On the MMSE Estimation of Norm of a Gaussian Vector under Additive White Gaussian Noise with Randomly Missing Input Entries

Samrat Mukhopadhyay

Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, INDIA

Abstract

This paper considers the task of estimating the \(l_2\) norm of a \(n\)-dimensional random Gaussian vector from noisy measurements taken after many of the entries of the vector are missed and only \(K (0 \leq K \leq n)\) entries are retained and others are set to 0. Specifically, we evaluate the minimum mean square error (MMSE) estimator of the \(l_2\) norm of the unknown Gaussian vector performing measurements under additive white Gaussian noise (AWGN) on the vector after the data missing and derive expressions for the corresponding mean square error (MSE). We find that the corresponding MSE normalized by \(n\) tends to 0 as \(n \to \infty\) when \(K/n\) is kept constant. Furthermore, expressions for the MSE is derived when the variance of the AWGN noise tends to either 0 or \(\infty\). These results generalize the results of Dytso et al. [1] where the case \(K = n\) is considered, i.e. the MMSE estimator of norm of random Gaussian vector is derived from measurements under AWGN noise without considering the data missing phenomenon.

Keywords: MMSE estimation; Additive White Gaussian Noise (AWGN); Missing data.

1. Introduction

In this paper we consider the estimation of the \(l_2\) norm of a random vector with missing entries under measurement noise. In particular, the measurement model that we consider in this paper is the following:

\[
Y = BX + N,
\]

where \(X \sim N(0, I_n)\), \(N \sim N(0, \sigma^2 I_n)\), with \(X\) being the input vector and \(N\) the measurement noise vector, which are independent. Furthermore, \(B = \text{diag}(B_1, \cdots, B_n)\), where \(B_i \in \{0, 1\}\) and \(\sum_{i=1}^{n} B_i = K\), for a given non-negative integer \(K \leq n\). The data missing process is assumed to be independent of \(X\) and \(N\), i.e. \(B\) is independent of \(X, N\). The block diagram of the measurement process is shown in Fig. 1.

The problem of estimating \(l_2\) norm of a vector from noisy measurements can have different applications in domains such as distributed computing [2], wireless systems [3] and wireless networks [4], where different nodes in the system might want to estimate the \(l_2\) norm of some vector of interest, either in a centralized or distributed manner in order to perform different tasks such as secure transmission, selection of best transmitting node etc. Dytso et al. [1] have studied the optimal MMSE estimator for the \(l_2\) norm of a random Gaussian vector in additive white Gaussian noise when no entry of the input vector is missing. However, the problem of missing input data entries at random instances is quite common in many signal processing and communication applications [5, 6, 7, 8, 9, 10]. Such missing input...
data model is also prevalent in information and coding theory where deletion channels \([11]\) are used to model a communication channel where sequence of input bits are deleted at random locations. Furthermore, the measurement model of Eq. \((1)\) is also useful in the context of applications where storage is scarce so that only a few of the input vectors can be loaded while the others have to be considered 0.

In this work we find explicit expressions for the minimum mean square error (MMSE) estimator of \(\|X\|_2\) given the measurements according to the model \((1)\). In particular, we obtain an expression for the conditional expectation of \(\|X\|_2\) given the measurement vector \(Y\). We further characterize the mean square error (MSE) for this optimal estimator for \(0 \leq K \leq n\) and demonstrate using numerical simulations that even when \(K\) is small, the MSE is close to the MSE corresponding to the case \(K = n\). Our main contribution can be summarized as the generalization of the results of \([1]\) by characterizing the MMSE estimator in the presence of the data missing phenomenon where only \(0 \leq K \leq N\) entries are retained and showing that even with small \(K\), the MMSE estimation of a Gaussian vector under AWGN noise is pretty accurate, especially for large input size \(n\).

In the following, let \([n] := \{1, \ldots, n\}\). For any vector \(x \in \mathbb{R}^n\) and subset \(S \subset [n]\), \(x_S\) denotes the vector \(x\) restricted to \(S\), i.e., \(x_S\) contains those entries of \(x\) that are indexed by \(S\). Let \(\text{supp}(x)\) denote the support of \(x\), i.e., \([x]_i = 0\) if \(i \notin \text{supp}(x)\). We denote by \(I_n\), the identity matrix of size \(n \times n\) and \(I_S\) the submatrix containing the columns of \(I_n\) indexed by \(S\). We denote by \(P(\cdot)\), \(E[\cdot]\) the probability and expectation operators, respectively, and by \(E_R[\cdot]\) we denote expectation with respect to a random variable \(R\). Let \(\delta(\cdot)\) denote the Dirac delta function, \(F_\lambda\) denote the generalized hypergeometric function \([12]\), defined as

\[
F_\lambda(a_1, \ldots, a_p; b_1, \ldots, b_q; x) = \sum_{k=0}^\infty \frac{(a_1)_k \cdots (a_p)_k}{(b_1)_k \cdots (b_q)_k} \frac{x^k}{k!},
\]

where the Pochhammer symbol \([12]\) \((a)_k\) is defined as \((a)_k = a(a+1) \cdots (a+k-1)\) for any integer \(k \geq 1\), and \((a)_0 = 1\). We denote by \(f_X(\cdot)\) the probability density function of any continuous random variable \(X\). \(B(\cdot, \cdot)\) denotes the beta function defined as \(B(m, n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}\), where \(\Gamma(\cdot)\) is the gamma function.

2. MMSE Estimator and Its Characterization

In this section we present our results on the MMSE estimator of \(\|X\|_2\) given the measurement \(Y\) according to the measurement model \((1)\). The following Theorem presents a closed form expression of the MMSE estimator.

**Theorem 2.1.** Let \(X, Y, B, N\) be as described in the measurement model \((1)\). Let \(P(\text{supp}(B) = S) = 1/\binom{n}{K}\) for any \(S \subset [n]\) with \(|S| = K\) and \(J = n - K\). Then,

\[
E[\|X\|_2 | Y = y, B = b] = E_{U_1, U_2} \left[ \sqrt{2} \frac{\sigma^2}{\sigma^2 + 1} U_1 + U_2 \right] = \sqrt{2} \left( \frac{\sigma^2}{\sigma^2 + 1} \right)^{\frac{J+1}{2}} e^{\frac{-\|y\|_2^2}{2\sigma^2(\sigma^2 + 1)}} \sum_{l=0}^{\infty} a_l \left( \frac{\|y_S\|_2^2}{2\sigma^2(\sigma^2 + 1)} \right)^l,
\]

where, with \(\text{supp}(b) = S\), \(U_1, U_2\) are independent; \(U_1\) has a noncentral \(\chi^2\) distribution with noncentrality parameter \(\lambda = \frac{\|y_S\|_2^2}{\sigma^2}\) and degrees of freedom \(K\), while \(U_2\) has a central \(\chi^2\) distribution with degrees of freedom \(J\). Consequently,

\[
E[\|X\|_2 | Y = y] = \frac{\sqrt{2}}{\binom{n}{K}} \sum_{l=0}^{\infty} a_l e^{\frac{-\|y_S\|_2^2}{2\sigma^2(\sigma^2 + 1)}} \left( \frac{\|y_S\|_2^2}{2\sigma^2(\sigma^2 + 1)} \right)^l
\]

where \(Z_S = \frac{\|y_S\|_2^2}{\sigma^2(\sigma^2 + 1)}\), and

\[
a_l = \frac{1}{l!} \sum_{k=0}^{\infty} \binom{l/2}{k} \Gamma\left( \frac{l/2 + k + l}{2} \right) \frac{1}{k!(\sigma^2 + 1)^k}.
\]

**Proof.** The proof is deferred to Section[3.1].

\[\square\]
Remark 2.1. The Theorem 2.1 provides several alternate descriptions of the MMSE estimator. First of all, while Eq. (2) enables one to evaluate the conditional expectation \( \mathbb{E} [\|X\|_2^2 | Y = y, B = b] \) by Monte-Carlo simulations, Eq. (3) can be used to find the estimator by evaluating terms of a series expansion. Furthermore, while Eq. (4) is a simple consequence of Eq. (3), Eq. (5) gives another interpretation of the MMSE estimator as an expectation over randomly chosen subsets \( S \subset [n] \) such that \( |S| = K \).

The conditional MMSE estimator \( \mathbb{E} [\|X\|_2^2 | Y = y, B = b] \) is evaluated for different values of \( \|y_S\|_2 \), where \( |S| = K \) and the results are plotted in Fig. 2 where it can be observed that even with small number of actual measurements \( K(K = 1, 2) \), the MMSE estimator is pretty close to the one obtained when actual measurement of all the \( n \) coordinates are considered.

We now proceed to calculate the MSE corresponding to the MMSE estimator derived in Eq. (4). The following Theorem presents an expression for the MSE for the estimator in Eq. (4), which is now referred to as \( \text{mse} \).

**Theorem 2.2.** Let \( X, Y, B, N \) be described as in model (1) and \( \mathbb{P}(\text{supp}(B) = S) = 1/(\binom{n}{K}) \) for any \( S \subset [n] \) with \( |S| = K \) and \( J = n - K \). Then,

\[
\text{mse} = n - 2 \frac{\sigma^2}{\sigma^2 + 1} \sum_{r_{\min} \leq r \leq K} \binom{n}{r} \binom{K}{n-K} a_i a_j h_{ij}(r),
\]

where \( r_{\min} = \max\{0, 2K - n\} \) and

\[
h_{ij}(r) = \begin{cases} \frac{\sigma^2 \Gamma(i+j+\frac{1}{2})\Gamma(i+\frac{3}{2})}{\Gamma(i+j+1)} & r = 0, \\ \frac{\sigma^2 \Gamma(i+j+2r+1)\Gamma(i+j+2r-k+l+\frac{1}{2})}{(\sigma^2+1)^{K-r}(\sigma^2+2)^{r+K-r}} & 1 \leq r \leq K-1, \\ \frac{\sigma^2 \Gamma(i+j+2r+1)\Gamma(i+j+2r-k+l)}{\Gamma(i+j+2r+1)^{K-r}} & r = K. \end{cases}
\]

**Proof.** The proof is deferred to Section 3.2. \( \square \)

Remark 2.2. If \( K = n, J = 0 \) and \( r_{\min} = n \), from Eq. (7), \( \text{mse} = n - \frac{2\sigma^2}{\sigma^2 + 1} \sum_{i,j \geq 0} a_i a_j h_{ij}(n) \). Using the expression of \( a_i \) from Eq. (6) and \( h_{ij}(K) \) from Eq. (8), it can be verified that we obtain the same expression as that of Theorem 2 of (1). This establishes the result of Theorem 2.2 as a generalization of Theorem 2 of (1).

The MSE for the estimators for different values of \( K, n \) are plotted in Fig. 3 where it is observed that even with small \( K \) the \( \text{mse} \) is pretty close to the \( \text{mse} \) with \( K = n \) which again.

We conclude this section with some results on the asymptotic behavior of the MMSE:

**Theorem 2.3.** Let \( X, Y, B, N \) be described as in model (1) and \( \mathbb{P}(\text{supp}(B) = S) = 1/(\binom{n}{K}) \) for any \( S \subset [n] \) with \( |S| = K \) and \( J = n - K \). Then,
Consequently, the expression for \( \text{mmse} \) associated to partial entries of \( X \) is
\[
\sigma \to \infty
\]
which is interesting since it shows that even with small ratio \( f \) of \( X \), the estimator reduces to
\[
\text{mmse} = n - (n + J) \frac{\Gamma\left(\frac{\alpha+1}{2}, \frac{\beta}{2}\right)}{\Gamma\left(\frac{\alpha+1}{2}, \frac{\beta}{2}\right)} \frac{\Gamma\left(\frac{\alpha}{2}, \frac{\beta}{2}, -\frac{1}{2}, n+J+1, \frac{\alpha+1}{2}, 1\right)}{\Gamma\left(\frac{\alpha}{2}, \frac{\beta}{2}, -\frac{1}{2}, n+J+1, \frac{\alpha+1}{2}, 1\right)}.
\]

1. For fixed \( K, n \), \( \lim_{\sigma \to 0} \text{mmse} = n - (n + J) \frac{\Gamma\left(\frac{\alpha+1}{2}, \frac{\beta}{2}\right)}{\Gamma\left(\frac{\alpha+1}{2}, \frac{\beta}{2}\right)} \frac{\Gamma\left(\frac{\alpha}{2}, \frac{\beta}{2}, -\frac{1}{2}, n+J+1, \frac{\alpha+1}{2}, 1\right)}{\Gamma\left(\frac{\alpha}{2}, \frac{\beta}{2}, -\frac{1}{2}, n+J+1, \frac{\alpha+1}{2}, 1\right)}.

2. For fixed \( K, n \), \( \lim_{\sigma \to \infty} \text{mmse} = n - 2 \left( \frac{\Gamma\left(\frac{\alpha+1}{2}, \frac{\beta}{2}\right)}{\Gamma\left(\frac{\alpha+1}{2}, \frac{\beta}{2}\right)} \right)^2 \).

3. For fixed \( \sigma > 0 \) and \( K, n \) \( \to \infty \) with \( K/n = p \in [1/n, 1] \), \( \lim_{n \to \infty} \frac{\text{mmse}}{n} = 0 \).

**Proof.** The proof is deferred to Section 3.3. \[\Box\]

**Remark 2.3.** The first result shows that when \( \sigma \to 0 \), there is a nonzero \( \text{mmse} \) arising entirely due to the randomness associated to partial entries of \( X \) due to data missing. Note that if \( J = 0 \) so that no entries of \( X \) is missing, the expression for \( \text{mmse} \) reduces to 0 when \( \sigma \to 0 \), which is intuitively satisfying. The second result shows that when \( \sigma \to \infty \), the \( \text{mmse} \) is identical to the second result of Theorem 3 of [1]. The third result of Theorem 2.3 shows that as long as \( K \) is such that although it is large but the ratio of \( K/n \) is between \( 1/n \) to 1, the ratio \( \text{mmse}/n \) tends to 0, which is interesting since it shows that even with small ratio \( K/n \), the estimator yields near similar performance as the one which uses measurements of all coordinates.

3. Proof of Theorems

3.1. **Proof of Theorem 2.1**

Note that \( \mathbb{E} [\|X\|_2 | Y = y, B = b] = \int_{\mathbb{R}^N} f_{\|X\|_2, Y|B}(t, y | b) dt = \int_{\mathbb{R}^N} f_{\|X\|_2, Y|B}(t, y | b) f_X(x) dx = \int_{\mathbb{R}^N} f_Y|B,Y|B(y | b, x) f_X(x) dx \),

which follows from the observation that given \( X \), \( \|X\|_2 \) is fixed, so that it is independent of \( Y \). Now, given \( B \) such that \( \text{supp}(B) = S \) with \( |S| = K \), \( J = n - K \), we get \( Y_{SC} | X = N_{SC} \) and \( Y_S | X = X_S + N_S \), so that \( Y_S | X \sim X(X_S, \sigma^2 I_S), Y_{SC} | X \sim N(0_{SC}, \sigma^2 I_{SC}) \) and \( Y_S \sim N(0_S, (\sigma^2 + 1)I_S) \), \( Y_{SC} \sim N(0_{SC}, \sigma^2 I_{SC}) \). Consequently,

\[
f_{\|X\|_2,Y|B}(t, y | b) = e^{-\frac{\|y_{SC}\|^2}{2\sigma^2}} \int_{\mathbb{R}^N} e^{-\frac{\|y_S-y\|_2^2}{2\sigma^2}} \delta(\|x\|_2 - t) dx
\]

\[
= e^{-\frac{\|y_{SC}\|^2}{2\sigma^2}} \int_{\mathbb{R}^N} e^{-\frac{\|y_S-y\|_2^2}{2\sigma^2}} \delta(\|x\|_2 - t) dx
\]

\[
= \left(\frac{2\pi(\sigma^2 + 1)}{\Gamma(\frac{\sigma^2}{\sigma^2 + 1})}\right)^{K/2} \left(\frac{2\pi}{\Gamma(\frac{\sigma^2}{\sigma^2 + 1})}\right)^{1/2} \int_{\mathbb{R}^N} \left(\frac{2\pi}{\Gamma(\frac{\sigma^2}{\sigma^2 + 1})}\right)^{K/2} \left(\frac{2\pi}{\Gamma(\frac{\sigma^2}{\sigma^2 + 1})}\right)^{1/2} \delta(\|x\|_2 - t) dx
\]

\[
= f_Y|B(y | b) f_Z(t),
\]
where the penultimate step above uses the following identity:
\[
\frac{\|y_S - x_S\|^2}{2\sigma^2} + \frac{\|x_S\|^2}{2} = \frac{\|x_Sc\|^2}{2} + \frac{\sigma^2 + 1}{2\sigma^2} \|x_S - \frac{y_S}{\sigma^2 + 1}\|^2 + \frac{\|y_S\|^2}{2(\sigma^2 + 1)},
\]
and we define the random variable \(Z = \|\overline{X}_2\|_2\), \(\overline{X} \sim \mathcal{N}(\mu, \Sigma)\), with
\[
\mu = \left[ \frac{y_S}{\sigma + 1} \right], \quad \Sigma = \left[ \frac{\sigma^2 + 1}{\sigma + 1} I_S \right],
\]

Therefore,
\[
E[\|X\|_2 | Y = y, B = b] = \int_0^\infty t f_{Z}(t) dt = E[Z].
\]

Note that one can write \(Z^2 = \frac{\sigma^2}{\sigma + 1} \|X_1\|^2 + \|X_2\|^2\), where \(X_1 \sim \mathcal{N}\left(\frac{y_S}{\sqrt{\sigma(\sigma + 1)}}, I_S\right)\) and \(X_2 \sim \mathcal{N}(0_{SC}, I_{SC})\).

Therefore, \(Z^2 = \frac{\sigma^2}{\sigma + 1} U_1 + U_2\) with the descriptions of \(U_1, U_2\) provided in Theorem 2.1 and consequently Eq. (2) follows.

To further simplify Eq. (2), writing \(V = Z^2\), one obtains, for all \(v \geq 0\),
\[
f_V(v) = f_{\sigma^2/\sigma + 1} u_1(v) \circ f_{\sigma^2/\sigma} u_2(v) = \frac{\sigma^2 + 1}{\sigma^2} \int_0^v f_{U_1} \left( \frac{(\sigma^2 + 1)u}{\sigma^2} \right) f_{U_2}(v - u) du.
\]

Now we recall that for all \(u \geq 0\),
\[
f_{U_1}(\lambda, K; u) = e^{-\nu/2} f_1(\nu; K/2; \lambda u/2) e^{-u/2} u^{K/2 - 1} 2^{K/2} \Gamma(K/2),
\]
\[
f_{U_2}(J; u) = \frac{e^{-u/2} u^{J/2 - 1}}{\Gamma(J/2)^2}. \frac{\nu}{2\sigma^2}. \frac{\nu}{2\sigma^2}.
\]

where we define, for \(\alpha, \beta, \gamma, \nu, \epsilon \geq 0\),
\[
R(\alpha, \beta, \gamma, \nu, \epsilon) = \int_0^1 e^{-\alpha x} x^{\beta - 1}(1 - x)^{\gamma - 1} f_1(\nu; \epsilon x) dx.
\]

In Appendix A the following is established:
\[
R(\alpha, \beta, \gamma, \beta, \epsilon) = e^{-\alpha} B(\beta, \gamma) \sum_{k,l \geq 0} \frac{\alpha^k \beta^l}{k! l! (\beta + \gamma)_{k+l}},
\]

Therefore, one obtains,
\[
f_V(v) = \frac{\left(\frac{\sigma^2 + 1}{\sigma^2}\right)^{K/2}}{2^{n/2} \Gamma(J/2) \Gamma(K/2)} e^{-\nu/2} B(K/2, J/2) \sum_{k,l \geq 0} \frac{\left(\frac{\nu}{2\sigma^2}\right)^{k}}{(n/2)_{k+l}} \int_0^1 e^{(\nu/2)(1 - x)^{-1}} F_1(\nu; x, x) dx.
\]

Consequently,
\[
E[\|X\|_2 | Y = y, B = b] = \int_0^\infty t \cdot 2t f_V(t^2) dt
\]
\[
= \int_0^\infty 2t^2 \frac{\left(\frac{\sigma^2 + 1}{\sigma^2}\right)^{K/2}}{2^{n/2} \Gamma(n/2)} \sum_{k,l \geq 0} t^{2(k+l)} \left(\frac{\nu}{2\sigma^2}\right)^{2l} \frac{(J/2)_k}{(n/2)_{k+l}} dt.
\]
Then we find,

\[ \frac{(J/2)_{k}}{(n/2)_{k+l}} \int_{0}^{\infty} t^{n+2(k+l)} e^{-t^{2}/(n/2)} dt. \tag{20} \]

For any \( a, b > 0 \), it is straightforward to deduce that \( \int_{0}^{\infty} t^{n} e^{-bt^2} dt = \frac{\Gamma(\frac{n+1}{2})}{2b^{n/2}} \), so that,

\[ E[\|X\|_2 \mid Y = y, B = b] = \frac{\sum_{i,j=0}^{\infty} a_{i,j} E[Y] E[Z_{S} Z_{T} e^{-Z_{S} - Z_{R}} | Y]}{2 \sum_{i,j=0}^{\infty} a_{i,j} \sum_{S,T : |S| = |T| = K} E[Y] Z_{S} Z_{T} e^{-Z_{S} - Z_{R}}]. \tag{23} \]

which results in Eq. (3) after plugging in the expressions of \( a_{i}, l \geq 0 \) from Eq. (6). Thereafter, taking expectation over the realizations of \( B \), the Eqs. (4) and (5) follow.

### 3.2. Proof of Theorem 2.2

The expression for MSE of the estimator \( E[\|X\|_2 \mid Y = y] \) is

\[ \min_{a \in \mathbb{R}} E[\|X\|_2 - E[Z_{S} e^{-Z_{S}} | Y]] = n - E[\|X\|_2 - E[Z_{S} e^{-Z_{S}} | Y]]. \tag{22} \]

Using Eq. (5), it follows that,

\[ E[\|X\|_2 - E[Z_{S} e^{-Z_{S}} | Y]] = 2 \sum_{i,j=0}^{\sum_{i,j=0}^{\infty} a_{i,j} \sum_{S,T : |S| = |T| = K} E[Y] Z_{S} Z_{T} e^{-Z_{S} - Z_{R}}]. \]

For any two subsets \( S, T \subset [n] \) such that \( |S| = |T| = K \), let \( P = S \setminus T, Q = T \setminus S, R = S \cap T \). If \( |R| = r, \) then \( P, Q \) are both \( K - r \)-size, and \( Z_{S} = Z_{P} + Z_{R}, Z_{T} = Z_{Q} + Z_{R} \). Since \( Z_{S} = \frac{|Y_{S}|^2}{2\sigma^2(p+1)} \) and \( Y_{S} \sim N(0, (\sigma^2 + 1)I_{S}) \), it follows that if \( 1 \leq r \leq K - 1 \), \( Z_{P}, Z_{Q}, Z_{R} \) are all independent and \( 2\sigma^2 Z_{P} \sim \chi^2_{K-r}, 2\sigma^2 Z_{Q} \sim \chi^2_{K-r}, 2\sigma^2 Z_{R} \sim \chi^2_{r} \). Similarly, if \( r = K, Z_{P} = Z_{Q} = 0 \) and \( 2\sigma^2 Z_{R} \sim \chi^2_{K} \). Let \( h_{ij}(r) \) denote \( E[Z_{R} e^{-(Z_{S} + Z_{R})}] \) when \( |S \cap T| = r, 0 \leq r \leq K \). Then we find,

\[ h_{ij}(0) = E[Z_{R} e^{-Z_{R}}] \]

\[ = \frac{\sigma^{2K} \Gamma(i + \frac{K}{2}) \Gamma(j + \frac{K}{2})}{(\sigma^2 + 1)^{K+i+j} \Gamma(\frac{K}{2})}, \]

\[ h_{ij}(K) = E[Z_{R} e^{-Z_{R}}] \]

\[ = \frac{\sigma^{2K} \Gamma(i + \frac{K}{2}) \Gamma(j + \frac{K}{2})}{(\sigma^2 + 1)^{K+i+j} \Gamma(\frac{K}{2})}, \]

\[ h_{ij}(r) = E[Z_{R} e^{-Z_{R}}] \]

\[ = \frac{\sigma^{2K} \Gamma(s + \frac{r}{2}) \Gamma(t + \frac{K-r}{2}) \Gamma(i + j - (s + t) + \frac{K}{2})}{(\sigma^2 + 1)^{s+t+K} \Gamma(\frac{K}{2})} \]

where we have used the fact that if \( aU \sim \chi^2_{n}(a > 0) \), for \( p, q \geq 0, \)

\[ E[U^{p} e^{-u^{q}}] = \int_{0}^{\infty} u^{p} e^{-u^{q}} du = \frac{\Gamma(p + \frac{q}{2})}{\Gamma(\frac{q}{2})^{2}}, \]

Therefore, \( \sum_{S \subset [n]} E[Z_{R} e^{-(Z_{S} + Z_{R})}] = \sum_{r=0}^{K} N_{r} h_{ij}(r) \), where \( N_{r} (0 \leq r \leq K) \) is the number of pairs of subsets \( S, T \) of \([n]\) such that \( |S| = |T| = K \) and \( |S \cap T| = r \). An expression of \( N_{r} \) can be obtained in the following...
way. First note that one can choose a subset $S \subset [n]$ with $|S| = K$ in $\binom{n}{K}$ distinct ways. Now, for each such subset $S$, one can create a subset $T \subset [n]$ with $|T| = K$, $|S \cap T| = r$, by selecting $r$ elements from $S$ and $K - r$ elements from $[n] \setminus S$. Note that when $K - r \leq n - K$, for each $S$, this can be done in $\binom{r}{K-r}$ ways. However, if $K - r > n - K$, then one cannot find a set $T$ with $|T| = K$ and $|T \cap S| = r$. Therefore, the total number of ways of choosing such $S, T$, is found to be $N_r = \binom{n}{K-r} \binom{r}{K-r}$ when $r \geq r_{\min}$ where $r_{\min} = \max(0, 2K - n)$, and $N_r = 0$ if $r < 2K - n$. Consequently, using Eq. (23) along with the expression of $h_{ij}(r)$ from Eq. (6) as well as the expression of $N_r$, $0 \leq r \leq K$, one obtains,

$$
\mathbb{E} \left[ (\|X\|_2 \mid Y = y)^2 \right] = \frac{2}{\sigma^2} \left( \frac{n}{\sigma + 1} \right)^{J+1} \sum_{r_{\min} \leq r \leq K} \binom{K}{r} \binom{r}{K-r} a_i a_j h_{ij}(r).
$$

(28)

Using $\mathbb{E} \left[ \|X\|_2^2 \right] = n$ and plugging the above expression in Eq. (22) one obtains Eq. (7).

3.3. Proof of Theorem 2.3

3.3.1. Finding $\lim_{\sigma \rightarrow 0} \text{mmse}$ and $\lim_{\sigma \rightarrow \infty} \text{mmse}$

To evaluate $\lim_{\sigma \rightarrow 0} \text{mmse}$ as well as $\lim_{\sigma \rightarrow \infty} \text{mmse}$, we first observe from Eq. (12) that one can express the MMSE estimator, given $B = b$, as the expectation of the non-negative random variable $Z$, such that we can decompose $Z^2$ in the following way (see the description of $Z^2$ following Eq. (12)) $Z^2 = \|X_2\|_2^2 + \|X_3\|_2^2$, where $X_2 \sim N(0, \text{I}_{J, J})$ and $X_3, X_4 \sim N(0, \text{I}_{J, J})$, where $S = \text{supp}(b)$ and $X_2$ and $X_3$ are independent. This follows since $Y_S \sim N(0, \text{I}_{(\sigma^2 + 1)I_J})$. Therefore, the second term in the expression of the mmse in (22) can be expressed as $\mathbb{E}_S \left[ (\mathbb{E}_{X_2 - X_3} [Z])^2 \right]$. Now observe that $Z^2 \overset{a.s.}{\rightarrow} Z_1^2 = \|X_2\|_2^2 + \|X_4\|_2^2$ as $\sigma \rightarrow 0$, and $Z^2 \overset{a.s.}{\rightarrow} Z_2^2 = \|X_2\|_2^2 + \|X_3\|_2^2$, as $\sigma \rightarrow \infty$. It follows that $Z \overset{a.s.}{\rightarrow} |Z_1|$ as $\sigma \rightarrow 0$ and $Z \overset{a.s.}{\rightarrow} |Z_2|$ as $\sigma \rightarrow \infty$. Furthermore, $Z^2 \leq W^2 = \|X^2\|_2^2 + 2 \left( \|X_2\|_2^2 + \|X_3\|_2^2 \right)$, where $\mathbb{E} [W^2] = J + 4K < \infty$. Therefore, $Z \leq |W|$, with $\mathbb{E} [W] < \infty$. Therefore, by the dominated convergence theorem, $\mathbb{E} [Z] \rightarrow \mathbb{E} [Z_1]$ as $\sigma \rightarrow 0$ and $\mathbb{E} [Z] \rightarrow \mathbb{E} [Z_2]$ as $\sigma \rightarrow \infty$.

Now note that $Z_2^2$ is distributed as a $\chi^2_2$ random variable, so that $\mathbb{E}_{X_2, X_3} \left[ |Z_2|^2 \right] = \frac{\sqrt{2J(1/2)}}{\Gamma(1/2)}$, which implies from Eq. (22) that $\lim_{\sigma \rightarrow \infty} \text{mmse} = n - 2 \left( \frac{\Gamma(1/2)}{\Gamma(1/2)} \right)^2$. On the other hand,

$$
\lim_{\sigma \rightarrow 0} \text{mmse} = n - \mathbb{E}_S \left[ \mathbb{E}_{X_2, X_3} \left[ |Z_1| \mid X_4 \right] \right].
$$

(29)

In order to evaluate the RHS of Eq. (29), let us define the random variables $V_1 = \|X_2\|_2^2$, $V_2 = \|X_4\|_2^2$, so that $V_1, V_2$ are distributed as $\chi^2_2$ and $\chi^2_K$ random variables respectively, and $|Z_1| = \sqrt{V_1 + V_2}$. Then one obtains,

$$
\mathbb{E}_{X_2, X_3} \left[ |Z_1|^2 \right] = \mathbb{E}_{V_2} \left[ \left( \mathbb{E}_{V_1} \left[ \sqrt{V_1 + V_2} \right] \right)^2 \right] = \int_0^\infty \left( \int_0^\infty \frac{e^{-v_1/2} J^{1/2-1}}{v_1^{1/2} \Gamma(1/2)} dv_1 \right) \frac{e^{-v_2/2} J^{1/2-1}}{v_2^{1/2} \Gamma(1/2)} dv_2 = \frac{16I(J, K)}{\Gamma(1/2)^4 \Gamma(1/2)^2},
$$

(30)

where

$$
I(J, K) = \int_0^\infty \int_0^\infty \int_0^\infty \sqrt{(x^2 + y^2)(z^2 + y^2)} e^{-(x^2+y^2)(x^2+y^2)(x^2+y^2)} dx dy dz,
$$

(31)

which is obtained by using the transformation of variables $v_1 = x^2$, $v_2 = y^2$, $v_3 = z^2$. Now, using the transformation of variables $x = r \cos a$, $y = r \sin a \cos b$, $z = r \sin a \sin b$ one can obtain,

$$
I(J, K) = \left( \int_0^\infty e^{-r^2} r^{n+J+1} dr \right) \int_0^{\pi/2} \int_0^{\pi/2} \sqrt{1 - \sin^2 a \sin^2 b \sin^2 a \cos a} J^{-1}(\cos b) K^{-1}(\sin b) J^{-1} K^{-1} dadb.
$$

(32)
Now using \( \sqrt{1-x} = \sum_{l \geq 0} \frac{(-\frac{1}{2})^l}{l!} x^l \) for \( |x| \leq 1 \),
\[
\int_0^{\pi/2} \int_0^{\pi/2} \sqrt{1 - \sin^2 a \sin^2 b \sin(a)^n (\cos a)^{J-1} (\cos b)^{K-1} (\sin b)^{J-1} da db}
= \sum_{l=0}^{\infty} \left( \frac{-1}{2} \right)^l \frac{\Gamma \left( \frac{n+2l+1}{2} \right) \Gamma \left( \frac{J+2l}{2} \right) \Gamma \left( \frac{K+2l}{2} \right)}{2 \Gamma \left( \frac{n+2l+J+K}{2} \right)} \frac{\Gamma \left( \frac{n+1}{2} \right) \Gamma \left( \frac{J+1}{2} \right) \Gamma \left( \frac{K+1}{2} \right)}{4 \Gamma \left( \frac{n+J+K+1}{2} \right) \Gamma \left( \frac{J+K+1}{2} \right)} 
\sum_{l=0}^{\infty} \frac{(-1)^l}{l!} \left( \frac{\Gamma \left( \frac{J+1}{2} \right) \Gamma \left( \frac{K+1}{2} \right)}{4 \Gamma \left( \frac{n+J+K+1}{2} \right) \Gamma \left( \frac{J+K+1}{2} \right)} \right)^l.
\]

Therefore, from Eqs. 30, 32, 33 and using the fact that \( \int_0^\infty e^{-r^2} r^{n+J+1} dr = \frac{1}{2^n} \Gamma \left( \frac{n+J+1}{2} \right) \) along with some further easy simplifications, one obtains,
\[
\lim_{\sigma \to 0} \mathbb{m} \mathbb{m} \mathbb{m} = n - (n + J) \frac{\Gamma \left( \frac{n+J+1}{2} \right)}{B \left( \frac{n+J+1}{2} \right)} \left( \frac{1}{2} \right)^{J+1} \sum_{l=0}^{\infty} \frac{(-1)^l}{l!} \left( \frac{\Gamma \left( \frac{J+1}{2} \right) \Gamma \left( \frac{K+1}{2} \right)}{4 \Gamma \left( \frac{n+J+K+1}{2} \right) \Gamma \left( \frac{J+K+1}{2} \right)} \right)^l.
\]

3.3.2. Finding \( \lim_{n \to \infty} \mathbb{m} \mathbb{m} \mathbb{m} \mathbb{m} / n \)

We now proceed to find an upper bound of \( \mathbb{m} \mathbb{m} \mathbb{m} \mathbb{m} \mathbb{m} \mathbb{m} \) by keeping the \( K/n \) fixed to \( p (1/n \leq p \leq 1) \). First observe that one can rewrite Eq. 34 in the following form:
\[
\mathbb{E} \left[ (\mathbb{E} \| X \|_2 \mid Y = y) \right] = 2 \left( \frac{\sigma^2}{\sigma^2 + 1} \right)^{J+1} \mathbb{E} \left[ (\mathbb{E} \| L_S \|_2 \mid Y = y) \right]^{J+1},
\]
where \( L_S \) is Poisson distributed with parameter \( Z_S \), conditioned on the set \( S \). Since \( \mathbb{m} \mathbb{m} \mathbb{m} = n - \mathbb{E} \left[ (\mathbb{E} \| X \|_2 \mid Y = y) \right] \),
to find an upper bound on \( \mathbb{m} \mathbb{m} \mathbb{m} \mathbb{m} \mathbb{m} \mathbb{m} \mathbb{m} \), it is enough to find a lower bound of \( 2 \left( \frac{\sigma^2}{\sigma^2 + 1} \right)^{J+1} \mathbb{E} \left[ (\mathbb{E} \| L_S \|_2 \mid Y = y) \right]^{J+1} \), such that \( K/n = p \).

Now, from Eq. 35, it can be easily observed that \( L_S \) is Poisson distributed with parameter \( \frac{\Gamma \left( \frac{n+J+1}{2} \right)}{B \left( \frac{n+J+1}{2} \right) \Gamma \left( \frac{n+J+K+1}{2} \right) \Gamma \left( \frac{J+K+1}{2} \right)} \), where \( W \) is a Poisson random variable (independent of \( L_S \)) with parameter \( 1/(\sigma^2 + 1) \). For large \( n \), applying the Stirling’s approximation \( \Gamma(n) \approx n! e^{-n} \sqrt{2\pi n} \), one obtains, for arbitrary non-negative integers \( L, l \),
\[
\sqrt{\frac{\frac{n+1}{2} + L + l - 1}{\left( \frac{n+1}{2} + L + l - 1 \right)^{\frac{1}{2}}} \left( 1 + \frac{1}{\frac{n+1}{2} + L + l - 1} \right)^{\frac{1}{2}}} \approx \left( \frac{n}{2} + L + l \right)^{1/2}.
\]
Therefore, for large \( n \),
\[
L_S \approx e^{1/(\sigma^2 + 1)} \mathbb{E}_W \left[ \frac{\left( \frac{n+1}{2} + L + l \right)^{1/2}}{\left( \frac{n+1}{2} + L + l - 1 \right)^{1/2}} \right] W \left[ \frac{\left( \frac{n}{2} + L + W \right)^{1/2}}{\frac{n}{2} + L + W} \right].
\]

Therefore, for large \( n \),
\[
\mathbb{m} \mathbb{m} \mathbb{m} \mathbb{m} \mathbb{m} \approx \mathbb{m} \mathbb{m} - \mathbb{E} \left[ (\mathbb{E} \| L_S \|_2 \mid Y = y) \right]^{J+1} H_n, \text{ where } H_n = \mathbb{E} \left[ (\mathbb{E} \| X \|_2 \mid Y = y) \right]^{J+1}.
\]
It is now enough to find a lower bound of \( H_n \) with \( K/n = p \). Before proceeding further, we note that one can apply the Taylor’s expansion to get a quadratic expansion lower bound of the function \( \sqrt{1 + x} \) around \( a \) (for arbitrary \( a > 0 \))
as \( \sqrt{1 + x} \geq \sqrt{1 + a} + \frac{x-a}{2\sqrt{1+a}} - \frac{(x-a)^2}{8} \). Therefore, writing \( Q = L_S + W \),
\[
\mathbb{E}_L \left[ \frac{\left( \frac{n}{2} + L_S + W \right) \mid Z_S} \right] \geq \mathbb{E}_L \left[ \frac{\left( \frac{n}{2} + L_S \right) \mid Z_S} \right] + \frac{W}{2} - a - \frac{E_L \left[ \left( \frac{2L_S + 2W}{n} \mid Z_S \right)^2 \right]}{8}.
\]
after some calculations, therefore, using Jensen’s inequality, \( \frac{K}{2n\sigma^2(\sigma^2 + 1)} \parallel Y \parallel^2 \),

\[
E_S \left[ Z_S^2 \mid Y \right] = \frac{1}{K} \sum_{S \subseteq [n]; |S| = K} \frac{\|Y_S\|^2}{4\sigma^4(\sigma^2 + 1)} = \frac{1}{K} \sum_{S \subseteq [n]; |S| = K} \frac{\sum_{i \in S} Y_i^4 + 2 \sum_{i \neq j} Y_i^2 Y_j}{\left( \frac{n-1}{K} \right) \left( \frac{n-2}{K} \right) \sum_{i < j} Y_i^2 Y_j}.
\]

Furthermore, \( E \left[ Y_i^2 \right] = \sigma^2 + 1 \), \( E \left[ Y_i^4 \right] = 3(\sigma^2 + 1)^2 \), so that

\[
E_S \left[ Z_S^2 \right] = \frac{K}{2\sigma^2}, \quad E \left[ Z_S^2 \right] = \frac{3K + K(K-1)}{4\sigma^4} = \frac{K^2 + 2K}{4\sigma^4}.
\]

Therefore, determining the number of units \( H_n \geq \left( \frac{n\sigma^2}{\sigma^2 + 1} \right) (g_n(a))^2 \), where

\[
g_n(a) = \sqrt{1 + a} + \frac{2W}{n} - a - \sqrt{\frac{2W}{n}} - \frac{1-p}{\sigma^2} + \frac{2(1-p)}{\sigma^2} (\frac{1-p}{\sigma^2} - a) + \frac{2(1-p)(1+\sigma^2)}{\sigma^2 n} \]}

Since \( a \) was chosen to be an arbitrary positive real number, we can choose \( a = \frac{1}{\sigma^2} \), obtaining

\[
H_n \geq \left( \frac{n\sigma^2}{\sigma^2 + 1} \right) (g(1/\sigma^2))^2 = n \left( 1 - \frac{\sqrt{n\pi}}{4n\sigma^2} \right)^2.
\]

Therefore, for large \( n \) and large \( K \) such that \( K/n = p (1/n \leq p \leq 1) \),

\[
\frac{n\sigma^2}{\sigma^2 + 1} \geq 1 - \left( \frac{\sqrt{n\pi}}{4n\sigma^2} \right)^2,
\]

from which, taking \( n \to \infty \), we obtain \( \lim_{n \to \infty} \frac{n\sigma^2}{\sigma^2 + 1} = 0. \)
4. Conclusion

In this paper MMSE estimation of $l_2$ norm of a random $n$ dimensional Gaussian vector is considered from AWGN perturbed measurements taken after the input vector undergoes random data missing retaining only $K$ ($0 \leq K \leq n$) of its entries while the other entries are set to 0. The expressions of the MMSE estimator and its MSE are derived and several asymptotic results are derived which generally simplify the results in [1] where $K = n$ was considered. The results show that even with large missing entries and consequently small $K$, the MMSE estimator can still be pretty close to the one obtained without the missing data phenomenon.

Appendix A. Derivation of Eq. (18)

We use the series expansion of the confluent hypergeometric function $\alpha F_1$ to obtain,

$$ R(\alpha, \beta, \gamma, \nu, \epsilon) = \int_0^1 e^{-\alpha x} x^{\beta-1} (1-x)^{\gamma-1} \sum_{l \geq 0} \frac{(\epsilon x)^l}{(\nu)_l l!} dx \leq \int_0^1 e^{-\alpha x} x^{\gamma-1} (1-x)^{\beta-1} \sum_{l \geq 0} \frac{(\epsilon (1-x))^l}{(\nu)_l l!} dx $$

$$ = e^{-\alpha} \int_0^1 \sum_{k,l \geq 0} \frac{\alpha^k \epsilon^l}{k! l! (\nu)_l} x^{k+\gamma-1} (1-x)^{\beta+l-1} dx \leq e^{-\alpha} \sum_{k,l \geq 0} \frac{\alpha^k \epsilon^l}{k! l! (\nu)_l} \int_0^1 x^{k+\gamma-1} (1-x)^{\beta+l-1} dx $$

$$ = e^{-\alpha} \sum_{k,l \geq 0} \frac{\alpha^k \epsilon^l}{k! l! (\nu)_l} \frac{\Gamma(k+\gamma)\Gamma(\beta+l)}{\Gamma(\beta+\gamma+k+l)} \leq e^{-\alpha} \frac{\Gamma(\gamma)}{\Gamma(\beta+\gamma)} \sum_{k,l \geq 0} \frac{\alpha^k \epsilon^l}{k! l! (\nu)_l (\beta+\gamma)_{k+l}}. \quad (A.1) $$

where step $(\zeta_1)$ used the variable transformation $x \rightarrow 1-x$; step $(\zeta_2)$ used the series expansion of $e^{-\alpha x}$; step $(\zeta_3)$ used Fubini’s theorem to interchange the order of the integral and the summation; steps $(\zeta_4)$ and $(\zeta_5)$ used the properties of beta and gamma functions respectively. Putting $\beta = \nu$ in Eq. (A.1) results in the expression of Eq. (18)

References

[1] A. Dyso, M. Cardone, and H. V. Poor, “On estimating the norm of a gaussian vector under additive white gaussian noise,” IEEE Signal Processing Letters, vol. 26, no. 9, pp. 1325–1329, 2019.

[2] C. Dwork, A. Roth et al., “The algorithmic foundations of differential privacy,” Foundations and Trends® in Theoretical Computer Science, vol. 9, no. 3–4, pp. 211–407, 2014.

[3] X. Zhang, E. A. Jorswieck, B. Ottersten, and A. Paulraj, “User selection schemes in multiple antenna broadcast channels with guaranteed performance,” in 2007 IEEE 8th Workshop on Signal Processing Advances in Wireless Communications. IEEE, 2007, pp. 1–5.

[4] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, “Randomized gossip algorithms,” IEEE transactions on information theory, vol. 52, no. 6, pp. 2508–2530, 2006.

[5] J.-M. Chen and B.-S. Chen, “System parameter estimation with input/output noisy data and missing measurements,” IEEE Trans. Signal Process., vol. 48, no. 6, pp. 1548–1558, 2000.

[6] H. Fang, Y. Shi, and J. Wu, “Parameter estimation with missing input/output data,” in 2009 American Control Conference. IEEE, 2009, pp. 5061–5066.

[7] P.-L. Loh and M. J. Wainwright, “High-dimensional regression with noisy and missing data: Provably guarantees with non-convexity,” in Advances in Neural Information Processing Systems, 2011, pp. 2726–2734.

[8] Y. Chen, C. Caramanis, and S. Mannor, “Robust sparse regression under adversarial corruption,” in International Conference on Machine Learning, 2013, pp. 774–782.

[9] S. Mukhopadhyay, “Stochastic gradient descent for linear systems with sequential matrix entry accumulation,” Signal Processing, vol. 171, p. 107494, 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168420300372

[10] S. Mukhopadhyay and A. Mukherjee, “Imdms: An imputation based lms algorithm for linear system identification with missing input data,” IEEE Transactions on Signal Processing, pp. 1–1, 2020.

[11] M. Mitzenmacher et al., “A survey of results for deletion channels and related synchronization channels,” Probability Surveys, vol. 6, pp. 1–33, 2009.

[12] F. W. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, NIST handbook of mathematical functions hardback and CD-ROM. Cambridge university press, 2010.