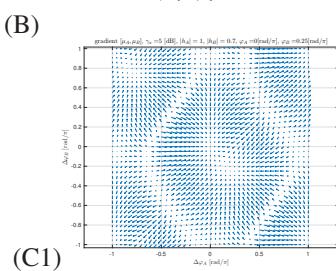
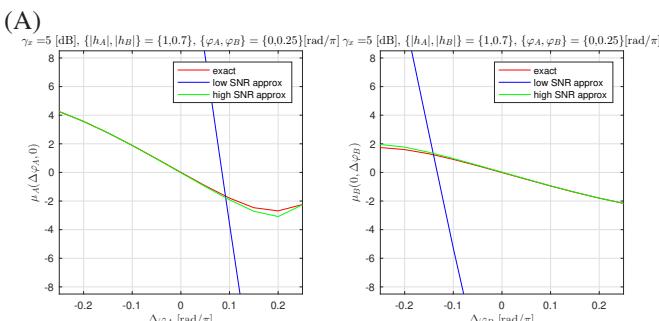
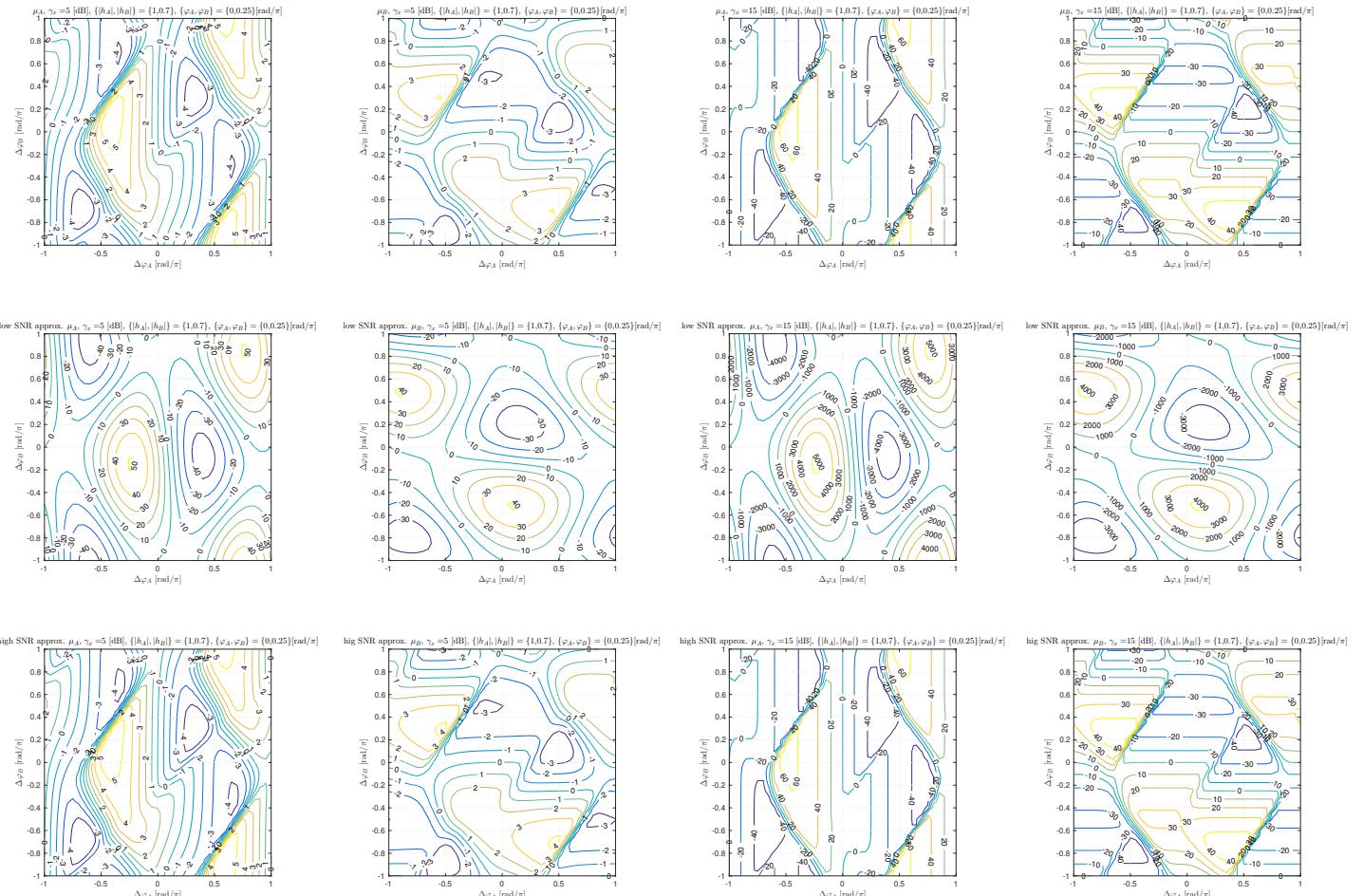


Figure 2. Equivalent detector gradient characteristics for $\gamma_x = 5$ [dB], $h_A = 1 e^{j0}$, $h_B = 0.7 e^{j0.25\pi}$. (A) gradient components $\mu_A(\Delta\phi)$, $\mu_B(\Delta\phi)$ (row #1) exact, (row #2) low SNR approximation, (row #3) high SNR approximation; (B) gradient components $\mu_A(\Delta\phi_A, 0)$, $\mu_B(0, \Delta\phi_B)$; (C1) gradient directions (exact); (C2) random received constellation points.

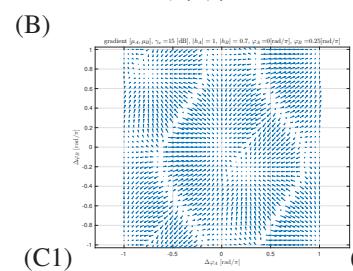
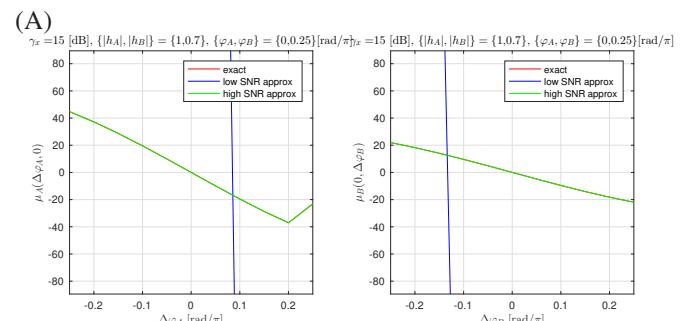


Figure 3. Equivalent detector gradient characteristics for $\gamma_x = 15$ [dB], $h_A = 1 e^{j0}$, $h_B = 0.7 e^{j0.25\pi}$. (A) gradient components $\mu_A(\Delta\phi)$, $\mu_B(\Delta\phi)$ (row #1) exact, (row #2) low SNR approximation, (row #3) high SNR approximation; (B) gradient components $\mu_A(\Delta\phi_A, 0)$, $\mu_B(0, \Delta\phi_B)$; (C1) gradient directions (exact); (C2) random received constellation points.

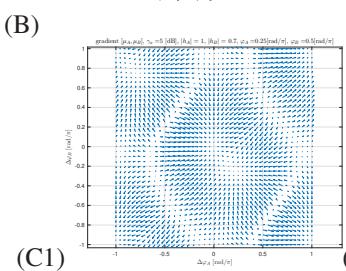
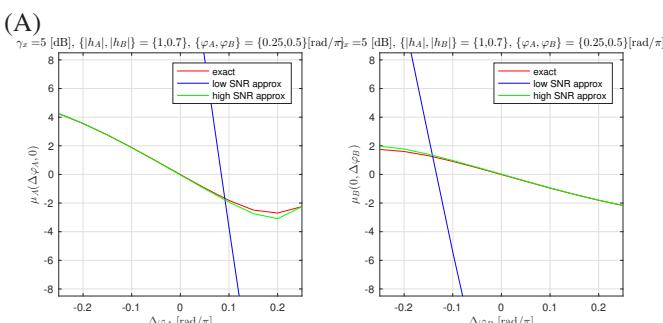
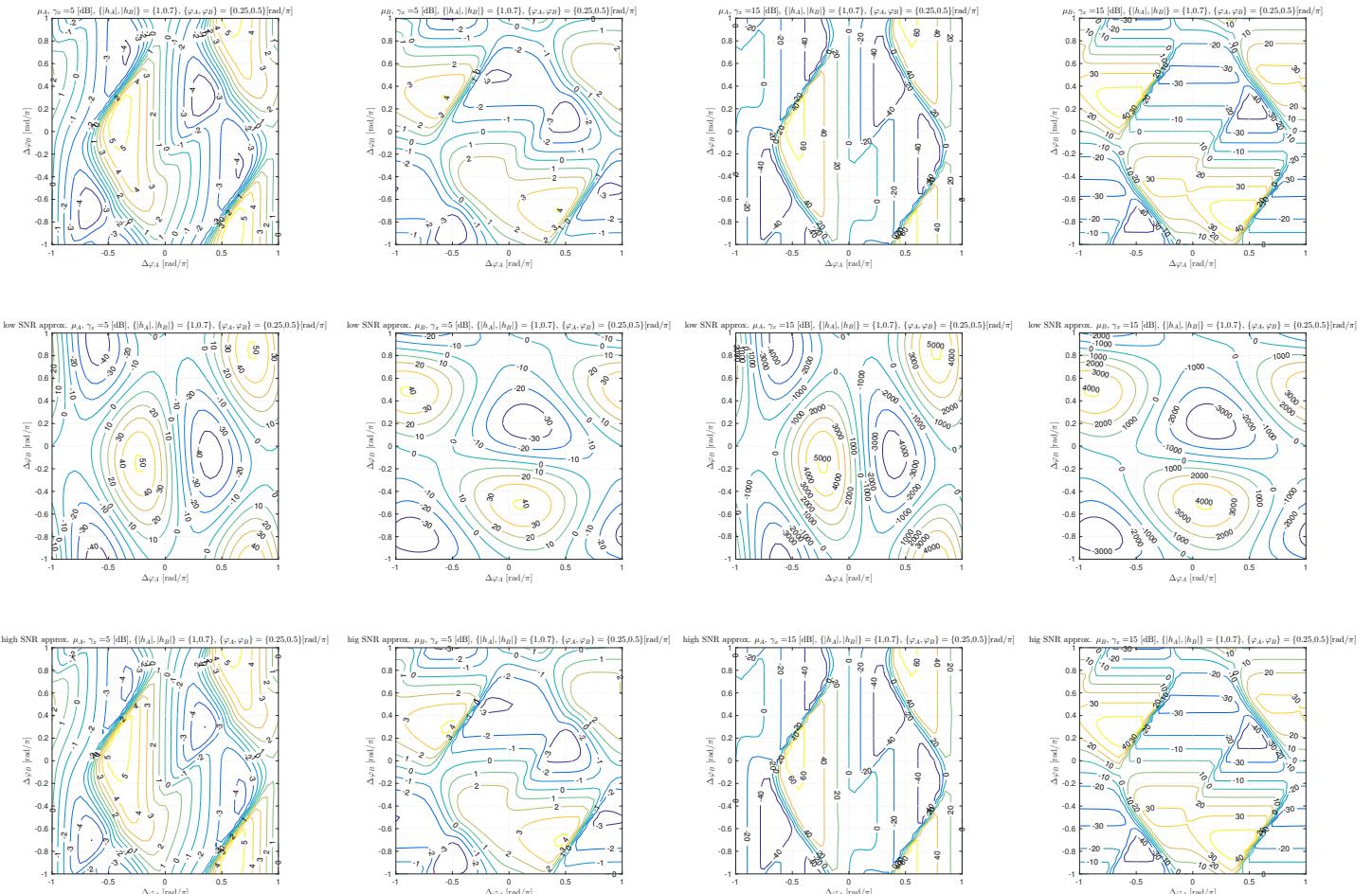


Figure 4. Equivalent detector gradient characteristics for $\gamma_x = 5, $h_A = 1e^{j0.25}$, $h_B = 0.7e^{j0.5\pi}$. (A) gradient components $\mu_A(\Delta\boldsymbol{\varphi})$, $\mu_B(\Delta\boldsymbol{\varphi})$ (row #1) exact , (row #2) low SNR approximation, (row #3) high SNR approximation; (B) gradient components $\mu_A(\Delta\varphi_A, 0)$, $\mu_B(0, \Delta\varphi_B)$; (C1) gradient directions (exact); (C2) random received constellation points.$

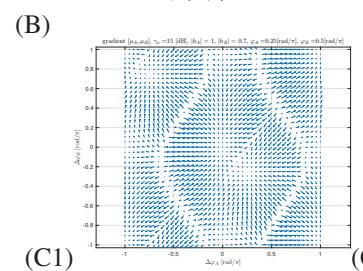
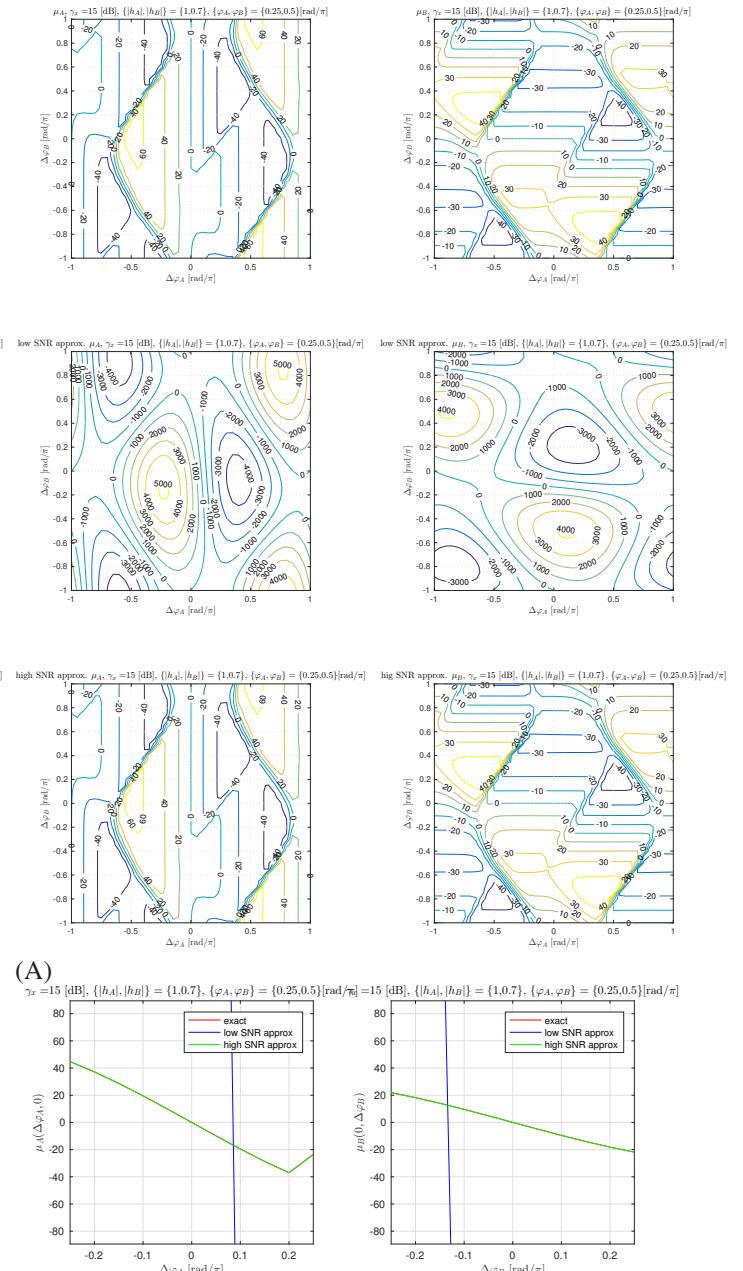


Figure 5. Equivalent detector gradient characteristics for $\gamma_x = 15, $h_A = 1e^{j0.25}$, $h_B = 0.7e^{j0.5\pi}$. (A) gradient components $\mu_A(\Delta\boldsymbol{\varphi})$, $\mu_B(\Delta\boldsymbol{\varphi})$ (row #1) exact , (row #2) low SNR approximation, (row #3) high SNR approximation; (B) gradient components $\mu_A(\Delta\varphi_A, 0)$, $\mu_B(0, \Delta\varphi_B)$; (C1) gradient directions (exact); (C2) random received constellation points.$

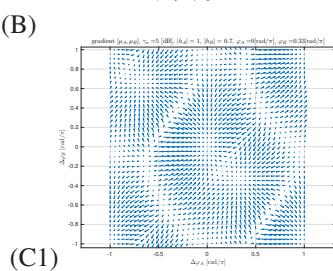
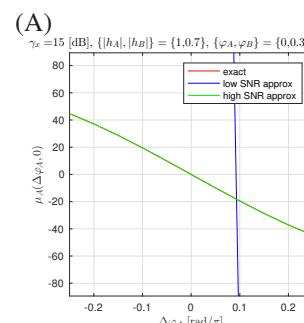
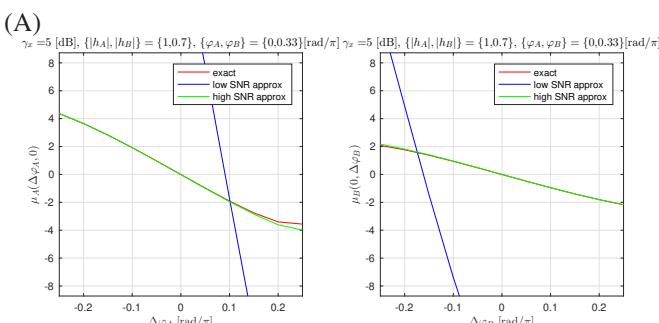
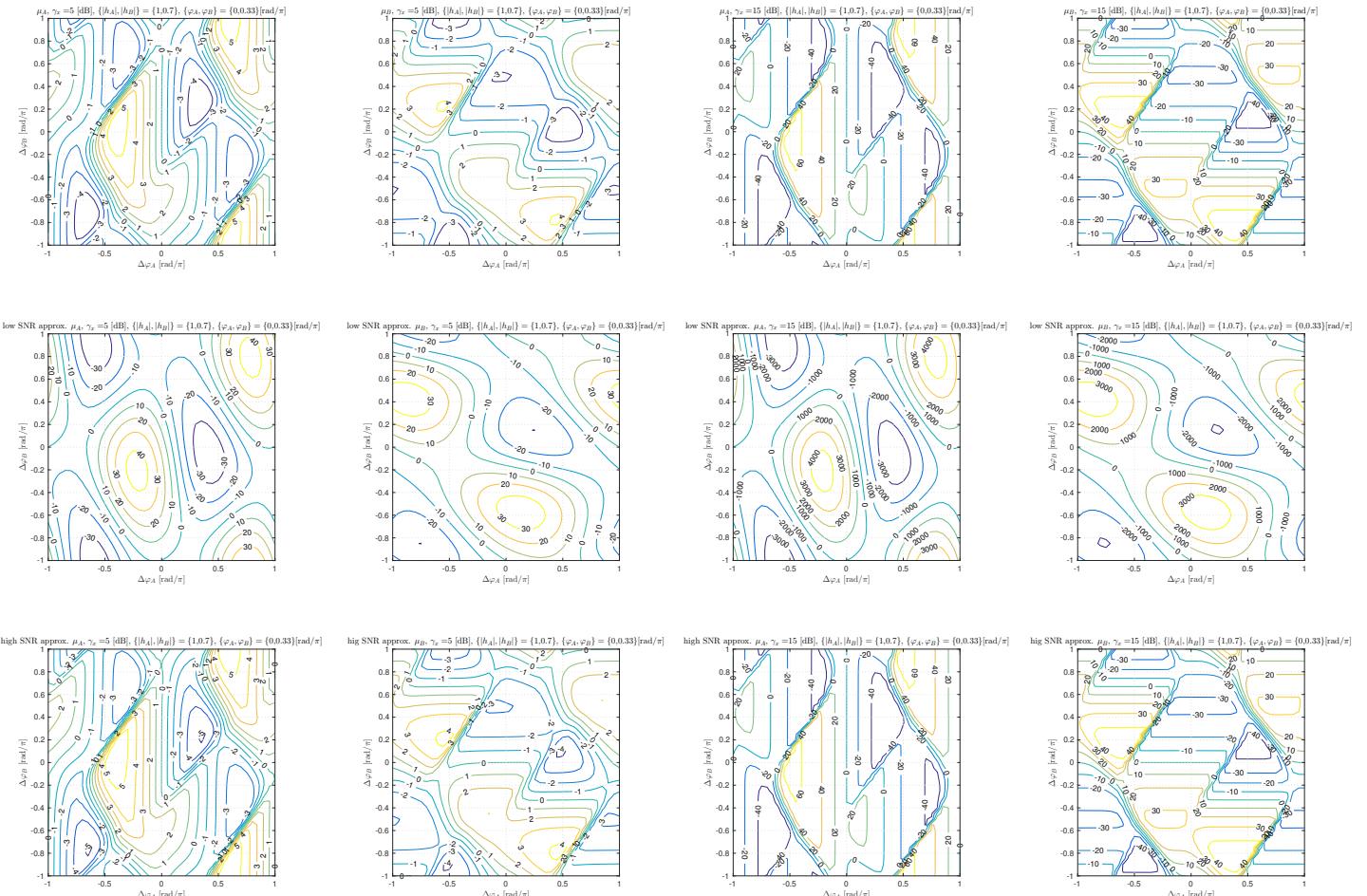


Figure 6. Equivalent detector gradient characteristics for $\gamma_x = 5$ [dB], $h_A = 1e^{j0}$, $h_B = 0.7e^{j0.33\pi}$. (A) gradient components $\mu_A(\Delta\phi)$, $\mu_B(\Delta\phi)$ (row #1) exact, (row #2) low SNR approximation, (row #3) high SNR approximation; (B) gradient components $\mu_A(\Delta\phi_A, 0)$, $\mu_B(0, \Delta\phi_B)$; (C1) gradient directions (exact); (C2) random received constellation points.

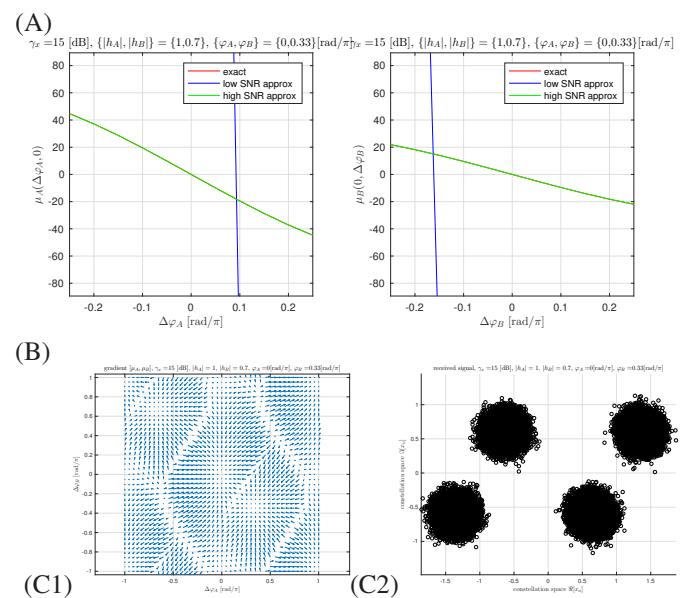
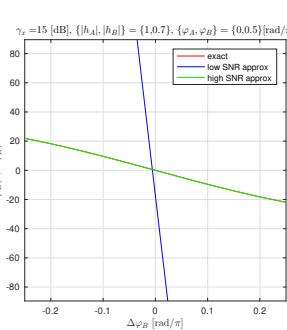
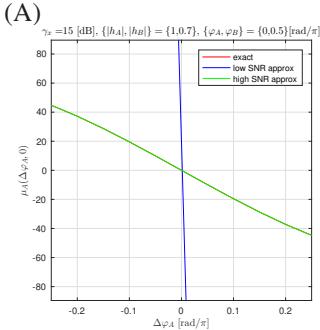
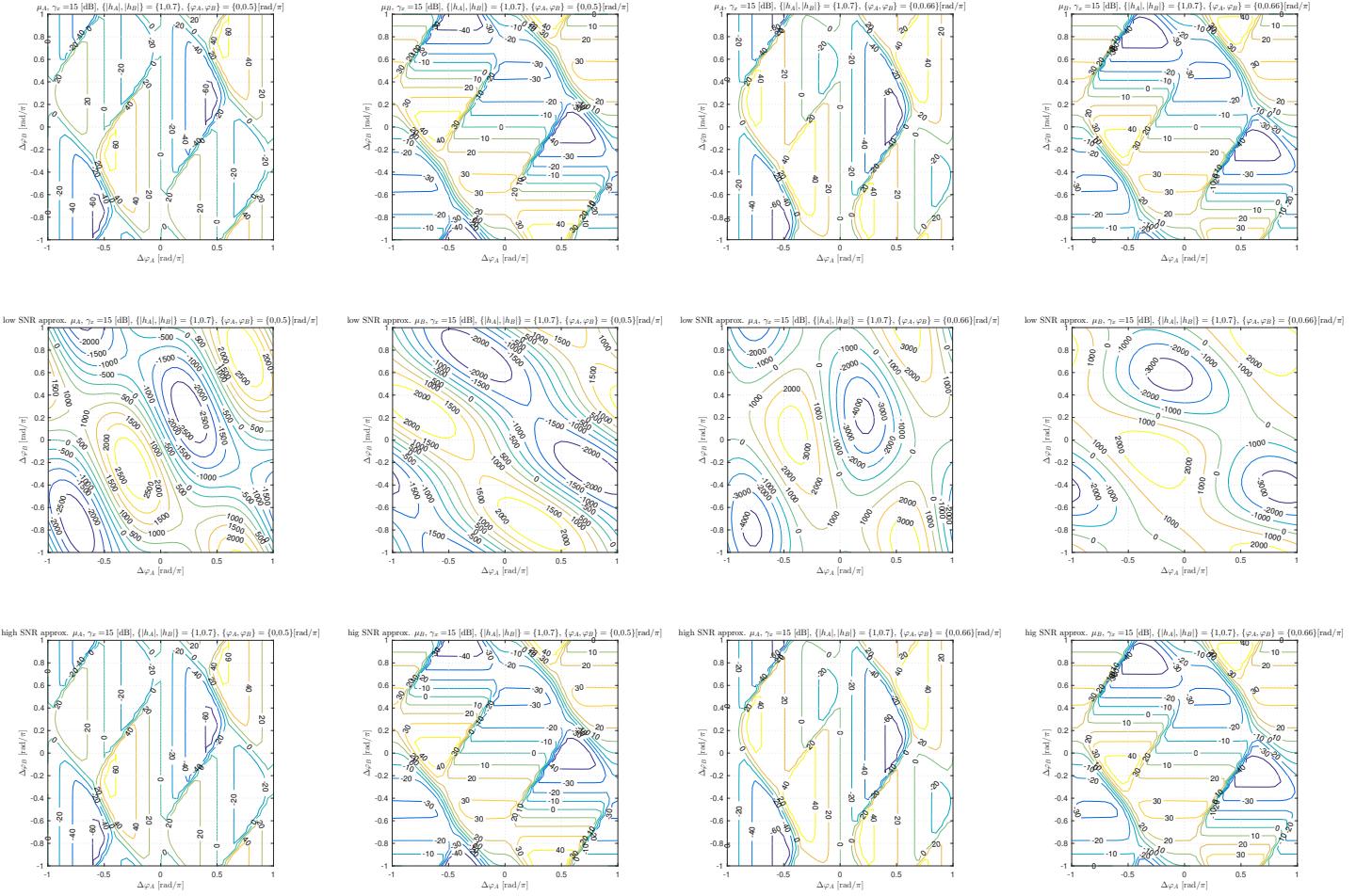
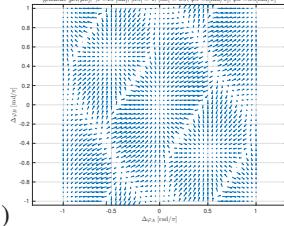


Figure 7. Equivalent detector gradient characteristics for $\gamma_x = 15$ [dB], $h_A = 1e^{j0}$, $h_B = 0.7e^{j0.33\pi}$. (A) gradient components $\mu_A(\Delta\phi)$, $\mu_B(\Delta\phi)$ (row #1) exact, (row #2) low SNR approximation, (row #3) high SNR approximation; (B) gradient components $\mu_A(\Delta\phi_A, 0)$, $\mu_B(0, \Delta\phi_B)$; (C1) gradient directions (exact); (C2) random received constellation points.

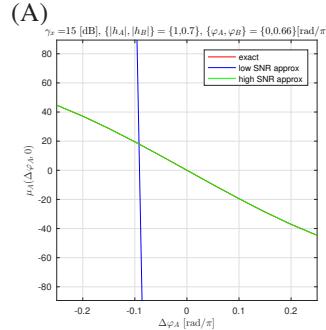
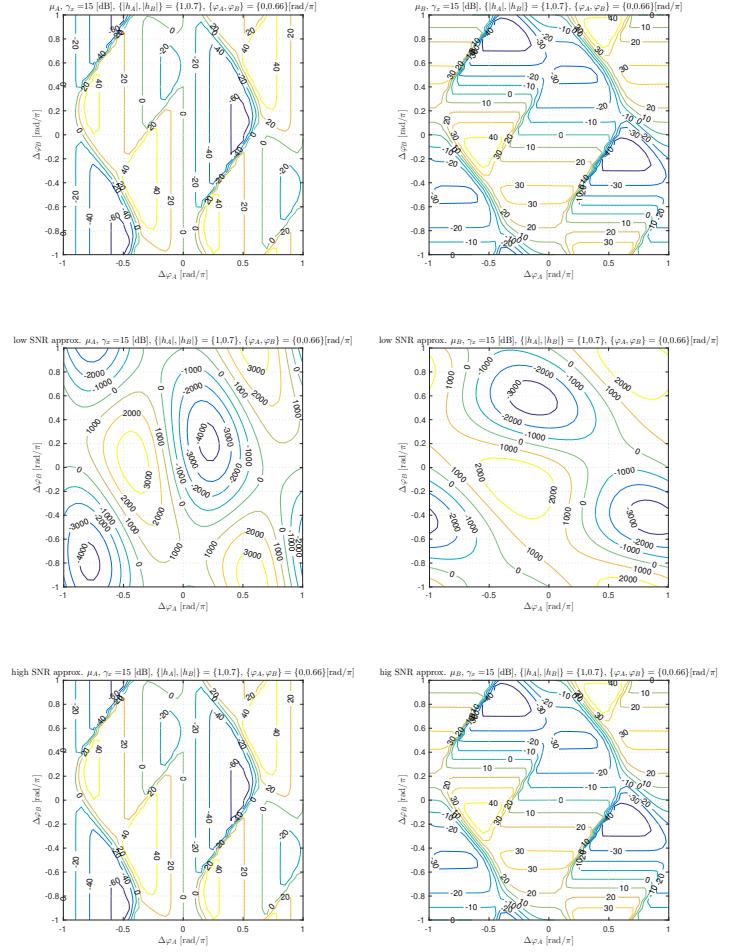


(B)

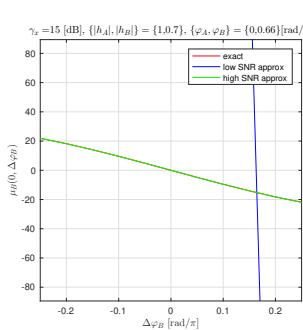


(C2)

Figure 8. Equivalent detector gradient characteristics for $\gamma_x = 15$ [dB], $h_A = 1e^{j0}$, $h_B = 0.7e^{j0.5\pi}$. (A) gradient components $\mu_A(\Delta\phi)$, $\mu_B(\Delta\phi)$ (row #1) exact , (row #2) low SNR approximation, (row #3) high SNR approximation; (B) gradient components $\mu_A(\Delta\phi_A, 0)$, $\mu_B(0, \Delta\phi_B)$; (C1) gradient directions (exact); (C2) random received constellation points.



(B)



(C1)

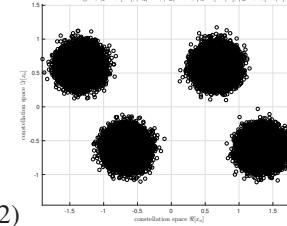


Figure 9. Equivalent detector gradient characteristics for $\gamma_x = 15$ [dB], $h_A = 1e^{j0}$, $h_B = 0.7e^{j0.66\pi}$. (A) gradient components $\mu_A(\Delta\phi)$, $\mu_B(\Delta\phi)$ (row #1) exact , (row #2) low SNR approximation, (row #3) high SNR approximation; (B) gradient components $\mu_A(\Delta\phi_A, 0)$, $\mu_B(0, \Delta\phi_B)$; (C1) gradient directions (exact); (C2) random received constellation points.