Soft-in soft-out detection using partial Gaussian approximation

GUO, Qinghua; HUANG, Defeng (David); NORDHOLM, Sven; XI, Jiangtao; LI, Ping

Published in:
IEEE Access

Published: 01/01/2014

Document Version:
Final Published version, also known as Publisher’s PDF, Publisher’s Final version or Version of Record

Publication record in CityU Scholars:
Go to record

Published version (DOI):
10.1109/ACCESS.2014.2322078

Publication details:
https://doi.org/10.1109/ACCESS.2014.2322078

Citing this paper
Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights
Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission
Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy
Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.
Soft-In Soft-Out Detection Using Partial Gaussian Approximation

QINGHUA GUO1,2, (Member, IEEE), DEFENG (DAVID) HUANG2, (Senior Member, IEEE), SVEN NORDHOLM3, (Senior Member, IEEE), JIANGTAO XI1, (Senior Member, IEEE), AND LI PING4, (Fellow, IEEE)
1School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW 2522, Australia
2School of Electrical, Electronic and Computer Engineering, University of Western Australia, Perth, WA 6009, Australia
3Department of Electrical and Computer Engineering, Curtin University, Perth, WA 6102, Australia
4Department of Electronic Engineering, City University of Hong Kong, Hong Kong

Corresponding author: Q. Guo (qguo@uow.edu.au)

The work of Q. Guo, D. Huang, S. Nordholm, and J. Xi was supported in part by the Australian Research Council’s DECRA under Grant DE120101266 and in part by the Discovery Project under Grant DP1093000 and Grant DP110100736. The work of Li Ping was supported by the Research Grants Council, Hong Kong, under Project CityU 118014. Part of this work was presented in 2012 International Conference on Wireless Communications and Signal Processing (WCSP 2012).

ABSTRACT This paper concerns the soft-in soft-out detection in a coded communication system, where the transmitted symbols are discrete valued, and the exact a posteriori probability (APP) detection often involves prohibitive complexity. By using the properties of Gaussian functions, an approximate approach to the APP detection is devised with the idea that, in the computation of the APP of each symbol, the remaining symbols are distinguished based on their contributions to the APP of the concerned symbol, and the symbols with less contributions are approximated as (continuous) Gaussian variables [hence the name partial Gaussian approximation (PGA)] to reduce the computational complexity. The connection between the PGA detector and the reduced dimension maximum a posteriori detector (RDMAP) is investigated. It is shown that, PGA is equivalent to RDMAP, but it has a complexity much lower than that of RDMAP, i.e., PGA can be regarded as an efficient implementation of RDMAP. In addition, the application of PGA in intersymbol interference (ISI) channel equalization is also investigated. We show that PGA allows further significant complexity reduction by exploiting the circulant structure of the system transfer matrix, which makes PGA very attractive in handling severe ISI channels with large memory length.

INDEX TERMS Soft-in soft-out detection, turbo equalization, Gaussian functions, LMMSE detection, RDMAP.

I. INTRODUCTION

Recently, the implementation of the soft-in soft-out (SISO) detector for iterative detection in a coded communication system has received tremendous attention as a turbo (iterative) receiver can achieve huge performance gain over a conventional non-iterative receiver [1]–[15]. The structure of a coded system with a turbo receiver is shown in Fig. 1, where the receiver consists of a SISO detector and a SISO decoder which work in an iterative manner. The task of the SISO detector is to compute the extrinsic log-likelihood ratio (LLR) for each code bit with the output extrinsic LLRs from the decoder as a priori information [1]–[15].

Although we focus on iterative detection, the developed detector in this work can also be employed in a non-iterative receiver.

The optimal implementation of the detection is the a posteriori probability (APP) detector [1]. However, the exponential complexity of the APP detector is often too
As shown in Section III.C, PGA can be regarded as an efficient implementation of RDMAP for a general system transfer matrix, e.g., in MIMO detection. With much lower complexity compared with RDMAP, PGA enjoys the same performance advantage as RDMAP in MIMO detection in [19]. Moreover, we will also show another appealing feature of PGA, i.e., it allows a further significant complexity reduction in ISI channel equalization by exploiting the circulant structure of the system transfer matrix. This makes PGA very attractive in handling severe ISI channels with large memory length.

This paper is organized as follows. Problem description and the exact APP detector are introduced in Section II. In Section III, PGA is derived by using some useful properties of the Gaussian functions listed in Appendix A, and the connection between RDMAP and PGA is investigated. The implementation of PGA for a circulant system transfer matrix in ISI channel equalization is detailed in Section IV. Simulation results are provided in Section V, followed by conclusions in Section VI.

The notations used in this paper are as follows. Lower and upper case letters denote scalars. Bold lower and upper case letters represent column vectors and matrices, respectively. The probability density function (PDF) of a continuous random variable and the probability mass function of a discrete random variable are represented by $p(\cdot)$ and $P(\cdot)$, respectively. We use $\propto$ to denote equality of functions up to a scale factor, and use $I_N$ to denote an $N \times N$ identity matrix. The superscptions $\cdot^T$ and $\cdot^H$ denote the transpose and conjugate transpose, respectively. We use $\mathcal{N}(\mathbf{x}, \mathbf{m}, \mathbf{V})$ to denote a complex multivariate Gaussian function with $\mathbf{x}$ as variable and $\mathbf{m}$ and $\mathbf{V}$ as parameters, where $\mathbf{V}$ is a Hermitian matrix.

II. THE EXACT APP SISO DETECTOR

As shown in Fig. 1, the turbo receiver consists of a SISO detector and a SISO decoder, which are connected by an interleaver and its corresponding deinterleaver. The SISO detector and the SISO decoder exchange extrinsic information iteratively to achieve performance improvement with the iteration [1]–[15]. We focus on the implementation of the SISO detector in this work.

A. PROBLEM DESCRIPTION

The received signal in the coded system shown in Fig. 1 can be represented as

$$\mathbf{r} = \mathbf{Hx} + \mathbf{w} \quad (1)$$

where $\mathbf{r}$ denotes a length-$Z$ observation vector, $\mathbf{H}$ denotes a $Z \times N$ system transfer matrix (e.g., a MIMO channel matrix or an ISI channel matrix), $\mathbf{w}$ denotes a length-$Z$ circularly symmetric additive white Gaussian noise (AWGN) vector with PDF $\mathcal{N}(\mathbf{w}; 0, \sigma^2 I_Z)$, and $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$ is mapped from an interleaved code sequence $\mathbf{c}$, i.e., each $x_n \in \mathcal{A} = \{\alpha_1, \alpha_2, \ldots, \alpha_{2^Q}\}$ ($|\mathcal{A}| = 2^Q$) corresponds to a length-$Q$ subsequence of $\mathbf{c}$ denoted by $c_n = [c_{n,1}, c_{n,2}, \ldots, c_{n,Q}]^T$. The task of the detector is to compute the log-likelihood

As shown in Section III.C, PGA can be regarded as an efficient implementation of RDMAP for a general system transfer matrix, e.g., in MIMO detection. With much lower complexity compared with RDMAP, PGA enjoys the same performance advantage as RDMAP in MIMO detection in [19]. Moreover, we will also show another appealing feature of PGA, i.e., it allows a further significant complexity reduction in ISI channel equalization by exploiting the circulant structure of the system transfer matrix. This makes PGA very attractive in handling severe ISI channels with large memory length.

This paper is organized as follows. Problem description and the exact APP detector are introduced in Section II. In Section III, PGA is derived by using some useful properties of the Gaussian functions listed in Appendix A, and the connection between RDMAP and PGA is investigated. The implementation of PGA for a circulant system transfer matrix in ISI channel equalization is detailed in Section IV. Simulation results are provided in Section V, followed by conclusions in Section VI.

The notations used in this paper are as follows. Lower and upper case letters denote scalars. Bold lower and upper case letters represent column vectors and matrices, respectively. The probability density function (PDF) of a continuous random variable and the probability mass function of a discrete random variable are represented by $p(\cdot)$ and $P(\cdot)$, respectively. We use $\propto$ to denote equality of functions up to a scale factor, and use $I_N$ to denote an $N \times N$ identity matrix. The superscptions $\cdot^T$ and $\cdot^H$ denote the transpose and conjugate transpose, respectively. We use $\mathcal{N}(\mathbf{x}, \mathbf{m}, \mathbf{V})$ to denote a complex multivariate Gaussian function with $\mathbf{x}$ as variable and $\mathbf{m}$ and $\mathbf{V}$ as parameters, where $\mathbf{V}$ is a Hermitian matrix.

II. THE EXACT APP SISO DETECTOR

As shown in Fig. 1, the turbo receiver consists of a SISO detector and a SISO decoder, which are connected by an interleaver and its corresponding deinterleaver. The SISO detector and the SISO decoder exchange extrinsic information iteratively to achieve performance improvement with the iteration [1]–[15]. We focus on the implementation of the SISO detector in this work.

A. PROBLEM DESCRIPTION

The received signal in the coded system shown in Fig. 1 can be represented as

$$\mathbf{r} = \mathbf{Hx} + \mathbf{w} \quad (1)$$

where $\mathbf{r}$ denotes a length-$Z$ observation vector, $\mathbf{H}$ denotes a $Z \times N$ system transfer matrix (e.g., a MIMO channel matrix or an ISI channel matrix), $\mathbf{w}$ denotes a length-$Z$ circularly symmetric additive white Gaussian noise (AWGN) vector with PDF $\mathcal{N}(\mathbf{w}; 0, \sigma^2 I_Z)$, and $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$ is mapped from an interleaved code sequence $\mathbf{c}$, i.e., each $x_n \in \mathcal{A} = \{\alpha_1, \alpha_2, \ldots, \alpha_{2^Q}\}$ ($|\mathcal{A}| = 2^Q$) corresponds to a length-$Q$ subsequence of $\mathbf{c}$ denoted by $c_n = [c_{n,1}, c_{n,2}, \ldots, c_{n,Q}]^T$. The task of the detector is to compute the log-likelihood

As shown in Section III.C, PGA can be regarded as an efficient implementation of RDMAP for a general system transfer matrix, e.g., in MIMO detection. With much lower complexity compared with RDMAP, PGA enjoys the same performance advantage as RDMAP in MIMO detection in [19]. Moreover, we will also show another appealing feature of PGA, i.e., it allows a further significant complexity reduction in ISI channel equalization by exploiting the circulant structure of the system transfer matrix. This makes PGA very attractive in handling severe ISI channels with large memory length.

This paper is organized as follows. Problem description and the exact APP detector are introduced in Section II. In Section III, PGA is derived by using some useful properties of the Gaussian functions listed in Appendix A, and the connection between RDMAP and PGA is investigated. The implementation of PGA for a circulant system transfer matrix in ISI channel equalization is detailed in Section IV. Simulation results are provided in Section V, followed by conclusions in Section VI.

The notations used in this paper are as follows. Lower and upper case letters denote scalars. Bold lower and upper case letters represent column vectors and matrices, respectively. The probability density function (PDF) of a continuous random variable and the probability mass function of a discrete random variable are represented by $p(\cdot)$ and $P(\cdot)$, respectively. We use $\propto$ to denote equality of functions up to a scale factor, and use $I_N$ to denote an $N \times N$ identity matrix. The superscptions $\cdot^T$ and $\cdot^H$ denote the transpose and conjugate transpose, respectively. We use $\mathcal{N}(\mathbf{x}, \mathbf{m}, \mathbf{V})$ to denote a complex multivariate Gaussian function with $\mathbf{x}$ as variable and $\mathbf{m}$ and $\mathbf{V}$ as parameters, where $\mathbf{V}$ is a Hermitian matrix.

II. THE EXACT APP SISO DETECTOR

As shown in Fig. 1, the turbo receiver consists of a SISO detector and a SISO decoder, which are connected by an interleaver and its corresponding deinterleaver. The SISO detector and the SISO decoder exchange extrinsic information iteratively to achieve performance improvement with the iteration [1]–[15]. We focus on the implementation of the SISO detector in this work.

A. PROBLEM DESCRIPTION

The received signal in the coded system shown in Fig. 1 can be represented as

$$\mathbf{r} = \mathbf{Hx} + \mathbf{w} \quad (1)$$

where $\mathbf{r}$ denotes a length-$Z$ observation vector, $\mathbf{H}$ denotes a $Z \times N$ system transfer matrix (e.g., a MIMO channel matrix or an ISI channel matrix), $\mathbf{w}$ denotes a length-$Z$ circularly symmetric additive white Gaussian noise (AWGN) vector with PDF $\mathcal{N}(\mathbf{w}; 0, \sigma^2 I_Z)$, and $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$ is mapped from an interleaved code sequence $\mathbf{c}$, i.e., each $x_n \in \mathcal{A} = \{\alpha_1, \alpha_2, \ldots, \alpha_{2^Q}\}$ ($|\mathcal{A}| = 2^Q$) corresponds to a length-$Q$ subsequence of $\mathbf{c}$ denoted by $c_n = [c_{n,1}, c_{n,2}, \ldots, c_{n,Q}]^T$. The task of the detector is to compute the log-likelihood

As shown in Section III.C, PGA can be regarded as an efficient implementation of RDMAP for a general system transfer matrix, e.g., in MIMO detection. With much lower complexity compared with RDMAP, PGA enjoys the same performance advantage as RDMAP in MIMO detection in [19]. Moreover, we will also show another appealing feature of PGA, i.e., it allows a further significant complexity reduction in ISI channel equalization by exploiting the circulant structure of the system transfer matrix. This makes PGA very attractive in handling severe ISI channels with large memory length.

This paper is organized as follows. Problem description and the exact APP detector are introduced in Section II. In Section III, PGA is derived by using some useful properties of the Gaussian functions listed in Appendix A, and the connection between RDMAP and PGA is investigated. The implementation of PGA for a circulant system transfer matrix in ISI channel equalization is detailed in Section IV. Simulation results are provided in Section V, followed by conclusions in Section VI.

The notations used in this paper are as follows. Lower and upper case letters denote scalars. Bold lower and upper case letters represent column vectors and matrices, respectively. The probability density function (PDF) of a continuous random variable and the probability mass function of a discrete random variable are represented by $p(\cdot)$ and $P(\cdot)$, respectively. We use $\propto$ to denote equality of functions up to a scale factor, and use $I_N$ to denote an $N \times N$ identity matrix. The superscptions $\cdot^T$ and $\cdot^H$ denote the transpose and conjugate transpose, respectively. We use $\mathcal{N}(\mathbf{x}, \mathbf{m}, \mathbf{V})$ to denote a complex multivariate Gaussian function with $\mathbf{x}$ as variable and $\mathbf{m}$ and $\mathbf{V}$ as parameters, where $\mathbf{V}$ is a Hermitian matrix.
ratio (LLR) for each code bit \(c_{n,q}\) (without taking coding into account), which can be expressed as [1]–[5]

\[
L(c_{n,q}) = \ln \frac{P(c_{n,q} = 0 \mid r)}{P(c_{n,q} = 1 \mid r)} = \ln \frac{\sum_{x_n \in A^0_q} P(x_n \mid r)}{\sum_{x_n \in A^1_q} P(x_n \mid r)}
\]

(2)

where \(A^0_q (A^1_q)\) denotes the subset of all \(a_i \in A\) corresponding to a binary subsequence with the \(q\)th bit given by 0 (1). In a turbo receiver as shown in Fig. 1, the extrinsic LLR [1]–[5]

\[
q
\]

to a binary subsequence with the

\(A\)

\(x\)

\(P(x_n \mid r)\)

will be the input to the decoder, where \(L^d(c_{n,q})\) is the output extrinsic LLR of the decoder in the last iteration.

It can be seen that the key task of the detector is to compute the APP \(P(x_n \mid r)\) for each symbol \(x_n\), which is the focus of this work.

### B. THE EXACT APP DETECTOR

According to Bayes’ rule,

\[
P(x_n \mid r) = \sum_{\tilde{x}_n} P(x_n \mid \tilde{x}_n) \propto \sum_{\tilde{x}_n} P(x_n)p(r \mid x) \tag{4}
\]

where the length-\((N − 1)\) vector \(\tilde{x}_n\) consists of the elements of \(x\) except \(x_n\). Given \(x, r\) is Gaussian distributed, i.e., \(p(r \mid x) = \mathcal{N}(r; Hx, \sigma^2I)\). Hence (4) can be rewritten as

\[
P(x_n \mid r) \propto \sum_{\tilde{x}_n} P(x_n)\exp \left[ -\frac{(r - Hx)^H(r - Hx)}{\sigma^2} \right]. \tag{5}
\]

In (5), assuming that the interleaved code bits are independent of each other, we have [1]–[5]

\[
P(x) = \prod_i P(x_i) \tag{6}
\]

\[
P(x_i) = \prod_q P(c_{i,q}) \tag{7}
\]

where \(P(c_{i,q})\) can be calculated based on the LLRs from the decoder.

By using (5), we can calculate the APPs \(\{P(x_n = a_i \mid r), \forall i\}\) up to some constant. Although the constant can be found through normalization, we do not need to care about it in the calculation of the LLR using (2) as it will be cancelled out. Due to the high-dimensional summation involved in (5), the computational complexity of the APP detector is \(\mathcal{O}(N|A|^N)\) per symbol (assuming that \(Z\) and \(N\) are in the same order), which is prohibitive even for a moderate \(N\).

Note that, when (1) is used to model the impact of an ISI channel \(h = [h_0, h_1, \ldots, h_L]^T\) on the transmitted signal \(x\), i.e., \(r\) is the convolution of \(x\) and \(h\), corrupted by AWGN \(w\), the system transfer matrix \(H\) is a Toeplitz matrix. In this case, the APP detector can be implemented using the BCJR algorithm with complexity \(\mathcal{O}(|A|^L)\) [18]. However, when \(L\) is large (e.g., in broadband wireless communications or underwater acoustic communications), the use of the BCJR-based APP detector is still impractical due to the high complexity. In the following, we consider approximate approaches to the computation of \(P(x_n \mid r)\).

### III. APPROXIMATE APP DETECTION USING PARTIAL GAUSSIAN APPROXIMATION

#### A. AN ALTERNATIVE REPRESENTATION FOR THE EXACT APP (5)

Define vector \(a = (H^H H)^{-1}H^H r\). It is easy to verify that

\[
(r - Hx)^H(r - Hx) = (x - a)^H H^H H(x - a) - r^H (H(H^H H)^{-1}H^H - I_Z)r. \tag{8}
\]

As a result,

\[
\exp \left[ -\frac{(r - Hx)^H(r - Hx)}{\sigma^2} \right] = \varphi g(x) \tag{9}
\]

where

\[
g(x) = \exp \left[ -\frac{(x - a)^H H^H H(x - a)}{\sigma^2} \right] \tag{10}
\]

and

\[
\varphi = \exp \left[ r^H (H(H^H H)^{-1}H^H - I_Z)r \right] / \sigma^2 \tag{11}
\]

is independent of \(x\).

Now we reformulate the APP in (5) using (9). Noting that the likelihood function \(p(r \mid x)\) in (4) is a function of \(x\) (rather than \(r\) which is fixed), we can safely ignore \(\varphi\) in (9). Hence (5) can be rewritten as

\[
P(x_n \mid r) \propto \sum_{\tilde{x}_n} P(x_n)g(x) \tag{12}
\]

\[
\propto \sum_{\tilde{x}_n} P(x_n)\mathcal{N}(x; a, A) \tag{13}
\]

where \(A = \sigma^2(H^H H)^{-1}\). To obtain (13), we have used

\[
g(x) \propto \mathcal{N}(x; a, A). \tag{14}
\]

Here, we emphasize that \(\mathcal{N}(x; a, A)\) in (13) and (14) just denotes a function of \(x\) as shown in (46), and it does not mean that \(x\) is a Gaussian random vector.

Although the exact computation of the APP \(P(x_n \mid r)\) using (13) still involves high-dimensional summation, it enables us to derive low-complexity approximate approaches based on the properties of Gaussian functions listed in Appendix A.

#### B. PARTIAL GAUSSIAN APPROXIMATION

If we treat \(x\) as a Gaussian random vector with (a priori) PDF \(\mathcal{N}(x; m, V)\) (instead of a discrete random vector with a priori probability mass function \(P(x)\)), the summation in (13) is changed to integration accordingly. So we have

\[
P_{FGA}(x_n \mid r) \propto \int_{\tilde{x}_n} \mathcal{N}(x; m, V)\mathcal{N}(x; a, A)d\tilde{x}_n \tag{15}
\]
where we note that $P_{\text{FGA}}(x_n | \mathbf{r})$ is an approximation to $P(x_n | \mathbf{r})$ due to the above Gaussian approximation. In (15), the a priori covariance matrix $\mathbf{V}$ is a diagonal matrix as $[x_i]$ are independent of each other. A straightforward way to determine $m_i$ (the $i$th element of $\mathbf{m}$) and $v_i$ (the $i$th diagonal element of $\mathbf{V}$) is to use the first and second moments of $x_i$, i.e.,

$$m_i = \sum_{\alpha_i \in \mathcal{A}} \alpha_i P(x_i = \alpha_k)$$

(16)

$$v_i = \sum_{\alpha_i \in \mathcal{A}} |\alpha_i - m_i|^2 P(x_i = \alpha_k).$$

(17)

Although (15) involves high-dimensional integration, it can be efficiently calculated based on the properties of Gaussian functions shown by (50)–(52) and (49) in Appendix A. The complexity is about $O(N^2)$ per symbol (assuming that $N$ and $Z$ are in the same order), which is much lower than that of calculating (5) exactly with complexity $O(N|\mathcal{A}|^N)$. As all the symbols are treated as Gaussian variables in this approach, we call it full Gaussian approximation approach.

Simulation results in Section V show that the full Gaussian approximation approach delivers very poor performance. This motivates us to develop the partial Gaussian approximation approach, i.e., only part of the symbols are approximated as Gaussian variables.

To devise PGA, a primary question we need to answer is which symbols should be approximated as Gaussian variables. It is interesting that the alternative representation for the APP $P(x_n | \mathbf{r})$, i.e., (12) or (13) developed in Section III.A, provides a direct suggestion for us.

Now we re-examine (12) and (10). By expanding the exponent of $g(x)$ in (10) to a scalar form, it can be found that some of the non-zero off-diagonal elements of $\mathbf{H}^T \mathbf{H}$ generate the cross terms between $x_n$ and $[x_i, i \neq n]$. It is obvious that if all of these cross terms are zeroes (i.e., $\mathbf{H}^T \mathbf{H}$ is a diagonal matrix), $[x_i, i \neq n]$ make no contributions to the APP of $x_n$. Due to the non-zero cross terms, $[x_i, i \neq n]$ may have different contributions to the APP of $x_n$. Although it is hard to evaluate their contributions exactly, we can use the following approach to distinguish them roughly. Define $\mathbf{p}_n$ as the $n$th column vector of matrix $\mathbf{H}^T \mathbf{H}$. We can find that the $i$th element of $\mathbf{p}_n$ is the coefficient of the cross term between $x_n$ and $x_i$ up to a constant. We distinguish the contributions of different symbols based on the magnitudes of the elements of $\mathbf{p}_n$, i.e., the larger the magnitude of an element, the larger the contribution of its corresponding symbol to the APP of $x_n$. According to this, from $[x_i, i \neq n]$, we can select $M$ most important symbols to $x_n$. We keep the $M$ most important symbols as discrete variables without any approximation, but treat the remaining $N - M - 1$ symbols as Gaussian variables. Including $x_n$ itself, the number of discrete variables is $M + 1$, and they are denoted by

$$\mathbf{x}^D_n = \mathbf{S}_n \mathbf{x}$$

(18)

where $\mathbf{S}_n$ is the corresponding $(M + 1) \times N$ selecting matrix for $x_n$. The remaining $N - M - 1$ symbols to be treated as Gaussian variables are denoted by $\mathbf{x}^G_n$, and their means and variances can be computed using (16) and (17). We denote the mean vector and (diagonal) covariance matrix of $\mathbf{x}^D_n$ by $\mathbf{m}^D_n$ and $\mathbf{V}^D_n$, respectively.

Based on the above partial Gaussian approximation, part of the summation in (13) will be changed to integration, and the APP of $x_n$ can be approximated as

$$P_{\text{PGA}}(x_n | \mathbf{r}) \propto \sum_{\mathbf{x}^G_n} P(x_n^D) h(x_n^D)$$

(19)

where $x_n^D$ consists of the elements in $x_n^D$ except $x_n$, and

$$h(x_n^D) = \int \mathcal{N}(\mathbf{x}_n^G; \mathbf{m}_n^G, \mathbf{V}_n^G) \mathcal{N}(\mathbf{x}; \mathbf{a}, \mathbf{A}) d\mathbf{x}_n^G.$$  

(20)

It can be seen that, assuming $h(x_n^D)$ is available, (19) involves an $M$-dimensional summation. To reduce the complexity, we may set $M$ to be a small integer. In the following, we discuss how to compute $h(x_n^D)$ in (20).

Note that (20) involves the multiplication of two Gaussian functions with different dimensions (noting that $x_n^G$ is a subvector of $\mathbf{x}$), and hence (50)–(52) in Appendix A cannot be directly applied. To overcome this problem, we rewrite $h(x_n^D)$ as

$$h(x_n^D) = \frac{\mathcal{N}(\mathbf{x}_n^G; \mathbf{m}_n^G, \mathbf{V}_n^G)}{\int \mathcal{N}(\mathbf{x}_n^G; \mathbf{m}_n^G, \mathbf{V}_n^G) \mathcal{N}(\mathbf{x}; \mathbf{a}, \mathbf{A}) d\mathbf{x}_n^G}$$

(21)

where $\mathbf{m}_n^G$ and $\mathbf{V}_n^G$ are the mean vector and (diagonal) covariance matrix of $\mathbf{x}^G_n$, respectively, and

$$\mathcal{N}(\mathbf{x}; \mathbf{a}, \mathbf{A}) \propto \mathcal{N}(\mathbf{x}; \mathbf{c}, \mathbf{C}),$$

(22)

where

$$\mathbf{C} = (\mathbf{V}^{-1} + \frac{1}{\sigma^2} \mathbf{H}^T \mathbf{H})^{-1},$$

(23)

$$\mathbf{c} = \mathbf{V}^{-1} \mathbf{m} + \frac{1}{\sigma^2} \mathbf{H}^T \mathbf{r}.$$  

(24)

Then from (21) we have

$$h(x_n^D) \propto \frac{\int \mathcal{N}(\mathbf{x}; \mathbf{c}, \mathbf{C}) d\mathbf{x}_n^G}{\mathcal{N}(\mathbf{x}_n^G; \mathbf{m}_n^G, \mathbf{V}_n^G)}$$

(25)

$$\propto \frac{\mathcal{N}(\mathbf{x}_n^G; \mathbf{m}_n^G, \mathbf{V}_n^G)}{\mathcal{N}(\mathbf{x}_n^G; \mathbf{S}_n \mathbf{c}, \mathbf{S}_n \mathbf{C} \mathbf{S}_n^T)}.$$  

(26)

Here we note that the multiplication of Gaussian functions in (22) is straightforward as there are no common elements between $x_n^G$ and $\mathbf{x}_n^G$, which is different from that in (50).
where (25) is from (49), and \( z_n \) and \( Z_n \) in (26) can be computed based on (53)–(55), i.e.,

\[
Z_n = (S_nC_n\Sigma_n^{-1} - (V_n^D)\Sigma_n^{-1})^{-1}
\]

(27)

\[
z_n = Z_n(S_nC_n\Sigma_n^{-1}S_n) - (V_n^D)\Sigma_n^{-1}m_n^D.
\]

(28)

Finally, by using (26), the approximate APP of \( x_n \) in (19) can be represented as

\[
P_{PGA}(x_n | r) \propto \sum_{\hat{x}_n^D} P(x_n^D | r)\exp \left[ -x_n^D - z_n \right] Z_n^{-1}(x_n^D - z_n).
\]

(29)

The steps of PGA are summarized as follows:

- **Step 1.** Calculate the mean vector \( \mathbf{m} \) and (diagonal) covariance matrix \( \mathbf{V} \) of \( \mathbf{x} \), and then \( \mathbf{c} \) and \( \mathbf{C} \) using (23) and (24).

- **Step 2.** Choose the \( M \) most important symbols for each \( x_n \) based on the \( n \)th column vector of \( H_n^D \mathbf{H}_n \) and calculate \( Z_n \) and \( z_n \) for each \( x_n \) using (27) and (28).

- **Step 3.** Calculate the approximate APP for each \( x_n \) using (29).

Before analyzing the complexity of PGA, we note that different \( x_n \) has different \( x_n^D \) and \( x_n^G \), and hence different \( Z_n \) and \( z_n \). That is why we put the subscript \( n \) for these variables. The computational complexity involved in Step 1 is \( O(N^3 + 2N^2) \) or \( O(N^3) \) with the assumption that \( Z \) and \( N \) are in the same order. We note that the results of Step 1 can be shared by all the symbols, so the complexity per symbol is \( O(N^3) \). The complexity of Step 2 and that of Step 3 depend on the value of \( M \), which are \( O(M + 1) \) and \( O(M + 1)^2 |A|^{M+1} \) per symbol (noting that \( Z_n^{-1} \) required in Step 3 is already available in Step 2), respectively. If \( M \) is much smaller than \( N \) (e.g., \( M = 2 \) will be used in our simulations), the complexity of the algorithm can be significantly lower than that of the exact APP approach (\( O(N |A|^N) \) per symbol).

It can be easily found that the exact APP approach is a special case of PGA with \( M = N - 1 \) (i.e., without Gaussian approximation). The well-known LMMSE detector in [2], [3], and [5] is also its special case with \( M = 0 \) as shown by Proposition 1.

**Proposition 1:** When \( M = 0 \) (i.e., all the symbols except \( x_n \) are treated to be Gaussian), PGA is equivalent to the LMMSE detector in [2], [3], and [5].

**Proof:** See Appendix B.

### C. CONNECTION WITH RDMAP IN [19]

In [19], the original model (11) is rewritten in the following form with respect to each \( x_n^A \)

\[
r = H_n^A x_n^A + H_n^B x_n^B + w
\]

(30)

where elements in \( \mathbf{x} \) are grouped into \( x_n^A \) and \( x_n^B \). \( x_n \) is included in \( x_n^A \), and \( H_n^A \) and \( H_n^B \) consist of the column vectors in \( \mathbf{H} \). which correspond to \( x_n^A \) and \( x_n^B \), respectively. \( x_n^A \) is called reduced-dimension signal vector and \( x_n^B \) is the interference vector. In RDMAP, to minimize the noise enhancement due to noise whitening, the elements in \( x_n^A \) and \( x_n^B \) are also determined by the channel correlation matrix \( H_n^D H_n^B \), which results in \( x_n^A = x_n^D \) and \( x_n^B = x_n^G \). Thereafter we use

\[
r = H_n^D x_n^D + H_n^G x_n^G + w
\]

(31)

to replace (30) to facilitate the investigation of the connection between PGA and RDMAP.

In [19], the interference and noise \( H_n^G x_n^G + w \) is treated as a colored noise with mean \( H_n^G m_n^G \) and covariance matrix

\[
R_n = H_n^G V_n^G (H_n^G)^H + \sigma^2 \mathbf{I}.
\]

(32)

The colored noise is whitened by removing its mean and then applying the noise-whitening filter given by

\[
W_n = \Sigma_n^{-1/2} Q_n^H
\]

(33)

where \( \Sigma_n \) and \( Q_n \) are obtained from the eigenvalue decomposition of \( R_n \)

\[
R_n = Q_n \Sigma_n Q_n^H.
\]

(34)

i.e., \( \Sigma \) is a diagonal matrix and

\[
Q_n Q_n^H = \mathbf{I}.
\]

(35)

The whitened observation can be represented as

\[
p_n = W_n (r - H_n^G m_n^G).
\]

(36)

The residual noise at the filter output can be expressed as

\[
W_n (H_n^G x_n^G - H_n^G m_n^G + w) = p_n - W_n H_n^D x_n^D.
\]

(37)

Under the assumption that the noise is Gaussian, i.e.,

\[
p_n - W_n H_n^D x_n^D \sim \mathcal{N}(0, \mathbf{I}),
\]

(38)

the APP of \( x_n \) can be represented as,

\[
P_{RDMAP}(x_n | r) \propto \sum_{\check{x}_n^D} P(x_n^D | r)\exp \left[ -||p_n - W_n H_n^D x_n^D||^2 \right]
\]

(39)

where \( \check{x}_n^D \) consists of the elements of \( x_n^D \) except \( x_n \). Once the right hand side of (39) is available, the LLRs can be calculated using (2). RDMAP needs to evaluate (39) for each \( x_n \). In contrast, PGA evaluates (29) for each \( x_n \).

As we can see from the above that the derivation for PGA is very different from that for RDMAP, and the final results, i.e., (29) for PGA and (39) for RDMAP are in completely different forms. However, in RDMAP, the whitened interference plus noise \( W_n (H_n^G x_n^G - H_n^G m_n^G + w) \) is assumed to be Gaussian, which can be reached by treating \( x_n^G \) as a Gaussian vector as well. As a result, PGA and RDMAP have the same fundamental assumption that part of the elements in \( \mathbf{x} \) (i.e., \( x_n^A \)) are treated as Gaussian variables and the remaining elements in \( \mathbf{x} \) (i.e., \( x_n^B \)) are kept as discrete random variables without approximation. It can be shown that these two approaches are equivalent through a tedious proof in Appendix C.
**Proposition 2:** PGA is equivalent to RDMAP.

**Proof:** See Appendix C.

**Remarks:**
- Although PGA is equivalent to RDMAP, the complexity of PGA is much lower than that of RDMAP. Note that, in RDMAP, the noise-whitening filter \( W_n \) is constructed based on the eigenvalue decomposition of \( R_n \), and \( W_n \) needs to be computed for each \( x_n \) with complexity \( O(N^3) \) per symbol. It can be shown that, the total complexity of RDMAP is \( O(N^3 + O(N(N-M-1)|A|^{M+1})) \) per symbol. In contrast, the total complexity of PGA is \( O(N^2) + O((M+1)^2|A|^{M+1}) \), which is significantly lower than that of RDMAP especially when \( N \) is large and \( M \) is much smaller than \( N \). Hence, PGA can be regarded as an efficient implementation of RDMAP.
- It is interesting that both approaches use the channel correlation matrix \( \mathbf{H}^H \mathbf{H} \) to determine the subsets \( x_n^D \) and \( x_n^G \) for each \( x_n \), but the arguments are very different. The argument in RDMAP is minimizing the noise enhancement due to the noise whitening. In contrast, that \( \mathbf{H}^H \mathbf{H} \) appears in the new APP representation (12) (see \( g(x) \) in (10)) in PGA directly suggests the use of \( \mathbf{H}^H \mathbf{H} \) to distinguish the contributions of \( x_n, i \neq n \) to the APP of \( x_n \), thereby determining the subsets \( x_n^D \) and \( x_n^G \).
- In RDMAP, \( x_n^G \) is treated as interference to \( x_n^D \), and interference cancellation and whitening are used in (36) to handle the impact of the interference. The proof of the equivalence between RDMAP and PGA in Appendix C reveals that the same interference cancellation and whitening operations are hidden in PGA. With (71) and (78) in Appendix C, \( z_n \) in PGA (see (28) and (29)) can be represented as

\[
\begin{align*}
\mathbf{z}_n &= \mathbf{Z}_n (\mathbf{H}_n^H)^{1/2} \mathbf{R}_n^{-1}(\mathbf{r} - \mathbf{H}_n^G \mathbf{m}_n^G) \\
&= \mathbf{z}_n (\mathbf{H}_n^H)^{1/2} \mathbf{Q}_n \mathbf{Z}_n^{-1/2} \mathbf{W}_n (\mathbf{r} - \mathbf{H}_n^G \mathbf{m}_n^G),
\end{align*}
\]

where (33) and (34) are used in the above derivation. The term \( \mathbf{W}_n (\mathbf{r} - \mathbf{H}_n^G \mathbf{m}_n^G) \) in (40) indicates the hidden operations of interference cancellation and whitening in PGA.

**IV. IMPLEMENTATION FOR CIRCULANT** \( \mathbf{H} \)

The previous section reveals that (for a general system transfer matrix) PGA is equivalent to RDMAP but with much lower complexity. However, in ISI channel equalization where the size of the system transfer matrix may be very large (i.e., \( N \) is very large), the complexity of PGA (and RDMAP) in the previous section may still be a serious concern.

With the the cyclic prefixing technique, \( \mathbf{H} \) in model (1) turns into a circulant matrix \([15],[16]\). In the following, we will show another appealing feature of PGA, i.e., it allows a further significant complexity reduction by exploiting the circulant structure of \( \mathbf{H} \). In contrast, it is difficult to exploit the circulant structure in RDMAP.

A useful property of a circulant matrix is that it can be diagonalized by the discrete Fourier transform (DFT) matrix, i.e., \( \mathbf{F} \mathbf{H}^H = \mathbf{D} \) or equivalently

\[
\mathbf{H} = \mathbf{F}^H \mathbf{D} \mathbf{F}
\]

where \( \mathbf{H} \) is an \( N \times N \) circulant matrix, \( \mathbf{F} \) is a normalized DFT matrix (i.e., its \( (m,n) \)-th element is given by \( N^{-1/2} e^{-j2\pi mn/N} \), where \( j = \sqrt{-1} \)), \( \mathbf{D} \) is a diagonal matrix whose diagonal elements are given by \( \sqrt{N} \mathbf{f}_h \), and \( \mathbf{f}_h \) is the first column vector of \( \mathbf{H} \). As in the implementation of the LMMSE detector in the frequency domain, we approximate the diagonal covariance matrix \( \mathbf{V} \) to be a scaled identity matrix (see \([11],[13],[22]\)), i.e.,

\[
\mathbf{V} \approx \alpha \mathbf{I}_N
\]

where \( \alpha \) is the average of the diagonal elements of \( \mathbf{V} \).

1) IMPLEMENTATION OF STEP 1

With (41) and (42), \( \mathbf{C} \) in (23) can be rewritten as

\[
\mathbf{C} = \mathbf{F}^H \mathbf{A} \mathbf{F}
\]

where \( \mathbf{A} = (\alpha^{-1} \mathbf{I}_N + \frac{1}{\sigma^2} \mathbf{D}^H \mathbf{D})^{-1} \) is a diagonal matrix. Hence \( \mathbf{C} \) is a circulant and Hermitian matrix and its first column vector is given by

\[
\rho = N^{-1/2} \mathbf{F}^H \eta
\]

where the length-\( N \) vector \( \eta \) consists of the diagonal elements of \( \mathbf{A} \). Vector \( \mathbf{c} \) in (24) can be represented as

\[
\mathbf{c} = \mathbf{F}^H \left( \alpha^{-1} \mathbf{f}_m + \frac{1}{\sigma^2} \mathbf{D}^H \mathbf{F} \mathbf{r} \right).
\]

By using the FFT algorithm, the complexity of Step 1 can be reduced from \( O(N^2) \) per symbol to \( O(\log N) \) per symbol.

2) IMPLEMENTATION OF STEP 2

We can find that \( \mathbf{H}^H \mathbf{H} \) whose \( n \)-th column is used to select the most important \( M \) symbols to \( x_n \) is also a circulant and Hermitian matrix. Hence its first column vector can be found using the FFT algorithm (its other columns are cyclic shifts of the first column). As \( \mathbf{C} \) is a circulant Hermitian matrix, it is not hard to verify that the \( \mathbf{S}_n \mathbf{S}_n^H \) in (27) is independent of \( n \). We use the approximation that \( \mathbf{V}_n^D \approx \alpha \mathbf{I}_{M+1} \), so \( \mathbf{Z}_n \) in (27) is also independent of \( n \). This means that \( (\mathbf{S}_n \mathbf{S}_n^H)^{-1} \) and \( \mathbf{Z}_n \) only need to be calculated once and can be shared by all the \( N \) symbols. The complexity of calculating \( \mathbf{Z}_n \) is \( O((M+1)^2) \). The total complexity of this step is \( O((logN)+O((M+1)^2)) \) per symbol.

The main complexity of the whole approach is on Step 3, which is \( O((M+1)^2|A|^{M+1}) \) per symbol. To make the complexity low, we can set \( M \) to be a small integer. In Section V, we will show that the proposed detector PGA with \( M = 2 \) can achieve significant performance gain over the frequency domain LMMSE detector (which is equivalent to PGA with \( M = 0 \)).
V. SIMULATION RESULTS

PGA can be applied to MIMO detection (assuming a general channel matrix $H$) and ISI channel equalization. However, we focus on ISI channel equalization in this section as PGA is equivalent to RDMAP (i.e., PGA enjoys the same performance as RDMAP but with much lower complexity) and the simulations about MIMO detection and comparisons with other detectors such as the LMMSE detector and group MAP detector have been provided in [19].

Consider a single-carrier block transmission system over ISI channels, where the cyclic prefixing technique is used, i.e., $H$ is a circulant matrix. We assume that a rate-1/2 nonsystematic convolutional code with generator $(5, 7)_{8}$ is employed, and the APP decoder is implemented using the BCJR algorithm. The sequence of the code bits is interleaved and mapped to a symbol sequence using QPSK with Gray mapping. Then the symbol sequence is divided into length-512 blocks, which are cyclic prefixed before transmission over an ISI channel. In all the simulations the number of iterations is 10. For PGA, we set $M = 2$.

Example A: In this example, we use the Proakis’ 5-tap ISI channel with coefficients $[0.227, 0.460, 0.688, 0.460, 0.227]$ (noting that the energy of the channel is 1), which incurs severe distortion on the transmitted signal as shown in [20] by the frequency response of the channel. The number of the information bits is set to 4096. The BER performance of the turbo receiver with different detectors is shown in Fig. 2. We can see from this figure that the detector with full Gaussian approximation (denoted by “Full Approx.”) delivers very poor performance. The MAP detector (equivalent to PGA with $M = N - 1$) delivers the best performance at the cost of high complexity. It can also be seen that the receiver with PGA significantly outperforms the frequency domain LMMSE detector (which is equivalent to PGA with $M = 0$), and approaches closely the AWGN bound (the performance of the system over AWGN channel) in the relatively high $E_b/N_0$ range.

Example B: In this example, we test the detector using the power line reference channel 3 (REF-CH3), which is rated as a “bad” channel in [23]. The parameters of REF-CH3 are given in [23]. The transfer function of REF-CH3 is shown in Fig. 3, where we restrict the maximum frequency to 8.5 MHz due to the high attenuation of the channel with the increase of the frequency. The impulse response of the channel is represented by 128 channel coefficients. The energy of the ISI channel is normalized to 1 to facilitate performance comparison against the AWGN bound. The length of the information sequence is set to 7680. The BER performance is shown in Fig. 4. It can be seen that the full Gaussian approximation approach again exhibits poor performance. Compared with the LMMSE detector, PGA achieves a significant performance gain.

VI. CONCLUSION

We have presented an approximate APP detector with the partial Gaussian approximation, in which part of the symbols
are treated as Gaussian variables to reduce the computational complexity. We also revealed that PGA is equivalent to RDMAP in [19] but with much lower complexity, i.e., PGA can be regarded as an efficient implementation of RDMAP. We have shown that, by exploiting the circulant structure of the channel matrix, the complexity of PGA can be significantly reduced further, making it very attractive in severe ISI channel equalization. Simulation results demonstrated that PGA can achieve significant performance gain over the LMMSE detector, and provide a good trade-off between complexity and performance.

**APPENDIX A**

**PROPERTIES OF GAUSSIAN FUNCTIONS**

We use $\mathcal{N}(x; m, V)$ to denote a complex multivariate Gaussian function with $x$ as variable and $m$ and $V$ as parameters, i.e.,

$$
\mathcal{N}(x; m, V) = \frac{1}{\pi^N \det(V)} \exp \left[ - (x - m)^H V^{-1} (x - m) \right]
$$

(46)

where $N$ is the length of $x$, and $V$ is a Hermitian matrix.

**A. MARGINAL DISTRIBUTION OF A COMPLEX GAUSSIAN VECTOR**

Assume that $x$ is a length-$N$ complex Gaussian random vector with PDF

$$
p(x) = \mathcal{N}(x; m, V)
$$

(47)

and $x_i$ with length $M (M \leq N)$ is a sub-vector of $x$, i.e.,

$$
x_i = Sx
$$

(48)

where $S$ is an $M \times N$ selecting matrix (i.e., each of its row vectors is a row of the identity matrix $I_N$). It can be shown that the distribution of $x_i$

$$
p(x_i) = \int \mathcal{N}(x; m, V) dx_r
$$

$$
= \mathcal{N}(x_i; Sm, SVS^T)
$$

(49)

where $x_r$ consists of the remaining elements in $x$ after removing those in $x_i$.

**B. MULTIPLICATION OF TWO MULTIVARIATE GAUSSIAN FUNCTIONS**

The multiplication of two Gaussian functions is another Gaussian function [21], i.e.,

$$
\mathcal{N}(x; m_a, V_a) \mathcal{N}(x; m_b, V_b) \propto \mathcal{N}(x; m_c, V_c)
$$

(50)

where

$$
V_c = (V_a^{-1} + V_b^{-1})^{-1}
$$

(51)

$$
m_c = V_c^{-1} m_a + V_b^{-1} m_b.
$$

(52)

From the above, we have

$$
\mathcal{N}(x; m_c, V_c) \propto \mathcal{N}(x; m_b, V_b)
$$

(53)

and

$$
V_b = (V_c^{-1} - V_a^{-1})^{-1}
$$

(54)

$$
m_b = V_b (V_c^{-1} m_c - V_a^{-1} m_a).
$$

(55)

**APPENDIX B**

**PROOF OF PROPOSITION 1**

Proof: When $M = 0$, $S_n$ is a row vector (i.e., the $n$th row of the identity matrix $I_N$), and $x_n^D$ only contains $x_n$. In this case, $Z_n$ and $z_n$ in (27) and (28) are both scalars, and can be represented as

$$
Z_n = \frac{v_n C_n}{v_n - C_n}
$$

(56)

$$
z_n = \frac{c_n v_n - C_n m_n}{v_n - C_n}
$$

(57)

where $C_n$ is the $n$th diagonal element of $C$ and $c_n$ is the $n$th element of $c$. Using the matrix inversion lemma [17], and (23) and (24), we have

$$
C_n = v_n - v_n^2 h_n^H A^{-1} h_n
$$

(58)

$$
c_n = m_n C_n / v_n + v_n h_n^H A^{-1} (r - Hm + m_n h_n)
$$

(59)

where $h_n$ is the $n$th column of $H$, and $A = HVH^H + \sigma^2 I$. Substituting (58) and (59) into (56) and (57), we have

$$
Z_n = \frac{1 - v_n h_n^H A^{-1} h_n}{h_n^H A^{-1} h_n}
$$

(60)

$$
z_n = \frac{h_n^H A^{-1} (r - Hm + m_n h_n)}{h_n^H A^{-1} h_n - v_n (h_n^H A^{-1} h_n)^2}
$$

(61)

As $x_n^D = x_n$, and $Z_n$ and $z_n$ are both scalars, (29) can be rewritten as

$$
P_{PGA}(x_n | r) \propto \frac{P(x_n) \exp \left[ - \frac{|x_n - z_n|^2}{Z_n} \right]}{\frac{|h_n^H A^{-1} (r - Hm + m_n h_n) - x_n h_n^H A^{-1} h_n|^2}{h_n^H A^{-1} h_n - v_n (h_n^H A^{-1} h_n)^2}}
$$

(62)

From (62) and (2), it is easy to show that the extrinsic LLR $L^e(c_{n,q})$ in (3) and that in [2] are exactly the same.

**APPENDIX C**

**PROOF OF PROPOSITION 2**

Proof: Eqn (39) in RDMAP can be rewritten as

$$
P_{RDMAP}(x_n | r) \propto \sum_{x_n^D} P(x_n^D) \exp \left[ - (r_n - H_n x_n^D)^H R_n^{-1} (r_n - H_n x_n^D) \right]
$$

(63)

where

$$
r_n = r - H_n^G m_n^G.
$$

(64)
By comparing (63) with (29), we need to prove that
\[
\exp \left[ - (x_n^D - z_n) H_n^{-1} (x_n^D - z_n) \right]
\approx \exp \left[ - (r_n - H_n^D x_n^D)^T R_n^{-1} (r_n - H_n^D x_n^D) \right]
\]  
(65)
with respect to \( x_n^D \). It is not hard to show that
\[
\exp \left[ - (r_n - H_n^D x_n^D)^T R_n^{-1} (r_n - H_n^D x_n^D) \right]
\approx \exp \left[ - (x_n^D - e_n)^T (H_n^D)^T R_n^{-1} H_n^D (x_n^D - e_n) \right]
\]  
(66)
with respect to \( x_n^D \), where
\[
e_n = ((H_n^D)^T R_n^{-1} H_n^D)^{-1} (H_n^D)^T R_n^{-1} r_n.
\]  
(67)
From (65) and (66), we only need to prove that
\[
Z_n^{-1} = (H_n^D)H R_n^{-1} H_n^D
\]  
(68)
and
\[
z_n = e_n.
\]  
(69)

Now we prove (68). Recall the definition of matrix \( C \) in (23). Using the matrix inversion lemma [17], we have
\[
S_n C_n^T = S_n (V^{-1} + \frac{1}{\sigma^2} H^T H)^{-1} S_n^T
\]
\[
= S_n V S_n^T - S_n V H^T (H V H^T + \sigma^2 I)^{-1} H V S_n^T
\]
\[
= V^D - V^D (H_n^D)^T (R_n + H_n^D V^D (H_n^D)^T)^{-1} H_n^D V^D
\]
\[
= ((H_n^D)^T R_n^{-1} H_n^D + (V^D)^{-1})^{-1}.
\]  
(70)
Using (27) and (70), we have
\[
Z_n^{-1} = (S_n C_n^T)^{-1} - (V^D)^{-1}
\]
\[
= (H_n^D)^T R_n^{-1} H_n^D.
\]  
(71)
Next, we prove (69). For the simplicity of notation, we define matrix \( A = H V H^T + \sigma^2 I \). With the matrix inversion lemma [17], we have
\[
(H_n^D)^T R_n^{-1} = (H_n^D)^T A^{-1} H_n^D (V_n^D)^{-1}
\]
\[
= (I - (H_n^D)^T H_n^D)^{-1} (V_n^D)^{-1}
\]
\[
= (V_n^D V_n^D (H_n^D)^T A^{-1} H_n^D V_n^D)^{-1} V_n^D (H_n^D)^T H_n^D A^{-1}
\]
\[
= (S_n C_n^T)^{-1} S_n V H A^{-1}.
\]  
(72)
Using the matrix inversion lemma [17] also yields
\[
V H^T A^{-1} = \frac{1}{\sigma^2} C H
\]  
(73)
Substituting (73) to (72), we have
\[
(H_n^D)^T R_n^{-1} = \frac{1}{\sigma^2} (S_n C_n^T)^{-1} S_n C H
\]  
(74)
We use \( P_n \) to denote the selecting matrix for \( x_n^G \), i.e., \( P_n x = x_n^G \). Then, we have
\[
S_n C P_n^T = S_n V P_n^T - S_n V H^T A^{-1} H P_n^T
\]
\[
= -S_n V H^T A^{-1} H S_n^T V
\]  
(75)
where we note that \( S_n V P_n^T = 0 \). It is not hard to show that
\[
S_n C V^{-1} m = S_n C [S_n^T P_n^T]^{-1} (V_n^D)^{-1} m_n^G.
\]  
(76)
Based on (72), (75), and (76), we have
\[
(S_n C S_n^T)^{-1} S_n C V^{-1} m - (V_n^D)^{-1} m_n^G
\]
\[
= (S_n C S_n^T)^{-1} S_n C P_n^T (V_n^D)^{-1} m_n^G
\]
\[
= - (S_n C S_n^T)^{-1} S_n V H^T A^{-1} H m_n^G
\]
\[
= - (H_n^D)^T H_n^D m_n^G.
\]  
(77)
Using (24), (28), (68), (74), and (77), we have
\[
z_n = Z_n \left( (S_n C S_n^T)^{-1} S_n C (V^{-1} m + \frac{1}{\sigma^2} H^T r - (V_n^D)^{-1} m_n^D) \right)
\]
\[
= (H_n^D)^T R_n^{-1} (H_n^D)^T (r - H_n^D m_n^G)^{-1}
\]
\[
\approx e_n.
\]  
(78)
REFERENCES

QINGHUA GUO (S’07–M’08) received the B.E. degree in electronic engineering and the M.E. degree in signal and information processing from Xidian University, Xi’an, China, in 2001 and 2004, respectively, and the Ph.D. degree in electronic engineering from the City University of Hong Kong, Hong Kong, in 2008.

He is currently with the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW, Australia, and also with the School of Electrical, Electronic, and Computer Engineering, University of Western Australia, Perth, WA, Australia. His research interests include statistical signal processing and telecommunications. He was a recipient of the Australian Research Council’s inaugural Discovery Early Career Researcher Award.

DEFENG (DAVID) HUANG (M’01–S’02–M’05–SM’07) received the B.E.E.E. and M.E.E.E. degrees in electronic engineering from Tsinghua University, Beijing, China, in 1996 and 1999, respectively, and the Ph.D. degree in electrical and electronic engineering from the Hong Kong University of Science and Technology (HKUST), Hong Kong, in 2004.

He is currently a Professor with the School of Electrical, Electronic and Computer Engineering, University of Western Australia, Perth, WA, Australia. He serves as an Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS, and also served as an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2005 to 2011).

SVEN NORDHOLM (SM’04) is a Professor with Curtin University, Bentley, WA, Australia. He received the M.Sc.E.E. (Civilingenjör) degree, the Licentiate of Engineering degree, and the Ph.D. degree in signal processing from Lund University, Lund, Sweden, in 1983, 1989, and 1992, respectively.

He has been with the Curtin University of Technology, Perth, WA, since 1999. From 1999 to 2002, he was the Director of the Australian Telecommunications Research Institute and a Professor with Curtin University of Technology. From 2002 to 2009, he was the Director of the Signal Processing Laboratory, Western Australian Telecommunication Research Institute, Perth, a joint institute between the University of Western Australia and Curtin University. Since 2009, he has been a Professor of Electrical and Computer Engineering at Curtin University. From 2012 to 2013, he served as the Head of the Department. He has also been the Chief Scientist and Co-Founder of a startup company Sensear, Perth, which provides voice communication in extreme noise conditions. He is also the Founder and Director of a hearing aid company Hearmore, Perth. He is an Associate Editor of EURASIP Journal on Advances in Signal Processing and Journal of the Franklin Institute. He is a member of the IEEE Technical Committee for Audio and Acoustic Signal Processing.

JIANGTAO XI (M’95–SM’06) received the B.E. degree from the Beijing Institute of Technology, Beijing, China, in 1982, the M.E. degree from Tsinghua University, Beijing, in 1985, and the Ph.D. degree from the University of Wollongong, Wollongong, NSW, Australia, in 1996, all in electrical engineering.

He was a Post-Doctoral Fellow with the Communications Research Laboratory, McMaster University, Hamilton, ON, Canada, from 1995 to 1996, a Technical Staff Member at Bell Laboratories, Lucent Technologies Inc., Murray Hill, NJ, USA, from 1996 to 1998. He was the Chief Technical Officer at TCL IT Group Company, Huizhou, China, from 2000 to 2002. In 2003, he rejoined the University of Wollongong, Wollongong, NSW, Australia, as a Senior Lecturer, where he is currently a Professor and the Head of the School of Electrical, Computer and Telecommunications Engineering. His research interests include signal processing and its applications in various areas, such as optoelectronics, instrumentation and measurement, and telecommunications.

LI PING (S’87–M’91–SM’06–F’10) received the Ph.D. degree from University of Glasgow, Glasgow, U.K., in 1990.

He was a Lecturer with the Department of Electronic Engineering, University of Melbourne, Parkville, VIC, Australia, from 1990 to 1992, and was a Research Staff with Telecom Australia Research Laboratories, Melbourne, Australia, from 1993 to 1995. Since 1996, he has been with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong, where he is currently a Chair Professor of Information Engineering. He was a recipient of the Institution of Electrical Engineers J. J. Thomson Premium in 1993, the Croucher Foundation Award in 2005, and the British Royal Academy of Engineering Distinguished Visiting Fellowship in 2010. He served as a member of the Board of Governors for the IEEE Information Theory Society from 2010 to 2012.