Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 14, No. 1, pp. 51–61, 2018

DISTRIBUTED MAXIMUM LIKELIHOOD ESTIMATION FOR CENSORED DEPENDENT QUANTIZED DATA

Hongzeng Wang, Qingling Zhang* and Jian Wu

Abstract: In wireless sensor networks, dependent quantized data are widely studied. We consider the problem of parameters estimation for dependent quantized communication data where some of the data are censored to the Fusion Center (FC). We propose a distributed imputation method to fill in the censored values, and introduce a two-step maximum likelihood estimation (MLE) method to estimate the unknown parameters of the sensor system. The lower bound of the variance of the estimator is discussed to show the asymptotic efficiency. Numerical example of the constant false alarm rate (CFAR) detection system demonstrate the effectiveness of the proposed scheme.

Key words: *wireless sensor networks, maximum likelihood estimation, censored data, distributed imputation*

Mathematics Subject Classification: *94A12, 65C50*

1 Introduction

Dependent quantized data are often encountered in wireless sensor networks (WSNs) when the communication bandwidth and the power are limited. One application of WSNs is in radar and satellite-based remote-sensing systems to sonar and seismology [6]. More examples can be found in [15] and [17], such as air-traffic control, military command and control and weather prediction. System identification is a primary problem in WSNs, and parameter estimation based on dependent quantized data is wildly applied to identify these systems.

Several methods have been suggested in the literature of the system identification and estimation in recent years, some of the methods were based on dependent quantized data. For example, In [3], the authors reviewed the estimating methods applied to the quantized data, and studied the problem of dithering noise at the sensors. In [4], an expectation maximization algorithm and Quasi-Newton optimization method were suggested to solve the problem of estimating the parameter of a linear system, including the single input single output (SISO) case and multiple input multiple output (MIMO) case. The authors in [13] studied an MLE based approach for the local estimation problem assuming that the distributed sensors are not independent and the problem of Copula selection was also considered. In [16], the authors suggested a quadratic programming-based method for identification of finite impulse response (FIR) system from quantized data. In [18], the authors presented a two-stage distributed algorithm with the running average technique achieving the centralized

*Corresponding author. The research of the second author is supported by the National Natural Science Foundation of China (Grant No. 61673099).

sample mean estimate in a distributed manner. Parallel sensors with multiple groups of quantizers is a better design of the networks than with single quantizer. For the reason that estimating the parameters with single quantizer is sensitive to the choice of thresholds due to the uncertainty of probability density function(PDF) of the networks [16]. So the work [11] proposed a robust distributed MLE approach with multiple groups of quantizers for the quantized data.

Prior researches on the system identification with quantized data often focused on complete data to estimate the unknown parameters of the system. But in some extremely adverse working conditions, such as in high temperature, high pressure or electromagnetic interference surroundings, the quantizers will be broken or become no response. Then the FC will receive censored quantized communication data, which will lead to greater bias and variance of the estimator of the unknown parameter, and the system will be identified by mistake. In [2, 7, 9] and [12] authors discussed the problem of distribution estimation using auxiliary information in survey data or complex survey data. Motivated by the complex data analysis in practice, we investigate the problem of parameter estimation in parallel sensor network system with multiply groups of quantizers, with which dependent quantized data are censored.

Censored data is a kind of missing data, in this case the mechanism leading to missing data may not be under control of the system. The analysis of the data needs to take account of this information to avoid biased result [5]. Our main contribution is that we propose a two-step MLE of censored dependent quantized data by filling in the missing values. We analytically derive the asymptotic efficiency of the estimator. Numerical example shows that the estimator based on our method has less mean-square error (MSE) than those obtained by the censored data.

The outline of the paper is as follows: in Section 2 the problem formulation is given, in Section 3 the method of imputation and procedure of two-step estimation are proposed, in Section 4 a numerical example of distributed CFAR detection system is discussed, in Section 5 conclusions are made, in Section 6 lemmas and regulation conditions to be used to prove theorem 3.1 are listed.

2 Problem Formulation

We focus on the problem of the parallel sensor network system. Suppose the total of L sensors have a joint observation population $(X_1, \dots, X_L)^T$, which has a given family of a joint PDF:

$$\{p(x_1, \dots, x_L|\theta)\}_{\theta \in \Theta \subseteq \mathbb{R}^s}, \tag{2.1}$$

where θ is the unknown s-dimensional deterministic parameter vector including not only the marginal parameters but the dependence parameters as well. Let $\mathbf{X}_{N\times L}$ denote the total of L sensors with the N observation samples:

$$\mathbf{X} = (\mathbf{X_1}, \dots, \mathbf{X_L}),$$

where $\mathbf{X_i} = (X_{1i}, \dots, X_{Ni})^T$, $i = 1, \dots, L$, denote the N observation samples of the ith sensor.

In many practical situation, $\mathbf{X}_{N\times L}$ needed to be typically quantized to minimize the utilization of communication resource of the sensor network system [8]. In this paper we consider the binary modulation. Let $(Q_1, \dots, Q_L)^T$ denote the quantized population, we have

$$(Q_1, \dots, Q_L)^T = (I_1(X_1), \dots, I_L(X_L))^T, \tag{2.2}$$

where $I_i(X_i) : X_i \to \{0,1\}$ for $i = 1,\ldots,L$ is the indicator quantized function of the ith sensor.

Based on (2.1) and (2.2), $(Q_1,\ldots,Q_L)^T$ has a joint probability mass function (PMF),

$$f_Q(q_1,\ldots,q_L|\theta) = \int \cdots \int p(x_1,\ldots,x_L|\theta)dx_1 \cdots dx_L, \tag{2.3}$$

and we have the total of L sensors with the N quantized observation samples,

$$\mathbf{Q} \triangleq (\mathbf{Q}_1,\ldots,\mathbf{Q}_L),$$

where

$$\mathbf{Q}_i \triangleq (Q_{1i},\ldots,Q_{Ni})^T, i = 1,\ldots,L,$$

and

$$Q_{ni} \triangleq I_i(X_{ni}), i = 1,\ldots,L, n = 1,\ldots,N.$$

By (2.3), we obtain the log likelihood function of the parameter θ given $\mathbf{Q}$,

$$l(\theta|\mathbf{Q}) \triangleq \log \prod_{n=1}^{N} f_Q(Q_{n1},\ldots,Q_{nL}|\theta).$$

Thus, the MLE method can be adopted to obtain the estimator $\hat{\theta}$ by

$$\hat{\theta} = \arg\max_{\theta} l(\theta|\mathbf{Q}).$$

It can be proved that $\hat{\theta}$ is a consistent and asymptotically efficient estimator of the unknown parameter θ^0 [1], [14].

Let $I_i^{(j)}(X_i) : X_i \to \{0,1\},\ i = 1,\ldots,L,\ j = 1,\ldots,J$ denote the J groups of different indicator quantized functions. The quantized observation samples are $\{(Q_{n_j1}^{(j)},\ldots,Q_{n_jL}^{(j)})\}_{n_j=1}^{N_j}$, where N_j is sample size of the jth quantizer.

The ML estimator of the parameter θ is given by

$$\hat{\theta}_R = \arg\max_{\theta} \log \prod_{j=1}^{J} \prod_{n_j=1}^{N_j} f_Q^{(j)}(Q_{n_j1}^{(j)},\ldots,Q_{n_jL}^{(j)}|\theta), \tag{2.4}$$

where $f_Q^{(j)}(\cdot|\theta), j = 1,\ldots,J$ are the joint PMF of the quantized population $(Q_{n_j1}^{(j)},\ldots,Q_{n_jL}^{(j)})$, obtained similarly to (2.3). The robust estimator $\hat{\theta}_R$ is a consistent and asymptotically efficient estimator of the unknown parameter θ^0 [11].

3 MLE for Censored Dependent Quantized Data

In this section, we consider the problem of the censored data with multiple groups of quantizers. In some extreme conditions, the quantizers will be broken or receive nonresponse, and the FC will receive the quantized data with censoring. As known the mechanism of the censored data, we suppose an imputation (fill in the missing values) procedure to estimate the unknown parameter θ^0.

The idea of imputation comes from the mechanism that the $J-groups$ of quantizers are not independence. If not all of the quantizers receive nonresponse at the same time, we can fill in the censored values by using the dependent distribution of the quantizers. There are some basic methods of imputation, known as *mean imputation*, *Hot deck imputation*, *substitution imputation* (see e.g. [5]), all of them suggest to fill in the censored values by the observed data. In this problem, the quantizers may have a high probability of nonresponse, thus the observed data will be too little for imputation by applying above methods. We suppose a distributed imputation method as follows.

We write $\mathbf{Q}_1 = (\mathbf{Q}_{obs}, \mathbf{Q}_{cen})$, where $\mathbf{Q}_{obs}$ denotes the observed values, and $\mathbf{Q}_{cen}$ denotes the censored values. Based on (2.3), let $f_Q(q_c|\theta) = f_Q(q_{obs}, q_{cen}|\theta)$ denotes the joint PMF of $\mathbf{Q}_{obs}$ and $\mathbf{Q}_{cen}$. We can obtain the marginal PMF of $\mathbf{Q}_{obs}$ by

$$f_{obs}(q_{obs}|\theta) = \int f_Q(q_{obs}, q_{cen}|\theta) dq_{cen}.$$

Suppose $N_{(j)}, j = 1, \dots, J$ denote the sample size of the jth nonresponse quantizer. We have $0 \leqslant N_{(1)} \leqslant N_{(2)} \leqslant \dots N_{(J)}$. We can obtain the first step estimator with the observed values by the MLE method. The log likelihood function of the observed values is

$$L(\theta|q_{obs}) = \log \prod_{j=1}^{J} \prod_{n_j=1}^{N_{(j)}} f_{obs}^{(j)}(q_{n_j 1}^{(j)}, \dots, q_{n_j L}^{(j)}|\theta),$$

where $f_{obs}^{(j)}(\cdot|\theta)$ are the marginal PMF of the observed population $(Q_{n_j 1}^{(j)}, \dots, Q_{n_j L}^{(j)})$. Let $\hat{\theta}_{obs}$ denote the solution of MLE with the observed values. According to [11], $\hat{\theta}_{obs}$ is a consistent and asymptotically efficient estimator of the unknown parameter θ^0.

Let $f_{cen}(q_{cen}|\theta)$ denote the marginal PMF of the censored values. The PMF of $f_{cen}(q_{cen}|\theta)$ can be obtained by integrating out the observed data,

$$f_{cen}(q_{cen}|\theta) = \int f_Q(q_{obs}, q_{cen}|\theta) dq_{obs}. \tag{3.1}$$

Then, we have an approximate marginal PMF of $\mathbf{Q}_{cen}$ as $\mathbf{Q}_{cen} \dot{\sim} f_{cen}(q_{cen}|\hat{\theta}_{obs})$. Based on the mechanism of censored data and the discussion above, we fill in the missing values by the following algorithm,

Algorithm:

1: **Input:** $\hat{\theta}_{obs}, Q_{obs}, I$
2: **Output:** Q_{im}
3: Generate random numbers with parameter $\hat{\theta}_{obs}$,
   $X \sim F(x|\hat{\theta}_{obs})$
4: Compute the quantized data $Q^{(j)} = I^{(j)}(X), j = 1, \dots, J$
5: Check Q
   if $Q^{(j)} == Q_{obs}^{(j)}$, j in $obs(1, \dots, J)$
   then $Q_{im}^{(j)} = Q^{(j)}$, j in $cen(1, \dots, J)$
6: end

We write $\mathbf{Q}_2 = (\mathbf{Q}_{obs}, \mathbf{Q}_{im})$, where $\mathbf{Q}_{im}$ denotes the imputed values based on the suggest algorithm. The log likelihood function of the sample, which are combined by observed values together with the imputed values is,

$$L(\theta|\mathbf{Q}_2) = \log \prod_{j=1}^{J} \prod_{n_j=1}^{N_{(J)}} f_Q^{(j)}(q_{n_j 1}^{(j)}, \dots, q_{n_j L}^{(j)}|\theta), \tag{3.2}$$

where $f_Q^{(j)}(\cdot|\theta)$ is the joint PMF of the jth quantizer with the complete samples. We can obtain the second step estimator by maximum the log likelihood function. We have the following theorem.

Theorem 3.1. *Assume that $I^{(j)}(x), j = 1, \ldots, J$ are the quantizers of different sensor, $f_Q^{(j)}(\cdot|\theta)$ is the joint PMF of the jth quantizer, which satisfy the regularity conditions listed in the appendix of section 6, θ^0 is the unknown deterministic parameter vector, $\hat{\theta}$ is the solution of MLE with (3.2). Then,*

$$\sqrt{N}(\hat{\theta} - \theta^0) \xrightarrow{L} \mathcal{N}\left[0, \left(\frac{1}{J}\sum_{j=1}^{J}\mathcal{I}(\theta^0; I^{(j)}(y))\right)^{-1}\right], \tag{3.3}$$

where $N = J \cdot N_{(J)}$, and $\mathcal{I}(\theta^0; I^{(j)}(y))$ is the Fisher information matrix for one quantized sample. That is, $\hat{\theta}$ is a consistent and asymptotically efficient estimator of θ^0.

Proof. We use Taylor expansion on first derivative of the log likelihood function around the true value θ^0.

$$\frac{\partial l(\theta|\overrightarrow{Q}^{(1)}, \ldots, \overrightarrow{Q}^{(J)})}{\partial\theta} = D^1(\theta^0) + D^2(\theta^0)(\theta - \theta^0) + \frac{1}{2}D^3(\theta - \theta^0; \theta^*)(\theta - \theta^0), \tag{3.4}$$

where,

$$D^1(\theta^0) = \left.\frac{\partial l(\theta|\overrightarrow{Q}^{(1)}, \ldots, \overrightarrow{Q}^{(J)})}{\partial\theta}\right|_{\theta^0},$$

$$D^2(\theta^0) = \left.\frac{\partial^2 l(\theta|\overrightarrow{Q}^{(1)}, \ldots, \overrightarrow{Q}^{(J)})}{\partial\theta^2}\right|_{\theta^0},$$

and

$$D^3(\theta - \theta^0; \theta^*) = \left((\theta - \theta^0)'\left[\left.\frac{\partial^3 l(\theta|\overrightarrow{Q}^{(1)}, \ldots, \overrightarrow{Q}^{(J)})}{\partial\theta^3}\right|_{\theta^*}\right]\right).$$

with θ^* between θ and θ^0.

As $\hat{\theta}$, is the solution of (3.4), we have

$$0 = \left.\frac{\partial l(\theta|\overrightarrow{Q}^{(1)}, \ldots, \overrightarrow{Q}^{(J)})}{\partial\theta}\right|_{\hat{\theta}} = D^1(\theta^0) + D^2(\theta^0)(\hat{\theta} - \theta^0) + \frac{1}{2}D^3(\hat{\theta} - \theta^0; \theta^*)(\hat{\theta} - \theta^0), \tag{3.5}$$

then,

$$\sqrt{N}(\hat{\theta} - \theta^0) = \frac{-1}{\sqrt{N}}D^1(\theta^0)\left[\frac{1}{N}D^2(\theta^0) + \frac{1}{2N}D^3(\theta - \theta^0; \theta^*)\right]^{-1} \tag{3.6}$$

Firstly, for $D^1(\theta^0)$, we have

$$E\left[\left.\frac{\partial \log f^{(j)}(\overrightarrow{Q}^{(1)}, \ldots, \overrightarrow{Q}^{(J)}|\theta)}{\partial\theta}\right|_{\theta^0}\right] = 0,$$

and

$$B_{N_j}^2 \triangleq Cov\left[\left.\frac{\partial \log f^{(j)}(\overrightarrow{Q}^{(1)}, \ldots, \overrightarrow{Q}^{(J)}|\theta)}{\partial\theta}\right|_{\theta^0}\right] = N_j\mathcal{I}(\theta^0, I^{(j)}).$$

By applying Lindeberg's central limit theorem and lemma 6.3, we have,

$$\frac{1}{\sqrt{N}} D^1(\theta^0) \xrightarrow{L} \mathcal{N}(0, B_n^2) \tag{3.7}$$

where, $B_n^2 = \sum_{j=1}^J B_{N_j}^2$.

Secondly, for every $j = 1, \ldots, J$,

$$\begin{aligned}
& E\left[\frac{\partial^2 \log f^{(j)}(\overrightarrow{Q}^{(j)}|\theta)}{\partial \theta^2}\bigg|_{\theta^0}\right] \\
= \; & E\left[\frac{\partial^2 f^{(j)}(\overrightarrow{Q}^{(j)}|\theta^0)}{\partial \theta^2}\frac{1}{f^{(j)}(\overrightarrow{Q}^{(j)}|\theta^0)} - \left(\frac{\partial f^{(j)}(\overrightarrow{Q}^{(j)}|\theta_0)}{\partial \theta}\frac{1}{f^{(j)}(\overrightarrow{Q}^{(j)}|\theta^0)}\right)^2\right] \\
= \; & -E\left(\frac{\partial f^{(j)}(\overrightarrow{Q}^{(j)}|\theta_0)}{\partial \theta}\frac{1}{f^{(j)}(\overrightarrow{Q}^{(j)}|\theta^0)}\right)^2 \\
= \; & -N_j \mathcal{I}(\theta^0, I^{(j)})
\end{aligned} \tag{3.8}$$

By the properties of the law of large number,

$$\frac{1}{N_{(j)}}\frac{\partial^2 l(\theta|\overrightarrow{Q}^{(j)})}{\partial \theta^2}\bigg|_{\theta^0} \xrightarrow{P} -\mathcal{I}(\theta^0, I^{(j)}(y)),$$

then

$$\begin{aligned}
\frac{1}{N} D^2(\theta^0) &= \sum_{j=1}^{J} \frac{N_{(J)}}{N}\frac{1}{N_{(J)}}\frac{\partial^2 l(\theta|\overrightarrow{Q}^{(j)})}{\partial \theta^2}\bigg|_{\theta^0} \\
&\xrightarrow{P} \frac{1}{J}\sum_{j=1}^{J}\mathcal{I}(\theta^0, I^{(j)}(y)) \\
&= -B_N^2
\end{aligned} \tag{3.9}$$

Thirdly, based on Theorem 4.17 in [10] p290, $\hat{\theta} \xrightarrow{P} \theta^0$, and Regularity Condition (A6), that the three times differentiation of the log likelihood functions are bounded by integrable functions. Noticed that θ^* is between θ and θ^0, we have

$$\frac{1}{2N} D^3(\theta - \theta^0; \theta^*) \xrightarrow{P} 0 \tag{3.10}$$

Finally, applying Slutsky's lemma and (3.6), (3.7), (3.9), (3.10),we have

$$\sqrt{N}(\hat{\theta} - \theta^0) \xrightarrow{L} \mathcal{N}\left[0, \; \left(\frac{1}{J}\sum_{j=1}^{J}\mathcal{I}(\theta^0; I^{(j)}(y))\right)^{-1}\right]$$

□

Remark 3.2. The covariance matrix $\left(\frac{1}{J}\sum_{j=1}^J \mathcal{I}(\theta^0, I^{(j)}(y))\right)^{-1}$ is the Cramer-Rao lower bound. By the suggested technology of imputation, we obtain complete values and the estimator in (3.3). The asymptotic covariance matrix is greater than that of complete values and less than that of censored values.

4 Numerical Example

In this section, we consider the problem of estimating the parameters in the distributed CFAR detection system [15]. We assume the two-sensor system has a joint PDF,

$$(X_1, X_2) \sim p(x_1, x_2|\theta),$$

where

$$p(x_1, x_2|\theta) = c(F_1(x_1|\theta_1), F_2(x_2|\theta_2)|\theta_0)p_1(x_1|\theta_1)p_2(x_2|\theta_2),$$

and

$$c(v_1, v_2|\theta_0) = \Big[\max\{v_1^{-\theta_0} + v_2^{-\theta_0+1}; 0\}\Big]^{-1/\theta_0}$$

is the Clayton copula density function with unknown parameters θ_0. Assume $X_i, i = 1, 2$ has the marginal distribution of $Gamma(\theta_i, \beta), i = 1, 2$ with an unknown shape parameter $\theta_i, i = 1, 2$, and a rate parameter $\beta = 4$, $F_i(x_i|\theta_i)$ and $p_i(x_i, |\theta_i)$ are the cumulative distribution function(CDF) and PDF of X_i. We write $\theta = (\theta_0, \theta_1, \theta_2)$, which is the parameter to be estimated. Moreover, there are four groups of quantizers in each sensor, denote as

$$I^{(j)}(x) = (I_1^{(j)}(x_1), I_2^{(j)}(x_2)) = (I[x_1 - h_j], I[x_2 - h_j]), j = 1, 2, 3, 4,$$

where h_j will be chosen by the prior information of the sensors.

Table 1: MEAN, MSE of θ_0

$\theta_0 = 1.076$		$p = 0.001$	$p = 0.01$	$p = 0.05$	$p = 0.1$
ORACLE	MEAN	1.078	1.115	1.430	2.359
	MSE	.0072	.0041	5.944	45.18
NAIVE	MEAN	1.112	1.212	2.268	4.774
	MSE	.1095	.4985	60.46	436.4
IMPUTE	MEAN	1.094	1.167	2.178	4.611
	MSE	.0360	.3877	36.22	366.5

Table 2: MEAN, MSE of θ_1

$\theta_1 = 4$		$p = 0.001$	$p = 0.01$	$p = 0.05$	$p = 0.1$
ORACLE	MEAN	4.000	4.001	4.008	4.012
	MSE	2.5×10^{-5}	.0286	.1384	.2800
NAIVE	MEAN	3.999	3.997	4.012	4.024
	MSE	.0279	.0930	.3411	.5874
IMPUTE	MEAN	4.000	4.003	4.009	4.021
	MSE	.0066	.0450	.2145	.4242

Table 3: MEAN, MSE of θ_2

$\theta_2 = 5$		$p = 0.001$	$p = 0.01$	$p = 0.05$	$p = 0.1$
ORACLE	MEAN	5.001	5.001	5.019	5.032
	MSE	.0037	.0350	.1751	.3316
NAIVE	MEAN	5.002	5.002	5.040	5.072
	MSE	.0139	.1170	.4411	.4861
IMPUTE	MEAN	5.002	5.001	5.031	5.062
	BIAS	.0081	.0581	.2819	.5441

Table 1 - Table 3 present the MEAN and MSE of the parameter θ, based on 5000 Monte Carlo runs. Let $\theta = (1.076, 4, 5)$, where $\theta_0 = 1.076$ corresponds to Spearman's $\rho = 0.5$, which is a dependence measurement of (X_1, X_2). If X has Gamma distribution $X \sim Gamma(\alpha, \beta)$, the expected value of X is $E(X) = \alpha\beta$. According to the assumption that the two sensors have different marginal distributions $X_1 \sim Gamma(4, 4)$ and $X_2 \sim Gamma(5, 4)$, we have $E(X_1) = 16, E(X_2) = 20$. The thresholds of quantizers $h = (h_1, h_2, h_3, h_4)$ should be chosen by prior information of the censors, so we assume that $h = (25, 20, 15, 10)$.

Let p denote measurement of the nonresponse probability of the quantizes, and $p = (0.001, 0.01, 0.05, 0.1)$ simulate the four different conditions that cause the quantizers do not work. Because of the mechanism of the censored data, the sample size $N_{(j)}, j = 1, 2, 3, 4$ of each quantizer are $(1000, 100, 20, 10)$ in average. Without loss of generality, let the non-response probability of the quantizes be equal. In Table 1 - Table 3, ORACLE denote the complete sample data of size $N_{(J)}$ of each quantizer, NAIVE denote the observed censored data and IMPUTE denote the observed data together with the imputed data also at size $N_{(J)}$.

The result in Table 1 - Table 3 illustrate that (1) By proposed method, the bias of the ML estimators based on 5000 M. C. runs are going down together with the nonresponse probability of quantizes becoming small, for the estimator $\hat{\theta}$ is asymptotically efficient. (2) In the same level of nonresponse probability of the quantizes, MSE of IMPUTE data is greater than the value of ORACLE and less than the value of NAIVE, which is consistent to the discussion in Remark 3.2.

5 Conclusion

In this paper, we have investigated the problem of parameter estimate in system identification by the censored dependent quantized data. According to the knowledge of mechanism of censored data, we have considered a technology of distribution imputation named as distribution imputation to fill in the censored quantized values. Then we have proposed a two-step MLE method to estimate the unknown parameter of the system. We have also discussed the asymptotical efficient of the ML estimator. Simulation results of the CFAR detection system showed that both the bias and variance of the estimator became smaller than those of the censored data.

References

[1] G. Casella and R.L. Berger, *Statistical Inference*, 2^{nd} edition Duxbury, New York, 2001.

[2] J. Chen and R.R. Sitter, A pseudo empirical likelihood approach to the effective use of auxiliary information in complex surveys, *Stat. Sinica*, 9 (1999) 385–406.

[3] F. Gustafsson and R. Karlsson, Statistical results for system identification based on quantized observations, *Automatica* 45 (2009) 2794–2801.

[4] B.I. Godoy, G.C. Goodwin, J.C. Agüero, D. Marelli and T. Wigren, On identification of fir systems having quantized output data, *Automatica* 47 (2011) 1905–1915.

[5] R.J.A. Little and D.B.Rubin, *Statistical analysis with missing data*, John Siley & Sons, New York, 1987.

[6] V. Megalooikonomou and Y. Yesha, Quantizer design for distributed estimation with communication constraints and unknown observation statistics, *IEEE T. Commun.* 48 (2000) 181–184.

[7] J.F. Muñoz, E. Álvarez and M. Rueda, Optimum design-based ratio estimators of the distribution function, *J. Appl. Stat.* 41 (2014) 1395–1407.

[8] N.G. Nair, F. Fagnani, S. Zampieri and R.J. Evans, Feedback control under data rate constraints: an overview, *P. IEEE* 95 (2007) 108–137.

[9] J.N.K. Rao, J.G. Kovar and H.J. Mantel, On estimating distribution functions and quantiles from survey data using auxiliary information, *Biometrika* 77 (1990) 365–375.

[10] J. Shao, *Mathematical Satistics*, 2^{nd} edition Springer Press, New York, 2007.

[11] X.J. Shen, P. K. Varshney and Y.M. Zhu, Robust distributed maximum likelihood estimation with dependent quantized data, *Automatica* 50 (2014) 169–174.

[12] P.N. Silva and C.J. Skinner, Estimating distribution functions with auxiliary information using poststratification, *J. Off. Stat.* 11 (1995) 277–294.

[13] F.A. Sundaresan and P. K. Varshney, Location estimation of a random signal source based on correlated sensor observations, *IEEE T. Signal. Proces.* 59 (2011) 787–799.

[14] A.W. Van der Vaart, *Asymptotic Statistics*, Cambridge University Press, New York, 2000.

[15] P.K. Varshney, *Distributed Detection and Data Fusion*, Springer-Verlag, New York, 1997.

[16] J.D. Wang and Q.H. Zhang, Identification of fir systems based on quantized output measurements: a quadratic programming-based method, *IEEE T. Automat. Contr.* 60 (2015) 1439–1444.

[17] L.Y. Wang, G.G. Yin, J.F. Zhang and Y.L. Zhao, *System Identification with Quantized Observations*, Birkhäuser, Boston, 2010.

[18] S.Y. Zhu, Y.C. Soh and L.H. Xie, Distributed parameter estimation with quantized communication via running average, *IEEE T. Signal. Proces.* 63 (2015) 4634–4646.

6 Appendix

In this appendix we list the lemmas and regularity conditions, that applied in the proof of Theorem 3.1.

Lemma 6.1 (Slutsky). *Let X_n, X and Y_n random vectors or variables, if $X_n \xrightarrow{L} X$ and $Y_n \xrightarrow{P} c$ (c is a constant). Where " $\xrightarrow{L}$ " means convergence in distribution, and " $\xrightarrow{P}$ " means convergence in probability. Then*

(1) $X_n + Y_n \xrightarrow{L} X + c$;

(2) $X_n Y_n \xrightarrow{L} cX$;

(3) $X_n/Y_n \xrightarrow{L} X/c \quad (c \neq 0)$.

Lemma 6.2 (Lindeberg's Central Limit Theorem)**.** *Suppose X_n is a sequence of independent random vectors with finite variances. Let $a_i = E(X_i), b_i^2 = Cov(X_i), B_n^2 = \sum_{i=1}^n b_i^2$, if for every vector $\varepsilon > 0$*

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{i=1}^n E[(X_i - a_i)'(X_i - a_i)\mathbf{1}_{(|X_i - a_i| > \varepsilon' B_n)}] = 0, \tag{6.1}$$

then,

$$\sum_{i=1}^n (X_i - a_i) \xrightarrow{L} \mathcal{N}(0, B_n^2). \tag{6.2}$$

In Lemma 6.2, the equation (6.1) is known as Lindeberg's condition, and $\mathbf{1}_{(\cdot)}$ is the indicator function.

Lemma 6.3. *Let X_n are the independent random vectors in Lemma 6.2, if there exist constants K_n, such that*

$$\max_{1\le i\le n} |X_i| \le K_n \tag{6.3}$$

and

$$\lim_{n\to\infty} \frac{K_n}{B_n} = 0 \tag{6.4}$$

then (6.2) is satisfied.

Proof. (Proof of Lemma 6.3) For any $\varepsilon > 0$, if n is sufficiently large , we have $2K_n \le \varepsilon' B_n$, and

$$(X_i - a_i) \le 2K_n, i = 1, \dots, n$$

thus

$$\{(X_i - a_i) \le \varepsilon' K_n\} = \Omega, i = 1, \dots, n$$

where Ω is the sample space of population X. Suppose $F_i(x)$ is the CDF of X_i, we have

$$\begin{aligned}
& \lim_{n\to\infty} \frac{1}{B_n^2} \sum_{i=1}^n E[(X_i - a_i)'(X_i - a_i)\mathbf{1}_{(|X_i - a_i| \le \varepsilon' B_n)}] \\
= & \lim_{n\to\infty} \frac{1}{B_n^2} \sum_{i=1}^n \int_{(|X_i - a_i| \le \varepsilon' B_n)} (x_i - a_i)'(x_i - a_i) dF_i(x) \\
= & \lim_{n\to\infty} \frac{1}{B_n^2} \sum_{i=1}^n \int_\Omega (x_i - a_i)'(x_i - a_i) dF_i(x) = 1
\end{aligned}$$

Thus, Lindeberg's condition is satisfied, and we have $\sum_{i=1}^n (X_i - a_i) \xrightarrow{L} \mathcal{N}(0, B_n^2)$. □

Regularity Conditions: For every $j = 1, \dots, J$.

(A1) The quantized data $Q^{(j)}, j = 1, \dots, J$ of different sensors are independent. The joint PMF of the jth quantizer is $f_Q^{(j)}(\cdot\,|\theta)$, and $Q^{(j)} \sim f_Q^{(j)}(q|\theta)$.

(A2) The parameter is identifiable, that is, if $\theta \ne \theta'$, then $f_Q^{(j)}(q|\theta) \ne f_Q^{(j)}(q|\theta')$.

(A3) The densities $f_Q^{(j)}(q|\theta)$ have common support, and $f_Q^{(j)}(q|\theta)$ is differentiable in θ.

(A4) The parameter space Θ contains an open set, of which the true parameter value θ^* is an interior point.

(A5) The density $f_Q^{(j)}(q|\theta)$ is three times differentiable with respect to θ, and $\int f_Q^{(j)}(q|\theta)dq$ can be differentiated three times under the integral sign.

(A6) For all $\theta \in \Theta$, there exists a function $M_{ikl}^{(j)}(q)$ such that

$$\left| \frac{\partial^3}{\partial\theta_i \partial\theta_k \partial\theta_l} \log f_Q^{(j)}(q|\theta) \right| \leq M_{ikl}^{(j)}(q),$$

with $E_{\theta_0}[M^{(j)}(Q)] < \infty$, for all i, k, l.

Manuscript received 22 May 2016
revised 16 January 2017
accepted for publication 18 January 2018

Hongzeng Wang
Department of Mathematics
Northeastern University Shenyang
110819, P. R. China
E-mail address: wanghongzeng@mail.neu.edu.cn

Qingling Zhang
Department of Mathematics
Northeastern University Shenyang
110819, P. R. China
E-mail address: qlzhang@mail.neu.edu.cn

Jian Wu
Department of Mathematics
Northeastern University Shenyang
110819, P. R. China
E-mail address: wujian@mail.neu.edu.cn