On Estimating the Norm of a Gaussian Vector under Additive White Gaussian Noise

Alex Dytso†, Martina Cardone*, H. Vincent Poor†
† Princeton University, Princeton, NJ 08544, USA, Email: {adytso, poor}@princeton.edu
* University of Minnesota, Minneapolis, MN 55404, USA, Email: cardo089@umn.edu

***Abstract*—This letter considers the task of estimating the norm of an n-dimensional Gaussian random vector given a noisy/perturbed observation of it. In particular, the focus is on the case of additive Gaussian noise perturbation, which is assumed to be independent of the original vector. First, an expression for the optimal estimator is derived, and then the corresponding minimum mean square error (MMSE) is computed. The regime of large vector size is also analyzed, and it is shown that the MMSE normalized by n equals zero when $n \to \infty$.**

I. Introduction

In this letter, we consider an estimation framework that seeks to estimate the ℓ^2-norm of a random vector based on a noisy observation of the vector itself. More specifically, the observed n-dimensional random vector $\mathbf{Y}$ is obtained by passing an input $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$ through an additive white Gaussian noise channel. In other words, $\mathbf{Y} = \mathbf{X} + \mathbf{N}$, where $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n\sigma^2)$, and independent of $\mathbf{X}$. The goal is to estimate $\|\mathbf{X}\|$ given the observation of $\mathbf{Y}$.

The problem of estimating the norm of a random vector given a noisy observation of it, can be applied in several scenarios. For instance, in *distributed computing*, a master node might want to distribute the effort of computing the magnitude/norm of a massive vector across some worker machines which operate in parallel. However, this vector might contain sensitive/confidential data (e.g., clinical/genomic health), and hence it has to be perturbed [1], [2], [3] before being distributed to the worker machines. In this case, the level of noise perturbation (i.e., captured by σ^2) can be chosen so that the corresponding MMSE is below a certain threshold. Another scenario where our estimation framework can find applicability is in the context of *wireless systems*. The strength/norm estimation of the noisy channel can indeed be leveraged to improve the performance of the system, for instance by selecting the "best" users for transmission [4], and efficiently adapting the transmission data rate [5]. Moreover, in the context of *wireless sensor networks*, different sensors can independently measure different noisy vectors $\mathbf{Y}$ (e.g., capturing a noisy signal) and the final estimate of the norm of $\mathbf{X}$ (e.g., signal strength) can be obtained in a centralized way (using a common fusion center) or in a fully decentralized fashion [6], [7].

Depending on the particular optimality criterion used, there are different approaches to optimally estimating the norm of $\mathbf{X}$. Specifically, consider the following estimators:

The work of A. Dytso and H. V. Poor was supported in part by the U.S. National Science Foundation under Grant CCF-0939370. The work of M. Cardone was supported in part by the U.S. National Science Foundation under Grant CCF-1849757.

$$L_{\text{mle}}(\mathbf{y}) = \|\mathbf{y}\|, \tag{1}$$

$$L_{\text{plug-in-1}}(\mathbf{y}) = \|\mathbb{E}[\mathbf{X}|\mathbf{Y}=\mathbf{y}]\| = \frac{1}{1+\sigma^2}\|\mathbf{y}\|, \tag{2}$$

$$L_{\text{plug-in-2}}(\mathbf{y}) = \sqrt{\mathbb{E}[\|\mathbf{X}\|^2|\mathbf{Y}=\mathbf{y}]} = \sqrt{n\frac{\sigma^2}{1+\sigma^2} + \frac{1}{(1+\sigma^2)^2}\|\mathbf{y}\|^2}. \tag{3}$$

The estimator in (1) is the maximum likelihood estimator (MLE) of $\|\mathbf{X}\|$, which can be derived using standard techniques from the literature [8, Ch. 7]. The estimator in (2) is a plug-in estimator that first computes the Bayesian estimator for $\mathbf{X}$ and then applies the norm function. Finally, the estimator in (3) is a plug-in estimator that first estimates (in an MMSE fashion) the norm squared $\|\mathbf{X}\|^2$ and then applies the square-root function. Due to Jensen's inequality we have that $L_{\text{plug-in-1}}(\mathbf{y}) \leq \mathbb{E}[\|\mathbf{X}\||\mathbf{Y}=\mathbf{y}] \leq L_{\text{plug-in-2}}(\mathbf{y})$. In other words, $L_{\text{plug-in-1}}(\mathbf{y})$ underestimates the true value and $L_{\text{plug-in-2}}(\mathbf{y})$ overestimates the true value.

To asses the performance of the aforementioned estimators, we first derive an expression for the optimal estimator, i.e., the conditional expectation of $\|\mathbf{X}\|$ given the observation of $\mathbf{Y}$, and then compute the corresponding MMSE. Towards this end, we use properties of Gaussian and Poisson random variables, such as the Poisson representation of the probability density function (PDF) of the non-central chi-squared random variable. Moreover, we consider the large n regime, and show that the MMSE normalized by n vanishes as $n \to \infty$.

The derived optimal quantities are then used to argue that the plug-in estimator $L_{\text{plug-in-2}}$ in (3) is a reasonable choice for the estimation of $\|\mathbf{X}\|$. First, while $L_{\text{plug-in-2}}$ does overestimate the true value, the overestimation is relatively small and in fact it decreases with the dimension n. Second, $L_{\text{plug-in-2}}$ is relatively simple to implement compared to the optimal estimator. Third, by using the derived optimal MMSE, we numerically show that the MSE of the plug-in estimator $L_{\text{plug-in-2}}$ is close to the fundamental lower bound. Finally, we numerically demonstrate that the MLE in (1) and $L_{\text{plug-in-1}}$ in (2) should only be used in the small noise regime.

II. Optimal Estimator and its MMSE

In this section, we present our results on estimating the ℓ^2-norm of $\mathbf{X}$ given the observation of $\mathbf{Y}$. The next theorem presents an expression for the optimal MMSE estimator, i.e., the conditional expectation of $\|\mathbf{X}\|$ given $\mathbf{Y}$.

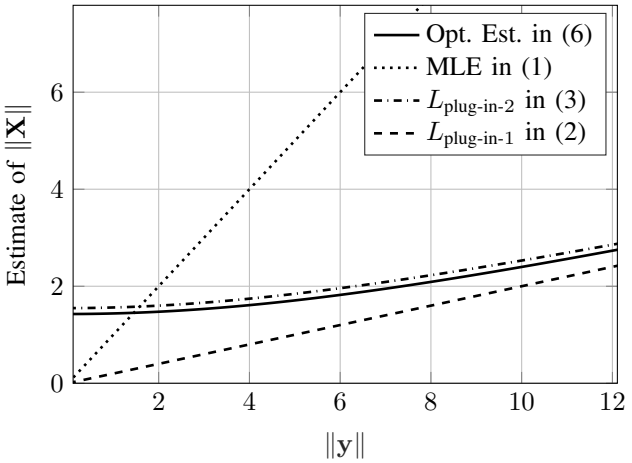

Fig. 1: Comparison of estimators of $\|\mathbf{X}\|$ in (1), (2) and (3) with the optimal estimator in (6). Here $n = 3$ and $\sigma = 2$.

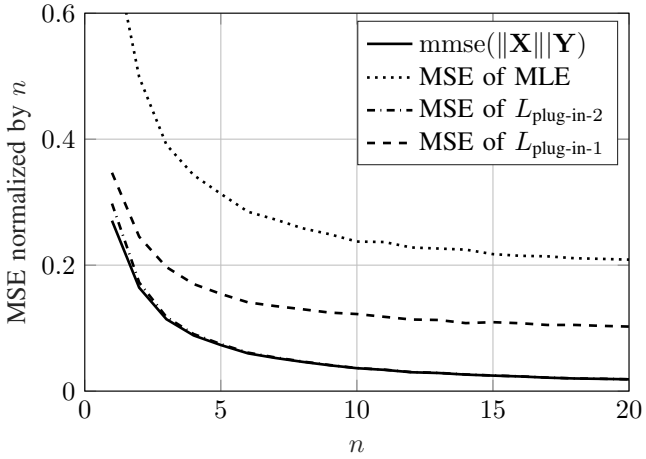

Fig. 2: Comparison of MSEs of estimating $\|\mathbf{X}\|$ with the estimators in (1), (2) and (3) and the optimal estimator in (6). The noise parameter is $\sigma = 1$.

Theorem 1. *Let* $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$*, and* $\mathbf{Y} = \mathbf{X} + \mathbf{N}$ *with* $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n\sigma^2)$*. Then,*

$$\begin{aligned}
&\mathbb{E}\left[\|\mathbf{X}\| \mid \mathbf{Y} = \mathbf{y}\right] \\
=&\frac{\sqrt{\sigma^2}\frac{1}{(2\pi)^{\frac{n}{2}}}}{f_{\mathbf{Y}}(\mathbf{y})}\mathbb{E}\left[\sqrt{V_{\frac{\|\mathbf{y}\|^2}{\sigma^2}}}\exp\left(-\frac{\sigma^2}{2}V_{\frac{\|\mathbf{y}\|^2}{\sigma^2}}\right)\right] &(4)\\
=&\sqrt{\frac{2\sigma^2}{1+\sigma^2}}\mathrm{e}^{-\frac{\|\mathbf{y}\|^2}{2\sigma^2(1+\sigma^2)}}\sum_{k=0}^{\infty}a_k\left(\frac{\|\mathbf{y}\|^2}{2\sigma^2(1+\sigma^2)}\right)^k &(5)\\
=&\frac{\Gamma(\frac{n+1}{2})\sqrt{2\sigma^2}\mathrm{e}^{-\frac{\|\mathbf{y}\|^2}{2\sigma^2(1+\sigma^2)}}}{\Gamma(\frac{n}{2})\sqrt{1+\sigma^2}}F_{1,1}\left(\frac{n+1}{2},\frac{n}{2};\frac{\|\mathbf{y}\|^2}{2\sigma^2(1+\sigma^2)}\right), &(6)
\end{aligned}$$

where V_λ *is a random variable distributed according to a chi-squared distribution with parameter* λ*, and*

$$a_k = \frac{\Gamma\left(k+\frac{n+1}{2}\right)}{k!\Gamma\left(k+\frac{n}{2}\right)}, \tag{7}$$

with $\Gamma(\cdot)$ *being the gamma function, and where* $F_{1,1}(\cdot,\cdot;\cdot)$ *is the confluent hypergeometric function [9, Ch. 13].*

Remark 1. *Note that Theorem 1 provides three different forms of the optimal estimator. The one in* (4) *is particularly suitable for running Monte Carlo simulations. The one in* (5) *can be easily approximated by truncating the sum to a sufficiently large number of terms. Finally, the one in* (6) *can be computed by standard off-the-shelf packages for hypergeometric functions.*

In Fig. 1 we compare the optimal estimator in (6) with the MLE and plug-in estimators in (1)-(3), respectively. Using the explicit representations of the optimal estimator (see Theorem 1), we can obtain an explicit expression for the MMSE of estimating the ℓ^2-norm of $\mathbf{X}$ from $\mathbf{Y}$.

Theorem 2. *Let* $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$*, and* $\mathbf{Y} = \mathbf{X} + \mathbf{N}$ *with* $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n\sigma^2)$*. Then, for* a_k *defined in* (7)

$$\begin{aligned}
\text{mmse}\left(\|\mathbf{X}\| \mid \mathbf{Y}\right) &= \mathbb{E}\left[\|\mathbf{X}\|^2\right] - \mathbb{E}\left[\left(\mathbb{E}[\|\mathbf{X}\| \mid \mathbf{Y}]\right)^2\right] \\
&= n - \frac{2\sigma^{2+n}}{1+\sigma^2}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\frac{a_k a_m}{(\sigma^2+2)^{k+m+\frac{n}{2}}}\frac{\Gamma(\frac{n}{2}+k+m)}{\Gamma(\frac{n}{2})}.
\end{aligned}$$

The MMSE in Theorem 2 provides the fundamental limit on the recover of $\|\mathbf{X}\|$ under the square error. In Fig. 2, we use the characterization of the MMSE in Theorem 2, to show that the MSE obtained by the suboptimal estimator $L_{\text{plug-in-2}}(\mathbf{y})$ in (3) has comparable performance. Moreover, in Fig. 2 we also compare the MSEs of the MLE in (1) and $L_{\text{plug-in-1}}$ in (2).

We conclude this section by providing some asymptotic analysis of the MMSE in Theorem 2.

Theorem 3. *Let* $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$*, and* $\mathbf{Y} = \mathbf{X} + \mathbf{N}$ *with* $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n\sigma^2)$*. Then, we have the following asymptotic results:*

- $\sigma \to 0$*, then* $\text{mmse}\left(\|\mathbf{X}\| \mid \mathbf{Y}\right) = 0$*;*
- $\sigma \to \infty$*, then* $\text{mmse}\left(\|\mathbf{X}\| \mid \mathbf{Y}\right) = \mathbb{V}[\|\mathbf{X}\|]$ *where* $\mathbb{V}[\|\mathbf{X}\|] = n - \left(\sqrt{2}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\right)^2$*; and*
- $n \to \infty$*, then* $\lim_{n\to\infty}\frac{\text{mmse}(\|\mathbf{X}\|\|\mathbf{Y})}{n} = 0$*, for all* $\sigma > 0$*.*

III. Proof of Main Results

A. Proof of Theorem 1

We here derive the optimal estimator in Theorem 1. We have

$$\begin{aligned}
&\mathbb{E}\left[\|\mathbf{X}\| \mid \mathbf{Y} = \mathbf{y}\right] \\
\overset{(a)}{=}&\int_0^\infty t\frac{f_{\mathbf{Y},\|\mathbf{X}\|}(\mathbf{y},t)}{f_{\mathbf{Y}}(\mathbf{y})}\,\mathrm{d}t \\
\overset{(b)}{=}&\int_0^\infty t\frac{\frac{1}{(2\pi)^{\frac{n}{2}}}\mathrm{e}^{-\frac{t^2}{2}}\frac{1}{\sqrt{\sigma^2}}f_{\text{Chi}}\left(\frac{t}{\sqrt{\sigma^2}};n,\frac{\|\mathbf{y}\|}{\sqrt{\sigma^2}}\right)}{f_{\mathbf{Y}}(\mathbf{y})}\,\mathrm{d}t \\
\overset{(c)}{=}&\frac{\sqrt{\sigma^2}\frac{1}{(2\pi)^{\frac{n}{2}}}\mathbb{E}\left[U_{\frac{\|\mathbf{y}\|}{\sqrt{\sigma^2}}}\exp\left(-\frac{\sigma^2}{2}U^2_{\frac{\|\mathbf{y}\|}{\sqrt{\sigma^2}}}\right)\right]}{f_{\mathbf{Y}}(\mathbf{y})} \\
\overset{(d)}{=}&\frac{\sqrt{\sigma^2}\frac{1}{(2\pi)^{\frac{n}{2}}}\mathbb{E}\left[\sqrt{V_{\frac{\|\mathbf{y}\|^2}{\sigma^2}}}\exp\left(-\frac{\sigma^2}{2}V_{\frac{\|\mathbf{y}\|^2}{\sigma^2}}\right)\right]}{f_{\mathbf{Y}}(\mathbf{y})},
\end{aligned} \tag{8}$$

where the labeled equalities follow from: (a) Bayes' rule; (b) is derived below in (10) with $f_{\text{Chi}}(x;n,\lambda)$ being the PDF of the non-central chi random variable, i.e.,

$$f_{\text{Chi}}(x;n,\lambda)=\mathrm{e}^{-\frac{x^2+\lambda^2}{2}}\left(\frac{x}{\lambda}\right)^{\frac{n}{2}}\lambda I_{\frac{n}{2}-1}(\lambda x),\ x>0, \quad (9)$$

where n specifies the degrees of freedom, $I_v(z)$ is the modified Bessel function of the first kind and $\lambda=\sqrt{\sum_{i=1}^{n}\left(\frac{\mu_i}{\sigma_i}\right)^2}$, such that X_i are n independent, normally distributed random variables with means μ_i and variances σ_i^2; (c) doing a change of variable $\frac{t}{\sqrt{\sigma^2}}=u$ and noting that $f_{\text{Chi}}\left(u;n,\frac{\|\mathbf{y}\|}{\sqrt{\sigma^2}}\right)$ is the PDF of the non-central chi random variable $U_{\frac{\|\mathbf{y}\|}{\sqrt{\sigma^2}}}$ with parameter $\frac{\|\mathbf{y}\|}{\sqrt{\sigma^2}}$; and (d) using the transformation between chi and chi-squared random variables. The equality above in (b) follows from the following derivation:

$$\begin{aligned}
f_{\mathbf{Y},\|\mathbf{X}\|}(\mathbf{y},t)&\overset{(b1)}{=}\int_{\mathbb{R}^n}f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})f_{\|\mathbf{X}\||\mathbf{X}}(t|\mathbf{x})f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}\mathbf{x}\\
&=\int_{\mathbb{R}^n}\frac{\mathrm{e}^{-\frac{\|\mathbf{y}-\mathbf{x}\|^2}{2\sigma^2}}}{(2\pi\sigma^2)^{\frac{n}{2}}}\delta\left(t-\|\mathbf{x}\|\right)\frac{1}{(2\pi)^{\frac{n}{2}}}\mathrm{e}^{-\frac{\|\mathbf{x}\|^2}{2}}\,\mathrm{d}\mathbf{x}\\
&\overset{(b2)}{=}\int_{S_{n-1}}\frac{\mathrm{e}^{-\frac{\|\mathbf{y}-t\boldsymbol{\Theta}\|^2}{2\sigma^2}}}{(2\pi\sigma^2)^{\frac{n}{2}}}\frac{1}{(2\pi)^{\frac{n}{2}}}\mathrm{e}^{-\frac{t^2}{2}}t^{n-1}\,\mathrm{d}\boldsymbol{\Theta}\\
&=\frac{\mathrm{e}^{-\frac{\|\mathbf{y}\|^2+t^2}{2\sigma^2}}}{(2\pi\sigma^2)^{\frac{n}{2}}}\frac{\mathrm{e}^{-\frac{t^2}{2}}}{(2\pi)^{\frac{n}{2}}}t^{n-1}\int_{S_{n-1}}\mathrm{e}^{\frac{t\boldsymbol{\Theta}^T\mathbf{y}}{\sigma^2}}\,\mathrm{d}\boldsymbol{\Theta}\\
&\overset{(b3)}{=}\frac{1}{(2\pi)^{\frac{n}{2}}}\mathrm{e}^{-\frac{t^2}{2}}\frac{1}{\sqrt{\sigma^2}}f_{\text{Chi}}\left(\frac{t}{\sqrt{\sigma^2}};n,\frac{\|\mathbf{y}\|}{\sqrt{\sigma^2}}\right), \quad (10)
\end{aligned}$$

where the labeled equalities follow from: (b1) Markov chain $\|\mathbf{X}\|\rightarrow\mathbf{X}\rightarrow\mathbf{Y}$ since $f_{\mathbf{Y},\|\mathbf{X}\||\mathbf{X}}(\mathbf{y},t|\mathbf{x})=f_{\|\mathbf{X}\||\mathbf{X}}(t|\mathbf{x})f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$, where we use the chain rule of the PDF, and the fact that, given $\mathbf{X}=\mathbf{x}$, then $\|\mathbf{X}\|=t$ is uniquely determined; (b2) changing the integration to spherical coordinates, where $S_{n-1}=\{\mathbf{x}:\|\mathbf{x}\|\leq 1\}$ is the n-dimensional unit-hypersphere, and using the sifting property of the delta function [10]; (b3) computing the integral as in (11) and using the PDF of the non-central chi random variable in (9). To compute the integral in (b3), we use the following steps:

$$\begin{aligned}
&\int_{S_{n-1}}\mathrm{e}^{\boldsymbol{\Theta}_n^T\mathbf{y}R}\mathrm{d}\boldsymbol{\Theta}_n\\
&\overset{(b3')}{=}\int_0^{2\pi}\int_0^{\pi}\mathrm{e}^{R\|\mathbf{y}\|\cos(\theta_1)}\\
&\quad\cdot\prod_{k=1}^{n-2}(\sin\theta_k)^{n-1-k}\,\mathrm{d}\theta_1\ldots\mathrm{d}\theta_{n-2}\mathrm{d}\theta_{n-1}\\
&\overset{(b3'')}{=}\int_0^{\pi}\mathrm{e}^{R\|\mathbf{y}\|\cos(\theta_1)}(\sin(\theta_1))^{n-2}\mathrm{d}\theta_1\int_{S_{n-2}}\mathrm{d}\boldsymbol{\Theta}_{n-1}\\
&\overset{(b3''')}{=}\frac{(2\pi)^{\frac{n}{2}}}{(R\|\mathbf{y}\|)^{\frac{n}{2}-1}}I_{\frac{n}{2}-1}(R\|\mathbf{y}\|), \quad (11)
\end{aligned}$$

where the labeled equalities follow from: (b3′) by noting that the Jacobian of the spherical transformation is given by

$$\mathrm{d}\boldsymbol{\Theta}_n=\prod_{k=1}^{n-1}(\sin\theta_k)^{n-1-k}\,\mathrm{d}\theta_1\ldots\mathrm{d}\theta_{n-2}\mathrm{d}\theta_{n-1}, \quad (12)$$

where $\theta_i\in[0,\pi), i\in[1:n-2]$, $\theta_{n-1}\in[0,2\pi)$, and by applying the definition of inner product; (b3″) observing that the integration over $\theta_2,...,\theta_{n-1}$ is an integral over the S_{n-2} sphere; and (b3‴) using the definition of modified Bessel function of the first kind, and the fact that

$$\int_{S_{n-2}}\mathrm{d}\boldsymbol{\Theta}_{n-1}=\frac{2\pi^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)}.$$

We now prove that the expression in (8) is equal to the conditional expected value provided in (5). Towards this end, we compute $\mathbb{E}\left[\sqrt{V_\lambda}\mathrm{e}^{-tV_\lambda}\right]$ as

$$\begin{aligned}
&\mathbb{E}\left[\sqrt{V_\lambda}\mathrm{e}^{-tV_\lambda}\right]\\
&\overset{(a)}{=}\int_0^{\infty}\sqrt{x}\mathrm{e}^{-tx}\frac{1}{2}\mathrm{e}^{-\frac{x+\lambda}{2}}\left(\frac{x}{\lambda}\right)^{\frac{n}{4}-\frac{1}{2}}I_{\frac{n}{2}-1}(\sqrt{\lambda x})\mathrm{d}x\\
&\overset{(b)}{=}\frac{\mathrm{e}^{-\frac{\lambda-\frac{\lambda}{1+2t}}{2}}}{(1+2t)^{\frac{n+1}{2}}}\int_0^{\infty}\sqrt{u}f_{\text{Chi-Sq}}\left(u;n,\frac{\lambda}{1+2t}\right)\mathrm{d}u\\
&=\frac{\mathrm{e}^{-\frac{\lambda-\frac{\lambda}{1+2t}}{2}}}{(1+2t)^{\frac{n+1}{2}}}\mathbb{E}\left[\sqrt{V_{\frac{\lambda}{1+2t}}}\right], \quad (13)
\end{aligned}$$

where the labeled equalities follow from: (a) using the PDF of the non-central chi-squared random variable, i.e., for $x>0$

$$f_{\text{Chi-Sq}}(x;n,\gamma)=\frac{1}{2}\mathrm{e}^{-\frac{x+\gamma}{2}}\left(\frac{x}{\gamma}\right)^{\frac{n}{4}-\frac{1}{2}}I_{\frac{n}{2}-1}(\sqrt{\gamma x}); \quad (14)$$

(b) change of variable $(1+2t)x=u$, and using the definition of $f_{\text{Chi-Sq}}\left(u;n,\frac{\lambda}{1+2t}\right)$ in (14). Next, we use the Poisson representation of the PDF of the non-central chi-squared random variable [11], i.e.,

$$f_{\text{Chi-Sq}}(u;n,\lambda)=\sum_{k=0}^{\infty}p\left(k;\frac{\lambda}{2}\right)f_{\text{Chi-Sq}}(u;n+2k,0),$$

where $p(k;x)=\frac{\mathrm{e}^{-x}x^k}{k!}$ $k=0,1,2,\ldots$ Let $V_{0,y}$ denote a centered chi-squared random variable of order y. Then,

$$\begin{aligned}
&\mathbb{E}\left[\sqrt{V_{\frac{\lambda}{1+2t}}}\right]=\sum_{k=0}^{\infty}p\left(k;\frac{\lambda}{2(1+2t)}\right)\mathbb{E}\left[\sqrt{V_{0,n+2k}}\right]\\
&=\sum_{k=0}^{\infty}\frac{\sqrt{2}\mathrm{e}^{-\frac{\lambda}{2(1+2t)}}\left(\frac{\lambda}{2(1+2t)}\right)^k}{k!}\frac{\Gamma\left(\frac{n+2k+1}{2}\right)}{\Gamma\left(\frac{n+2k}{2}\right)}, \quad (15)
\end{aligned}$$

where the last equation follows by using the expression for the moments of a centered chi-squared random variable, i.e.,

$$\mathbb{E}[X^m]=2^m\frac{\Gamma\left(m+\frac{k}{2}\right)}{\Gamma\left(\frac{k}{2}\right)},\ m>0.$$

Combining (13) and (15), we arrive at

$$\mathbb{E}\left[\sqrt{V_\lambda}\mathrm{e}^{-tV_\lambda}\right]=\mathrm{e}^{-\frac{\lambda}{2}}(1+2t)^{-\frac{n+1}{2}}\sqrt{2}\sum_{k=0}^{\infty}a_k\left(\frac{\lambda}{1+2t}\right)^k,$$

where $a_k=\frac{1}{2^k}\frac{\Gamma\left(k+\frac{n+1}{2}\right)}{k!\Gamma\left(k+\frac{n}{2}\right)}$. Finally, the proof of Theorem 1 is concluded by using the follows series:

$$F_{1,1}(a,b;x)=\sum_{k=0}^{\infty}\frac{\Gamma(a+k)\Gamma(b)}{\Gamma(a)\Gamma(b+k)}\frac{x^k}{k!},\ \min\{a,b,x\}>0.$$

B. Proof of Theorem 2

We here derive the MMSE in Theorem 2. We have

$$\begin{aligned}
&\mathbb{E}\left[(\mathbb{E}[\|\mathbf{X}\| \mid \mathbf{Y}])^2\right] \\
&\overset{(a)}{=} \frac{2\sigma^2}{1+\sigma^2}\mathbb{E}\left[\mathrm{e}^{-\frac{\|\mathbf{Z}\|^2}{\sigma^2}}\left(\sum_{k=0}^{\infty} a_k\left(\frac{\|\mathbf{Z}\|^2}{2\sigma^2}\right)^k\right)^2\right] \\
&= \frac{2\sigma^2}{1+\sigma^2}\mathbb{E}\left[\mathrm{e}^{-\frac{\|\mathbf{Z}\|^2}{\sigma^2}}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty} a_k a_m \frac{\|\mathbf{Z}\|^{2(k+m)}}{2^{k+m}\sigma^{2(k+m)}}\right] \\
&\overset{(b)}{=} \frac{2\sigma^{2+n}}{1+\sigma^2}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\frac{a_m a_k}{(\sigma^2+2)^{k+m+\frac{n}{2}}}\frac{\Gamma(\frac{n}{2}+k+m)}{\Gamma(\frac{n}{2})}, \qquad (16)
\end{aligned}$$

where the labeled equalities follow from: (a) using the optimal estimator derived in Theorem 1, and rescaling the Gaussian random vector $\mathbf{Y}$; and (b) changing the order of summation and integration via Tonelli's theorem, and using

$$\begin{aligned}
\mathbb{E}\left[\|\mathbf{Z}\|^k \mathrm{e}^{\frac{t\|\mathbf{Z}\|^2}{2}}\right] &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}_n}\|\mathbf{x}\|^k \mathrm{e}^{\frac{t\|\mathbf{x}\|^2}{2}}\mathrm{e}^{-\frac{\|\mathbf{x}\|^2}{2}}\mathrm{d}\mathbf{x} \\
&\overset{(b1)}{=} \frac{1}{(1-t)^{\frac{n}{2}}}\mathbb{E}\left[\|\mathbf{X}_G\|^k\right] \\
&= \frac{1}{(1-t)^{\frac{n}{2}}}\frac{1}{(1-t)^{\frac{k}{2}}}\mathbb{E}\left[\|\mathbf{Z}\|^{2\frac{k}{2}}\right] \\
&\overset{(b2)}{=} \frac{1}{(1-t)^{\frac{k+n}{2}}}2^{\frac{k}{2}}\frac{\Gamma\left(\frac{k+n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}, \qquad (17)
\end{aligned}$$

where the labeled equalities follow from: (b1) defining $\mathbf{X}_G \sim \mathcal{N}\left(\mathbf{0}, \frac{1}{1-t}\mathbf{I}_n\right)$; and (b2) using the $\frac{k}{2}$-th moment about zero of the chi-squared random variable with n degrees of freedom. Combining (16) with $\mathbb{E}\left[\|\mathbf{X}\|^2\right] = n$ concludes the proof of Theorem 2.

C. Proof of Theorem 3

Because of space limitations, we omit the detailed proof for $\sigma \to 0$ and $\sigma \to \infty$, which amounts to an application of the dominated convergence theorem. We focus on $n \to \infty$. We start by analyzing the following term, which is equivalent to $\frac{1+\sigma^2}{2\sigma^2}\frac{1}{n}\mathbb{E}\left[(\mathbb{E}[\|\mathbf{X}\| \mid \mathbf{Y}])^2\right]$ (because of step (a) in (16))

$$\begin{aligned}
&\frac{1}{n}\mathbb{E}\left[\mathrm{e}^{-\frac{\|\mathbf{Z}\|^2}{\sigma^2}}\left(\sum_{k=0}^{\infty} a_k\left(\frac{\|\mathbf{Z}\|^2}{2\sigma^2}\right)^k\right)^2\right] \\
&\overset{(a)}{=} \frac{1}{n}\mathbb{E}\left[\mathrm{e}^{-2U}\left(\sum_{k=0}^{\infty} a_k U^k\right)^2\right] \\
&\overset{(b)}{=} \frac{1}{n}\mathbb{E}\left[\left(\mathbb{E}[K! a_K|U]\right)^2\right] \\
&\overset{(c)}{\approx} \frac{1}{n}\mathbb{E}\left[\left(\mathbb{E}\left[\sqrt{K+\frac{n}{2}}|U\right]\right)^2\right], \qquad (18)
\end{aligned}$$

where the labeled equalities follow from: (a) letting $U = \frac{\|\mathbf{Z}\|^2}{2\sigma^2}$ with standard normal random variable $\mathbf{Z}$; (b) noting that $K|U$ is a Poisson random variable with parameter U; and (c) using the expression for a_K in (7) and observing that

$$\frac{\Gamma\left(K+\frac{n+1}{2}\right)}{\Gamma\left(K+\frac{n}{2}\right)} = C_{K,n}\sqrt{K+\frac{n}{2}}, \qquad (19)$$

where $C_{K,n} \approx 1$ for large enough n (independently of K) as a consequence of using Stirling's bounds [12] as follows:

$$\sqrt{\frac{\left(1+\frac{1}{n}\right)^n}{\mathrm{e}}}\mathrm{e}^{-\frac{1}{(n+2)(n+1)}} \le C_{K,n} \le \left(1+\frac{1}{n}\right)\mathrm{e}^{\frac{1}{(n+2)(n+1)}}.$$

We now show that

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\left[\left(\mathbb{E}\left[\sqrt{K+\frac{n}{2}}|U\right]\right)^2\right] = \frac{1}{2\sigma^2}+\frac{1}{2}. \qquad (20)$$

Towards this end, we derive an upper bound and a lower bound on the limit in the left-hand side of (20) and show that these bounds converge to the right-hand side of (20).

Jensen's inequality provides the upper bound

$$\frac{1}{n}\mathbb{E}\left[\left(\mathbb{E}\left[\sqrt{K+\frac{n}{2}}|U\right]\right)^2\right] \le \frac{1}{n}\mathbb{E}\left[\mathbb{E}\left[K+\frac{n}{2}|U\right]\right] = \frac{1+\sigma^2}{2\sigma^2},$$

where we used $\mathbb{E}[K|U] = U$ since $K|U$ is a Poisson random variable with parameter U. Next, we focus on the lower bound. We use the following bound which follows by using the Taylor expansion of $\sqrt{1+x}$ about a variable a

$$\sqrt{x+1} \ge \sqrt{a+1} + \frac{x-a}{2\sqrt{a+1}} - \frac{(x-a)^2}{8(a+1)^{\frac{3}{2}}} \qquad (21)$$

for any $a > 0$. With this, we obtain

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\mathbb{E}\left[\sqrt{K+\frac{n}{2}}|U\right] \\
&\ge \frac{1}{\sqrt{2}}\left(\sqrt{a+1} + \frac{\frac{2U}{n}-a}{2\sqrt{a+1}} - \frac{\frac{4}{n^2}(U^2+U) - \frac{4}{n}Ua + a^2}{8(a+1)^{\frac{3}{2}}}\right),
\end{aligned}$$

where we used the fact that $\mathbb{E}[K|U] = U$ and $\mathbb{E}[K^2|U] = U + U^2$. This, together with Jensen's inequality, leads to

$$\frac{1}{n}\mathbb{E}\left[\left(\mathbb{E}\left[\sqrt{K+\frac{n}{2}}|U\right]\right)^2\right] \ge \left(\frac{1}{\sqrt{2}}f(a)\right)^2,$$

where by using the fact that $\mathbb{E}[\|\mathbf{Z}\|^4] = n(n+2)$ we obtain

$$f(a) = \sqrt{a+1} + \frac{\frac{1}{\sigma^2}-a}{2\sqrt{a+1}} - \frac{\frac{4}{n^2}\left(\frac{n(n+2)}{4\sigma^4}+\frac{n}{2\sigma^2}\right) - \frac{2}{\sigma^2}a + a^2}{8(a+1)^{\frac{3}{2}}}.$$

Next, by taking the limit as $n \to \infty$, we obtain

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\left[\left(\mathbb{E}\left[\sqrt{K+\frac{n}{2}}|U\right]\right)^2\right] \ge \frac{1}{2}\left(1+\frac{1}{\sigma^2}\right),$$

where since a is arbitrary, we have chosen it to be $a = \frac{1}{\sigma^2}$. Thus, (20) holds. Now, observe that the term that we computed in (18) is equivalent to $\frac{1+\sigma^2}{2\sigma^2}\frac{1}{n}\mathbb{E}\left[(\mathbb{E}[\|\mathbf{X}\| \mid \mathbf{Y}])^2\right]$. This follows from step (a) in (16). We therefore obtain

$$\lim_{n\to\infty}\frac{\mathrm{mmse}\left(\|\mathbf{X}\| \mid \mathbf{Y}\right)}{n} = 1 - \lim_{n\to\infty}\frac{\mathbb{E}\left[(\mathbb{E}[\|\mathbf{X}\| \mid \mathbf{Y}])^2\right]}{n} = 0.$$

This concludes the proof of Theorem 3.

REFERENCES

[1] C. C. Aggarwal and P. S. Yu, *A General Survey of Privacy-Preserving Data Mining Models and Algorithms*. Springer US, 2008, pp. 11–52.

[2] L. Sweeney, "k-anonymity: A model for protecting privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 557–570, 2002.

[3] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, no. 3-4, pp. 211–407, Aug. 2014. [Online]. Available: http://dx.doi.org/10.1561/0400000042

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

[5] E. Björnson and B. Ottersten, "Pilot-based Bayesian channel norm estimation in Rayleigh fading multi-antenna systems," in *Proceedings of the Twentieth Nordic Conference on Radio Science and Communications*, 2008.

[6] 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.

[7] ——, "Gossip algorithms: design, analysis and applications," in *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, vol. 3, 2005, pp. 1653–1664 vol. 3.

[8] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Prentice Hall PTR, 1993.

[9] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical functions: With Formulas, Graphs, and Mathematical Tables*. Courier Corporation, 1970, vol. 9.

[10] R. N. Bracewell and R. N. Bracewell, *The Fourier Transform and its Applications*. McGraw-Hill New York, 1986, vol. 31999.

[11] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. John Wiley & Sons, 2009, vol. 197.

[12] H. Robbins, "A remark on Stirling's formula," *The American Mathematical Monthly*, vol. 62, no. 1, pp. 26–29, 1955.