Exact Closed-Form Expression for the Inverse Moments of One-sided Correlated Gram Matrices

Khalil Elkhalil, Abla Kammoun, Tareq Y. Al-Naffouri and Mohamed-Slim Alouini
Electrical Engineering Department, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia.
Email(s): khalil.elkhalil, abla.kammoun, tareq.alnaffouri, slim.alouini@kaust.edu.sa

***Abstract*—In this paper, we tackle the problem of computing the inverse moments of random Gram matrices with one side correlation. The problem is motivated by signal processing and wireless communications applications where such matrices naturally arise. For instance, we provide a closed-form expression for the inverse moments and show how it can exactly analyze the performance of the best linear unbiased estimator in terms of the mean square error.**

***Index Terms*—Inverse moments, Gram matrices, One-sided correlation, BLUE, mean square error.**

I. INTRODUCTION

Recent advances in spectral analysis of large random matrices has attracted a lot of interest in the computation of moments of random matrices [1,2]. These works are mainly driven by the potential of moments in the behavior understanding of certain scalar functionals of random matrices that naturally arise in signal processing and wireless communication applications. For instance, the work in [1] allows to use asymptotic moments in order to get some insights on the transmit power of multiple signal sources. Also, the authors in [3] analyzed the behavior of asymptotic moments in the aim of the design of low complexity receiver that compares to the linear minimum mean square error (LMMSE) in performance. Although relying on the asymptotic moments permits to have closed-form results on the performance and enables more tractable designs from which insights can be extracted easily, it fails to provide accurate results as long as finite dimensions are considered. As a matter of fact, asymptotic moments are not of much interest in the finite dimension case and thus the exact approach has to be considered alternatively.

In fact, as far as Gram random matrices are considered, the exact approach relies on the available expression of the marginal eigenvalues' density that has been limited to the computation of moments for Wishart random matrices [4,5]. As such, the case of random Gram matrices has never been thoroughly investigated to the best of the authors' knowledge.

This paper tackles this problem and thus provides an answer to the so-far unsolved problem of computing the inverse moments of Gram random matrices. More precisely, we consider the exact derivation of moments of random matrices taking the form $\mathbf{S} = \mathbf{H}^*\mathbf{\Lambda}\mathbf{H}$, where $\mathbf{H}$ is a $n \times m$ $(n > m)$ matrix with independent and identically distributed (i.i.d) zero-mean unit variance complex Gaussian random entries, and $\mathbf{\Lambda}$ is a deterministic $n \times n$ positive definite matrix. Our derivations are mainly based on the already derived Mellin transform [6]. However, the expressions provided in [6] are not straightforward to use since it make appear singularity issues as it will be shown in the course of the paper. Having the exact inverse moments in hands permits to revisit some problems in signal detection in signal processing, a problem that is of tremendous importance to the signal processing community. Interestingly, we show that we can exactly analyze the performance of the best linear unbiased estimator (BLUE) in terms of mean square error.

The remainder of this paper is organized as follows. In section II, we present the motivations behind the current work. In section III, we provide the main result of the paper with a detailed proof. In section IV, we provide some numerical results that validate our theoretical result. We then conclude the paper in section V.

Notations: Throughout this paper, we use the following notation: $\mathbb{E}\left(\mathbf{X}\right)$ stands for the expectation of a random quantity $\mathbf{X}$ and $\mathbb{E}_{\mathbf{X}}\left(f\right)$ stands for the expected value of f with respect to $\mathbf{X}$. Matrices are denoted by bold capital letters, rows and columns of the matrices are referred with lower case bold letters ($\mathbf{I}_n$ is the size-n identity matrix). Given a matrix $\mathbf{A}$, we use $\left[\mathbf{A}\right]_{i,j}$ to refer its (i,j)th entry and use $\mathbf{A}^t$ and $\mathbf{A}^*$ to respectively denote its transpose and Hermitian. When $\mathbf{A}$ is a square matrix, we respectively denote by $\operatorname{tr}\left(\mathbf{A}\right)$, $\det\left(\mathbf{A}\right)$ and $\|\mathbf{A}\|$ its trace, determinant and spectral norm. Finally, we denote by $\operatorname{diag}\left[a_1, a_2, \cdots, a_n\right]$ the diagonal matrix with diagonal elements, $a_1, a_2, \cdots, a_n$.

II. MOTIVATION (LINEAR ESTIMATION)

Estimating signals from sequence of observations has been extensively studied in the signal processing literature [7]–[9]. The problem can be solved if joint statistics relating the observations and the unknown signal are available. However, obtaining joint statistics is in general out of reach regarding the unknown nature of the signal or simply because these kind of information is unavailable. This limitation can be addressed by considering sub-optimal techniques for minimizing the *mean square error*. As an alternative, one can think of applying a linear transformation to the observed vector. While this technique is sub-optimal in nature, it provides a more tractable framework and allows to explicitly analyze the system performance.

Consider the case where the output and the input are related as follows

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}, \tag{1}$$

where $\mathbf{y} \in \mathbb{C}^{n\times 1}$ is the observed vector, $\mathbf{H} \in \mathbb{C}^{n\times m}$ the channel matrix, $\mathbf{x} \in \mathbb{C}^{m\times 1}$ the unknown signal vector and $\mathbf{z} \in \mathbb{C}^{n\times 1}$ the noise vector with covariance matrix $\mathbf{\Sigma}_z$. Here, $\mathbf{H}$ represents the channel matrix of a block fading model, where it is assumed to be constant during a given time interval (block) and changes independently from one block to the next. In this line, it is of tremendous importance to focus on average performances by taking expectation over $\mathbf{H}$. In what follows, we make the following assumptions

- $\mathbf{H}$ is a $(n \times m)$ matrix with *i.i.d* complex zero mean unit variance Gaussian random entries,
- $\mathbf{z}$ is a $(n \times 1)$ zero mean additive Gaussian noise with covariance matrix $\mathbf{\Sigma}_z = \mathbb{E}\{\mathbf{z}\mathbf{z}^*\}$, i.e. $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}_n, \mathbf{\Sigma}_z)$.

Let $n > m$, and assume that the noise covariance matrix $\mathbf{\Sigma}_z$ is perfectly known, then the best linear unbiased estimator (BLUE) [8] recovers $\mathbf{x}$ as

$$\begin{aligned}\hat{\mathbf{x}}_{blue} &= \left(\mathbf{H}^*\mathbf{\Sigma}_z^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^*\mathbf{\Sigma}_z^{-1}\mathbf{y} \\ &= \mathbf{x} + \left(\mathbf{H}^*\mathbf{\Sigma}_z^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^*\mathbf{\Sigma}_z^{-1}\mathbf{z} \\ &= \mathbf{x} + \mathbf{e}_{blue},\end{aligned} \tag{2}$$

where $\mathbf{e}_{blue} = \left(\mathbf{H}^*\mathbf{\Sigma}_z^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^*\mathbf{\Sigma}_z^{-1}\mathbf{z}$ is the residual error after applying the BLUE, where the covariance matrix of $\mathbf{e}_{blue}$ is given by $\mathbf{\Sigma}_{e,blue} = \mathbb{E}\{\mathbf{e}_{blue}\mathbf{e}_{blue}^*\} = \left(\mathbf{H}^*\mathbf{\Sigma}_z^{-1}\mathbf{H}\right)^{-1}$. Consequently, the BLUE average estimation error can be derived as follows

$$\begin{aligned}\mathbb{E}_{\mathbf{H}}\{\|\hat{\mathbf{x}}_{blue} - \mathbf{x}\|^2\} &= \mathbb{E}_{\mathbf{H}}\mathrm{tr}\left[\mathbf{\Sigma}_{e,blue}\right] \\ &= \mathbb{E}_{\mathbf{H}}\mathrm{tr}\left[\left(\mathbf{H}^*\mathbf{\Sigma}_z^{-1}\mathbf{H}\right)^{-1}\right].\end{aligned} \tag{3}$$

To the best of the authors knowledge, the quantity $\mathbb{E}_{\mathbf{H}}\mathrm{tr}\left[\left(\mathbf{H}^*\mathbf{\Sigma}_z^{-1}\mathbf{H}\right)^{-1}\right]$ has never been expressed in closed-form. This constitutes the main motivation of the current paper. In the sequel, we solve the general case where we derive closed-form expressions for the moments given by

$$\mu_{\mathbf{\Lambda}}(r) \triangleq \frac{1}{m}\mathrm{tr}\left[\mathbb{E}_{\mathbf{H}}\{\mathbf{S}^r\}\right], \quad 1 \le r \le p = \min(m, n-m). \tag{4}$$

where $\mathbf{S} = \mathbf{H}^*\mathbf{\Lambda}\mathbf{H}$ is a Gram matrix and $\mathbf{\Lambda}$ is a deterministic $(n \times n)$ positive definite matrix with distinct eigenvalues $(\theta_1 < \cdots < \theta_n)$.

III. Exact Closed-Form Expression for the Moments

In this section, we state the main contribution of this paper. Before stating the main result, we start by providing some useful definitions.

Lemma 1. *[6, Theorem 2] Let $\mathbf{S}$ be as in (4). Then,*

$$\begin{aligned}\mathcal{M}_{f_\lambda}(s) = L\sum_{j=1}^{m}\sum_{i=1}^{m}\mathcal{D}(i,j)\,\Gamma(s+j-1)\Bigg(&\theta_{n-m+i}^{n-m+s+j-2} \\ &- \sum_{l=1}^{n-m}\sum_{k=1}^{n-m}\left[\mathbf{\Psi}^{-1}\right]_{k,l}\theta_l^{n-m+s+j-2}\theta_{n-m+i}^{k-1}\Bigg),\end{aligned} \tag{5}$$

with $L = \frac{\det(\mathbf{\Psi})}{m\prod_{k<l}^{n}(\theta_l-\theta_k)\prod_{l=1}^{m-1} l!}$, $\Gamma(.)$ the Gamma function, $\mathbf{\Psi}$ the $(n-m)\times(n-m)$ Vandermonde matrix given by

$$\mathbf{\Psi} = \begin{bmatrix} 1 & \theta_1 & \cdots & \theta_1^{n-m-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \theta_{n-m} & \cdots & \theta_{n-m}^{n-m-1}\end{bmatrix}$$

and $\mathcal{D}(i,j)$ the $(i,j)-$cofactor of the $(m\times m)$ matrix $\mathcal{C}$ whose $(l,k)-$th entry is given by

$$\begin{aligned}\left[\mathcal{C}\right]_{l,k} = (k-1)!\Bigg(&\theta_{n-m+l}^{n-m+k-1} - \sum_{p=1}^{n-m}\sum_{q=1}^{n-m}\left[\mathbf{\Psi}^{-1}\right]_{p,q} \\ &\times \theta_{n-m+l}^{p-1}\theta_q^{n-m+k-1}\Bigg).\end{aligned}$$

From the Mellin transform expression, the inverse moments defined in (4) can be obtained by a crude substitution of s by $-r-1$ with $r \ge 0$. However, this is unfeasible since in some terms in the sum the Gamma function will be applied to negative integers on which it is not defined. This may lead to a quick conclusion that the inverse moments are infinite. However, relying on existing results on inverse moments of Wishart matrices, lead us to suspect the feasibility of deriving the inverse moments of Gram matrices. An intuitive way to deal with the divergent behavior of the Gamma function is to expect its contribution to be canceled out for some terms and to converge to a limit with others. To study this behavior, we resort to the study of $\mathcal{M}_{f_\lambda}(s-r+1)$ for infinitesimal values of s. Such an intuition is confirmed by the following lemma.

Lemma 2. *If $r \le n-m$, then the limit $\lim_{s\downarrow 0}\mathcal{M}_{f_\lambda}(s-r+1)$ exists and*

$$\mu_{\mathbf{\Lambda}}(-r) = \lim_{s\downarrow 0}\mathcal{M}_{f_\lambda}(s-r+1).$$

Proof: See the proof of Lemma 2 in [10]. ∎

The key idea now is to observe that the expression of $\mathcal{M}_{f_\lambda}(s-r+1)$ reveals that the sum over j makes appear two types of terms. The first one corresponds to the indices of j where the Gamma function is well defined ($-r+j-1$ is positive). The second, turns out to be more complex to analyze as it contains indices where the Gamma function is evaluated on negative values. Having said that, the Mellin transform $\mathcal{M}_{f_\lambda}(s-r+1)$ is decomposed as follows

$$\mathcal{M}_{f_\lambda}(s-r+1) = \mathcal{M}_1(s-r+1) + \mathcal{M}_2(s-r+1), \tag{6}$$

where

$$\mathcal{M}_1(s) = L\sum_{j=1}^{r}\sum_{i=1}^{m}\mathcal{D}(i,j)\,\Gamma(s+j-1)\Bigg(\theta_{n-m+i}^{n-m+s+j-2} - \sum_{l=1}^{n-m}\sum_{k=1}^{n-m}\left[\boldsymbol{\Psi}^{-1}\right]_{k,l}\theta_l^{n-m+s+j-2}\theta_{n-m+i}^{k-1}\Bigg).$$

$$\mathcal{M}_2(s) = L\sum_{j=r+1}^{m}\sum_{i=1}^{m}\mathcal{D}(i,j)\,\Gamma(s+j-1)\Bigg(\theta_{n-m+i}^{n-m+s+j-2} - \sum_{l=1}^{n-m}\sum_{k=1}^{n-m}\left[\boldsymbol{\Psi}^{-1}\right]_{k,l}\theta_l^{n-m+s+j-2}\theta_{n-m+i}^{k-1}\Bigg).$$

Let's start by handling the second term $\mathcal{M}_2(s-r+1)$ that gathers indices for which the Gamma function is evaluated on positive integers. The following lemma shows that the limit of the second term is zero as $s\downarrow 0$ which means that it does not contribute to the final expression of the moment.

Proposition 1. *The term $\mathcal{M}_2(s-r+1)$ vanishes as s goes to zero i.e,*

$$\lim_{s\to 0}\mathcal{M}_2(s-r+1)=0, \quad r=1,\cdots,m.$$

Proof: Note that

$$\begin{aligned}
&\lim_{s\to 0}\mathcal{M}_2(s-r+1)\\
&= L\sum_{j=r+1}^{m}\sum_{i=1}^{m}\mathcal{D}(i,j)\,\Gamma(-r+j)\\
&\times\Bigg(\theta_{n-m+i}^{n-m-r+j-1} - \sum_{l=1}^{n-m}\sum_{k=1}^{n-m}\left[\boldsymbol{\Psi}^{-1}\right]_{k,l}\theta_l^{n-m-r+j-1}\theta_{n-m+i}^{k-1}\Bigg)\\
&= L\sum_{j=r+1}^{m}\sum_{i=1}^{m}[\mathcal{D}]_{i,j}\,[\mathcal{C}]_{i,j-r}\\
&= L\sum_{j=r+1}^{m}\left[\mathcal{D}^t\mathcal{C}\right]_{j,j-r},
\end{aligned}$$

where $\mathcal{D}$ and $\mathcal{C}$ are as defined in Lemma 1. Since $\mathcal{D}$ is the cofactor of $\mathcal{C}$, then $\mathcal{D}^t\mathcal{C}=\det(\mathcal{C})\,\mathbf{I}_m$. Therefore, $[\mathcal{D}^t\mathcal{C}]_{j,j-r}=0$ for $j=r+1,\cdots,m$. ∎

Consequently, the final expression of the moment is totally governed by the limit of the first term $\mathcal{M}_1(s-r+1)$. To provide the final expression of its limit, we are going to need the following piece of notation.

$$\begin{aligned}
\mathbf{a}_j &= \left[\theta_1^{n-m-r+j-1},\theta_2^{n-m-r+j-1},\cdots,\theta_{n-m}^{n-m-r+j-1}\right]^t.\\
\mathbf{b}_i &= \left[1,\theta_{n-m+i},\cdots,\theta_{n-m+i}^{n-m-1}\right]^t.\\
\mathbf{D}_i &= \mathrm{diag}\Bigg[\log\left(\frac{\theta_{n-m+i}}{\theta_1}\right),\log\left(\frac{\theta_{n-m+i}}{\theta_2}\right),\\
&\qquad\cdots,\log\left(\frac{\theta_{n-m+i}}{\theta_{n-m}}\right)\Bigg].
\end{aligned}$$

Having defined the previous notations, we are now in position to state the following result

Proposition 2. *Let $p=\min(m,n-m)$, then for $1\le r\le p$ we have*

$$\begin{aligned}
&\lim_{s\to 0}\mathcal{M}_1(s-r+1)\\
&= L\sum_{j=1}^{r}\sum_{i=1}^{m}\mathcal{D}(i,j)\,\frac{(-1)^{r-j}}{(r-j)!}\mathbf{b}_i^t\boldsymbol{\Psi}^{-1}\mathbf{D}_i\mathbf{a}_j.
\end{aligned}$$

Proof: The handling of $\mathcal{M}_1(s-r+1)$ is delicate because it involves evaluation of the Gamma function at negative integers. Hopefully, a compensation effect occurs due to the multiplicative term in front of the Gamma function. The proof relies on a divide and conquer strategy that consists of decomposing $\mathcal{M}_1(s-r+1)$ into a sum of terms and then evaluating each term separately. To this end, we need to introduce the following notation.

$$\begin{aligned}
\boldsymbol{\Psi}_s &\triangleq \begin{bmatrix}\theta_1^s & \theta_1^{1+s} & \cdots & \theta_1^{n-m+s-1}\\ \vdots & \vdots & \ddots & \vdots\\ \theta_{n-m}^s & \theta_{n-m}^{1+s} & \cdots & \theta_{n-m}^{n-m+s-1}\end{bmatrix}.\\
\mathbf{a}_{s,j} &\triangleq \Bigg[\theta_1^{n-m+s-r+j-1},\theta_2^{n-m+s-r+j-1},\\
&\qquad\cdots,\theta_{n-m}^{n-m+s-r+j-1}\Bigg]^t.\\
\mathbf{b}_{s,i} &\triangleq \left[\theta_{n-m+i}^s,\theta_{n-m+i}^{1+s},\cdots,\theta_{n-m+i}^{n-m+s-1}\right]^t.\\
\mathbf{e}_k &\triangleq \left[\mathbf{zeros}(n-m-k-1),1,\mathbf{zeros}(k)\right]^t,\\
&\quad k=0,\cdots,n-m-1,
\end{aligned} \tag{7}$$

where $\mathbf{zeros}(k)$ is the zero vector of dimension k. Using the previously defined varaibles, we can rewrite $\mathcal{M}_1(s-r+1)$ as in (8) (on top of the next page).

The first term in equation (8) is equal to zero. This can be seen by noticing that $\boldsymbol{\Psi}_s\mathbf{e}_{r-j}=\mathbf{a}_{s,j}$ and $\mathbf{b}_{s,i}^t\mathbf{e}_{r-j}=\theta_{n-m+i}^{n-m+s-r+j-1}$. Thus, $\boldsymbol{\Psi}_s^{-1}\mathbf{a}_{s,j}=\mathbf{e}_{r-j}$ and consequently $\mathbf{b}_{s,i}^t\boldsymbol{\Psi}_s^{-1}\mathbf{a}_{s,j}=\theta_{n-m+i}^{n-m+s-r+j-1}$.
It remains thus to deal with the last two terms. Using a Taylor approximation of $\mathbf{b}_{s,i}$ as s approaching 0, we have

$$\begin{aligned}
\mathbf{b}_{s,i}-\mathbf{b}_i &= s\Bigg[\log(\theta_{n-m+i}),\theta_{n-m+i}\log(\theta_{n-m+i}),\\
&\qquad\cdots,\theta_{n-m+i}^{n-m-1}\log(\theta_{n-m+i})\Bigg]^t + o(s)\\
&= s\log(\theta_{n-m+i})\,\mathbf{b}_i + o(s).
\end{aligned} \tag{9}$$

To deal with the Gamma function evaluated at non positive integers, we rely on the result of the following lemma.

Lemma 3. *[11] For non positive arguments $-k$, $k=0,1,2,\cdots$, the Gamma function can be evaluated as*

$$\lim_{s\to 0}\frac{\Gamma(s-k)}{\Gamma(s)}=\frac{(-1)^k}{k!}, \tag{10}$$

where $\Gamma(s)=\frac{1}{s}+o(s)$ as s approaches 0.

Thus, $\Gamma(s-r+j)\underset{s\to 0}{=}\frac{(-1)^{r-j}}{s(r-j)!}+o(s)$. Therefore , as s

$$\begin{aligned}
\mathcal{M}_1(s-r+1) &= L\sum_{j=1}^{r}\sum_{i=1}^{m}\mathcal{D}(i,j)\,\Gamma(s-r+j)\left(\theta_{n-m+i}^{n-m+s-r+j-1}-\sum_{l=1}^{n-m}\sum_{k=1}^{n-m}\left[\boldsymbol{\Psi}^{-1}\right]_{k,l}\theta_l^{n-m+s-r+j-1}\theta_{n-m+i}^{k-1}\right)\\
&= L\sum_{j=1}^{r}\sum_{i=1}^{m}\mathcal{D}(i,j)\,\Gamma(s-r+j)\left(\theta_{n-m+i}^{n-m+s-r+j-1}-\mathbf{b}_i^t\boldsymbol{\Psi}^{-1}\mathbf{a}_{s,j}\right)\\
&= L\sum_{j=1}^{r}\sum_{i=1}^{m}\mathcal{D}(i,j)\,\Gamma(s-r+j)\left(\theta_{n-m+i}^{n-m+s-r+j-1}-\mathbf{b}_i^t\boldsymbol{\Psi}_s^{-1}\mathbf{a}_{s,j}\right)+L\sum_{j=1}^{r}\sum_{i=1}^{m}\mathcal{D}(i,j)\,\Gamma(s-r+j)\\
&\times\mathbf{b}_i^t\left(\boldsymbol{\Psi}_s^{-1}-\boldsymbol{\Psi}^{-1}\right)\mathbf{a}_{s,j}\\
&= L\sum_{j=1}^{r}\sum_{i=1}^{m}\mathcal{D}(i,j)\,\Gamma(s-r+j)\left(\theta_{n-m+i}^{n-m+s-r+j-1}-\mathbf{b}_{s,i}^t\boldsymbol{\Psi}_s^{-1}\mathbf{a}_{s,j}\right)+L\sum_{j=1}^{r}\sum_{i=1}^{m}\mathcal{D}(i,j)\,\Gamma(s-r+j)\\
&\times\left(\mathbf{b}_{s,i}^t-\mathbf{b}_i^t\right)\boldsymbol{\Psi}_s^{-1}\mathbf{a}_{s,j}+L\sum_{j=1}^{r}\sum_{i=1}^{m}\mathcal{D}(i,j)\,\Gamma(s-r+j)\,\mathbf{b}_i^t\left(\boldsymbol{\Psi}_s^{-1}-\boldsymbol{\Psi}^{-1}\right)\mathbf{a}_{s,j}
\end{aligned} \tag{8}$$

approaches 0, we have

$$\begin{aligned}
&\Gamma(s-r+j)\left(\mathbf{b}_{s,i}^t-\mathbf{b}_i^t\right)\boldsymbol{\Psi}_s^{-1}\mathbf{a}_{s,j}\\
&=\frac{(-1)^{r-j}\log(\theta_{n-m+i})}{(r-j)!}\mathbf{b}_i^t\boldsymbol{\Psi}^{-1}\mathbf{a}_j+o(s),
\end{aligned}$$

Finally, to deal with the last term, we use the following resolvent identity

$$\boldsymbol{\Psi}_s^{-1}-\boldsymbol{\Psi}^{-1}=\boldsymbol{\Psi}_s^{-1}\left(\boldsymbol{\Psi}-\boldsymbol{\Psi}_s\right)\boldsymbol{\Psi}^{-1}.$$

We also make use of the fact that as s approaches 0

$$\left(\boldsymbol{\Psi}-\boldsymbol{\Psi}_s\right)\underset{s\to 0}{=}-s\tilde{\boldsymbol{\Psi}}+o(s),$$

where

$$\tilde{\boldsymbol{\Psi}}=\boldsymbol{\Phi}\boldsymbol{\Psi},$$

with $\boldsymbol{\Phi}=\operatorname{diag}\left[\log(\theta_1),\log(\theta_2),\cdots,\log(\theta_{n-m})\right]$. Thus, as s approaches 0, we have

$$\begin{aligned}
&\Gamma(s-r+j)\,\mathbf{b}_i^t\left(\boldsymbol{\Psi}_s^{-1}-\boldsymbol{\Psi}^{-1}\right)\mathbf{a}_{s,j}\\
&\underset{s\to 0}{=}\frac{(-1)^{r+1-j}}{(r-j)!}\mathbf{b}_i^t\boldsymbol{\Psi}^{-1}\tilde{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1}\mathbf{a}_j+o(s).
\end{aligned}$$

Finally, we have the following limit

$$\begin{aligned}
&\lim_{s\to 0}\mathcal{M}_1(s-r+1)\\
&=L\sum_{j=1}^{r}\sum_{i=1}^{m}\mathcal{D}(i,j)\left[\frac{(-1)^{r-j}\log(\theta_{n-m+i})}{(r-j)!}\mathbf{b}_i^t\boldsymbol{\Psi}^{-1}\mathbf{a}_j\right.\\
&\left.+\frac{(-1)^{r+1-j}}{(r-j)!}\mathbf{b}_i^t\boldsymbol{\Psi}^{-1}\tilde{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1}\mathbf{a}_j\right].
\end{aligned}$$

This expression can be further simplified by noticing that $\tilde{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1}=\boldsymbol{\Phi}$. Finally, we have

$$\begin{aligned}
&\lim_{s\to 0}\mathcal{M}_1(s-r+1)\\
&=L\sum_{j=1}^{r}\sum_{i=1}^{m}\mathcal{D}(i,j)\frac{(-1)^{r-j}}{(r-j)!}\mathbf{b}_i^t\boldsymbol{\Psi}^{-1}\Big[\log(\theta_{n-m+i})\,\mathbf{I}_{n-m}\\
&-\boldsymbol{\Phi}\Big]\mathbf{a}_j\\
&=L\sum_{j=1}^{r}\sum_{i=1}^{m}\mathcal{D}(i,j)\frac{(-1)^{r-j}}{(r-j)!}\mathbf{b}_i^t\boldsymbol{\Psi}^{-1}\mathbf{D}_i\mathbf{a}_j.
\end{aligned}$$

Thereby ending up the proof of the proposition. ■

Taking into account the results of the above propositions, the final expression of the inverse moments is given by the following theorem.

Theorem 1. *Let* $p=\min(m,n-m)$*, then for* $1\le r\le p$*, we have*

$$\mu_{\boldsymbol{\Lambda}}(-r)=L\sum_{j=1}^{r}\sum_{i=1}^{m}\mathcal{D}(i,j)\frac{(-1)^{r-j}}{(r-j)!}\mathbf{b}_i^t\boldsymbol{\Psi}^{-1}\mathbf{D}_i\mathbf{a}_j.$$

IV. Numerical Validation and Discussion

To validate the findings of Theorem 1, we start by comparing the BLUE normalized mean square error with its corresponding exact value provided by the Theorem 1. Note that the normalized mean square error for the BLUE can be expressed as

$$\text{NMSE}_{(blue)}\triangleq\frac{1}{m}\mathbb{E}_{\mathbf{H}}\{\|\hat{\mathbf{x}}_{blue}-\mathbf{x}\|^2\}, \tag{11}$$

where $\mathbf{x}$ and $\mathbf{x}_{blue}$ are as defined in (1) and (2) respectively. Let's assume that the channel matrix $\mathbf{H}$ has a transmit correlation matrix given by $\boldsymbol{\Lambda}$, i.e. $\mathbf{H}=\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{W}$, where $\mathbf{W}$ has i.i.d zero-mean unit variance complex Gaussian random entries and $\boldsymbol{\Lambda}$ adopts the following structure

$$[\boldsymbol{\Lambda}]_{i,j}=J_0\left(\pi\,|i-j|^2\right), \tag{12}$$

where $J_0(.)$ is the zero-order Bessel function of the first kind. This kind of matrices is widely used to model the correlation

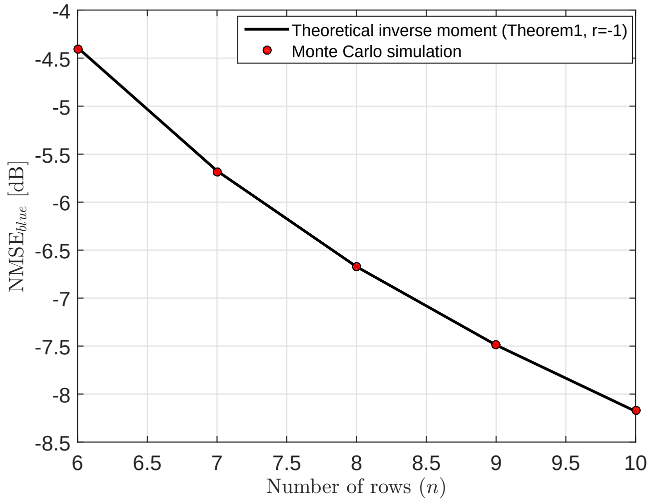

Figure 1: BLUE average estimation error for $m = 3$: Monte Carlo simulation (5×10^4 realizations) versus theory (Theorem 1)

between transmit antennas in a dense scattering environment [6,12]. For simplicity, we also assume that $\mathbf{\Sigma}_z = \mathbf{I}_n$[1]. Based on (3), the NMSE$_{(blue)}$ is thus given by

$$\text{NMSE}_{(blue)} = \mu_{\mathbf{\Lambda}}(-1). \tag{13}$$

As shown in Figure 1, the theoretical formula of the inverse moment obtained in theorem 1 exactly matches the performance given by the Monte Carlo simulation.

To further validate our theoretical findings, we compare $\mu_{\mathbf{\Lambda}}(r)$ with the normalized asymptotic moments derived in [10, Theorem 2] and the empirical moments obtained by simulation. It is clear from Figure 2 that the theoretical inverse moments (Theorem 1) perfectly match the empirical moments (Monte Carlo simulation) for all moment orders, $r \in \{-4, -3, -2, -1\}$. On the other hand, we can notice that the accuracy of the asymptotic approach improves as n increases. However, the accuracy deteriorates as we increase n, that is the asymptotic provides better accuracy for lower order moments as compared to higher order moments.

V. Conclusion

This paper provides a closed-form expression to evaluate the inverse moments of one side correlated Gram matrices. In addition, the paper investigated a possible application for the derived result which is the exact evaluation of the performance of the BLUE in terms of mean square error.

References

[1] *Free Deconvolution for OFDM Multicell SNR Detection*, 2008.

[2] J. Yao, A. Kammoun, and J. Najim, "Eigenvalue estimation of parameterized covariance matrices of large dimensional data," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5893 –5905, nov. 2012.

[3] J. Hoydis, M. Debbah, and M. Kobayashi, "Asymptotic moments for interference mitigation in correlated fading channels," in *Information Theory Proceedings (ISIT)*, July 2011, pp. 2796–2800.

[1] This assumption does not affect our analysis and $\mathbf{\Sigma}_z$ can be taken different from $\mathbf{I}_n$ as long as the matrix $\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{\Sigma}_z^{-1}\mathbf{\Lambda}^{\frac{1}{2}}$ has distinct eigenvalues.

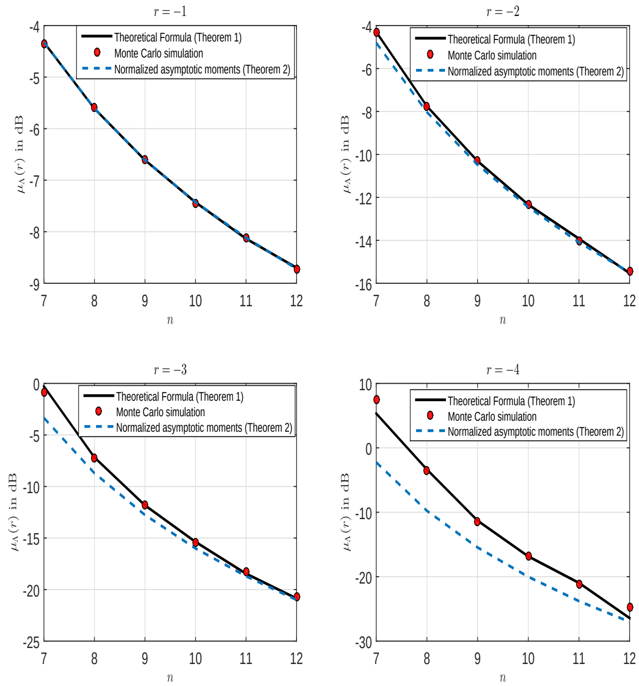

Figure 2: Inverse moments for $\mathbf{\Lambda}$ defined as in (12): A comparison between theoretical result (Theorem 1), normalized asymptotic moments [10, Theorem 2] and Monte Carlo simulations (10^5 realizations).

[4] D. Maiwald and D. Kraus, "Calculation of moments of complex Wishart and Complex Inverse Wishart Distributed Matrices," in *IEEE Proceeding Radar Sonar and Navigation*, 2000.

[5] G. Letac and H. Massam, "All Invariant Moments of the Wishart Distribution," *Scandinavian Journal of Statistics*, vol. 31, no. 2, pp. 295–318, Jun. 2004.

[6] G. Alfano, A. M. Tulino, A. Lozano, and S. Verdu, "Capacity of MIMO Channels with One-sided Correlation," *ISSSTA*, August 2004.

[7] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall, 2000.

[8] A. Sayed, *Fundamentals of Adaptive Filtering*. John Wiley & Sons, 2003.

[9] H. V. Poor, *An Introduction to Signal Detection and Estimation*. Springer-Verlag, 1988.

[10] K. Elkhalil, A. Kammoun, T. Y. Al Naffouri, and M.-S. Alouini, "Analytical Derivation of the Inverse Moments of One-sided Correlated Gram Matrices with Applications," *Submitted to IEEE Trans. on Signal Processing*. [Online]. Available: http://arxiv.org/pdf/1507.00629v1.pdf

[11] M. A. Chaudhry and S. M. Zubair, *On a Class of Incomplete Gamma Function with Applications*. Boca Raton-London-Ney York-Washington, D.C.: Chapman & Hall/CRC, 2002.

[12] X. Chen, "Antenna Correlation and Its Impact On Multi-Antenna System," *Progress In Electromagnetics Research B*, vol. 62, pp. 241–253, 2015.