Misspecified Cramér-Rao Bounds for Anomalous Clock Data in Satellite Constellations

Hamish McPhee
Research Laboratory TéSA
University of Toulouse
Toulouse, France
hamish.mcphee@tesa.prd.fr

Jean-Yves Tourneret
Research Laboratory IRIT
University of Toulouse
Toulouse, France
jean-yves.tourneret@toulouse-inp.fr

David Valat
Timing and Frequency
French Space Agency (CNES)
Toulouse, France
david.valat@cnes.fr

Jérôme Delporte
Timing and Frequency
French Space Agency (CNES)
Toulouse, France
jerome.delporte@cnes.fr

Yoan Grégoire
Navigation and Signal Processing
French Space Agency (CNES)
Toulouse, France
yoan.gregoire@cnes.fr

Philippe Paimblanc
Research Laboratory TéSA
Toulouse, France
philippe.paimblanc@tesa.prd.fr

***Abstract*—Robust estimation methods are useful in mitigating the impact of anomalies in clock data. Such anomalous clock data is assumed to be well modeled by a Student's t-distribution. This paper derives a lower bound on the performance of the misspecified Gaussian model using the theory of the Misspecified Cramér-Rao bound (MCRB). The results of these derivations are verified by analyzing the Mean Square Error (MSE) of the misspecified Gaussian Maximum Likelihood Estimator (MLE) when using data generated by the Student's t-distribution. The derived MCRB indicates a constraint on the MSE when assuming a Gaussian distribution. The MLE for the mean of the Student's t-distribution is obtained with an Expectation maximization algorithm and is shown to obtain a lower MSE than the MCRB and hence, the misspecified estimator. This indicates an improvement in performance if anomalous clock data is appropriately accounted for in the statistical model.**

I. Introduction

Timing requires a stable reference frame called a time scale to avoid erroneous results in navigation and positioning applications. The members of a swarm of satellites each have an onboard clock that is biased compared to the others, like the satellites in Global Navigation Satellite Systems (GNSS). Nanosatellites would be expected to have even less stable clocks than those used in GNSS satellites, meaning a greater likelihood of timing errors. This can already mean a reduction in the quality of the position determination within a satellite swarm unless the clocks are synchronized to a common reference that is robust to clock anomalies. If the swarm is to operate autonomously, i.e., without the aid of synchronization via GNSS satellites, then a reference should be created using only the clocks in the swarm. An upcoming scientific mission related to a swarm of nanosatellites orbiting the moon is a prime example of an application that will need such autonomous time synchronization [1].

When clock bias measurements contain outliers, the usual assumptions on the statistical distribution will no longer be valid. Specifically, the Maximum Likelihood Estimator (MLE) for the typical Gaussian assumption is not the most efficient in the presence of anomalies. The MLE defined for estimating the parameters of a heavy-tailed distribution is expected to be more robust to the impact of anomalies because the model accounts for potential outliers. In this work, we assess performance losses due to assuming a Gaussian distribution when the data is contaminated by outliers.

The true distribution chosen to represent contaminated data is the heavy-tailed Student's t-distribution. This distribution is selected because of existing work on the Cramér-Rao Bound (CRB) [2] and its ability to model outliers without assuming a shift in the mean [3]. The performance loss due to assuming a Gaussian distribution can be analyzed using the Misspecified Cramér-Rao Bound (MCRB), which is well-defined for parametric models but does not explicitly address specific cases of anomalous data [4, 5, 6, 7]. The derivations for the misspecified model investigated should coincide with the recent generalized results from [8], i.e., the MCRB for the mean assuming a Gaussian distribution is identical to Gaussian CRB, given by the variance of the data.

The rest of this paper is split into several sections that introduce the context, the derivations, and the results. Section II presents the clock timing data and the proposed Student's t-distribution for building a time scale. Section III presents the method to obtain the MCRB and the derived closed-form expressions for the investigated combination of true and misspecified distributions. The results are analyzed in Section IV, where the Mean Square Error (MSE) of the appropriate estimators verifies the derived expression and the preferred estimator.

II. Anomalous clock data

Only the difference in time between two satellites can be observed through the pseudorange equations made for each inter-satellite link. Consider a pair of satellites labeled A and B in a constellation. A simplified representation of their pseudorange equation is:

$$\rho_{AB}(t) = R(t) + cx_{AB}(t) + n_{AB}(t), \tag{1}$$

where $R(t)$ is the actual geometric distance between the two satellites at time t, c is the speed of light, $x_{AB}(t)$ is the time difference between the two satellite clocks at time t, and $n_{AB}(t)$ is the noise related to the measurement process including instrument delays and possible interferences. Measurements of geometric range and clock bias are obtainable by using a two-way transfer method and the fact that $\rho_{AB} = \rho_{BA}$. The measurement process can be impacted by anomalous observations or increased measurement noise on certain links, causing outliers in the data.

The clock biases between pairs of satellites provide observations of the onboard clock behaviors, which are used to build a common reference time for every satellite, i.e., a time scale. The time scale should be more stable than any individual clock in the constellation. The stability of the clocks is observed by comparing clock bias measurements with the clock predictions. The predictions are made from time $t-\tau$ to time t (where τ is the interval since the previous measurements were available) denoted as $\hat{x}_j(t|t-\tau)$. Then, $z_{ji}(t)$ is the residual between the measurement associated with satellites j and i and the prediction

$$z_{ji}(t) = \hat{x}_j(t|t-\tau) - [x_{ji}(t) + n_{ji}(t)], \tag{2}$$

where $x_{ji}(t) = x_j(t) - x_i(t)$ is the actual clock bias and measurement noise $n_{ji}(t)$ is included due to the noise in the pseudorange equation. The predictions use estimates of the clock frequency and frequency drift to model the dynamics of the clock phase since the previous measurements. Methods of estimating the clock frequency and drift are outside the scope of this work but are well-defined for existing time scale solutions [9]. The predictions are made for each of the independent clocks so that the clock residuals in (2) can be written in the following form:

$$z_{ji}(t) = x_i(t) + [\hat{x}_j(t|t-\tau) - x_j(t)] - n_{ji}(t). \tag{3}$$

For a fixed clock i, the clock residuals $z_{ji}(t)$ allow N independent observations of the absolute time of that clock $x_i(t)$. The MLE of $x_i(t)$ is directly related to the generation of a time scale [10]. Each observation has some stochastic component related to the prediction errors of the other clocks $(\hat{x}_j(t|t-\tau)-x_j(t))$ and the noise on each of the inter-satellite links $n_{ji}(t)$. This definition of clock residuals is consistent with existing time scale algorithms [9], although this work includes a contribution of adding non-negligible measurement noise.

Interference in the inter-satellite links and malfunctioning clocks cause outliers in the measurements of clock biases or clock predictions. A heavy-tailed distribution is used to model such anomalous clock data. For example, the residuals can follow a Student's t-distribution with parameters $\boldsymbol{\eta}_T = [\mu_T(t), \sigma_T^2(t), \nu(t)]^T$ [11]:

$$z_j(t) \sim T(\mu_T(t), \sigma_T^2(t), \nu(t)), \tag{4}$$

where μ_T is the location parameter, σ_T^2 is the scale parameter, and ν is the number of degrees of freedom. The dependence on time and the central satellite i is omitted for the rest of the article as we focus on the measurements made by a single satellite at a single instant in time. The location parameter is linked to the absolute time of the satellite making measurements i, $\mu_T(t) = x_i(t)$, and the other parameters are linked to the dispersion of the residual data. The joint Probability Density Function (PDF) of the univariate Student's t-distribution for a vector of N i.i.d. random variables $\mathbf{z} = (z_1, \cdots, z_N)^T$ that are contaminated by outliers is:

$$p(\mathbf{z}; \boldsymbol{\eta}_T) = \prod_{j=1}^{N} \frac{1}{\sigma_T\sqrt{\pi\nu}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \left(1 + \frac{1}{\nu}\left(\frac{z_j - \mu_T}{\sigma_T}\right)^2\right)^{-\left(\frac{\nu+1}{2}\right)}. \tag{5}$$

If the presence of the anomalies is neglected and a Gaussian distribution is used for $\mathbf{z}$ then the estimate of the absolute time does not provide a stable reference time. That is, the clock residuals are misspecified with parameter vector $\boldsymbol{\theta} = [\mu_G, \sigma_G^2]^T$:

$$z_j(t) \sim \mathcal{N}(\mu_G, \sigma_G^2). \tag{6}$$

The Gaussian assumption would then be an incorrect model for the random variable, providing a misspecified PDF with different location and scale parameters

$$q(\mathbf{z}; \mu_G, \sigma_G^2) = \prod_{j=1}^{N} \frac{1}{\sqrt{2\pi\sigma_G^2}} \exp\left[-\frac{1}{2}\left(\frac{z_j - \mu_G}{\sigma_G}\right)^2\right]. \tag{7}$$

The reference time computed for the constellation is expected to lose stability in the presence of anomalies if the Gaussian assumption is made. By deriving the MCRB for this case, the benefits of correctly fitting a heavy-tailed distribution to the anomalous data are revealed.

III. Misspecified Cramér Rao Bounds

For precision in the notations, the MCRB will be presented with subscripts to refer to the estimated parameters. For example $\mathbf{MCRB}_{\boldsymbol{\theta}}$ refers to the matrix that contains the bounds (independent and joint) for each parameter in $\boldsymbol{\theta}$. The MCRB is defined using the Kullback Leibler Divergence (KLD), which is a statistical "distance" between the true and assumed models as follows:

$$D_{\text{KL}}(p(\mathbf{z}; \boldsymbol{\eta}) || q(\mathbf{z}; \boldsymbol{\theta})) = E_p\left[\log\left(\frac{p(\mathbf{z}; \boldsymbol{\eta})}{q(\mathbf{z}; \boldsymbol{\theta})}\right)\right], \tag{8}$$

where the subscript of E_p indicates the expectation with respect to the true PDF p, which remains the general notation for declaring an arbitrary true distribution. To derive the MCRB, we must first estimate the pseudo-true parameters, which are defined as the parameters that minimize the KLD between the true and assumed models

$$\tilde{\boldsymbol{\theta}}_p = \arg\min_{\boldsymbol{\theta}} \{D_{\text{KL}}\} = \arg\min_{\boldsymbol{\theta}} \left\{E_p\left[\log\left(\frac{p(\mathbf{z}; \boldsymbol{\eta})}{q(\mathbf{z}; \boldsymbol{\theta})}\right)\right]\right\}. \tag{9}$$

The subscript p indicates that the pseudo-true parameters depend on the true distribution. The following pseudo-true values are obtained for the true Student's t-distribution that models observations with outliers:

$$\tilde{\mu}_p = \mu_T, \qquad \tilde{\sigma}_p^2 = \sigma_T^2 \frac{\nu}{\nu - 2}, \tag{10}$$

where the full derivations are detailed in Appendix A. The pseudo-true parameters are denoted as $\tilde{\boldsymbol{\theta}}_p = [\tilde{\mu}_p, \tilde{\sigma}_p^2]^T$ and the pseudo-true scale $\tilde{\sigma}_p^2$ is only defined for $\nu > 2$, just as the variance of Student t-distributed data is undefined.

The next steps to derive the MCRB involve the computation of two matrices:

$$\mathbf{A}(\tilde{\boldsymbol{\theta}}_p) = E_p \left[\left(\frac{\partial^2 \log(q(z_j;\boldsymbol{\theta}))}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right] \right)_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}_p}, \tag{11}$$

$$\mathbf{B}(\tilde{\boldsymbol{\theta}}_p) = E_p \left[\left(\left(\frac{\partial \log(q(z_j;\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \right) \left(\frac{\partial \log(q(z_j;\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}^T} \right) \right] \right)_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}_p}. \tag{12}$$

Note that the above matrices only depend on the marginal PDF of a single sample z_j. Under the assumption that each sample is i.i.d., which is relevant for clock measurements made "simultaneously" at a single time instant, the MCRB is defined as:

$$\mathbf{MCRB}_{\boldsymbol{\theta}}(p||q) = \frac{1}{N} \mathbf{A}(\tilde{\boldsymbol{\theta}}_p)^{-1} \mathbf{B}(\tilde{\boldsymbol{\theta}}_p) \mathbf{A}(\tilde{\boldsymbol{\theta}}_p)^{-1}. \tag{13}$$

The terms in (11) and (12) are provided in Appendix B. The results are then used to compute the MCRB of interest:

$$\mathbf{MCRB}_{\boldsymbol{\theta}}(p_T||q) = \begin{bmatrix} \frac{\tilde{\sigma}_p^2}{N} & 0 \\ 0 & \left(\frac{\nu-1}{\nu-4}\right) \frac{2\tilde{\sigma}_p^4}{N} \end{bmatrix}, \tag{14}$$

where the bound for σ_G^2 is undefined for $\nu \leq 4$, whereas the bound for the mean is defined as long as the variance for the Student's t-distribution is defined $\nu > 2$. Hence, estimating the scale parameter has scenarios with undefined estimation performance and is more sensitive to the intensity of outliers than estimating the location parameter. As expected, the misspecified bound for σ_G^2 also simplifies to the Gaussian CRB for $\nu \to \infty$ because the Student's t-distribution converges to a Gaussian.

For conciseness of the results, only the bounds for estimation of the mean are assessed in this article with a value of the shape parameter corresponding to significant outliers $\nu = 3$ that cause the other bound to be undefined. This number of degrees of freedom is not uncommon for data following a heavy-tailed Student's t-distribution.

In this work, the MCRB derivation has been completed for a specific example of a real, elliptically symmetric distribution that is compatible with modeling anomalies. The derived MCRB is not necessarily of the same form for all other types of heavy-tailed noises, especially those that are skewed. The MCRB can be compared to the CRB to show the gain in MSE that is achievable when the presence of outliers is correctly specified. The CRB for the location parameter of the univariate Student's t-distribution is [2]

$$\mathrm{CRB}_{\mu_T} = \left(\frac{\nu+3}{\nu+1}\right) \frac{\sigma_T^2}{N}, \tag{15}$$

Figure 1 demonstrates the equivalence of the MCRB derived in this work and the Gaussian CRB. The bounds are displayed such that the variance of the data is normalized for each distribution. This ensures that the resulting bounds are appropriately compared for a common dataset with an unknown statistical model. The Gaussian CRB and the MCRB are higher than the CRB for the Student's t-distribution. This is because the estimator that correctly assumes the Student's t-distribution reduces the impact of the outliers that cause inflated variance in the data. Hence, correctly specifying the type of distribution is preferred over assuming a Gaussian model if the true model is a Student's t-distribution. This is verified in the next section by testing an appropriate estimator.

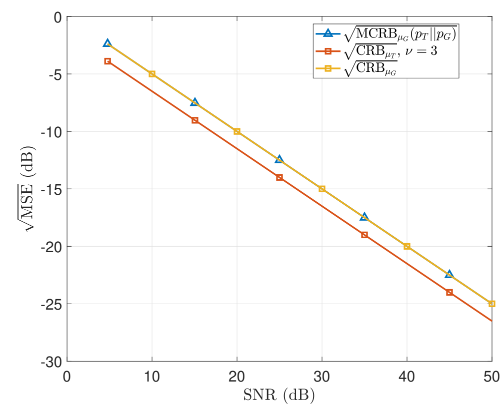

Fig. 1: Comparison of the derived MCRB and relevant CRBs, distributions are parameterized with $\boldsymbol{\theta} = [0, 1]$ and $\boldsymbol{\eta} = [0, 1/3, 3]$. The scale of the Student's t-distribution is reduced to result in a distribution with unit variance, equivalent to the Gaussian. The shape parameter $\nu = 3$ symbolizes a distribution with heavy tails but with defined variance.

IV. Convergence of Maximum Likelihood Estimator Mean Square Errors to Bounds

Instead of using real clock data, the analysis is conducted with random variables produced with the given distributions, and estimations are made under the corresponding correct and misspecified assumptions. To validate the derived MCRB, the performance of specific estimators is evaluated alongside the relevant bounds. The Gaussian MLE is considered a misspecified MLE (MMLE). The MLE for the Student's t-distribution is not closed-form and requires an iterative Expectation Maximization (EM) algorithm. As this paper already includes substantial derivations related to the bounds, the derivation of the EM algorithms is omitted. More details on the EM algorithm for the Student's t-distribution are available in the literature [12].

Figure 2 shows that the MSE of an estimator assuming a Student's t-distribution when the data is truly Gaussian converges to the MCRB and therefore the Gaussian CRB (blue diamonds converge to light blue triangles). When outliers are present and the model is truly a Student's t-distribution, the

MSE of the Gaussian MLE (red circles) also converges to the MCRB, verifying the derivation. According to these results, there is no drawback in terms of MSE performance when assuming a Student's t-distribution to estimate the mean when the data is nominal.

The MSE of the correctly specified MLE for anomalous data has an MSE (black crosses) lower than the MCRB, showing an improvement in accuracy obtained by correctly specifying the heavy-tailed model. The MSE of the correctly specified estimator converges to the CRB for the Student's t-distribution with a low number of degrees of freedom, corresponding to large and frequent outliers. This validates the gain available in properly accounting for anomalies in the assumed statistical model. Since the shape parameter is linked to the weight of

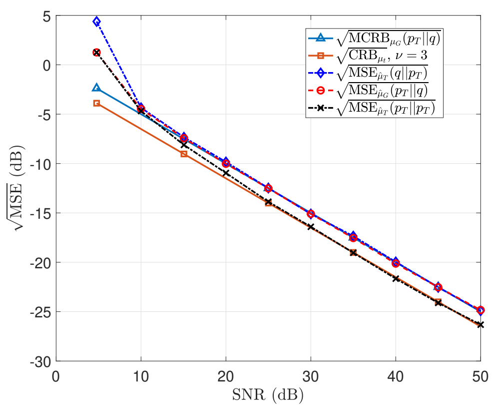

Fig. 2: MSE for each of the estimators are presented with $(p_1||p_2)$, where p_1 is the true distribution and p_2 is the assumed distribution. Distributions are parameterized with $\boldsymbol{\theta} = [0, 1]$ and $\boldsymbol{\eta} = [0, 1/3, 3]$ and 2000 Monte Carlo iterations are used to determine the MSE of the estimators.

the tails of the distribution, it is more difficult to converge to the MLE for even lower values of ν. The bound and estimator performance should still be investigated for the more extreme cases of anomalies where the number of degrees of freedom is less than 3. The EM algorithm does not always provide an asymptotically efficient estimator but it is expected to perform better than the MCRB.

V. Conclusion

The derived MCRB agrees with recent results in the literature, i.e., the MCRB assuming a Gaussian distribution is identical to the Gaussian CRB. This result is verified because the MLE that assumes a Gaussian distribution when the true model for the data is a Student's t-distribution obtains an MSE equivalent to the MCRB. The estimator based on the Student's t-distribution achieves lower MSE than the MCRB, and hence is preferred over the misspecified Gaussian estimator. The performance of the EM algorithm should be further investigated in future work to ensure the CRB is obtainable with a more restrictive number of degrees of freedom. Correctly assuming a Student's t-distribution for data with anomalies is theoretically proven to improve upon the Gaussian assumption. Future work can be conducted to investigate other types of distributions that can model anomalies and decide on a preferable model for real clock data. It could also be interesting to investigate the assumption of one heavy-tailed distribution when the true distribution follows a different outlier-compatible model, e.g., a Gaussian mixture of normal and contaminated observations.

Appendix A
Derivation of Pseudo-True Parameters.

The pseudo-true parameters $\tilde{\boldsymbol{\theta}} = [\tilde{\mu}, \tilde{\sigma}^2]^T$ are the parameters of the assumed distribution that minimize the KLD from the true distribution (see (9)). The cost function for finding the pseudo-true parameters is simplified to only include the parameters of the assumed distribution:

$$\tilde{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left\{ -E_p \left[\log \left(q(\mathbf{z}; \boldsymbol{\theta}) \right) \right] \right\}. \tag{16}$$

Substituting the log-likelihood function for the Gaussian distribution leads to:

$$-E_p \left[\log \left(q(\mathbf{z}; \boldsymbol{\theta}) \right) \right] = \frac{N}{2} E_p \left[\log \left(2\pi\sigma_G^2 \right) \right] \tag{17}$$

$$+ \frac{1}{2} \sum_{i=1}^{N} E_p \left[\left(\frac{z_i - \mu_G}{\sigma_G} \right)^2 \right]. \tag{18}$$

The mean and variance of a random variable that follows the Student's t-distribution are known to be:

$$E_p[z_i] = \mu_T, \mathrm{var}(z) = E_p[z_i^2] - E_p[z_i]^2 = \sigma_T^2 \frac{\nu}{\nu - 2}. \tag{19}$$

The objective function to be optimized can be written in terms of the mean and variance of the true distribution:

$$-E_p \left[\log \left(q(\mathbf{z}; \boldsymbol{\theta}) \right) \right] = \frac{N}{2} \log \left(2\pi\sigma_G^2 \right)$$

$$+ \frac{N}{2\sigma_G^2} \left(\mathrm{var}(z_i) + \left(E_p[z_i] - \mu_G \right)^2 \right). \tag{20}$$

The value of μ_G that minimizes the above cost function is obtained when $\mu_G = E_p \left[z_i \right]$. Therefore,

$$\tilde{\mu}_p = E_p \left[z_i \right] = \mu_T, \tag{21}$$

where the generalized derivation is in terms of $E_p[\cdot]$, meaning that any true distribution p should have the same form for the pseudo-true parameter. For the pseudo-true scale parameter, the following result is obtained:

$$-\frac{\partial}{\partial\sigma_G^2} E_p \left[\log \left(q(\mathbf{z}; \boldsymbol{\theta}) \right) \right] = \frac{N}{2\sigma_G^2} \tag{22}$$

$$- \frac{N}{2(\sigma_G^2)^2} \left(\mathrm{var}(z_i) + \left(E_p[z_i] - \mu_G \right)^2 \right). \tag{23}$$

Substituting the pseudo-true parameter for μ_G, one obtains:

$$\frac{N\sigma_G^2}{2} - \frac{N}{2} \mathrm{var}(z_i) = 0. \tag{24}$$

The resulting pseudo-true parameter is the sample variance of the Student's t-distribution:

$$\tilde{\sigma}_p^2 = \mathrm{var}(z_i) = \sigma_T^2 \frac{\nu}{\nu - 2}. \tag{25}$$

With the above result, we can conclude that the mean and variance of the Student's t-distribution are the pseudo-true parameters that minimize the KLD between a Gaussian distribution and a Student's t-distribution. As the derivation method is the same for other true distributions p, the same conclusion on the pseudo-true parameters when assuming a Gaussian distribution can be used to simplify future derivations.

APPENDIX B
MCRB COMPUTATION

The log-likelihood function of N i.i.d. Gaussian random variables in $\mathbf{z} = (z_1, \cdots, z_N)^T$ is well known, as are the derivatives of this function. Hence, the derivatives required to evaluate (11) and (12) are trivial. Only the moments of the true distribution are required, which have been shown to be related to the pseudo-true parameters. The matrix $\mathbf{A}$ is written in terms of the pseudo-true scale parameter or as shown above, the variance of the true distribution

$$\mathbf{A} = \begin{bmatrix} -\left(\tilde{\sigma}_p^2\right)^{-1} & 0 \\ 0 & -\frac{1}{2}\left(\tilde{\sigma}_p^2\right)^{-2} \end{bmatrix}. \tag{26}$$

The other matrix requires computation of higher order moments of the true distribution, i.e.,

$$B_{1,1} = \left(E_p \left[\left(\frac{z_i - \mu_G}{\sigma_G^2} \right)^2 \right] \right)_{\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}_p}, \tag{27}$$

$$B_{1,2} = \left(E_p \left[-\frac{z_i - \mu_G}{2(\sigma_G^2)^2} + \frac{1}{2} \frac{(z_i - \mu_G)^3}{(\sigma_G^2)^3} \right] \right)_{\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}_p}, \tag{28}$$

$$B_{2,2} = \left(E_p \left[\left(-\frac{1}{2\sigma_G^2} + \frac{1}{2} \frac{(z_i - \mu_G)^2}{(\sigma_G^2)^2} \right)^2 \right] \right)_{\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}_p}. \tag{29}$$

The first term is already known from the computation of $\mathbf{A}$, since $B_{1,1} = -A_{1,1}$. Due to the symmetry of the true distribution and the fact that $\tilde{\mu} = \mu_T$, $B_{1,2} = B_{2,1} = 0$. The term $B_{2,2}$ requires further derivation as it will have a unique form for the Student's t-distribution. Using $E_p[(z_i - \mu_G)^2] = \sigma_G^2$ leads to

$$B_{2,2} = \left(-\frac{1}{4(\sigma_G^2)^2} + \frac{1}{4(\sigma_G^2)^4} E_p \left[(z_i - \mu_G)^4 \right] \right)_{\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}_p}. \tag{30}$$

By substituting the pseudo-true parameters in (30), the expression simplifies and we can make use of the fourth-order central moment of the Student's t-distribution $E_p\left[(z_i - \mu_T)^4\right] = \frac{3\nu^2}{(\nu-2)(\nu-4)}\sigma_T^4$, hence:

$$B_{2,2}(p_T||q) = -\frac{1}{4(\tilde{\sigma}_p^2)^2} + \frac{\left(\sigma_T^2 \frac{\nu}{\nu-2}\right)^2}{4(\tilde{\sigma}_p^2)^4} \frac{3(\nu-2)}{(\nu-4)}. \tag{31}$$

The above can be simplified knowing that the pseudo-true scale is equal to the variance of the Student's t-distribution:

$$\mathbf{B} = \begin{bmatrix} \left(\tilde{\sigma}_p^2\right)^{-1} & 0 \\ 0 & \left(\frac{\nu-1}{2(\nu-4)}\right)\left(\tilde{\sigma}_p^2\right)^{-2} \end{bmatrix}. \tag{32}$$

All terms besides $B_{2,2}$ are the same in both $\mathbf{A}$ and $\mathbf{B}$ for any other type of symmetrical true distribution when assuming the Gaussian model. The final computation of the MCRB is a simple matrix multiplication, where the inverse of $\mathbf{A}$ is trivial because it is diagonal

$$\mathbf{MCRB}_{\boldsymbol{\theta}}(p_T||q) = \frac{1}{N} \begin{bmatrix} \tilde{\sigma}_p^2 & 0 \\ 0 & \left(\frac{2(\nu-1)}{(\nu-4)}\right)\left(\tilde{\sigma}_p^2\right)^2 \end{bmatrix}. \tag{33}$$

REFERENCES

[1] B. Cecconi, M. Dekkali, C. Briand, *et al.*, "NOIRE study report: Towards a low frequency radio interferometer in space," in *Proc. IEEE Aerospace Conference*, (Big Sky, MT, USA), pp. 1–19, IEEE, Mar. 2018.

[2] R. Piché, "Cramér-rao lower bound for linear filtering with t-distributed measurement noise," in *Proc. 19th International Conference on Information Fusion*, (Heidelberg, Germany), 2016.

[3] R. A. Maronna, R. D. Martin, and V. J. Yohai, *Robust Statistics: Theory and Methods*. Wiley Series in Probability and Statistics, Wiley, 1 ed., Mar. 2006.

[4] Q. H. Vuong, *Cramér-Rao bounds for misspecified models*. working paper 652, Div. of the Humanities and Social Sci., Caltech, Pasadena, USA, 1986.

[5] C. D. Richmond and L. L. Horowitz, "Parameter bounds on estimation accuracy under model misspecification," *IEEE Transactions on Signal Processing*, vol. 63, no. 9, pp. 2263–2278, 2015.

[6] S. Fortunati, F. Gini, and M. S. Greco, "The misspecified Cramér-Rao bound and its application to scatter matrix estimation in complex elliptically symmetric distributions," *IEEE Transactions on Signal Processing*, vol. 64, no. 9, pp. 2387–2399, 2016.

[7] S. Fortunati, F. Gini, M. S. Greco, and C. D. Richmond, "Performance bounds for parameter estimation under misspecified models: Fundamental findings and applications," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 142–157, 2017.

[8] S. Fortunati and L. Ortega Espluga, "On the efficiency of misspecified gaussian inference in nonlinear regression: application to time-delay and doppler estimation." 2023. hal-04349291.

[9] P. Tavella and C. Thomas, "Comparative study of time scale algorithms," *Metrologia*, vol. 28, pp. 57–63, jan 1991.

[10] H. McPhee, J.-Y. Tourneret, D. Valat, *et al.*, "A Robust Time Scale Based on Maximum Likelihood Estimation," in *Proc. of the 54th Annual Precise Time and Time Interval Systems and Applications Meeting*, (Long Beach, California), pp. 61–75, Feb. 2023.

[11] K. L. Lange, R. J. A. Little, and J. M. G. Taylor, "Robust statistical modeling using the t distribution," *Journal of the American Statistical Association*, vol. 84, no. 408, pp. 881–896, 1989.

[12] F. Doğru, Y. M. Bulut, and O. Arslan, "Doubly reweighted estimators for the parameters of the multivariate t-distribution," *Communications in Statistics - Theory and Methods*, vol. 47, 03 2018.