國立交通大學

電信工程研究所

博士論文

稀少取樣下之組合式訊號測不準表示法研究與其在訊號平均值估計之應用

An Efficient Representation of Uncertainty Measurement for Combined Signals on Small Sampling Size Condition and its Application to Signal Mean Estimation

研究生： 羅文輝

指導教授： 陳信宏博士

中華民國九十九年七月

稀少取樣下之組合式訊號測不準表示法研究與其在訊號平均值估計之應用

An Efficient Representation of Uncertainty Measurement for Combined Signals on Small Sampling Size Condition and its Application to Signal Mean Estimation

研究生：羅文輝　　Student: Wen-Hui Lo
指導教授：陳信宏　博士　　Advisor: Dr. Sin-Horng Chen

國立交通大學

電信工程研究所

博士論文

A Dissertation Submitted to Institute of
Communication Engineering
College of Electrical and Computer Engineering
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in
Communication Engineering
Hsinchu, Taiwan
2010 年 7 月

推薦函

中華民國九十九年六月十四日

一、 事由：本校電信工程研究所博士班研究生羅文輝提出論文以參加國立交通大學博士班論文口試。

二、 說明：本校電信工程研究所博士班研究生羅文輝已完成本校電信工程研究所規定之學科課程及論文研究之訓練。

有關學科部分，羅君已修滿十八學分之規定（請查閱學籍資料）並通過資格考試。

有關論文部分，羅君已完成其論文初稿，相關之論文亦分別發表或即將發表於國際期刊（請查閱附件）並滿足論文計點之要求。總而言之，羅君已具備國立交通大學電信工程研究所應有之教育及訓練水準，因此特推薦
羅君參加國立交通大學電信工程研究所博士班論文口試。

交通大學電信工程研究所教授 陳信宏 陳信宏

稀少取樣下之組合式訊號測不準表示法研究與其在訊號平均值估計之應用

研究生：羅文輝　　　　　　　指導教授：陳信宏　博士

國立交通大學電信工程研究所

中文摘要

在許多訊號處理應用中，量測訊號時往往無法單獨得出某種訊號成分，所以對於組合式訊號(combined signals)之量測(measurement)與表示有其基本必要性。現今對於組合式訊號之量測以及其相對不確定性(uncertainty)之表示及分析方法上，僅對於特定條件下之輸出訊號可進行詮釋。在最新JCGM 101 (The Joint Committee for Guides in Metrology) 2008年公開文獻中對於上述組合式訊號之量測與表示，仍然承襲過往GUM (Guide to the Expression of Uncertainty in Measurement)之範疇，以衍生分布(propagation of distribution)之模型描述組合訊號，量測結果則以JCGM 101所建議之表示法為標準，此表示法之成員有：平均值(mean)、標準誤(standard uncertainty)、母體涵蓋率之相對應覆蓋區間(coverage interval)、以及此覆蓋區間之端點(endpoints)位置。對於其中屬於標準誤之部分，JCGM 以 law of uncertainty of propagation (LUP)之概念處理組合式訊號輸出之聯合標準誤(associated standard uncertainty)，但是對於輸出型態之不確定性可能影響平均值和覆蓋區間之估計卻未提出較佳之克服方法。故本研究之主要範圍在於界定稀少取樣資料下之組合式訊號以最小估計誤差前提下之測不準現象最佳表示模型。

JCGM 所遺留下的基本問題在於組合訊號之平均值使用算術平均數(sample mean)計算，覆蓋區間則只能針對近似對稱之分布進行計算。有鑒於此，本研究針對JCGM於組合式訊號之量測問題所留下之難題提出可行的解決方式，並且以Monte Carlo method進行驗證提出以下幾種量測表示之優化解決方法：(1)首先確認組合式訊號之輸出型態為一近似常態分布之窗型機率密度函數型態(quasi-normal signals with asymptotic window-shape distribution, QSAW)；(2)提出適合所有分布型態之覆蓋區間之*pdf*表示式，以*pdf*解釋偏態母體中所定義之the probably shortest CI在asymptotically symmetry *pdf*下就是the shortest CI；(3)將覆蓋區間之意義延伸至統計覆蓋區間(statistical coverage interval)，並且以 truncated

normal 機率密度函數為基礎之聯合機率密度函數(variably truncated normal joint probability density function)模擬統計覆蓋區間，並進而估計組合訊號之平均值；(4)在以quantile為基礎之前提下，提出非線性quantile estimation之方法，藉以改良對QSAW組合式訊號之平均值估計；(5)運用使用於強健式統計法的“the asymptotic minimax principle”來改進對QSAW訊號之平均值估計；(6)使用the quantile mapping invariance (QMI) principle來增進quantile-based平均值估計器之效能，並將其應用至由取樣訊號所估計之相關矩陣訊號求取eigenvalue上限之問題。

實驗證明本研究所提出之嶄新數學架構模型可以完美補強JCGM在組合訊號中之描述不足部分。

An Efficient Representation of Uncertainty Measurement for Combined Signals on Small Sampling Size Condition and its Application to Signal Mean Estimation

Student: Wen-Hui Lo Advisor: Dr. Sin-Horng Chen

Institute of Communication Engineering, National Chiao Tung University
Hsinchu, Taiwan, Republic of China

Abstract

In many signal processing applications, to measure and represent combined signals is a necessary and essential work because it is generally difficult to obtain individual components of a combined signal. So far, there are only few attempts on analyzing the measurement and/or representing the uncertainty of some special combined signals. JCGM (the Joint Committee for Guides in Metrology) coordinated the publication of measurement standard since 1995 and followed the GUM's (Guide to the Expression of Uncertainty in Measurement) suggestion to publish a standard, JCGM 101, to outline the representation of combined signals by an additive model which models a combined signal as the result of the propagation of different input source signals. The suggested format of JCGM 101 includes the following four items: mean, standard uncertainty, coverage interval (CI) and its two endpoints. The JCGM standard uses the law of uncertainty of propagation to evaluate the associated standard uncertainty of a combined signal. But it does not provide the way to explore the effects of the output uncertainty on mean and coverage interval estimations. This motivates us in this study to exploit the optimal representation of the uncertainty of combined signals based on the minimal estimation error criterion under small sample size condition.

One basic problem of the JCGM standard is the use of sample mean to estimate the mean of a combined signal. It therefore neglects the uncertainty resulted from the rough mean estimation when the sample size is small. Another problem is that it evaluates the coverage interval based on the assumption of asymptotically symmetric distribution. This study proposes several approaches to attacking these problems and examines them by the Monte Carlo simulations. Items studied include: (1) We verify that the output of a combined signal distributes like a quasi-normal signal with asymptotic window-shape distribution (QSAW). (2) We derive a unified probability density function (*pdf*) for CI to eliminate the need of skewness recognition before the evaluation of CI. (3) We extend the CI representation to the statistical CI representation and form the variably truncated normal joint probability density function. A robust quantile-based mean estimator is accordingly proposed. (4) We try a nonlinear modification of the proposed quantile-based mean estimator and verify its robustness with specially focusing on the case when the *pdf* of the combined signal approximates a rectangular *pdf*. (5) We follow the robust statistical method using "the asymptotic minimax principle" to refine the sample mean. (6) We employ the quantile mapping invariance (QMI) principle to improve the efficiency of the quantile-based mean estimator and apply it to the task of finding the upper bound of eigenvalues from the correlation matrix calculated from sparse observed samples.

We believe that the proposed unified representation of CI and its application to the quantile-based mean estimation are very promising and can contribute to extend the usage of the JCGM standard.

致謝

一件事情的完成往往都有幕後的功臣，這篇論文亦不例外。我感謝父親羅靖南先生從小指引出這個明確的目標期待我來完成，雖然他沒能親眼看到我完成博士論文，但是我會決定攻讀博士學位的確是受他的影響很大。另外也要感謝母親彭怡敏女士這些年來對於日常生活起居無微不至的照顧，讓我能夠健康快樂的走過人生旅途中重要的關卡。

受業期間指導教授 陳信宏博士悉心指導與幫忙亦由衷表示謝意，因爲他給予自由的研究環境使得我的創造力在沒有外力干預的情況下得以開發，最後僥倖能夠在量測的領域中留下些微的足跡。

另外亦感謝王逸如教授在博士生涯期間對於觀念引導分析之指正、振宇及智合兩位學弟經常在個人事項處理上給予幫助，亦表示感謝。阿德和 Barking 及希群平日在博士班研究議題上亦經常提供議題切磋在此一併表示感謝之意。碩士班學妹妞妞、舒舒和 Puma 常在我疲勞之虞解悶亦爲研究生涯中注入生活泉源。宥余、皓翔、承燁、財祿、文良、啓全、宜樵.....等學弟共同學習回憶亦令人陶醉。

博士論文口試期間獲得清華大學電機系王小川教授、暨南國際大學電機系魏學文教授及交通大學電信工程研究所蘇育德教授、唐振寰教授及王蒞君教授之不吝指正，亦在此一併致謝。

1896

ACRONYMS

BLUE	the best linear unbiased estimation
cdf	cumulative distribution function
CI	coverage interval
CLT	central limit theorem
GLI	Gauss Legendre integration
i.i.d	independent and identical distribution
MLE	maximum likelihood estimation
MSE	mean square error
MMSE	minimum mean square error
MLL	marginal log likelihood
pdf	probability density function
QMLE	quantile-based maximum likelihood estimation
QSQ	quasi symmetric quantiles
UBE	upper bound of the eigenvalues
VTNJ *pdf*	variably truncated normal joint *pdf*

NOTATION

$p_?(.)$	pdf or conditional pdf for a certain variable
$\Pr(.)$	probability
x	random variable of normal distribution

$f_?(.)$	*pdf* of a certain random variable
$f_x(.)$	*pdf* of the normal population of random variable x with mean u and standard deviation σ; i.e., $f_x(x)=N(u,\sigma^2)$
$F_x(.)$	*cdf* of the normal population of random variable x; $F_x(x)=\int_{-\infty}^{x} f_x(y)dy$
u	population mean
σ	standard deviation of population
$x_{i:n}, 1\le i\le n$	the ranked random variable resulting from sorting the samples of x
n	sample size
$u[0,1]$	standard uniform distribution in $[0,1]$
ξ	random sequence of the standard normal distribution
$\xi_{i:n}, 1\le i\le n$	order statistics random variable generated from the ranked random variable ξ of the standard normal *pdf*
X_n	random sequence of length n
$\mathrm{E}_?[.]$ or $\mathrm{E}_{(?)}[.]$	expectation operator
$Cov[\cdot,\cdot]$	covariance operator
$Min[.]$	take the minimum value in set
I	identity vector
B	covariance matrix
L	likelihood
r	range
c	coverage

$U(.)$	unit step function
$Z(Cc_t, n)$	normalized factor for the fixed coverage point Cc_t, t is the sampling index for Gauss-Legendre Integration
η_j	root of the Hermite polynomials expanded coverage the order of Hermite polynomials
$[a, b]$	the interval for interval estimation of coverage
$w_{Hm}(\gamma_i)$	the roots of the i-th Hermite polynomial
$P_v(.)$	the v-th Legendre polynomial
r_s	the random variable of range on standard normal *pdf*
$\Phi(.)$	*cdf* of standard normal distribution
$\overline{x}$	if no emphasis, it is the sample mean or average of the truncated data
$\overline{x^2}$	mean of square
$w_{P_v}(\kappa_t)$	the weighting coefficient of the t-th root of the v th order Legendre polynomial

Contents

Abstract iii

致謝 v

ACRONYMS vi

NOTATION vi

Contents ix

List of Tables xii

List of Figures xii

Chapter 1: Introduction 1

1.1 Motivation 1

1.2 Stating the Function for Coverage Interval 2

1.3 Goal and Scope 3

Chapter 2: Paper Review 6

2.1 Coverage Interval 6

2.2 Mean Estimation 8

2.3 The Method Suggested by Cohen 10

2.4 Sample Mean Estimator 12

2.5 Quantizing the Combined Signal 14

2.6 The Issue of Application to the Finding of UBE 16

Chapter 3: The Probability Density Function of Coverage Interval 19

3.1 Introduction 19

3.2 A New Method to Formulate the *pdf* of CI 20

3.3 Evaluation of the *pdf* of CI by Simulations 24

3.4 A Realization of the Statistical CI 29

3.5 Extension of Statistical CI to the VTNJ *pdf* 32

Chapter 4 The Analytical Mean Estimator for Truncated Normal Distribution on Sparse Data Condition 36

4.1 Introduction 36

4.2 The Proposed Method 36

4.3 Standard Normal Transform for the VTNJ *pdf* Computation 39

4.3.1 Derive the Variably Truncated Normal Joint Distribution Estimator (VTNJE) 40

4.3.2 Algebraic Closed From for Parameter Estimation 41

4.4 Experiments 42

4.4.1 Test the Results with Consistency to Sample Mean under the QMI Principle—Case of the Default Percentile 42

4.4.2 Test the Results with Consistency to Sample Mean under the QMI Principle—Case of Realistic Percentile 44

4.4.3 Comparison of the Different Estimators 44

4.5 Conclusions 46

Chapter 5: The VTNJ Estimator Tested with the Combined Signals 47

5.1 Introduction 47

5.2 Robust Interval Detection for Small Sample Size 48

5.2.1 Test VTNJE for Combined Quantities 50

5.2.2 Test VTNJE for Different Uncertainty Ratio 52

5.3 Conclusions 53

5.4 Appendix: Derivation a Closed Form for VTNJE 53

Chapter 6 The Asymptotic Minimax Optimization for Mean Estimation of Combined Signals 58

6.1 Establish the Minimax Structure 59

6.1.1 Convex Set 59

6.1.2 Asymptotic Efficient near the Minimal Average of MSEs..........61

6.1.3 Minimax Structure for the Objective Function 61

6.2 QMLE optimization on MMSE of the Two Endpoints of Range, (QSQ) 62

6.2.1 Test the Robustness of Q2MMSE-CLT for Different Uncertainty Ratio 63

6.2.2 An Advanced Refinement of the QMLE..........64

6.3 Change the Variable to Obtain a Nonlinear Estimator for Mean Estimation..........67

6.4 Conclusions..........69

6.5 Appendix: Derivation of the Quantile-based Mean Estimator..........69

Chapter 7 An Efficient Representation for Combined Signal Activity Detection in Sparse Data Condition 71

7.1 The Simplified Quantile-based Mean Estimator 71

7.2 Simulation Results for UBE Finding with QMLE-QSQ Mean Estimation..........76

7.3 Conclusions..........78

Chapter 7 Conclusions and Further Works..........79

Publication List 81

List of Tables

Table 1: the multi-root and single-root solutions of the coverage constraint equation26

Table 2: Some *cdf* values of CI for single-root representation, multi-root representation and realistic case by simulation 27

Table 3: Single-root solutions using different initial conditions 28

Table 4: The confidence level $1-\alpha$ of two integration intervals for different sample sizes 32

Table 5：Performance of realistic QMI analysis 44

Table 6: Comparison with different estimators in association with bad initial searching points (Unit: MSE)...45

Table 7: Comparison with different estimators in association with good initial searching points (Unit: MSE)...45

Table 8: Computation results for the uncertainty ratio, UR =1.5, with 1,000 trials, normalized by $u_c^2(x)/n$, 4 mixing signals (Unit: Normalized MSE) 50

Table 9: Table of confusion for the conditions of combined mean and combined standard uncertainty...62

List of Figures

Fig. 1: Measurement is the front stage of signal processing for quantifying data recognized...3

Fig. 2: Un-determined properties of the output *pdf* resulting from combining different input *pdf*s. 4

Fig. 3: This study reverses the traditional direction for CI estimation respect to the asymptotically symmetry *pdf* and further extending CI for mean estimation 5

Fig. 4: An example of the QSAW signal with zero-mean R*N distribution for some uncertainty ratio (UR)...15

Fig. 5: Profile-conditional *pdf* by the sampling strategy. k is a constant................21

Fig. 6: A conceptual diagram shows the use of the bisection method to establish the two endpoints for CI. Here, $[a_{1,}b_1]$ is an effective interval for root finding and c represents the center (midpoint) of any new effective interval. 23

Fig. 7: The *pdf* (top) and *cdf* (buttom) of CI for normal random variable with experiment setting of coverage=0.95 and sample size=15. 24

Fig. 8: Histogram of CI generated by simulation using 10,000 trials. The experiment setting is coverage=0.95 and sample size=15. 25

Fig. 9: The multi-root and single-root representations of the *pdf* and *cdf* of CI simulated using the standard normal distribution of input with sample size $n = 15$, and coverage=0.95. 27

Fig. 10: The *cdf*s of CI using single-root representation, multi-root representation, and simulations 28

Fig. 11: Uniform *pdf* for the random variable s 30

Fig. 12: The *pdf* of coverage for some small sample size n 31

Fig. 13: Relation of variables' interference model 37

Fig. 14: Relative quantile mapping invariance based on their percentiles. Dash-line represents the original normal *pdf* and solid-line represents the standard normal *pdf*. 40

Fig. 15: Comparison of the conventional sample mean estimator and two coverage-based mean estimators. 43

Fig. 16: Experimental results for VTNJE using two output patterns of UR=3.7 and 10.4. Left: the distributions of output quantities. Right: MSEs normalized to $u_c^2(x)/n$ 51

Fig. 17: Performance comparison for six estimators using the first output pattern with UR= 3.7. 52

Fig. 18: Performance comparison for four estimators using the output quantity of 4-mixture combined quantities with different UR. The average MSE is normalized to $u_c^2(x)/n$ 53

Fig. 19: Standard normal *pdf* combined with its CLT *pdf* and QSQ *pdf* for sample size=11. We plot the QSQ of equal variance rectangular *pdf* $[-\sqrt{3},\sqrt{3}]$ as the blue solid line. 59

Fig. 20: MSE of QMLE versus difference=$\left(\mu_{ps}-\mu\right)$ for three normal distributions. Note that $\xi_{1:n}$ is calculated using true standard deviation σ 60

Fig. 21: Conditions A and C. y axis is normalized to $u_c^2(x)/n$ 63

Fig. 22: Conditions B and D. y axis normalized to $u_c^2(x)/n$ 63

Fig. 23: Histogram of 50,000 combined quantities for different URs. x-axis is the output of combined quantities and y-axis is the frequency count 64

Fig. 24: Gain performance for the different URs. The unit is $u_c^2(x)/n$ 64

Fig. 25: Good sample mean tested with the convex sets, normalized by $u_c^2(x)/n$, sample size is 15, 4 combined quantities...65

Fig. 26: Originally left biased of bad sample mean tested with the convex sets, double y-axes, normalized by $u_c^2(x)/n$, sample size is 15, 4 combined quantities.......66

Fig. 27: Originally right biased of bad sample mean tested with the convex sets, double y-axes, normalized by $u_c^2(x)/n$, sample size is 15, 4 combined quantities...67

Fig. 28: Refined Q2MMSE-CLT with the enlarging standard uncertainty, y-axis is normalized by $u_c^2(x)/n$, 4 combined quantities...67

Fig. 29: Nonlinear estimators compared to the sample mean estimator for different uncertainty ratio (UR), y axis is normalized to the combined standard uncertainty, $u_c^2(x)/n$ 69

Fig. 30: The *pdf*s of QSQ of the MP pair (i.e., normal and rectangular distributions) for the case of μ=0, σ=1, and n=11. Dots on x-axis are expectation values of QSQ. 73

Fig. 31: The performance of *QMLE-QSQ* mean estimation for different degree of deviation of the transform-domain quantile used to the true one. Note that y axis is normalized by $u_c^2(x)/n$. 75

Fig. 32: The mean estimation performance of *QMLE-QSQ* using two different quantile-expectation (QE) replacement schemes for two QSAW signals: (a) UR=3.7 and (b) UR=10.4. The y axis is normalized by $u_c^2(x)/n$. Note that the MSEs of *QMLE-QSQ* and sample mean are $\frac{n}{10000\cdot u_c^2(x)}\sum_{i=1}^{10000}(\frac{\mu_1^*(i)+\mu_n^*(i)}{2}-\mu)^2$ and $\frac{n}{10000\cdot u_c^2(x)}\sum_{i=1}^{10000}(\overline{\mu}-\mu)^2$, respectively..76

Fig. 33: The performance of UBE finding using the proposed QMLE-QSQ mean estimator 78

Chapter 1: Introduction

1.1 Motivation

The observation of nature which constitutes an experiment will almost inevitably take the form of a measurement. Measurement is represented as the precision type related as whether the experiment is effective, or in the other words, how much is taken about its confidence corresponding to the experiment.

Does the measurement merely have the purpose of standing for a qualitative conclusion? Such a question causes the focus of the meaning of any experiment whenever it is significant not only for someone's special idea but also lay themselves open to all the frailties of human judgments. That is, confidence report is needed in the formal measurement report. According to the requirement of duplicate verifications for the results of any new approach, the workers expect to convey the experimental results to someone else based on the condition of laboratory or field testing invariantly so that the level of confidence must be also included in the measurement task. Besides, confidence plays the key role to support whether to accept the other's report so as to avoid performing a duplicate experiment. Thus the center problem for measurement task includes showing the confidence level about the results.

The best qualification of measurement is admitted as a statement of the result of human's observations with high confidence. Because of this fundamental role of measurement it is necessary to consider in some detail what a measurement practically is. That is, how much confidence does we believe in the observations? Why does the measurement task pay attention to the confidence factor associated with the practical experiment? According to the scientific revolution, we think that the "uncertainty principle" brings the reason for any measurement event, especially in micro-electronics ones. For the reason to overcome the uncertainty representation, the Physics Laboratory of National Institute of Standards and Technology (NIST) conducts the standards and measurement method for electronic, optical and radiation technology for US. and takes the general Type A or Type B expression as the report for measurement task.

NIST keeps the policy based on the approach to expressing uncertainty in measurement recommended by the CIPM and the evaluation given in the Guide to the Expression of Uncertainty in Measurement (GUM), which was prepared by individuals nominated by the BIPM, IEC, ISO, or OIML. GUM is the most authorized reference on the general application to express measurement uncertainty till 1995. After that time, the Joint Committee for Guides in Metrology (JCGM) collects the above document and releases the new methods and standards for measurement. Thus this study will keep the work to follow the document published by JCGM as the reference. Although JCGM spent a long time for the general expression of uncertainty measurement, there are some occasions not included for practical applications and we focus on those which measure the combined signals on sparse data condition.

1.2 Stating the Function for Coverage Interval

Signal processing is a basic technique to process the sensoring signal and further sends the processed signal to the next stage or outputs it. In addition to choose a proper singal processing technique, we also need some other tools to check the properties of the input signal, such as coverage interval (CI), normal range, and reference interval, in order to determine whether the input signal is quantified to take the utility. CI is the predicative interval including a measured random quantity based on a pre-specifyied proportion of population. It is frequently applied to the cases with normal population assumption where they take the minimal CI to replace all other possible values of CI. The principal function of CI is to state the confidence and uncertainty about the measured quantities. It defines the prediction interval of values where 95% of the population fall into as suggested by JCGM [33]. For instances, we may reject the outlier data from the measured signal if the data are away from the mean value grater than 2 times of the standard deviation. A risk representation can also be applied by the way of CI to make a reject decision on sampling data if its value is out of the CI extent.

CIPM: International Committee for Weights and Measures

BIPM: International Bureau of Weights and Measures

IEC: International Electrotechnical Commission

ISO: International Organization for Standardization

OIML: International Organization of Legal Metrology

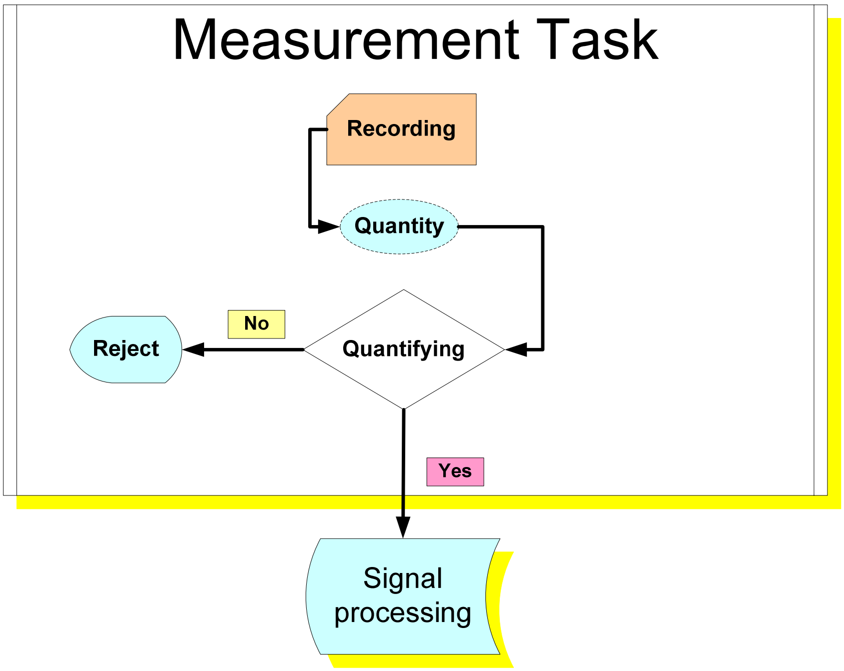

Fig. 1: Measurement is the front stage of signal processing for quantifying data recognized

Although CI is used as the standard item for the JCGM format of measurement tasks, there still have some shortcomings not being overcomed so far. The most commomly encountered problem is that CI is usually evaluated based on the assumption that the population has a asymptotically symmetric *pdf*, but we know this is not always appropriate, especially at the occasion of combined signal. The other CI computation method is the non-parametric method which is constructed basing on the percentile evaluated by the expectation of order statistics [1]; that is, we may take the quantile mapping to the corresponding percentile as the desired endpoint. The main difficulty of using CI for combined signal is that we don't know whether the symmetry property of the output signal is valid when applying the CI computng algorithm. There are still other statistical techniques, such as logarithm transform and Box-Cox transformation, suggested for enhancing the symmetry properties of the analyzed signal and the outliers examining are also necessary.

1.3 Goal and Scope

Combined signal is one of the most popular measured signals for the practical usage and is widely applied to the field test as well as to the industry production. In GUM, a combined signal is represented by an additive model in which the *pdf* of the output signal is modeled as the result of a propagation of input *pdf*s. Some special areas concern the measurement task of combined signals and treat it as an integration of the affecting factors caused from the environment.

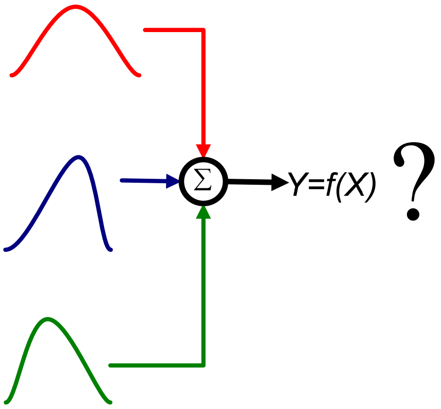

Fig. 2: Un-determined properties of the output *pdf* resulting from combining different input *pdf*s.

In accordance with the report expression of JCGM 101, coverage interval (CI) with its two endpoints, mean value and standard uncertainty are the three members of its main concern. They are also the main concern of this study. Due to the fact that the output of combined signal is random, we think the best description for CI representation is to formulate its *pdf*. Issues addressed are briefing as follows. First, we are interested in the formulation to unify the CI representations for skew and non-skew *pdf*s. Conventionally, different approaches are employed for these two types of *pdfs* to calculate their respective CI. Besides, we are also interested in the truncated probability density function normalized to its coverage. The non-skew, asymptotically symmetric *pdf* draws our special attention because it is the typical output *pdf* shape of combined signals. Moreover, the usual evaluation of asymptotically symmetric CI involves the interval composing of an upper quantile (half coverage) and a lower quantile (half coverage) with respect to the mean value. A robust CI estimation needs accurate quantile and mean formulations. This is the rule followed in the past studies so is the current study. There are a few exceptions to the rule. One is that we can consider giving a robust CI before the mean estimation, and this may leads to a good performance for mean estimation. A study will hence be conducted to try to use the traditional coverage interval to assist in the mean estimation. The issue is that if we are giving a more accurate coverage interval, can we make some progress on improving the mean estimation? Besides, we will introduce three new approaches of mean estimation and compare them with the classical sample mean estimator. They include a quantile-based mean estimator using the coverage interval, a nonlinear mean estimator and a robust statistical one using the minimax principle. Lastly, we will shape the proposed quantile-based mean estimator to a quasi-symmetric quantile-based one and use it in an application to find the upper bound of the maximum eigenvalue (UBE), to examine the usage of the robust JCGM expression in measurement.

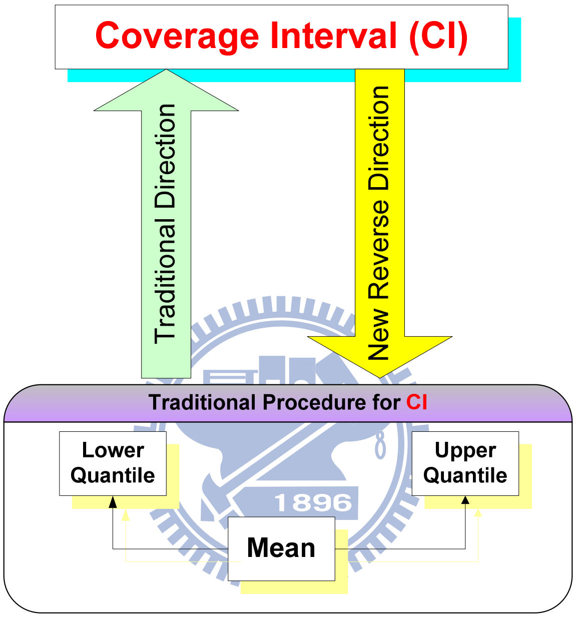

Fig. 3: This study reverses the traditional direction for CI estimation respect to the asymptotically symmetry *pdf* and further extending CI for mean estimation

Chapter 2: Paper Review

We review some literatures related to the three main topics discussed in the dissertation. They include coverage interval which is a member of JCGM expression, mean estimator, and finding UBE, which is an application of mean estimation. The sampling size requirement will be especially concerned in the following discussions.

2.1 Coverage Interval

Coverage interval (CI) is originally regarded as a parameter to represent the uncertainty of measurement. Fotowicz [2] proposed an analytic method to calculate CI from the distribution of the output of combined quantities, formed by taking the convolutions of the *pdf*s of its constituents which were assumed to be rectangular mixing with one of Student's t-, triangular, or normal distributions. It made some progress in the realization of CI without using complex numerical computations. Nadarajah [3] continued to extend the algorithm and applied it to a wide range of usage with higher degree of freedom. In those studies, CI was always used as a confidence measurement in the sampling plan.

CI is affected by coverage constraint realistically. If we turn to a different viewpoint relating to the coverage problem, the "statistical CI" is also a good tool to describe the uncertainty. Wilks [4] proposed a statistical CI, defined by

$$Pr\{p_x[(T_1,T_2)]\geq\beta\}\geq 1-\alpha\ , \qquad (2\text{-}1)$$

to describe the probability that a random variable x includes a β-content proportion of the population or more in the interval $[T_1,T_2]$ is greater than the threshold $1-\alpha$. The statistical CI has been proved to represent a certain confidence level [45]. In those past studies, the confidence level was usually obtained by the Monte Carlo simulations [5,6].There were some previous studies concerning the issue of randomness of coverage. The early topic was called "the random division of an interval", which means the range may be cut as many small sub-ranges which can be added to calculate the coverage [7,8].

The representation of CI can be categorized into two classes: parametric and non-parametric CI. Lin et al. [9] suggested using a non-parametric formulation to calculate CI when the population *pdf* is unknown. Chen [47] suggested that, while adopting the parametric CI approach, it had better take the minimum of all possible values of CI for computational simplification. In the past, CI was mainly applied to the cases of resource constrained for the original population. For instances, a clinical chemistry experiment first applies tests to healthy people to create a CI, and then takes the same test to a patient and collects the outputs. If the outputs are out of the CI, it implies that the patient has got a disease. In medical engineering, to collect large samples containing all the records of patients is a time-consuming task so that we should sometimes take a sampling plan of small sample size. Thus data sparseness is inevitable in this kind of application because the process of collecting data is time-consuming and expensive.

The use of CI is popular for the chemical substances in biological fluid for reference population [10], and for some other related fields of measurement. The International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) [11] has published a series of recommendations for the advanced utilizations of CI. IFCC defined the percentile between 0.025~0.975 as the standard CI of 95% reference interval, and suggested that the best population (reference values) size had better be greater than 120 so that a high confidence reliability can be guaranteed. IFCC made more rules and standards for the reference interval estimation and computation, but without further addressing the issue of the influence of sample size. This study discusses the CI problem concerning the size of sample data and deals with how to control the categories of influences if the sample size is far less than 120. We will take a new viewpoint to analyze the effect of sample size on CI. Actually, it is not necessary to formulate CI from the viewpoint of the aggregation method. If we evaluate the two endpoints of CI separately, we may consider estimating CI with the quantiles based on order statistics. The quantile-based estimator [12] was recently proposed by Heathcote et al. It performed very well for the response time estimation and showed high efficiency to the parameter estimation for some distributions.

2.2 Mean Estimation

The second topic we are interested is the mean estimation of population. We will try to use CI in the mean estimation basing on the asymptotically symmetric *pdf* assumption. In this case, mean is the midpoint of the two endpoints of CI and we truly believe that a more accurate estimation for CI will lead to a more accurate estimation for mean value.

In parameter estimation of using normally-distributed sparse data, there are two popular methods: the best linear unbiased estimation (BLUE) method and the maximum likelihood estimation (MLE) method. Balarkrishnan and Cohen [13], Lloyd [14], and Teichroew [15] proposed the BLUE method for parameter estimation of normal random variables by using order statistics. BLUE is a weighted least-square algorithm basing on the Gauss-Markov least-square theorem. It was popularly used for sparse data analysis. It is known that BLUE is unbiased and more efficient if it takes the censoring sampling scheme. We briefly discuss BLUE as follows.

Let x be a normal random variable with *pdf* $f_x(x) = N(u, \sigma^2)$. Assume that there are n independent observed samples $x_1, \cdots, x_n$ of x. Let $x_{1:n}, \cdots, x_{n:n}$ be the ranked samples of $x_1, \cdots, x_n$ in increasing order. The BLUE estimator is formulated as the sum of products of the observed data and properly-chosen coefficients. By performing the standard normal transformation, $\xi_i = (x_i - u)/\sigma$, to the observed data and sorting them in increasing order, we have

1896

$$X_n = [x_1, \cdots, x_n]^T$$

$$\xi = [\xi_1, \cdots, \xi_n]^T$$

$$E\{\xi_{i:n}\} = \rho_{i:n}$$

$$Cov\{\xi_{i:n}, \xi_{j:n}\} = \beta_{i,j:n}$$

$$\begin{aligned} E\{x_{i:n}\} &= u + \sigma\xi_{i:n} \\ E\{X_n\} &= u\mathrm{I} + \sigma\xi \end{aligned} \qquad (2\text{-}2)$$

$$\mathrm{I}_n = [1, \cdots, 1]^T_{n\times 1}$$

$$B = \sigma^2 I \tag{2-3}$$

for $1 \le i, j \le n$ and $i < j$, where I_n is an n-dimensional all-1 vector. Consider the generalized variance:

$$\left(X_n - u\mathrm{I}_n - \sigma\xi\right)^T B^{-1} \left(X_n - u\mathrm{I}_n - \sigma\xi\right) \tag{2-4}$$

Minimizing it with respect to u and σ, we obtain.

$$\begin{aligned} u\mathrm{I}_n^T B^{-1}\mathrm{I}_n + \sigma\mathrm{I}_n^T B^{-1}\xi = \mathrm{I}_n^T B^{-1} X_n \\ u\xi^T B^{-1}\mathrm{I}_n + \sigma\xi^T B^{-1}\xi = \xi^T B^{-1} X_n \end{aligned} \tag{2-5}$$

The solution of Eq.(2-5) is

$$u^* = \left\{ \frac{\xi^T B^{-1}\xi\mathrm{I}_n^T B^{-1} - \xi^T B^{-1}\mathrm{I}_n\xi^T B^{-1}}{(\xi^T B^{-1}\xi)(\mathrm{I}_n^T B^{-1}\mathrm{I}_n) - (\xi^T B^{-1}\mathrm{I}_n)^2} \right\} X_n = -\xi^T \Delta X_n = \sum_{i=1}^{n} \alpha_{1:i} x_{i:n} \tag{2-6}$$

$$\sigma^* = \frac{\mathrm{I}_n^T B^{-1}\mathrm{I}_n\xi^T B^{-1} - \mathrm{I}_n^T B^{-1}\xi\mathrm{I}_n^T B^{-1}}{(\xi^T B^{-1}\xi)(\mathrm{I}_n^T B^{-1}\mathrm{I}_n) - (\xi^T B^{-1}\mathrm{I}_n)^2} X_n = \mathrm{I}_n^T \Delta X_n = \sum_{i=1}^{n} \alpha_{2:i} x_{i:n} \tag{2-7}$$

where u^* and σ^* are the estimated parameters, and $\alpha_{1:i}$ and $\alpha_{2:i}$ are weighting coefficients. These coefficients have been tabulated by Sarhan and Greenberg [16,17] with entries in the 1956 tables being given for sample size up to 10 and in 1962 up to 20.

Generally speaking, BLUE performs well in small sample size. But it needs a table to look up, and this is a shortcoming. The other technique used is the MLE method which is often applied to the truncated normal distribution in sparse data condition. Cohen [54] derived the singly truncated and doubly truncated maximum likelihood estimators and found that they outperformed BLUE when the sample size was greater than 20. Cohen recognized the sparse data problem as a truncated normal *pdf* and defined its likelihood by

$$L = \left(\frac{UnitStep(x - x_{1:n}) - UnitStep(x - x_{1:n} - r)}{\sqrt{2\pi}\sigma(F_x(x_{1:n} + r) - F_x(x_{1:n}))} \right)^n \exp\left(-\sum_{i=1}^{n} \frac{(x_i - u)^2}{2\sigma^2}\right) \tag{2-8}$$

If we take the transformations of $\xi_{1:n} = (x_{1:n} - u)/\sigma$ and $\xi_{n:n} = (x_{n:n} - u)/\sigma$ and differentiate the resulting log-likelihood function with respect to u and σ, we obtain the following two equations.

$$\frac{n(\phi_{\xi}(\xi_{1:n})-\phi_{x}(\xi_{n:n}))}{\sigma(\Phi_{\xi}(\xi_{n:n})-\Phi_{\xi}(\xi_{1:n}))}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-u)$$

$$\sigma^2\left\{\frac{(\xi_{1:n}\phi_{\xi}(\xi_{1:n})-(\xi_{n:n})\phi_{\xi}(\xi_{n:n}))}{(\Phi_{\xi}(\xi_{n:n})-\Phi_{\xi}(\xi_{1:n}))}+1\right\}=\frac{1}{n}\sum_{i=1}^{n}(x_i-u)^2 \tag{2-9}$$

where ϕ and Φ are the standard normal *pdf* and *cdf*, respectively. By defining two new random variables

$$\Theta_L=\frac{\phi_{\xi}(\xi_{1:n})}{\Phi_{\xi}(\xi_{1:n}+r_s)-\Phi_{\xi}(\xi_{1:n})}$$

and

$$\Theta_R=\frac{\phi_{\xi}(\xi_{1:n}+r_s)}{\Phi_{\xi}(\xi_{1:n}+r_s)-\Phi_{\xi}(\xi_{1:n})},$$

we obtain the following two equations

$$H_1(\xi_{1:n},\xi_{n:n})=\frac{\overline{x}-x_{1:n}}{r}=\frac{\Theta_L-\Theta_R-\xi_{1:n}}{\xi_{n:n}-\xi_{1:n}} \tag{2-10}$$

$$H_2(\xi_{1:n},\xi_{n:n})\Rightarrow\frac{S^2}{r^2}=\frac{1+\xi_{1:n}\Theta_L-\xi_2\Theta_R-(\Theta_L-\Theta_R)^2}{(\xi_{n:n}-\xi_{1:n})^2} \tag{2-11}$$

2.3 The Method Suggested by Cohen

Cohen proposed a method to estimate mean and variance of normally distributed random variable. Let $\xi_L=\frac{x_L-u}{\sigma}$ and $\xi_R=\frac{x_R-u}{\sigma}$, where x_L and x_R are the left and right truncation points, respectively. The standard deviation can be estimated by:

$$\sigma=\frac{x_R-x_L}{\xi_R-\xi_L}.$$

The method first models all data samples by a truncated normal distribution shown below:

$$f_T(x)\Rightarrow\frac{f_x(x)UnitStep(x-x_{1:n})-UnitStep(x-x_{1:n}-r)}{F_x(x_{1:n}+r)-F_x(x_{1:n})} \tag{2-12}$$

where the left truncation point x_L is replaced by the minimum order random variable $x_{1:n}$ and so is to the right truncation point x_R replaced by $x_{n,n}$. It then defines a likelihood function by

$$L(x;u,\sigma,x_{1:n},r)=\prod_{i=1}^{n}(f_T(x_i;u,\sigma,x_{1:n},r)) \tag{2-13}$$

By taking $\frac{\partial}{\partial u}\{L(x;u,\sigma,x_{1:n},r)\}=0$, it obtains

$$\frac{n(f_x(x_{1:n})-f_x(x_{1:n}+r))}{\sigma(F_x(x_{1:n}+r)-F_x(x_{1:n}))}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-u) \tag{2-14}$$

It then takes the standard normal transformations for the two endpoints of ranked samples, $x_{1:n}$ and $x_{n:n}$, to obtain

$$\xi_{1:n}=\frac{x_{1:n}-u}{\sigma} \text{ and } \xi_{n:n}=\frac{x_{n:n}-u}{\sigma}.$$

The corresponding CI in the transform domain is $r_s=\xi_{n:n}-\xi_{1:n}$. It is noted that the *cdf*, $\Phi_\xi(\xi)$, of the transformed random variable is related to the *cdf*, $F_x(x)$, of the original random variable by $F_x(x_{1:n})=\Phi_\xi(\xi_{1:n})$ and $F_x(x_{n:n})=\Phi_\xi(\xi_{n:n})$. By denoting $\Theta_L=\frac{\phi_\xi(\xi_{1:n})}{\Phi_\xi(\xi_{1:n}+r_s)-\Phi_\xi(\xi_{1:n})}$ and $\Theta_R=\frac{\phi_\xi(\xi_{1:n}+r_s)}{\Phi_\xi(\xi_{1:n}+r_s)-\Phi_\xi(\xi_{1:n})}$, it has

$$\bar{x}-u=\sigma(\Theta_L-\Theta_R) \tag{2-15}$$

By taking $\frac{\partial}{\partial \sigma}\{L(x;u,\sigma,x_{1:n},r)\}=0$, it obtains

$$\frac{-n(x_{1:n}f_x(x_{1:n})-(x_{1:n}+r)f_x(x_{1:n}+r))}{\sigma(F_x(x_{1:n}+r)-F_x(x_{1:n}))}-\frac{n}{\sigma}+\frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-u)^2=0 \tag{2-16}$$

Eq.(2-16) can be further simplified and expressed by

$$\begin{gathered}\sigma^2\left\{\frac{(\xi_{1:n}\phi_\xi(\xi_{1:n})-(\xi_{n:n})\phi_\xi(\xi_{n:n}))}{(\Phi_\xi(\xi_{n:n})-\Phi_\xi(\xi_{1:n}))}+1\right\}=\frac{1}{n}\sum_{i=1}^{n}(x_i-u)^2 \\ \sigma^2\{\xi_1\Theta_L-\xi_2\Theta_R+1\}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2+\frac{1}{n}\sum_{i=1}^{n}(\bar{x}-u)^2=S^2+(\bar{x}-u)^2\end{gathered} \tag{2-17}$$

where $\bar{x}$ and S^2 are mean and variance of the data samples. If σ is known, the above equation can be solved by an iterative procedure.

If σ is unknown, Cohen suggested to solve the following two equations:

$$H_1(\xi_{1:n},\xi_{n:n})=\frac{\overline{x}-x_{1:n}}{r}=\frac{\Theta_L-\Theta_R-\xi_{1:n}}{\xi_{n:n}-\xi_{1:n}} \tag{2-18}$$

$$H_2(\xi_{1:n},\xi_{n:n})\Rightarrow\frac{S^2}{r^2}=\frac{1+\xi_{1:n}\Theta_L-\xi_{n:n}\Theta_R-(\Theta_L-\Theta_R)^2}{\left(\xi_{n:n}-\xi_{1:n}\right)^2} \tag{2-19}$$

where $w=r$, $v_1=\overline{x}-x_{1:n}$, and r is the range of the data samples. Eqs.(2-18) and (2-19) can be solved by the Newton and Raphson method. But, it is time-consuming unless good initial values are provided. Alternatively, Cohen [18] proposed the following iterative procedure to solve them:

$$\xi_{1:n}^{(i+1)}=\frac{-\xi_{n:n}^{(i+1)}\left\{\dfrac{\overline{x}-x_{1:n}}{r}\right\}+\left(\Theta_L^{(i)}-\Theta_R^{(i)}\right)}{\left(1-\dfrac{\overline{x}-x_{1:n}}{r}\right)} \tag{2-20}$$

$$\xi_{n:n}^{(i+1)}=A+Br\left(\overline{x}-x_{1:n}-r\right) \tag{2-21}$$

$$A=\left(\Theta_L^{(i)}-\Theta_R^{(i)}\right)$$

$$B=A+\frac{C+\sqrt{C^2+4\dfrac{S^2}{r^2}}}{2S^2}$$

$$C=A\frac{\overline{x}-x_{1:n}}{r}+\Theta_R^{(i)}$$

S : variance of the test sequence
r : range of the test sequence
i : iteration index

2.4 Sample Mean Estimator

In the past, sample mean is widely used in the mean value estimation for any signal no matter what its original *pdf* is. The main reason of using sample mean is that it is not only a uniformly minimum variance unbiased estimator (UMVUE) but also a random variable of the central limit theorem (CLT). In this study, we will propose a new mean estimator basing on the proposed CI representation and compare its performance with the traditional sample mean estimator [19]. Our study will specially focus on the mean value estimation problem for the output of combined quantities in the sparse data condition. Bowen [20] has pointed out that CLT may be explained as the sum of independent variables with the characteristic function formed by the product of the component characteristic functions. If we can ignore the unbiased requirement, there exist some biased estimators that outperform sample mean. Stearls [21] and Gleser [22] discussed a new approach to giving coefficients of variation of sample mean. Ashok et al. [23] further proposed a realistic method to adjust the coefficients of variation of sample mean to improve its performance.

Up till now, if we want to predict the mean value of combined quantities accurately, the only way is to take the sample mean on heavy observations. In practical applications of measurement, the basic volume required for one digit accuracy is 10^6 observations for 95% coverage interval [24]. If, there are not enough samples to support this rule, a medium- or small-size sampling plans should be taken. Besides, the good property of UMVUE for sample mean may be ineffective for the case of combined quantities which is of quasi-normal distribution. This is because the property of UMVUE is derived from the maximum likelihood estimation (MLE) on the basis of the normal *pdf* assumption.

In this dissertation, a new method of mean value estimation, referred to as the quantile-based maximum likelihood estimator (QMLE), is proposed. The classical application of quantiles is the general usage of empirical quantiles. Koenker and Bassett [25] extended the empirical quantiles to the regression quantiles, which is specially useful for predicting the bounding information. Gilchrist [26] collected many studies about the estimation, validation, and statistical regression with quantile models. In the single quantile application, Giorgi and Narduzzi [27] gave the quantile estimation for the self-similar process.

In the proposed QMLE, the quantiles are determined by the maximum percentage of population, i.e. coverage, so that it is composed of a couple of quasi-symmetric quantiles (QSQ). According to the past studies, the coverage-constrained quantiles will obey the properties of symmetric quantiles whose variances asymptotically approach to the Cramer-Rao lower bound [28]. The symmetric quantiles were described with strict definition given in [28]. But we treat them in a more flexible way as the ranked variables of the first ordered sample $x_{1:n}$ and the last ordered sample

$x_{n:n}$. Hence the QSQ we considered are both empirical and quasi-symmetric quantiles. Lo and Chen [29,30] also derived good quantile-based estimators for the sparse data condition. In this study, we plan to derive the QMLE basing on the order statistics and expect that it can support not only the concept of empirical quantiles but also the quasi-symmetric quantiles. Otherwise, we would still need a quantile function defined below to link quantiles and MLE

$$Q(p) \equiv \Pr(X \le x_p) = p \tag{2-22}$$

Here, the value x_p is called the *p*-quantile of population.

2.5 Quantizing the Combined Signal

Generally speaking, the measured quantities are affected by unknown noise so that they are always expressed in random representation. In the past studies, Fotowicz [2] suggested using “uncertainty ratio” to represent the combined signal comprising at least one input quantity with rectangular distribution. Suppose z_i, $1 \le i \le N$, are independent signals and c_i, $1 \le i \le N$, are corresponding weighting coefficients, then the linearly combined output x can be expressed by:

$$x = c_1 z_1 + c_2 z_2 + \cdots + c_N z_N . \tag{2-23}$$

The *pdf* of *x* is an R*N distribution which is the convolution of a rectangular distribution and a normal distribution, and can be expressed by:

$$f_{RN}(x) = \frac{1}{K_c 2\sqrt{6\pi}\,(UR)} \int_{x-\sqrt{3}(UR)}^{x+\sqrt{3}(UR)} e^{\frac{-t^2}{2}}\, dt , \tag{2-24}$$

where

$$UR = \frac{|Max[u_i(x)]|}{\sqrt{u_c^2(x) - Max[u_i(x)]^2}} ; \tag{2-25}$$

$u_c^2(x) = \sum_{i=1}^{N} c_i^2 \sigma^2(z_i)$ is the approximate variance of the combined signal; $\sigma(z_i)$ is the standard deviation of z_i; K_c is a normalization constant; and $u_i(x)$ is the standard deviation of the *i*-th input random variable which is subject to the rectangular distribution. The endpoint of *p*-quantile for the R*N distribution can be expressed by

$$U_p = \mu + k_{RN}\sqrt{\sum_{i=1}^{N}(\frac{t(v)}{k_N}u_i(y))^2}\ , \tag{2-26}$$

where μ is the population mean, $k_{RN} = \sqrt{\frac{3}{(UR)^2+1}}(1+(UR)-2\sqrt{(UR)(1-c)})$ is a coverage factor, c is a coverage, $t(v)$ is a quantile of Student's t-distribution, k_N is the corresponding quantile of coverage factor, e.g. k_N=1.96 for c=95%, and N is the number of input quantities. If the distribution of the i-th input random variable coincides to be a normal, rectangular, Student's t-, or triangular distribution, then $t(v)/k_N = 1$.

Fig. 4 displays the R*N distribution for UR = 1, 2, 3, and 4. According to Fig. 4, we describe the measured signal of combining quantities by additive mixture model as quasi-normal signals with asymptotic window-shape distribution (QSAW). A common property of QSAW signals is that they are usually distributed flatter than the normal *pdf* in the central part and then sharply decaying to zero at both ends. As shown in Fig. 4, the *pdf* of a QSAW signal looks like a normal distribution for small UR and a rectangular distribution for large UR.

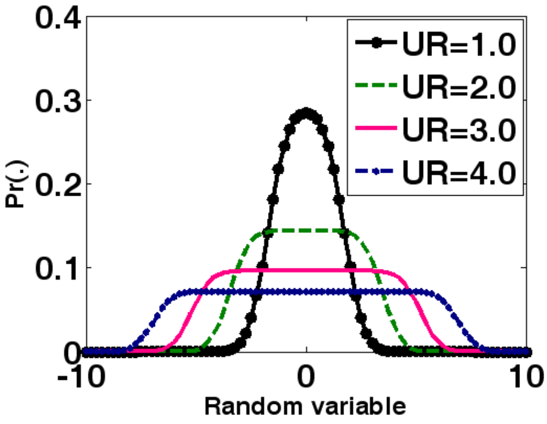

Fig. 4: An example of the QSAW signal with zero-mean R*N distribution for some uncertainty ratio (UR)

2.6 The Issue of Application to the Finding of UBE

In this study, we will consider the use of robust mean estimation in signal detection. Generally, the energy-based signal activity detection approach is robust to noise and may cost down the non-coherent detection within a communication receiver. Zeng et al. [31] showed the benefits of using the maximum eigenvalue as a result of energy representation on large sample size. Recently, compressive sampling (CS) [32] is an emerging research topic aiming at restoring a signal in an undersampled condition using special vector bases with prior knowledge of the signal. In addition to CS, eigen-analysis is also a popular technique to consider spanning a signal with sparse eigenvectors in which the prior knowledge of needing signal to be normally distributed is released. We will not only consider the combination of energy detection and sparse data sampling, but also fuse the demand of practical signal processing. For instance, measuring signal in a time-varying environment usually results in representing the measured signal as the output of combining quantities by an additive mixture model, as suggested and outlined in the manual published by JCGM [33]. Moreover, the combined quantities are usually resulted from the propagations of multi-source signals with different *pdf*s so that the representation for the *pdf* of the output random variable is not tractable.

Unexpectedly, the *pdf* of the maximum eigenvalue is too complex and inconvenient for computation [34] so that Ma and Zarowski [35] have tried to use the upper bound of the maximum eigenvalue, i.e., Dembo's bound, for an efficient signal representation. In the study, we are interested in using more accurate mean estimation to improve the finding of upper bound of eigenvalues (UBE) from sparse observed samples.

Since the environmental noise is usually time-varying or color, the traditional white-noise assumption is not realistic so that the mean value of noise can not always be regarded as zero. Hence this study proposes a new algorithm to evaluate the mean value in terms of noise combined with signal.

Let $x_{i:n}$, $1 \le i \le n$, represent the ranked random samples generated from the output of Eq.(2-23). In this study, we plan to estimate the mean value of a QSAW signal by a new quantile-based maximum likelihood estimator (QMLE) using only the quasi-symmetric quantiles (QSQ), i.e., the minimum sample, $x_{1:n}$, and the maximum sample, $x_{n:n}$. We will compare the performances of the QMLE and sample mean on mean estimation as well as on UBE finding.

There are two parts in our task: one is the QMLE mean estimation aiming at reducing the uncertainty of the estimated correlation matrix and another is the improved upper bound of eigenvalues finding. Conventionally, the mean value of a signal is estimated by sample mean which is UMVUE derived basing on the assumption of normally distributed observations. Although sample mean is a good mean estimator, there still exist some biased estimators that outperform it [23]. In mean estimation for quasi-normal signals, the non-parametric order statistics method was applied to overcome the mismatch between normal and quasi-normal data. In the study, we are interested in the special case of quantile application to mean estimation using the QSQ. The QSQ are determined by the maximum percentage of the observed samples covering the original population, i.e., the coverage which is the cumulative probability calculated between the two endpoints of range. There are good evidences to show that the symmetric property of QSQ is more efficient if they occupy either a very large or very small percentage of the population [36]. Lastly, the task of UBE finding is attractive because the maximum eigenvalue is an important cue of signal activity detection for fading channels with unknown dispersion [31] in multiple-input multiple-output (MIMO) systems [37]. Taparugssanagorn and Ylitalo [38] further indicated the upper bound of MIMO channel capacity being affected by the distribution of the maximum eigenvalue, which was evaluated by the covariance of short-term phase noise. Zhang and Ovaska [39] extended the eigenanalysis to singular value decomposition based on signal-to-noise ratio for the analog-to-digital converter, but their method is not realistic for the cyclostationary detection in spectrum reuse application. Wu et al. [40] proved that the well-trained eigenvector feature of vehicle sound signature was capable of vehicle recognition. UBE acts as the maximum eigenvalue owing to the fact that this representation has been well discussed for the case of deterministic covariance matrix with Hermitian, symmetric positive-definite, or Toeplitz property, Park and Lee [41] improved it by using the technique of series expansion. They proposed the following equations to find a better upper bound of maximum eigenvalue than the classical Dembo's bound:

$$R_{m\times m}=\begin{bmatrix} R_{(m-1)\times(m-1)} & b \\ b^H & a \end{bmatrix} \tag{2-27}$$

$$g_r(\varepsilon)=a-\varepsilon+r\sum_{i=0}^{r}\frac{1}{\varepsilon^{i+1}}b^H R^i_{(m-1)\times(m-1)}b+\frac{\eta_{m-1}\cdot b^H R^i_{(m-1)\times(m-1)}b}{\left(\varepsilon-\eta_{m-1}\right)\cdot\varepsilon^{r+1}}, \tag{2-28}$$

where $R_{m\times m}$ is the correlation matrix of the input signal, $r > 0$ is the order index, ε is an eigenvalue, η_{m-1} is the maximal eigenvalue of $R_{(m-1)\times(m-1)}$, b is an (m-1)-dimensional vector, and a is a scalar. Up till now, there are seldom studies devoting to the uncertainty analysis for the estimation of correlation matrix on sparse data condition. This study proposes the refreshing change-solution against the issue. It avoids the well-known heavy resampling and computation of the bootstrapping method [42] for small sample size. The main uncertainties of additive model result from the propagation of each source signal. In the reasoning for uncertainty of propagation, Denguir-Rekik et al.[43] fused the multiple marginal effects based on the multi-criteria for aggregated decision making. Ferrero and Salicone [44] addressed the issue of utilizing the random-fuzzy variable to fit the propagation of distribution.

Chapter 3: The Probability Density Function of Coverage Interval

3.1 Introduction

Coverage interval (CI) is an interval with two confidence extremes that covers a specified portion of the population. It has been intensively studied in recent years in biology, quality control, medical engineering, and some other research areas. CI is called reference interval in clinical chemistry [45] and is constructed based on the reference values belonging to the population. Motivated by the needs of processing data on small sample size condition for some newly developing areas, such as data mining for knowledge exploration and data representation for pattern recognition, this study deals with the problem of expressing CI under sparse data condition. The issue of applying CI representation to parameter estimation to against the large uncertainty caused by sparse data will also be addressed.

The International Organization for Standardization has issued a document, ISO GUM Suppl. 1: *Guide to the expression of uncertainty in measurement supplement 1* [24], to recommend applying CI as an expression of uncertainty measure to meet the recent trend of treating CI in a probabilistic way. The GUM method of evaluating and expressing uncertainty has been adopted widely by the industry. It can also be found from the manuals published by the Joint Committee for Guides in Metrology (JCGM) [46] that the probability assigned to the input quantity is important. But, a weakness of the probability assignment suggested by JCGM lies in the use of deterministic CI. A general way to represent CI, referred to as “parametric CI” [47], is based on defining a symmetric *pdf* for the input random variable. Alternatively, non-parametric CI representation is based on the empirical distribution of input data. It is usually applied to the case of skew distribution or to the case when the *pdf* is unknown. But, the dichotomy for CI representation is imperfect if a quasi-symmetric *pdf* is encountered. To solve the problem, a unified expression for the uncertainty representation of CI is proposed in this study.

Why should we need CI? It is well known that the information of an event can be represented as the logarithm of the reciprocal of its occurrence probability. It is the commonly used uncertainty measure of an individual event. Entropy is defined, from a macro view, as the expectation of the total information. Although both entropy and CI are macro view of sample data, entropy does not act like CI to provide a clear bounding message. This is analogous to the case of calculating the confidence interval of a parameter estimate. Confidence interval can show explicit bounding information for the estimated parameter.

The chapter is organized as follows. In Section 3.2, a new representation of CI is proposed. It adopts a new method to derive the *pdf* of CI. The effectiveness of the proposed CI representation is evaluated by simulations discussed in Section 3.3. A realization of the statistical CI is presented in Section 3.4. Section 3.5 describes an extension of the statistical CI to the variably truncated normal joint (VTNJ) *pdf*.

3.2 A New Method to Formulate the *pdf* of CI

In this study, we regard CI as a random variable representing the bounded range to meet the coverage constraint. We now derive the *pdf* of CI. According to the work based on the general *pdf* of order statistics [48], the *pdf* of range can be expressed in a non-parametric form by

$$\begin{aligned} f_{r|n}(r) &= \int_{-\infty}^{\infty} f_{r,x_{1:n}|n}(r, x_{1:n})\,dx_{1:n} \\ &= \int_{-\infty}^{\infty} n\cdot(n-1) f_x(x_{1:n}) f_x(x_{1:n}+r)(F_x(x_{1:n}+r)-F_x(x_{1:n}))^{n-2}\,dx_{1:n} \end{aligned}, \tag{3-1}$$

where $f_x(x)$ and $F_x(x)$ denote the *pdf* and *cdf* of random variable *x*, respectively; $x_{1:n}$ is the minimum order of ranked samples; *r* is the range of samples; and *n* is the sample size. It is known that the range *pdf* shown in Eq.(3-1) is accurate for all realistic cases.

We then perform the variable transformation to change the variable $x_{1:n}$ to *c* with range *r* being preserved, where $c = F_x(x_{1:n}+r) - F_x(x_{1:n})$ is the coverage. Suppose that there are *k* roots η_j, $1 \le j \le k$, of $x_{1:n}$ satisfying the coverage constraint equation $F_x(x_{1:n}+r)-F_x(x_{1:n})=Cc$ with a given constant coverage Cc. The joint distribution of r and c can then be expressed by

$$
\begin{aligned}
f_{r,c|n}(r,c=Cc) &= \sum_{j=1}^{k}\{f_{r,x_{1:n}|n}(r,x_{1:n})\times\frac{1}{\left|\begin{matrix}\frac{\partial r}{\partial r} & \frac{\partial r}{\partial x_{1:n}}\\ \frac{\partial c}{\partial r} & \frac{\partial c}{\partial x_{1:n}}\end{matrix}\right|_+}\}_{x_{1:n}=\eta_j}\\
&= \sum_{j=1}^{k}\{f_{r,\eta_j|n}(r,\eta_j)\times\frac{1}{\left|\begin{matrix}1 & 0\\ f_x(\eta_j+r) & f_x(\eta_j+r)-f_x(\eta_j)\end{matrix}\right|_+}\} \qquad (3\text{-}2)\\
&= \sum_{j=1}^{k}\left(f_{r,\eta_j|n}(r,\eta_j)\frac{1}{\left|f_x(\eta_j+r)-f_x(\eta_j)\right|}\right)
\end{aligned}
$$

It is worthwhile to note that the above expression for the joint *pdf* of r and c does not explicitly include the coverage variable c. Instead, c is implicitly included through the roots of the coverage constraint $F_x(\eta+r)-F_x(\eta)=Cc$ for each given sample of $c=Cc$.

We now take a new viewpoint, which is different from the traditional Bayes' theorem, to derive the conditional *pdf* $f_{r|c,n}(r)$. The general form of the Bayes' conditional *pdf* usually maps to a surface while our approach only needs some profiles in the same surface. The concept is shown in Fig. 5 and is realized by

$$f_{r|c=Cc,n}(r)=\frac{f_{r,c=Cc|n}(r,c=Cc)}{\int_{dr} f_{r,c=Cc|n}(r,c=Cc)}. \qquad (3\text{-}3)$$

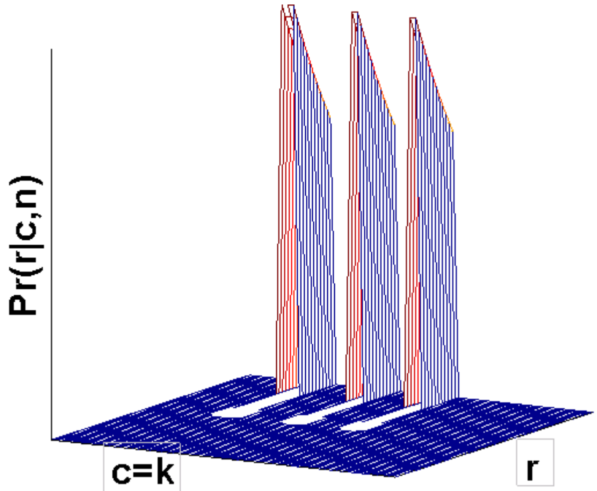

Fig. 5: Profile-conditional *pdf* by the sampling strategy. k is a constant.

A problem encountered in the implementation of Eq.(3-3) is how to expand the transcendental function $F_x(x_{1:n}+r)-F_x(x_{1:n})=Cc$ in order to find its roots. Generally, this can be accomplished by using the Fourier series expansion. But, due to the fact that Hermite polynomials can best fit the curve of normal distribution, we apply Hermite polynomial expansion to $F_x(x)$ in order to efficiently find the solutions of $F_x(x_{1:n}+r)-F_x(x_{1:n})=Cc$. Eq.(3-2) is then expressed by

$$f_{r,c=Cc|n}(r,c=Cc)=\sum_{j=1}^{k}\left(\frac{n(n-1)f_x(\eta_j)f_x(\eta_j+r)\left\{F_x(\eta_j+r)-F_x(\eta_j)\right\}^{n-2}}{\left|f_x(\eta_j+r)-f_x(\eta_j)\right|}\right), \qquad (3\text{-}4)$$

where $\eta_j \in R$ and $f_x(\eta_j+r)-f_x(\eta_j)\neq 0$ for $1\leq j\leq k$. The constraint that η_j must be real is to obey the output rule of Jacobian determinant.

Some modifications are still needed in practical consideration. The basic idea is to neglect some roots of $F_x(\eta+r)-F_x(\eta)=Cc$ which have very low occurrence probabilities. This is realized by setting two bounds for those roots. This is motivated by the general rule of excluding outliers via considering only data in the interval $[\mu-4\sigma,\mu+4\sigma]$ where μ and σ are the mean and standard deviation of the population. Normalization of Eq.(3-4) is also needed in order to make it obey the basic requirement for probability. The *pdf* of CI can then be expressed by

$$f_{r|c=Cc,n}(r)=\sum_{j=1}^{k'}\left(\frac{n(n-1)f_x(\eta_j)f_x(\eta_j+r)\left\{F_x(\eta_j+r)-F_x(\eta_j)\right\}^{n-2}}{\left|f_x(\eta_j+r)-f_x(\eta_j)\right|}\right)\cdot\frac{1}{Z(Cc,n)}, \qquad (3\text{-}5)$$

where $Z(Cc,n)$ is a normalization factor shown below

$$Z(Cc,n)=\int_{dr}\left\{\sum_{j=1}^{k'}\left[\frac{n(n-1)f_x(\eta_j)f_x(\eta_j+r)\left\{F_x(\eta_j+r)-F_x(\eta_j)\right\}^{n-2}}{\left|f_x(\eta_j+r)-f_x(\eta_j)\right|}\right]\right\};$$

$\eta_j, 1\leq j\leq k'$, are the roots of $F_x(x_{1:n}+r)-F_x(x_{1:n})=Cc$ that satisfy $\mu-4\sigma\leq\eta_j\leq\mu+4\sigma$. If $k'>1$, a root-finding procedure is applied to the Hermite polynomial expanded version of $F_x(x_{1:n}+r)-F_x(x_{1:n})=Cc$ for finding all roots in the interval, $[\mu-4\sigma,\mu+4\sigma]$.

Now, we demonstrate our method by exploiting the *pdf* of CI for the normal distribution shown below

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad -\infty < x < \infty. \tag{3-6}$$

The appropriate data structure to implement Eq.(3-5) is a two-dimensional matrix of r and η_j. Considering that the range variable r is also the abscissa of the *pdf* of CI, we arrange the data along the r direction in either an increasing or decreasing order. As referring to Fig. 6, it is more efficient if we apply the bisection method to determine the two endpoints for the roots-finding task. Once we decide the two endpoints, we can assume that all effective roots are inside the interval. This can greatly reduce the searching interval for r and guarantee that there exists at least one solution in the reduced searching interval. Then, for each r in the searching interval, we can find all solutions of η_j by directly solving the polynomial equation obtained by expanding the coverage constraint $F_x(x_{1:n}+r)-F_x(x_{1:n})-Cc$ using Hermite polynomials. Lastly, the *pdf* of CI is calculated by Eq.(3-5).

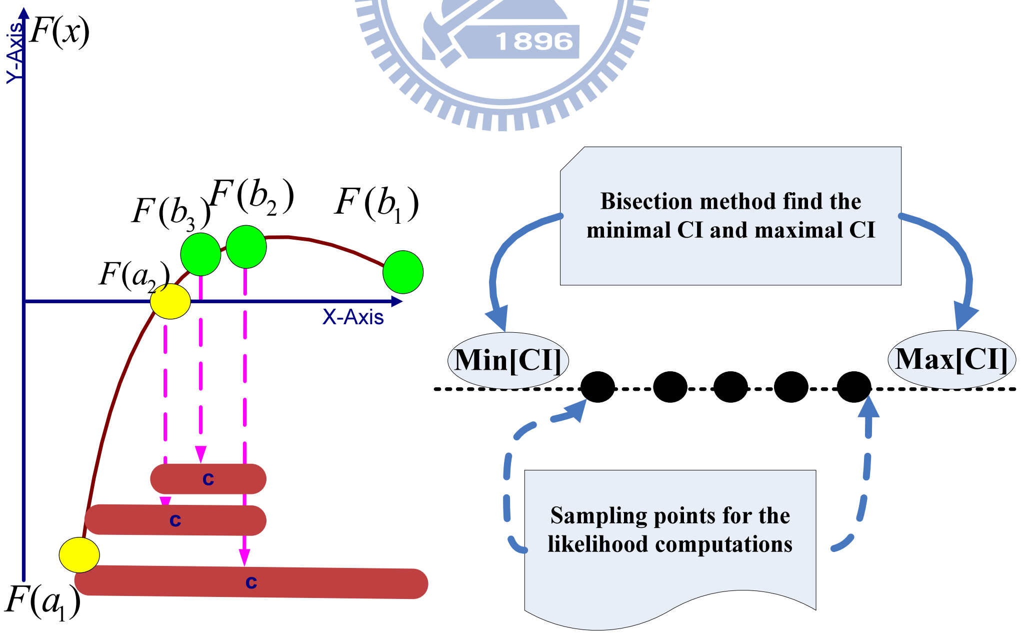
Y-Axis
$F(x)$
$F(b_3)$
$F(b_2)$
$F(b_1)$
$F(a_2)$
X-Axis
c
c
c
$F(a_1)$
Bisection method find the minimal CI and maximal CI
Min[CI]
Max[CI]
Sampling points for the likelihood computations

Fig. 6: A conceptual diagram shows the use of the bisection method to establish the two endpoints for CI. Here, $[a_1, b_1]$ is an effective interval for root finding and c represents the center (midpoint) of any new effective interval.

An example to demonstrate the effectiveness of the proposed method using the standard normal random variable, $N(x;0,1^2)$, with experiment setting of coverage=0.95 and sample size=15 is shown in Fig. 7. Here the top panel shows the *pdf* of CI while the bottom one is the *cdf*. It can be found from the figure that the *pdf* of CI looks like a narrow pulse located near its minimum value (i.e., $r = 3.92$). This result supports the idea of using the minimum case of CI to represent the whole *pdf* of CI as suggested by Chen et al. [47]. Moreover, by examining the *cdf* of CI shown in the bottom panel of Fig. 7, we find that about 70% of probability occurs at the minimum CI.

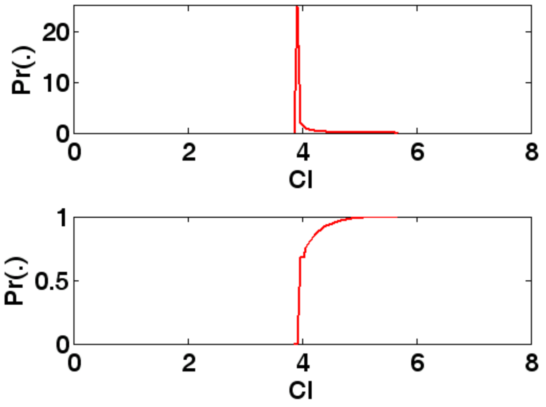

Fig. 7: The *pdf* (top) and *cdf* (buttom) of CI for normal random variable with experiment setting of coverage=0.95 and sample size=15.

3.3 Evaluation of the *pdf* of CI by Simulations

We have suggested using Hermite polynomials to expand the transcendental CI-constrained function, $F_x(x_{1:n}+r)-F_x(x_{1:n})=Cc$, for the best approach to deriving the *pdf* of CI. It is referred to as the multi-root representation of the *pdf* of CI. Now we want to evaluate the goodness-of-fit of the representation by simulations. The test inspects 10,000 trials. In each trial, 15 samples satisfying the constraint of coverage=0.95 with $\pm 10^{-3}$ error tolerance are generated to directly find CI by $x_{15:15}-x_{1:15}$. The histogram of CI is displayed in Fig. 8. It can be found from the figure that the empirical CI distribution fits well to the theoretical results shown in the top panel of Fig. 7.

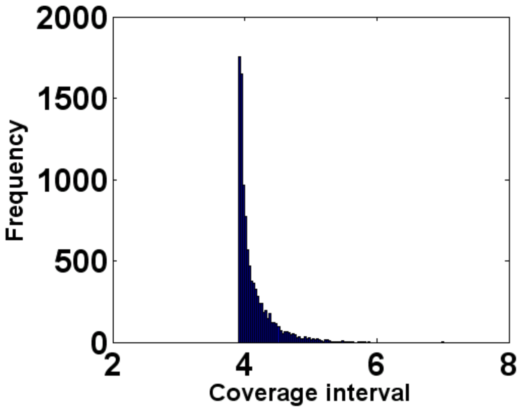

Fig. 8: Histogram of CI generated by simulation using 10,000 trials. The experiment setting is coverage=0.95 and sample size=15.

For further evaluating the proposed method, we compare it with another method basing on the Newton-Raphson root-finding algorithm. The method first uses the Newton-Raphson algorithm to find a root of $F_x(x_{1:n}+r)-F_x(x_{1:n})=Cc$ with initial searching point being set at the left endpoint (i.e., $\mu-4\sigma$), and then find the *pdf* of CI by substituting the root into Eq.(3-5). It is referred to as the single-root representation. Table 1 lists the multi-root and single-root solutions of the coverage constraint equation, $F_x(x_{1:n}+r)-F_x(x_{1:n})=Cc$, for the standard normal distribution with coverage=0.95. From Table 1, some observations are listed below:

Conditioned on our outlier rejection rules, there are only two roots in the interval $[\mu-4\sigma,\mu+4\sigma]$ no matter how the CI changes;

For all CIs, the single-root solutions are very close to the first roots η_1 of the multi-root solution;

For the minimum CI (i.e., 3.92), all roots of the multi-root solution and the single-root solution have nearly the same value;

For the multi-root solution, the sum of its second root and CI is very close to the negative of its first root. This is resulted from the symmetry property of the normal distribution.

Table 1: the multi-root and single-root solutions of the coverage constraint equation

CI	Multi-root solution		Single-root solution
	2nd root η_2	1st root η_1	η
3.92	-1.97	-1.95	-1.95
3.97	-2.15	-1.82	-1.82
4.02	-2.23	-1.78	-1.78
4.07	-2.32	-1.75	-1.75
4.12	-2.39	-1.73	-1.73
4.17	-2.45	-1.72	-1.72
4.22	-2.51	-1.71	-1.71
4.27	-2.57	-1.70	-1.70
4.32	-2.63	-1.69	-1.69
4.37	-2.69	-1.68	-1.68
4.42	-2.74	-1.68	-1.67
4.47	-2.80	-1.67	-1.67

Fig. 9 plots the multi-root and single-root representations of the *pdf* and *cdf* of CI. It can be found from the figure that the multi-root representation fits better to the realistic case (by simulation) shown in Fig. 8. This result justifies the appropriateness of using the Hermite polynomial expansion to help to find multiple roots of the coverage constraint equation for constructing the *pdf* of CI. Table 2 lists some *cdf* values of CI for single-root representation, multi-root representation, and simulation (calculated from Fig. 8) for r in the range of [3.92, 4.37]. It can be found from the table that the simulation results are all larger than the single-root representation, but smaller than the multi-root representation. This shows that the single-root representation is only a rough approximation of the probability distribution of CI. The phenomenon that the whole simulated *cdf* curve lies under that of the multi-root representation can be explained from the viewpoint of sampling. As shown in the top panel of Fig. 9, the theoretical $f_{r|c,n}(r)$ has very large point probability at $r_{\min}$. Since there exists a coverage tolerance of $\pm 10^{-3}$ set in the simulation, the peak-value case can not be generated every time. This makes the simulated *cdf* value at $r_{\min}$ degrade significantly so as to make its *cdf* curve lie under the curve of the multi-root representation which approaches the theoretical one (see Fig. 9).

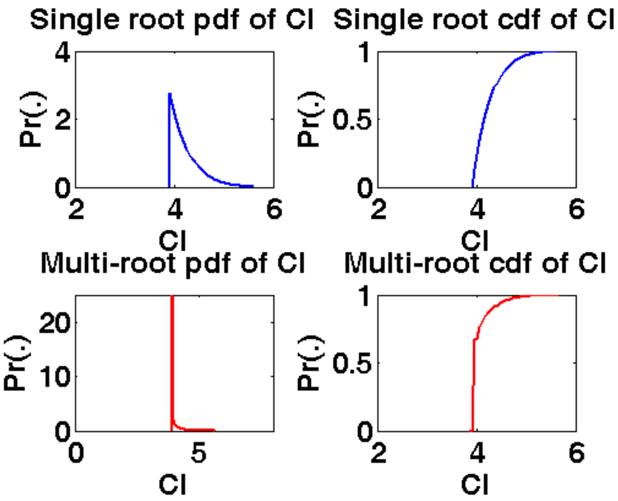

Fig. 9: The multi-root and single-root representations of the *pdf* and *cdf* of CI simulated using the standard normal distribution of input with sample size $n = 15$, and coverage=0.95.

Table 2: Some *cdf* values of CI for single-root representation, multi-root representation and realistic case by simulation

CI	*cdf* of CI		
	Single-root representation	Multi-root representation	Realistic case (Simulations)
3.91	0	0	0
3.97	0.13	0.68	0.36
4.02	0.24	0.72	0.50
4.07	0.34	0.76	0.59
4.12	0.43	0.79	0.66
4.17	0.51	0.83	0.71
4.22	0.58	0.86	0.76
4.27	0.63	0.88	0.79
4.32	0.69	0.90	0.82
4.37	0.73	0.91	0.85

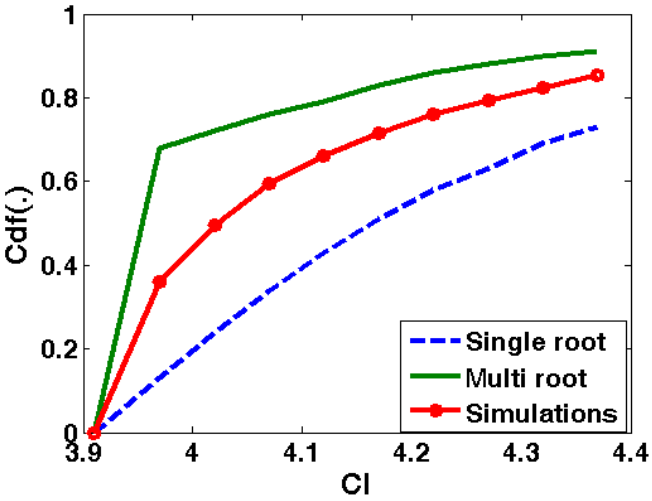

Fig. 10: The *cdf*s of CI using single-root representation, multi-root representation, and simulations

We then further examine their reliability. Table 3 lists the searching results of the single-root solution for some different initial conditions with coverage (=0.95) for standard normal distribution. It is clearly shown in the table that the Newton-Raphson method used for searching the single-root solution may fail with improper initial conditions. So the single-root representation is not always reliable. On the contrary, the method to find multiple roots via using the Hermite polynomial expansion of the coverage constraint function is always stable. So, the multi-root representation of the *pdf* of CI is reliable.

Table 3: Single-root solutions using different initial conditions

Test	Initial value		Final solution	
	η	r	η	r
1	-3	1	7.98	-1.47×10^{-6}
2	-2	2	-1.96	3.92
3	-1	3	-1.96	3.92
4	0	4	-1.96	3.92
5	1	5	-1.96	3.92
6	2	6	8.00	-4.78×10^{-7}

3.4 A Realization of the Statistical CI

Statistical CI is a wide-sense confidence interval representation which is expected to be stable for all sampling plans no matter how the sample size varies. Some computation skills were reported in related literatures as the non-parametric tolerance limits. We will further discuss the influences of CI caused by some properties of range and the corresponding endpoints. The issue has not been addressed yet. Now, we want to measure the confidence level in the statistical CI for the sparse data condition. The statistical CI [49], defined in ISO 3534 [50], was proposed for the concept of confidence level, but only few studies touched the realization algorithm [5,6]. Some other studies were related to the topic of non-parametric tolerance limit [51] which is similar to the statistical CI. Those past works discussed the coverage bound affected by the parameter estimation, inspected the quantile distribution, or described them from the non-parametric viewpoint to look the coverage variation. To extend those past works for further considering the effects of range and the minimum order of ranked samples on the coverage, we need to derive an explicit expression for these three random variables. We discuss the issue in detail as follows.

In practical Monte Carlo simulations, the general expression for statistical CI is typically rewritten, in Pearson's notation, using the incomplete Beta function [49] and expressed by

$$Pr\{p_x[(x_{i:n}, x_{n+1-i:n})] \geq c\} \geq 1-\alpha \quad \Rightarrow \quad 1 - I_c(n+1-2i, 2i) \geq 1-\alpha\,, \tag{3-7}$$

where

$$I_c(n+1-2i, 2i) = Beta(c, n+1-2i, 2i) / Beta(1, n+1-2i, 2i)$$

and

$$Beta(x, p, q) \equiv \int_0^x t^{p-1}(1-t)^{q-1}\,dt\,.$$

It means that the statistical CI of coverage greater than c, at minimal $1-\alpha$ confidence level is $[x_{i:n}, x_{n+1-i:n}]$.

For $i=1$, we can interpret Eq.(3-7) as the confidence level that the samples cover at least c portion of the population is $1-\alpha$. Since Pearson's notation is not convenient in practical realization, we derive a new polynomial form of the *pdf* of coverage to calculate the statistical CI (or confidence level) of c coverage in the sparse data condition. Let

$$s = \int_{-\infty}^{x} f_x(y)dy \,. \tag{3-8}$$

The variable s is subject to the standard rectangular distribution denoted by $\mathrm{rect}[0,1]$. If we take the variable transform to all ranked samples $x_{i:n}$ by Eq.(3-8), the new ranked variables $s_{i:n}$, $1 \le i \le n$, are related to $x_{i:n}$ by

$$s_{i:n} = \int_{-\infty}^{x_{i:n}} f_x(y)dy \tag{3-9}$$

for $1 \le i \le n$, where $0 < s_{i:n} < 1$. Since $s_{i:n}$ can be regarded as the ranked random variables of s, the new range variable can be calculated by

$$\begin{aligned} & r_{new} = s_{n:n} - s_{1:n} \\ & \Rightarrow \int_{-\infty}^{x_{n:n}} f_x(t)dt - \int_{-\infty}^{x_{1:n}} f_x(t)dt \\ & = \int_{x_{1:n}}^{x_{n:n}} f_x(t)dt = c(r) \end{aligned} \tag{3-10}$$

r_{new} : random variable of uniform distribution [0,1]

$x_{i:n}$:random variable of order statistic $s_{i:n}$,$1 \le i \le n$

$f_x(t)$:*pdf* of random variable x

$c(r)$ **:** coverage value of the relative range,

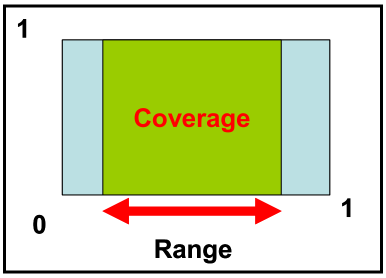

Fig. 11: Uniform *pdf* for the random variable s

As a matter of fact, r_{new} has the same coverage as the range of the original random variable x. Since $r_{new} = c$, Eq.(3-1) can be simplified to express the *pdf* of coverage by

$$f_{c|n}(c) = \int_{0}^{1-c} n(n-1)\cdot 1 \cdot 1\left(s_{1:n} + c - s_{1:n}\right)^{n-2} ds_{1:n} = n(n-1)c^{n-2}(1-c) \tag{3-11}$$

for $0 \le c \le 1$. The *cdf* of coverage can be accordingly expressed by

$$F_{c|n}(c) = c^{n-1}(n + (1-n)c) \tag{3-12}$$

for $0 \le c \le 1$. Due to the fact that the above derivation is true for any random variable, the *pdf* of coverage is distribution-free.

Fig. 12 displays the *pdf* of coverage calculated by Eq.(3-11) for some sample size n ranging from 5 to 20. It is clearly shown in the figure that the probability of coverage deviates away from 1 as the sample size decreases to a value less than 20. This means the common expectation in doing an experiment that the samples distribute like the original population becomes unrealistic as the sample size is less than 20. In other words, the samples are very likely to scatter in only a part of the population for a sparse data condition. We denote it as the short-tail problem.

It is well known that the Student's t-distribution is better than the normal distribution in terms of mean value estimation on the sparse data condition. The general reason is that the tail of a Student's t-distribution is shorter than that of a normal distribution. Applying the same rule, the short-tail phenomenon demonstrated in Fig. 12 needs a new approach to formulate it. It is worth pointing out that short-tail is always decided by the endpoints of the distribution where their values approach zero asymptotically. In our case the distribution of endpoints is simple to predict. From Fig. 7, we find that the two endpoints of CI are almost known when the coverage is high. In other words, if the coverage is known, the short tail of *pdf* can be roughly captured.

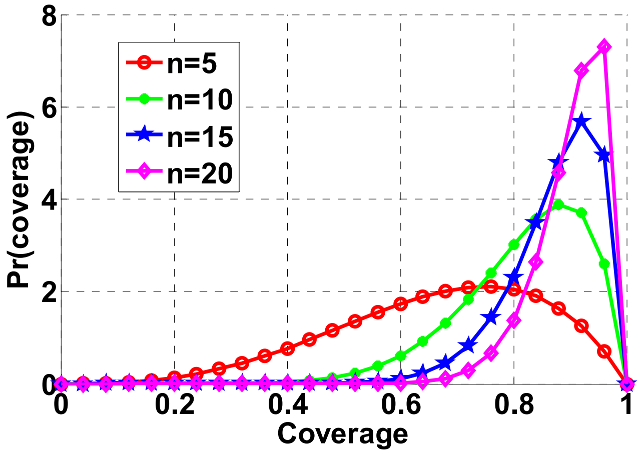

Fig. 12: The *pdf* of coverage for some small sample size n

Fig. 12 reveals that the *pdf* of coverage is not stable as the sample size is less than 20. So we had better know well the coverage variation due to the setting of the β-content level. This issue was addressed by Faulkenberry and Weeks [52].They formulated it as a precision control problem to avoid the increase of parameter uncertainty for the case of sparse data. They suggested that the confidence level $1-\alpha$ had better be set to a value smaller than the β-content (i.e., c).

Table 4 lists the values of confidence level $1-\alpha$ calculated according to Eq.(3-11) for two different types of integration interval. The normal interval $[0.95,1]$ is the general hypothesis testing requirement, while $[\frac{n-1}{n+1}-0.025, \frac{n-1}{n+1}+0.025]$ is the interval to compute the confidence level of uncertainty [52]. It is noted that $\frac{n-1}{n+1}$ is the expectation of coverage. We can see from Table 4 that the values of confidence level calculated using the normal interval of $[0.95,1]$ is smaller when the sample size is less than 20.

Table 4: The confidence level $1-\alpha$ of two integration intervals for different sample sizes

n	Integration interval for c	
	$[0.95,1]$	$[\frac{n-1}{n+1}-0.025, \frac{n-1}{n+1}+0.025]$
10	0.08	0.16
20	0.26	0.30
50	0.72	0.68
90	0.94	0.94
120	0.98	0.98

3.5 Extension of Statistical CI to the VTNJ *pdf*

We now draw some attentions to the two random variables $x_{1:n}$ and r which are related to statistical CI but do not appear in Eq.(3-11). The issue is treated by regarding the *pdf* of coverage shown in Eq.(3-11) as a marginal *pdf* for $x_{1:n}$ and r. As demonstrated by the CI shown in Fig. 7, only parts of $x_{1:n}$ and r near the minimal CI have effect on confidence level computation. We hence need a more precise description to relate the *pdf* of coverage with $x_{1:n}$ and r.The issue is addressed via transforming the statistical CI to an explicit joint *pdf* of $x_{1:n}$, r and c, and use it to compute the confidence level. We first decompose the joint *pdf* of the three random variables, $x_{1:n}$, r and c, into three terms by

$$f(x_{1:n}r,c\,|\,n)=f_{x_{1:n}|r,n}(x_{1:n})\cdot f_{r|c,n}(r)\cdot f_{c|n}(c) \tag{3-13}$$

The first term $f_{x_{1:n}|r,n}(x_{1:n})$ can be calculated from $f_{r|x_{1:n},n}(r)$ (see Eq.(3-1)) by applying the Beyes' rule. The other two terms, $f_{r|c,n}(r)$ and $f_{c|n}(c)$, have been formulated previously. So, according to Eq.(2-1), confidence level can be calculated from $f(x_{1:n}r,c\,|\,n)$ by

$$CL\equiv\int_{\beta}^{1}\int_{r_{\min}}^{r_{\max}}\int_{x_{1:n(L)}}^{x_{1:n(U)}} f(x_{1:n},r,c\,|\,n)dx_{1:n}drdc\,, \tag{3-14}$$

where β is the given β-content level (i.e., coverage), $[r_{\min},r_{\max}]$ is the interval of range corresponding to the coverage *c* in $[\beta,1]$, and $[x_{1:n(L)},x_{1:n(U)}]$ is the interval of $x_{1:n}$ corresponding to range *r* and coverage *c*. It is noted that the statistical CI estimates the probability of coverage greater than β so that Eq.(3-15) takes definite integration over $[\beta,1]$ for the coverage. The interval $[r_{\min},r_{\max}]$ is obtained by the previously mentioned bisection method (see Eq.(3-5)). Since $[r_{\min},r_{\max}]$ is the range of CI, any random interval $[T_1,T_2]$ in Eq.(2-1) will be a legal $[r_{\min},r_{\max}]$. We therefore need to estimate the random interval $[T_1,T_2]$ for a given *c*. The minimum order random variable $x_{1:n}$ is deterministic for some *pdf*s such as Pareto, Weibull and Lognormal. But in this study we consider the realistic and reasonable extent $[x_{1:n(L)},x_{1:n(U)}]$ in the integration. We therefore propose to perform the Hermite polynomial expansion on the coverage function, $F_x(x_{1:n}+r)-F_x(x_{1:n})$, and employ its values at some discrete points to find the relative roots mapping from r to $x_{1:n}$.

Based on above discussions, the confidence level can be calculated from Eq.(3-16) by the traditional Riemann sum-based integration method. However, in order to make a tradeoff between precision and computational complexity, we adopt an alternative approach to using Gauss-Legendre integration (GLI) [53] to realize Eq.(3-14). GLI is a popular method for computing definite integration based on the calculation of pre-determined known functions. For any piecewise continuous function, the task to calculate definite integral on the interval $[a,b]$ can be approximated by a weighted sum of Legendre‘s polynomials defined in the interval of $[-1,1]$. This is applicable to our situation because the *pdf* of coverage is a function given in Eq.(3-11). According to the order of integration, GLI should be applied to the output stage of $f_{r|c,n}(r)$, i.e. $f_{c|n}(c)$. Generally speaking, a GLI can be expressed by

$$\begin{aligned} \int_a^b g(x)dx &= \int_{-1}^{1} g\left(\frac{b-a}{2}\xi + \frac{b+a}{2}\right)\frac{(b-a)}{2}d\xi \\ &= \frac{b-a}{2}\sum_{\tau=1}^{v} w_v(\xi_\tau)\cdot g\left(\frac{b-a}{2}\xi_\tau + \frac{b+a}{2}\right) + R_v(\xi) \end{aligned} \tag{3-17}$$

where a and b are the endpoints of integration interval; ξ_τ, $-1 < \xi_\tau < 1$, is the τth root of the Legendre polynomial $P_v(\xi)$ with order v; $P_v(\xi) = \frac{1}{2^v v!}\frac{\partial^v}{\partial \xi^v}(\xi^2 - 1)^v$, for $v = 0,1,2,\cdots$; $g(x)$ is a known piecewise continuous function;

$$w_v(\xi_\tau) = \frac{(b-a)}{(1-\xi_\tau^2)(P_v'(\xi_\tau))^2} \tag{3-18}$$

is the weighting function; and

$$R_v(\xi) = \frac{2^{(2v+1)}\left(v!\right)^4}{(2v+1)\left((2v)!\right)^3} g^{(2v)}(\xi) \tag{3-19}$$

is the error term of the approximation.

The error term $R_v(\xi)$ of GLI shown in Eqs.(3-17)~(3-19) is proportional to the $2v$th-order derivative of the coverage *pdf* $f_{c|n}(c)$. It can be found from Eq.(3-17) that the number of discrete sampling points (the expansion order of Legendre polynomials) equals v. As shown in Eq.(3-11), the *pdf* of coverage can be expressed by a polynomial with order $n-1$. So, if $2v \ge n-1$, then $R_v(\xi)$ will theoretically become zero. This will result in an analytical closed form for the calculation of confidence level. Up till now, we have successfully formulated a direct computation of confidence level in term of the statistical CI defined in the past studies.

Chapter 4 The Analytical Mean Estimator for Truncated Normal Distribution on Sparse Data Condition

4.1 Introduction

In this chapter, we try to use the likelihood technology to perform the best mean estimator basing on the frequently used truncated normal distribution formed by normalizing the CI-truncation part of *pdf* to its corresponding coverage. Hence truncated normal distribution is usually applied to the sparse data condition when data collection is time-consuming or of high sampling cost. The study focuses on the mean estimation of normally distributed random variables under the sparse data constraint. Since the truncation or censoring scheme is usually adopted in sparse data estimation, our major goal is to improve the truncated normal estimator proposed by Cohen [54]. There are some shortcomings in Cohen's truncated normal estimator, including the need of looking-up tables for setting the positions of initial searching points, the need of a couple of endpoints to compute the standard deviation, the constraint that the expression of endpoints must be deterministic, and non-guarantee of convergence.

The study will use the *pdf* of coverage interval derived in Chapter 2 to construct a variably truncated normal joint (VTNJ) *pdf*, which considers coverage, coverage interval, the first order of ranked samples and the samples themselves. In addition, we reduce the computations of VTNJ *pdf* by employing the suggestion of Chen [55,47] about the parametric coverage interval to obtain a wide-sense parametric coverage estimator.

4.2 The Proposed Method

We use the concept of variably truncated normal distribution to cover the statistical CI in this study. Oour task is to estimate the mean of a random variable $\boldsymbol{x}$ with unknown normal distribution $f_x(x) = N(\mu, \sigma^2)$ from a set of n observed samples $\{x_i, 1 \le i \le n\}$ for $n \le 20$. We first rank these n samples in increasing order and denote them by $\{x_{i:n}, 1 \le i \le n\}$. The range and coverage of the sample set are then defined by $r = x_{n:n} - x_{1:n}$ and $c = F_x(x_{n:n}) - F_x(x_{1:n})$, respectively. Coverage is a macro view of random variable to carry global information of all observed samples. The general relation among coverage c, range r, the minimum order $x_{1:n}$, and samples X_n is shown in Fig. 13. In our basic assumption, we think the macro view random variables should be consistent to the result of micro view random variable. The dash-lines represent the interferences within the macro view random variables, while the solid-lines represent the interferences from the macro view to micro view random variables. A joint normal *pdf* of these four variables will be built in the following basing on Fig. 13 to compensate the coverage mismatch. We treat the distribution as a variably truncated normal joint (VTNJ) *pdf* to represent the randomness of the truncated points of a truncated normal distribution depending on coverage and sample size.

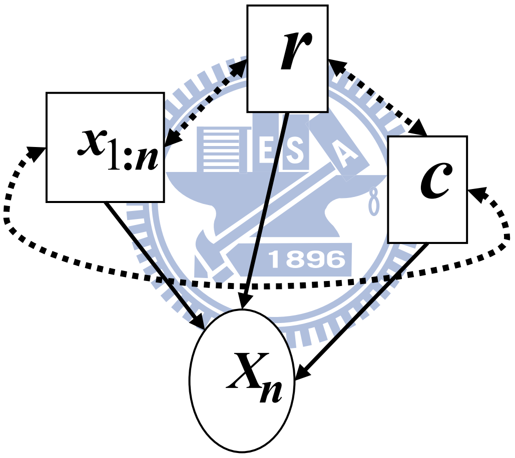

Fig. 13: Relation of variables' interference model

We first decompose $f_{x,x_{1:n},r,c;u,\sigma|n}(x,x_{1:n},r,c)$ into four conditional *pdf*s by

$$f_{x,x_{1:n},r,c;u,\sigma|n}(x,x_{1:n},r,c) = f_{x;u,\sigma|x_{1:n},r,c,n}(x) \cdot f_{x_{1:n}|r,n}(x_{1:n}) \cdot f_{r|c,n}(r) \cdot f_{c|n}(c) \tag{4-1}$$

where

$$f_{x;u,\sigma|x_{1:n},r,c,n}(x) = f_x(x)\frac{U(x-x_{1:n}) - U(x-x_{1:n}-r)}{Q(x_{1:n},r)}$$

is the truncated normal *pdf* depending on the sample size, the truncated points and the sample's coverage; and $Q(x_{1:n},r) = F_x(x_{1:n}+r) - F_x(x_{1:n})$ is the sample coverage.

We then derive the *pdf* of coverage. The *cdf* of coverage for small sample size can be expressed by Eq.(4-2) [56]

$$\Pr_{c|n}(C > c) = \sum_{k=0}^{n-2} \binom{n}{k} c^k (1-c)^{n-k} \tag{4-2}$$

We now simplify the coverage *pdf* as a polynomial of *c*. The derivation is given as follows.

$$\begin{aligned}
\Pr_{c|n}(C > c) &= ((1-c)^n n!\{-(\frac{1}{1-c})^n c\,\Gamma(n) + (-1)^n c(-\frac{c}{1-c})^n \Gamma(n) \\
&+(-1)^n(-\frac{c}{1-c})^n \Gamma(n+1) - (-1)^n c(-\frac{c}{1-c})^n \Gamma(n+1)\})/(c\,\Gamma(n)\Gamma(n+1)) \qquad (4\text{-}3)\\
&= \frac{n!(-c(-1+c^n)\Gamma(n) + (-1+c)c^n\Gamma(n+1))}{c\,\Gamma(n)\Gamma(n+1)} \\
&= \frac{n!}{\Gamma(n+1)} - \frac{c^{n-1} n!(c\,\Gamma(n) + \Gamma(n+1) - c\,\Gamma(n+1))}{\Gamma(n)\Gamma(n+1)} = 1 - nc^{n-1} + (n-1)c^n
\end{aligned}$$

where $\Gamma(.)$ denotes the Gamma function and $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$. Hence

$$f_{c|n}(c) = \frac{\partial}{\partial c}(1 - (\Pr_{c|n}(C > c))) = \frac{\partial}{\partial c}(nc^{n-1} - (n-1)c^n) = n(n-1)(c^{n-2} - c^{n-1}) \quad \text{for } 0 \le c \le 1 \tag{4-4}$$

It is worth to note that Eq.(4-4) is distribution-free because Pratt and Gibbsons [56] also proved it without assuming the distribution of the sampled random variable. Thus it is appropriately applied to any kind of *pdf*. As given in Chapter 2, Fig. 12 displays the coverage *pdf* for some small values of *n*. The figure shows the coverage distribution deviates away from 1 progressively and spreads wider as the sample size decreases from 20. We call this special phenomenon as distribution mismatch (DM) because it implicitly indicates that there exists a serious mismatch between the distributions of observed samples and the random variable when the sample size is small. The DM phenomenon reveals an important cue to the modeling of sparse data: coverage may serve as a confidence factor to indicate the appropriateness of observed data for robust parameter estimation. A higher value of coverage means a better match of the samples to its original normal distribution. To exploit the DM phenomenon, we treat coverage as a random variable and add it to the VTNJ *pdf*.

Eq.(4-1) can be established through Eqs.(3-13) and (3-17) by GLI which is a numerical technique for integration. Then, the VTNJ *pdf* can be implemented by the numerical technique to result in an interval estimation for the coverage fluctuation. In this study, the default settings are $a = E_{c|n}\left[c\right] - 0.005$ and $b = E_{c|n}\left[c\right] + 0.005$ to consider the interval estimation for variable coverage, c.

It will be perfect if we can use a fixed sampling number for GLI to reduce the error so as to make it approach to its minimum. As shown in Eq.(3-19), the error of GLI is related to the differential order of the integrated function. Obviously, its differential order is finite. From Eq.(3-19), if the GLI sampling number v meets the condition of $2v \geq n-1$, the estimation error $R_v(\xi)$ will be reduced to zero. In this case, GLI will approach to the theoretical optimal solution of no errors. Besides, Eq.(4-4) shows another important fact that the coverage *pdf* is independent of the distribution of the sampled random variable. So, we can claim that the *pdf* of coverage is distribution free. This property makes $f_{c|n}(c)$ freely connect to any kind of $f_{r|c,n}(r)$ by Chain rule.

4.3 Standard Normal Transform for the VTNJ *pdf* Computation

If we want to directly calculate the VTNJ *pdf* in $p_{r|c,n}(r)$, we will face the problem that the mean and standard deviation of the population must be known in advance. But this is unrealistic in our mission. We therefore adopt an alternative approach to construct a new bridge to conjoint with these variables. The idea is to transform the observed data into the standard normal domain. The suggestion is shown in Fig. 14. As shown in the figure, we transform the observed ranked samples into the domain of standard normal by $\xi_{i:n} = (x_{i:n} - u)/\sigma$. Each transform pair is marked with the same digit number. The range is also transformed by $r_s = \xi_{n:n} - \xi_{1:n}$. Notice that the transform is quantile mapping invariance (QMI) for the macro view random variables.

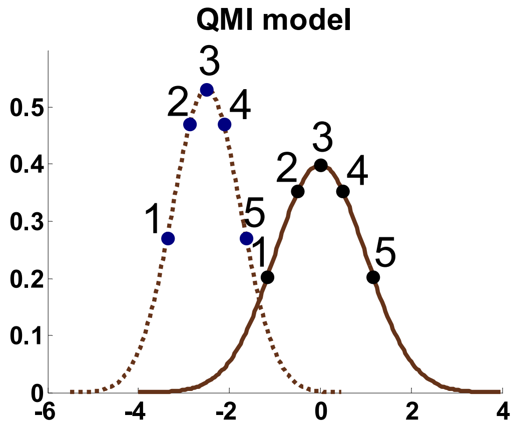

Fig. 14: Relative quantile mapping invariance based on their percentiles. Dash-line represents the original normal *pdf* and solid-line represents the standard normal *pdf*.

4.3.1 Derive the Variably Truncated Normal Joint Distribution Estimator (VTNJE)

We then apply GLI to the VTNJ *pdf* to obtain the marginal log likelihood defined by:

$$MLL(\cdot) \equiv \sum_{t=1}^{v} \left\{ \frac{b-a}{2} n(n-1)({Cc_t}^{n-2} - {Cc_t}^{n-1}) w_{P_v}(\kappa_t) \int_{dr_s} \int_{d\xi_{1:n}} G \right\} \tag{4-5}$$

where

$$G = \log\left\{ \left(\frac{1}{\sqrt{2\pi}\sigma \, \left(\Phi_{\xi}(\xi_{1:n}+r_s) - \Phi_{\xi}(\xi_{1:n})\right)} \right)^{n} \cdot \exp\left\{ -\sum_{i=1}^{n} \frac{(x_i-u)^2}{2\sigma^2} \right\} \right\} \cdot p_{\xi_{1:n}|r_s,n}(\xi_{1:n})$$
$$\cdot\, p_{r_s|c=Cc_t,n}(r_s)$$

The marginal log likelihood is complicated and computionally time-consuming. We suggested an idea to reduce its computation basing on the coverage interval. An example of the profile-conditional *pdf*, $f_{r|c,n}(r)$, is plotted in Fig. 7. It is to demonstrate the fact that if we would like to guarantee the coverage of the estimation to be large enough to greater than a lower bound, then there will be much more tolerance intervals qualified for solutions to reside. Let us return to Eq.(4-4) to inspect the *pdf* of coverage which is distribution-free. We find that its form is inconvenient for parameter estimation due to the no use of derivative operator. Fortunately, Chen [47] suggested that the *pdf* of coverage can be parametric if we constrain the coverage interval to be the minimum of all possible values.

4.3.2 Algebraic Closed From for Parameter Estimation

Let we apply the result of Fig. 7 to simplify Eq.(4-5). It can then be expressed as two quadric equations of variables σ and u respectively. Take the roots of these two quadric equations will result in the following solutions:

$$\sigma^* = \frac{B_\sigma \pm \sqrt{(B_\sigma)^2 + 4\left(\sum_{t=1}^{v} nD_t\right)C_\sigma}}{2\left(\sum_{t=1}^{v} nD_t\right)}, \tag{4-6}$$

where

$$B_\sigma = \left(\sum_{t=1}^{v} D_t \left(E_{\xi_{1:n}|c=Cc_t, Min\{r_s\}, n}\left\{\xi_{1:n}\right\}\right)\left(\sum_{i=1}^{n}(x_i - x_{1:n})\right)\right),$$

$$C_\sigma = \left(\sum_{t=1}^{v}\left(D_t \sum_{i=1}^{n}(x_i - x_{1:n})^2\right)\right),$$

$$D_t = (\frac{b-a}{2} n(n-1)(Cc_t^{\,n-2} - Cc_t^{\,n-1})\left(w_{P_v}(\kappa_t)\right));$$

and

$$u^* = \frac{-B_u \pm \sqrt{B_u^2 - 4(\sum_{t=1}^{v} D_t)C_u}}{2(\sum_{t=1}^{v} D_t)} \tag{4-7}$$

where

$$B_{u=}\sum_{t=1}^{v}\left\{D_t\left[\left(\overline{x} - x_{1:n}\right)\left(E_{\xi_{1:n}|c=Cc_t, Min\{r_s\}, n}\left\{\xi_{1:n}^2\right\}\right) - 2x_{1:n}\right]\right\},$$

$$C_u = \sum_{t=1}^{v}\left\{D_t\left[x_{1:n}^2 + \left(\overline{x}x_{1:n} - \overline{x^2}\right)\left(E_{\xi_{1:n}|c=Cc_t, Min\{r_s\}, n}\left\{\xi_{1:n}^2\right\}\right)\right]\right\}.$$

Here, $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the sample mean, $\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ is the mean of sample square, $w_{P_v}(\kappa_t) = \frac{(b\text{-}a)}{(1\text{-}\kappa_t^2)(P_v'(\kappa_t))^2}$ is the weighting coefficient of the t-th root of the v-th order Legendre polynomial, $[a,b]$ is the coverage estimation interval, and $\Phi_\xi(\xi)$

is the *cdf* of the standard normal distribution. The same strategy can be applied to the other endpoint $\xi_{n:n}$ via replacing $\xi_{1:n}$ by $\xi_{n:n}$. Then, the VTNJ *pdf* can be implemented by the numerical technique to result in an interval estimation for the coverage fluctuation. From Fig. 12, it clearly shows that the coverage is a random variable if the sample size is less than 20. Hence, we had better to set the most observed interval to inspect its randomness. Define the following $\beta\%$-inspection interval (β -II):

$\beta\%$-inspection interval is an interval estimation for the coverage random variable over the interval [a,b] with $a = E_{c|n}\left[c\right] - \beta/2,$ and $b = E_{c|n}\left[c\right] + \beta/2$.

We then aim at calculating the most possible happening probability.

4.4 Experiments

By checking Eqs.(4-6) and (4-7), we find that they are mainly affected by the sample mean, $\bar{x}$, and the individual ranked samples, $x_{i:n}, 1 \le i \le n$. Our strategy is to adjust the coverage to make it approach to the real coverage, generated from $\bar{x}$ and $x_{i:n}, 1 \le i \le n$. We examine two methods. One is to view the joint effect of $\bar{x}$ and $x_{i:n}$ under our suggestion of QMI (see Fig. 14). The other is to realize the QMI basing only on the real coverage. Its purpose is to see the differences between the sample mean without coverage estimation and VTNJE with coverage calibration.

4.4.1 Test the Results with Consistency to Sample Mean under the QMI Principle—Case of the Default Percentile

It is clear that coverage is a random variable based on Eq.(4-4) so that we should take the most observed samples. We first form an interval estimation for coverage by performing a coverage estimation from the expectation of order statistics by $E_{c|n}[c]$ and adding fluctuation of ± 0.005.

The VTNJE might work normally without the operations of looking up the tables so that it more convenient for the computer programming. We will compare it to the best estimator of sample mean. The test pattern is selected from the normal distribution, $\mathrm{N}(10,1^2)$. Two different conditions for sample mean are considered. One is to constrain the sample means in the interval of $-0.3\sigma+u \le \overline{x} \le 0.3\sigma+u$. It is referred to as the good sample mean case. The other is to constrain the sample means in the interval of $-2.3\sigma+u \le \overline{x} \le -1.3\sigma+u$ or $1.3\sigma+u \le \overline{x} \le 2.3\sigma+u$, and is referred to as the bad sample mean case due to its seriously skewness. Three estimators are compared: Scheme A represents the conventional sample mean estimator; and Scheme B is the coverage-based estimator defined below

$$u^* = u_p = x_{p:n} - \frac{\sum_{t=1}^{v}\left\{D_t\left[E_{\xi_{p:n}|c=Cc_t,Min\{r_s\},n}\left\{\xi_{p:n}\right\}\right]\right\}}{\sum_{t=1}^{v} D_t}\sigma_p \qquad (4\text{-}8)$$

where p is constrained to be either 1 or n which corresponded to the endpoints of the range;. If $p=1$, then the term $E_{\xi_{p:n}|c=Cc_t,Min\{r_s\},n}\left\{\xi_{p:n}\right\}$ can be computed by $(-1)E_{\xi_{1:n}|c=Cc_t,Min\{r_s\},n}\left\{\xi_{1:n}\right\}$. Scheme C is taking the result of Eq.(4-7). Those results are displayed in Fig. 15. It can be found from the figure that MSEs are very small for the case of good sample mean for all three estimators; while the MSEs are all large for the case of bad sample mean. This shows that the performance of VTNJE will asymptotically follow that of the sample mean. Those results also imply that very low MSE can be probably provided that the sample mean is near the population mean.

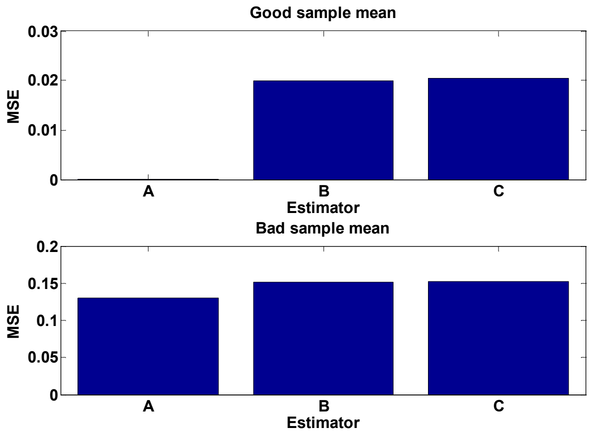

Fig. 15: Comparison of the conventional sample mean estimator and two coverage-based mean estimators.

4.4.2 Test the Results with Consistency to Sample Mean under the QMI Principle—Case of Realistic Percentile

In the test phase, we eliminate the effects caused by the QMI mapping mismatch for $\xi_{1:n}$ to $x_{1:n}$ or $\xi_{n:n}$ to $x_{n:n}$. In such a case, $\xi_{1:n}=(x_{1:n}-u)/\sigma$ and $\xi_{n:n}=(x_{n:n}-u)/\sigma$ are known. But, we pretend that we do not know u and σ. The fluctuation assumption for coverage is therefore not needed. So, the previous formulation can be simplified and expressed by

$$\sigma^{*}=\frac{\xi_{p:n}\left(\sum_{i=1}^{n}(x_i-x_{p:n})\right)}{2n}\pm\frac{\sqrt{\left(\xi_{p:n}\left(\sum_{i=1}^{n}(x_i-x_{p:n})\right)\right)^{2}+4n^{2}\left(\sum_{i=1}^{n}(x_i-x_{p:n})^{2}\right)}}{2n}, \qquad (4\text{-}9)$$

for $\sigma^{*}>0$ and

$$u^{*}=\frac{-\left(\left(\overline{x}-x_{p:n}\right)\left(\xi_{p:n}^{2}\right)-2x_{p:n}\right)}{2}\pm\frac{\sqrt{\left[\left(\overline{x}-x_{p:n}\right)\left(\xi_{p:n}^{2}\right)-2x_{p:n}\right]^{2}-4\left[x_{p:n}^{2}+\left(\overline{x}x_{p:n}-\left(\overline{x^{2}}\right)\right)\left(\xi_{p:n}^{2}\right)\right]}}{2} \qquad (4\text{-}10)$$

where p is constrained to be either 1 or n. Actually, Eq.(4-9) is equivalent to Eq.(4-10) because $u^{*}=x_{p:n}-\xi_{p:n}\sigma^{*}$. We generated 1,000 trials to examine the new estimator and used MSE as the score of comparison. The results are listed in Table 5.

Table 5：Performance of realistic QMI analysis

Item	sample mean	Realistic QMI
MSE	0.0765	0.0252

Notice that the MSE of realistic QMI was defined by $\frac{1}{1000}\sum\left(\frac{(u_1+u_n)}{2}-u\right)^{2}$, where u_1 and u_n were the estimated results for $x_{1:n}$ and $x_{n:n}$, respectively. It can be found from Table 5 that the realistic QMI mean estimator performed better than the sample mean estimator.

4.4.3 Comparison of the Different Estimators

We compared three different mean estimators in terms of their stabilities and efficiencies. They are the Cohen's method (shown in Eq.(2-18) to Eq.(2-21)), our

VTNJE and the sample mean which is the average of total samples. The test involves 5000 trials, and in each trial 13 samples submitted to the standard normal distribution, $N(0,1^2)$ are generated.

In the test, we apply three types of truncation intervals to force truncating the data outside them which they are [-2, 3], [-1.5, 1.75]. In such planning, we may easily to realize the performance between the Cohen, VTNJE and sample mean.

The formulation derived by Cohen request of the initial searching points so that we divided the initial searching condition into two classes, bad and good. The bad condition indicating the initial searching position for mean, u, is outside the interval, $[-2\sigma/\sqrt{n}+u, 2\sigma/\sqrt{n}+u]$ and good condition representing the initial searching position is inside the interval, $[-0.5\sigma/\sqrt{n}+u, 0.5\sigma/\sqrt{n}+u]$. Table 6 display the average of the square errors of 5000 trials for the bad initial conditions and Table 7 is the case of good initial condition. We find from these two tables that our VTNJE is stable and outperforms the Cohen's method. Besides, VTNJE performs slightly better than the sample mean.

Table 6: Comparison with different estimators in association with bad initial searching points (Unit: MSE)

Estimator	Truncation Interval		
	[-2,3]	[-1.8,2.5]	[-1.5,1.75]
Cohen	1.439	1.420	0.991
VTNJE	0.061	0.059	0.059
Sample mean	0.078	0.075	0.076

Table 7: Comparison with different estimators in association with good initial searching points (Unit: MSE)

Estimator	Truncation Interval		
	[-2.0,3.0]	[-1.8,2.5]	[-1.5,1.75]
Cohen	0.610	0.582	0.731
VTNJE	0.061	0.060	0.059
Sample mean	0.078	0.075	0.076

4.5 Conclusions

This study develops the variably truncated normal joint *pdf* to emulate the CI-truncation part of *pdf* normalize to its corresponding coverage. We have demonstrated the weakness of the Cohen's mean estimator using classical truncated normal distribution on its reliability when the sample size is less than 20. On the contrary, the proposed VTNJE using the truncated normal distribution derived based on the normalized-parametric coverage intervals is reliable and efficient.

We use Hermite polynomials to expand the coverage function accurately. It not only uses the high order polynomials to approach the real curve, but also guarantees the convergence for the condition when σ is known in advance (see Eqs.(4-9) and (4-10)).

VTNJE only needs one truncation point for estimation (see Eqs. (4-6) and (4-7)); thus it is superior to the original truncated normal estimator which needs a couple of endpoints to do iterations (see Eqs. (2-18) and (2-19)).

The third goodness of the VTNJ *pdf* is that it does not need any looking-up table for root-finding. It is expressed in an analytical closed form (see Eqs. (4-6) and (4-7)) and this feature may save time for computation. Furthermore, in the default QMI test, we have showed that our coverage-based mean estimator follows the sample mean so that the VTNJ *pdf* also solves the truncated normal problems with knowing only the possible information of the truncated points.

Lastly, we reformulate the equations for the case when σ is known. It works well if σ is known in our estimation process. In the original MLE formulation derived by Cohen, the solution-finding process often encounters the underflow problem. Since the coefficients of the variables are probability or cumulative probability of normal distribution. It is inconvenient for the realizing inverse function representation. Our truncated normal estimator outperforms the old one obviously.

Chapter 5: The VTNJ Estimator Tested with the Combined Signals

5.1 Introduction

In Chapter 3, we have shown that the VTNJ *pdf* can act as the statistical CI to function like a truncated *pdf*. So, we can regard CI as being embedded in the VTNJ *pdf*. We have also tried to use the VTNJ *pdf* in the case of normally distributed observed data for mean estimation. Now we want to further test the VTNJ estimator (VTNJE) for the case of sampling data of combined signal. In the early GUM recommendation [24, p.6], the uncertainty evaluation was considered as to construct a relation between the input quantities and the output quantities of combined signal. This style of uncertainty measurement is recognized as Type A expression in NIST [57]. Now, we want to test the realistic refined case for the CI estimation. Fotowicz [2,58] proposed an analytic method to estimate CI based on the assumption that the individual standard uncertainty was known. Note that we have mentioned that case in Eqs.(2-23) - (2-25).

From Eqs.(2-23) to (2-25), it is easy to realize that the standard uncertainties of input quantities must be known in advance. Fotowicz proved them on the basis of the Central Limit Theorem and concluded that if the distribution of the output of combined quantities is asymptotically symmetric, e.g. a normal distribution, then the output CI approaches the minimum of all possible values.

Although the output of combined quantities may not be of normal distribution, we still suggest using the normal distribution assumption to estimate its mean value based on the past experience. Since VTNJE is designed based on the small sample size condition, it is necessary to examine its robustness when the sample size is less than 20. Tests using different sample size ranging from 11 to 20 are therefore conducted. For each sample size, 1000 trials are tested. In each trial, we replace both the minimum order sample and the maximum order sample with their quantiles on the constrained uncertainty. In the following, two cases of the uncertainty loading tests in terms of VTNJE are examined. One is to test the random variable, coverage, versus the sample size and another is to compare the realistic case with six different estimators.

5.2 Robust Interval Detection for Small Sample Size

According to Eq.(3-11), we define the primary reliable quantiles calculated from the expectations of their ordered samples which may be computed from the general *pdf* of order statistics adding with a little variation. Taking the viewpoint from Fig. 12, the optional truncation points are regarded as some variations around the expectation of coverage. We define the possible truncation positions as the $\kappa\%$-inspection interval shown below:

$\kappa\%$-inspection interval is an interval estimation for the coverage random variable over the interval [a,b] with $a = E_{c|n}\left[c\right] - \kappa/2$, $b = E_{c|n}\left[c\right] + \kappa/2$, and $f_{c|n}(c)$ being given in Eq.(3-11).

From Eq.(3-11), it is easy to find that the expectation of coverage is $(n-1)/(n+1)$. From Fig. 7, the minimal CI has the maximal occurrence probability so that the right endpoint of percentile can be roughly determined as $n/(n+1)$ and the left endpoint is $1/(n+1)$ set based on the expectation of coverage. The default value for κ is set to be 0.2.

We now test the robustness of VTNJE by simulations. The experiment settings are described as follows. Let the output of combined quantities be composed of four independent random input quantities, including two normal distribution random variables, $z_1 \sim N(0.1, 1^2)$ and $z_2 \sim N(2.15, 1.5^2)$, and two rectangular distributions, $z_3 \sim \mathrm{rect}[-2\sqrt{3} - 1.05, 2\sqrt{3} - 1.05]$ and $z_4 \sim \mathrm{rect}[-4\sqrt{3} + 1.45, 4\sqrt{3} + 1.45]$. The output x is generated from four input quantities expressed by Eq.(5-1) and its uncertainty ratio UR is equal to 1.48 calculated by Eq.(2-25):

$$x = f(z_1, z_2, z_3, z_4) = z_1 + z_2 + z_3 + z_4 \tag{5-1}$$

We model the output quantity as a normal distribution shown below

$$N\left(\sum_{i=1}^{4} m(z_i), \sum_{i=1}^{4} u^2(z_i)\right), \tag{5-2}$$

where $m(z_i)$ and $u^2(z_i)$ are respectively means and square of standard uncertainties of individual input quantities. Since z_3 and z_4 are not normal distribution, we can not estimate the mean of x directly. Applying the law of uncertainty of propagation, the combined uncertainty $u_c(x)$ can be calculated by

$$u_c(x) = \sum_{i=1}^{4} (\frac{\partial f}{\partial z_i})^2 \cdot u^2(z_i) + 2\sum_{i=1}^{3} \sum_{j=i+1}^{4} \frac{\partial f}{\partial z_i} \frac{\partial f}{\partial z_j} u(z_i, z_j) \tag{5-3}$$

We combine the Fotowicz's equation with Eq.(2-26) to estimate the quantile which has been examined using 10^6 samples [2] . We generate 1,000 trials for each of the sample size in range of 11~20. Four estimators are compared. "VTNJE" represents the one using Eq.(4-7) by taking the average of the two outputs resulting from using the two input quantiles, $x_{1:n}$ and $x_{n:n}$. "VTNJE+Fotowicz" represents "VTNJE" with the two endpoints of samples being replaced by those calculated by Eq.(2-26). Because of the randomness of coverage resulting from the sparse data condition (see Fig. 12), the traditional 95% coverage interval is not appropriate for describing the variation of coverage. Thus, the quantiles of the endpoints are decided by the expectation of Eq.(3-11). "sample mean" is the conventional sample mean estimator. "sample mean+Fotowicz" is "sample mean" in terms of Fotowicz's quantiles.

The theoretical output mean can be approximated by $\sum_{i=1}^{4} m(z_i)$ (=2.650). The sample mean over 10^5 observations is near 2.678 by Monte Carlo simulation. The experimental results are shown in Table 8. From the table, we find that "VTNJE" stably outperform "sample mean" if the uncertainty ratio is greater than 1.5. We note that this conclusion happens only when the sample size is smaller than 20 (see Table 8).

Table 8: Computation results for the uncertainty ratio, UR =1.5, with 1,000 trials, normalized by $u_c^2(x)/n$, 4 mixing signals (Unit: Normalized MSE)

Sample size	Average mean square errors				Sample size	Average mean square errors			
	Sample mean	VTNJE	Sample mean + Fotowicz	VTNJE + Fotowicz		Sample mean	VTNJE	Sample mean + Fotowicz	VTNJE + Fotowicz
11	1.074	1.020	0.825	0.476	21	1.023	1.045	0.916	0.676
12	1.051	0.970	0.840	0.509	22	1.116	1.115	1.000	0.704
13	1.069	1.044	0.861	0.533	23	1.049	1.071	0.953	0.707
14	1.066	0.996	0.906	0.550	24	1.005	1.008	0.926	0.688
15	1.003	0.98	0.838	0.557	25	1.113	1.134	1.018	0.768
16	1.122	1.023	0.975	0.600	26	1.126	1.135	1.031	0.785
17	0.980	0.969	0.844	0.583	27	1.094	1.096	1.010	0.748
18	1.000	0.984	0.877	0.593	28	1.085	1.088	0.999	0.789
19	1.026	1.017	0.902	0.663	29	1.018	1.037	0.937	0.787
20	1.046	1.029	0.931	0.695	30	1.036	1.106	0.969	0.803

5.2.1 Test VTNJE for Combined Quantities

Two patterns of output of combined quantities are used to further test VTNJE. The first one is composed of four independent input random quantities, including two quantities of normal distribution, $z_1 \sim N(0.1, 1^2)$ and $z_2 \sim N(2.15, 1.5^2)$, and two quantities of rectangular distribution, $z_3 \sim \text{rect}\,[-2\sqrt{3}+0.15, 2\sqrt{3}+0.15]$ and $z_4 \sim \text{rect}[-10\sqrt{3}-0.1, 10\sqrt{3}-0.1]$. The second one is formed by changing z_4 of the first one to $\text{rect}\,[-28\sqrt{3}-0.1, 28\sqrt{3}-0.1]$ with the other three input quantities unchanged. They are mainly different by their uncertainty ratios: UR=3.7 for the first output quantity and UR=10.4 for the second one. Since VTNJE needs accurate quantiles, we predict the accurate quantiles for the two endpoints according to the order statistics in association with Eq.(5-2), $E[x_{p:n}]$.

Fig. 16 shows the experimental results of 3,000 trials for VTNJE using the two output patterns of UR=3.7 and 10.4. We find from the figure that VTNJE performs better for the pattern of higher UR whose shape of distribution is flatter than that of lower UR.

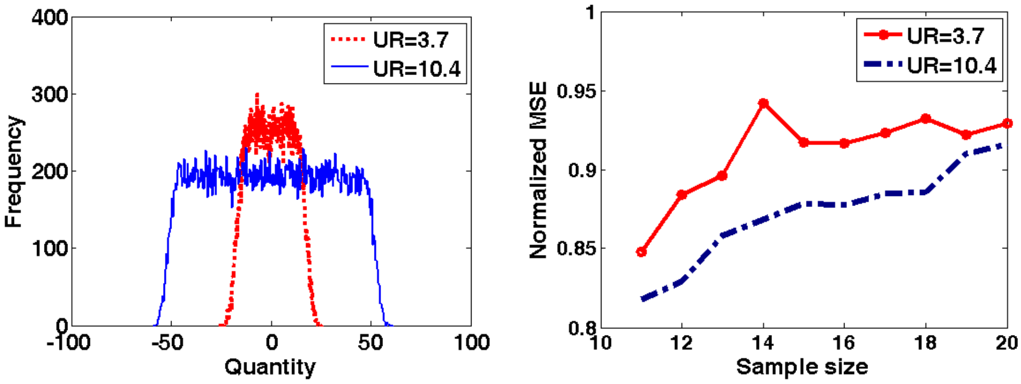

Fig. 16: Experimental results for VTNJE using two output patterns of UR=3.7 and 10.4. Left: the distributions of output quantities. Right: MSEs normalized to $u_c^2(x)/n$.

For further examining the efficiency of VTNJE, six mean estimators are tested. The first two are sample mean and VTNJE. The third one, denoted as sample mean+60% VR, is sample mean with the two endpoints substituted with the accurate quantiles varying with 60% VR. Here, VR denotes the basic variation unit of $u_c(x)/\sqrt{n}$. The fourth, VTNJE+60% VR, is VTNJE with input quantities having 60% VR on the expectation. The fifth, VTNJE+Parzen, is VTNJE fed with quantile estimated by Parzen estimator. Parzen [59] proposed a simple quantile estimator via smoothing adjacent neighbors:

$$F_n^{-1}(q) = n(\frac{i}{n} - q)x_{i-1:n} + n(q - \frac{i-1}{n})x_{i:n}, \quad i > \frac{n}{2} \tag{5-4}$$

where q is the percent of quantile and i is the sample index. If $i \le n/2$, reverse the order of weighting coefficients. The last, VTNJE+Fotowicz, is VTNJE using the quantiles estimated by the Fotowicz's algorithm. The experimental results are displayed in Fig. 17. It can be seen from the figure that VTNJE outperforms sample mean without any assumption. This is a great achievement as we recognize that sample mean is UMVUE for normal distribution. Although the R*N distribution is different in shape from the normal distribution, they are alike for UR<1. Moreover, both VTNJE+60% VR and VTNJE+Fotowicz perform even better. This shows that VTNJE can operate on the same level of Fotowicz's quantile even if it loads 60% combined uncertainty variation about the theoretical quantile. Lastly, VTNJE+Parzen performs not well in the sparse data condition. Here, we explain why VTNJE outperforms the sample mean estimator. The reason is that the quantiles of R*N distribution has a small scattering area corresponding to the equal standard uncertainty of quantiles in VTNJE. Besides, sample mean is a UMVUE only for the normal population, it is not the best mean estimator for the R*N distribution.

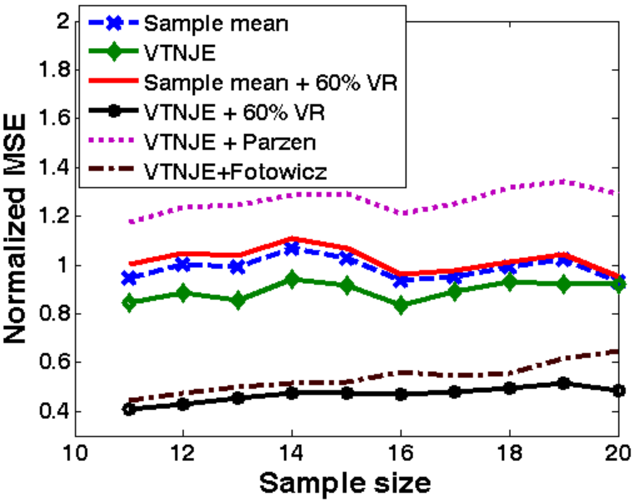

Fig. 17: Performance comparison for six estimators using the first output pattern with UR= 3.7.

5.2.2 Test VTNJE for Different Uncertainty Ratio

Lastly, we examine the performance of VTNJE for different uncertainty ratio. The experimental results are displayed in Fig. 18. First, we find from the figure that the average MSE of sample mean persists around its theoretic value of 1 according to the Central Limit Theorem. Here, average is taken over all sample sizes from 11 to 20. The average MSE of VTNJE decreases as UR increases and saturates to the value around 0.85 at UR near 6. Moreover, VTNJE outperforms sample mean when UR is greater than 2. As combined with the Fotowicz's algorithm, sample mean+Fotowicz outperforms VTNJE for small UR; and their performances are comparable for large UR. Lastly, VTNJE+Fotowicz performs best.

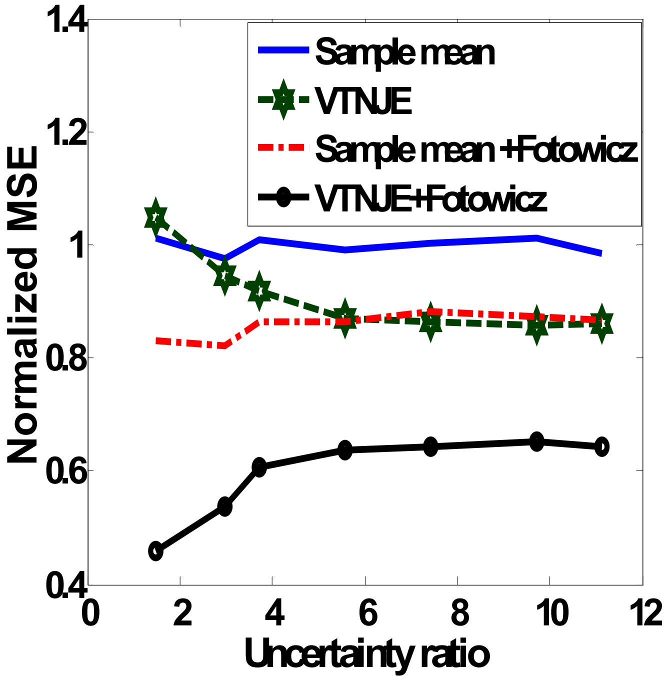

Fig. 18: Performance comparison for four estimators using the output quantity of 4-mixture combined quantities with different UR. The average MSE is normalized to $u_c^2(x)/n$.

5.3 Conclusions

The highly concentrated *pdf* of CI also provides a proper demonstration of the idea of the conventional parametric CI approach which suggests taking the minimum of all possible values of CI. The new formulation of the *pdf* of CI has been shown to serve as a unified framework of parametric and non-parametric CI representations. Lastly, we have discussed a new quantile-based VTNJE to estimate the mean value of the output of combined quantities. The VTNJE was shown to outperform the traditional sample mean estimator for the sparse data condition.

5.4 Appendix: Derivation a Closed Form for VTNJE

Our principal goal is to establish an analytical form of estimator for the truncated normal distribution so that some special skills can be applied to the whole schemes including marginal likelihood, withdraw certain terms in the derived equations and externally adding a certain factor in the equation. If the posterior analysis takes a good performance, then these schemes are right.

First, we take the integration of the log likelihood with respect to the three macro view random variables $x_{1:n}$, r, c to obtain the marginal log likelihood.

$$\begin{aligned} MLL(.) = &\sum_{t=1}^{v}(D_t \cdot (\int_{dr_s}\int_{d\xi_{1:n}}((-\frac{n}{2}\log 2\pi - n\log\sigma)\times Q))) \\ &+\sum_{t=1}^{v}(D_t \cdot (\int_{dr_s}\int_{d\xi_{1:n}}((-n\log(\Phi(\xi_{1:n}+r_s)-\Phi(\xi_{1:n}))-\sum_{i=1}^{n}\frac{(x_i-u)^2}{2\sigma^2})\times Q))) \end{aligned} \quad (5\text{-}5)$$

where,

$D_t = (\frac{b-a}{2} n(n-1)(Cc_t^{\ n-2} - Cc_t^{\ n-1})\times\left(w_{P_v}(\kappa_t)\right)$,

$Q = p_{\xi_{1:n}|r_s,n}(\xi_{1:n} \mid r_s, n)\cdot p_{r_s|c,n}(r_s \mid c = Cc_t, n)$,

and $p_{r_s|c,n}(r_s \mid c = Cc_t, n)$ is the profile-conditional *pdf* of r_s.

We then use a single truncation point to perform the estimation. Specifically, we employ the equation $u = x_{1:n} - \sigma\xi_{1:n}$ to obtain

$$\begin{aligned} MLL(.) = &\sum_{t=1}^{v}(D_t \cdot (\int_{dr_s}\int_{d\xi_{1:n}}((-\frac{n}{2}\log 2\pi - n\log\sigma)Q))) \\ &+\sum_{t=1}^{v}(D_t \cdot (\int_{dr_s}\int_{d\xi_{1:n}}\left(-n\log(\Phi(\xi_{1:n}+r_s)-\Phi(\xi_{1:n}))Q\right))) \\ &\quad+\sum_{t=1}^{v}(D_t \cdot (\int_{dr_s}\int_{d\xi_{1:n}}((-\sum_{i=1}^{n}\frac{(x_i-x_{1:n}+\sigma\xi_{1:n})^2}{2\sigma^2})Q))) \end{aligned} \quad (5\text{-}6)$$

We then ignore the second term in Eq.(5-6) i.e. $\sum_{t=1}^{v}\left\{D_t\cdot\left(\int_{dr_s}\int_{d\xi_{1:n}}\left(-n\log(\Phi(\xi_{1:n}+r_s)-\Phi(\xi_{1:n}))Q\right)\right)\right\}$. We have two reasons to make the decisions. One is that it is a transcendental function which is difficult to obtain an explicit expression for the variables. The second reason is that we have found $\Phi(\xi_{1:n}+r_s)-\Phi(\xi_{1:n})$ to be a coverage variable. Remember that we have derived the joint *pdf* of $x_{1:n}$, r, c. From Eq.(4-4), we find that the maximum power for coverage is $n-1$ in the *pdf* of coverage. So, the dominant term has been present in the *pdf* of coverage regardless whether the coverage variable, $\Phi(\xi_{1:n}+r_s)-\Phi(\xi_{1:n})$, is existing. Taking the expansion for the third term, we obain:

$$\begin{aligned} &\sum_{t=1}^{v}(D_t \cdot (\int_{dr_s}\int_{d\xi_{1:n}}((-\sum_{i=1}^{n}\frac{(x_i-x_{1:n}+\sigma\xi_{1:n})^2}{2\sigma^2})Q))) \\ &=\sum_{t=1}^{v}(D_t \cdot (\int_{dr_s}\int_{d\xi_{1:n}}((-\frac{1}{2}(\sum_{i=1}^{n}(\frac{(x_i-x_{1:n})}{\sigma})^2+2\frac{(x_i-x_{1:n})}{\sigma}\xi_{1:n}+\xi_{1:n}^{\ 2}))Q))) \end{aligned} \quad (5\text{-}7)$$

$$=\sum_{t=1}^{v}(D_t\cdot(-\frac{1}{2}(\sum_{i=1}^{n}(\frac{(x_i-x_{1:n})}{\sigma})^2+2\frac{(x_i-x_{1:n})}{\sigma}\cdot E_{\xi_{1:n},r_s|c=Cc_t,n}[\xi_{1:n}]+E_{\xi_{1:n},r_s|c=Cc_t,n}[\xi_{1:n}^2]))) \quad (5\text{-}8)$$

where $E_?[.]$ denotes the expectation operator.

When we withdraw the coverage term, the equation will become Eq.(5-8) and the other problem generated. If we inspect Eq.(5-8), it will be found that there is going to no any coverage interval term, r_s, to be left after integrating the variable r_s. This result violates Eq.(2-10) derived by Cohen. Since our VTNJ estimator is the extending work of his truncated normal estimator so that we should preserve the information for r_s. Thus we take the suggestion by Chen [47] to select the minimum coverage interval, $Min[r_s]$, representing the information for coverage interval, r_s. The new simplified equation is Eq.(5-9).

$$\begin{aligned}&\sum_{t=1}^{v}D_t(-\frac{n}{2}\log 2\pi-n\log\sigma)+\sum_{t=1}^{v}D_t(-\frac{1}{2}\sum_{i=1}^{n}(\frac{(x_i-x_{1:n})}{\sigma})^2\\&\quad-\sum_{i=1}^{n}\frac{(x_i-x_{1:n})}{\sigma}\cdot E_{\xi_{1:n}|c=Cc_t,Min[r_s],n}[\xi_{1:n}]-\frac{n}{2}E_{\xi_{1:n}|c=Cc_t,Min[r_s],n}[\xi_{1:n}^2])\end{aligned} \quad (5\text{-}9)$$

Taking the partial derivative of Eq.(5-9) with respect to σ and setting it to zero, i.e.,

$$\frac{\partial}{\partial\sigma}MLL(.)=\sum_{t=1}^{v}(D_t(-\frac{n}{\sigma}+\sum_{i=1}^{n}\frac{(x_i-x_{1:n})^2}{\sigma^3}+\sum_{i=1}^{n}\frac{(x_i-x_{1:n})}{\sigma^2}\cdot E_{\xi_{1:n}|c=Cc_t,Min[r_s],n}[\xi_{1:n}]))=0$$

$$=(\sum_{t=1}^{v}nD_t)\sigma^2-(\sum_{t=1}^{v}D_t(E_{\xi_{1:n}|c=Cc_t,Min[r_s],n}[\xi_{1:n}]\sum_{i=1}^{n}(x_i-x_{1:n})))\sigma-\sum_{t=1}^{v}(D_t\sum_{i=1}^{n}(x_i-x_{1:n})^2)=0$$

(5-10)

Solving Eq.(5-10), we obtain an estimate of the standard deviation of the population:

$$\sigma^*=\frac{B_\sigma\pm\sqrt{(B_\sigma)^2+4\left(\sum_{t=1}^{v}nD_t\right)C_\sigma}}{2\left(\sum_{t=1}^{v}nD_t\right)} \quad (5\text{-}11)$$

with $\sigma^*>0$, where

$$B_\sigma=\left(\sum_{t=1}^{v}D_t\left(\left(E_{\xi_{1:n}|c=Cc_t,Min[r_s],n}[\xi_{1:n}]\right)\sum_{i=1}^{n}(x_i-x_{1:n})\right)\right)$$

$$C_\sigma=\left(\sum_{t=1}^{v}\left(D_t\sum_{i=1}^{n}(x_i-x_{1:n})^2\right)\right)$$

$$D_t = (\frac{b-a}{2} n(n-1)({Cc_t}^{n-2} - {Cc_t}^{n-1}) \times \left(w_{P_v}(\kappa_{\mathrm{t}}) \right)$$

By substituting $\sigma = (x_{1:n} - u) / \xi_{1:n}$ into Eq.(5-5), we obtain

$$\begin{aligned} MLL(.) = & \sum_{t=1}^{v} (D_t \cdot (\int_{dr_s} \int_{d\xi_{1:n}} ((-\frac{n}{2} \log 2\pi - n \log(\frac{x_{1:n} - u}{\xi_{1:n}})) \times Q))) \\ & + \sum_{t=1}^{v} (D_t \cdot (\int_{dr_s} \int_{d\xi_{1:n}} ((-n \log(\Phi(\xi_{1:n} + r_s) - \Phi(\xi_{1:n})) - \sum_{i=1}^{n} \frac{(x_i - u)^2}{2(x_{1:n} - u)^2} \xi_{1:n}^2) \times Q))) \end{aligned} \tag{5-12}$$

Taking $\frac{\partial}{\partial u} MLL(.) = 0$, we obtain

$$\frac{\partial}{\partial u} MLL(.) = \sum_{t=1}^{v} (D_t \frac{n}{x_{1:n} - u}) - \sum_{t=1}^{v} (D_t \sum_{i=1}^{n} \frac{(x_i - x_{1:n})(x_i - u)}{\left(x_{1:n} - u\right)^3} E_{\xi_{1:n}|c=Cc_t, Min[r_s],n}[\xi_{1:n}^2]) = 0$$

$$= \sum_{t=1}^{v} (nD_t (x_{1:n} - u)^2) - \sum_{t=1}^{v} (D_t \sum_{i=1}^{n} ((x_i^2 - x_i u - x_{1:n} x_i + u x_{1:n}) E_{\xi_{1:n}|c=Cc_t, Min[r_s],n}[\xi_{1:n}^2])) = 0$$

$$= \sum_{t=1}^{v} (nD_t (x_{1:n} - u)^2) - \sum_{t=1}^{v} (D_t ((\sum_{i=1}^{n} x_i^2 - n\bar{x}u - n\bar{x}x_{1:n} + nux_{1:n}) E_{\xi_{1:n}|c=Cc_t, Min[r_s],n}[\xi_{1:n}^2])) = 0$$

$$= \sum_{t=1}^{v} (D_t \left(x_{1:n} - u\right)^2) - \sum_{t=1}^{v} (D_t ((\overline{x^2} - (\bar{x} - x_{1:n}) u - \bar{x}x_{1:n}) E_{\xi_{1:n}|c=Cc_t, Min[r_s],n}[\xi_{1:n}^2])) = 0 \tag{5-13}$$

$$\begin{aligned} & = \sum_{t=1}^{v} (D_t (x_{1:n}^2 - 2x_{1:n} u + u^2)) - \sum_{t=1}^{v} (D_t ((\overline{x^2})(E_{\xi_{1:n}|c=Cc_t, Min[r_s],n}[\xi_{1:n}^2]))) \\ & +(\sum_{t=1}^{v} (D_t ((\bar{x} - x_{1:n})(E_{\xi_{1:n}|c=Cc_t, Min[r_s],n}[\xi_{1:n}^2]))))u + \sum_{t=1}^{v} (D_t ((\bar{x}x_{1:n})(E_{\xi_{1:n}|c=Cc_t, Min[r_s],n}[\xi_{1:n}^2]))) = 0 \end{aligned} \tag{5-14}$$

With simple mathematical manipulations, the above equation can be simplified and expressed by

$$\begin{aligned} & = \sum_{t=1}^{v} \left(D_t\right) \mu^2 + (\sum_{t=1}^{v} (D_t ((\bar{x} - x_{1:n})(E_{\xi_{1:n}|c=Cc_t, Min[r_s],n}[\xi_{1:n}^2]) - 2x_{1:n})))\mu \\ & \quad + \sum_{t=1}^{v} (D_t (x_{1:n}^2 + (\bar{x}x_{1:n} - (\overline{x^2}))(E_{\xi_{1:n}|c=Cc_t, Min[r_s],n}[\xi_{1:n}^2]))) = 0 \end{aligned} \tag{5-15}$$

Solving Eq.(5-15), we obtain an estimator of μ:

$$\mu^{*}=\frac{-B_u\pm\sqrt{B_u^2-4(\sum_{t=1}^{v}D_t)C_u}}{2(\sum_{t=1}^{v}D_t)} \tag{5-16}$$

where

$$B_{u=}\sum_{t=1}^{v}\left(D_t\left(\left(\overline{x}-x_{1:n}\right)\left(E_{\xi_{1:n}|c=Cc_t,Min[r_s],n}[\xi_{1:n}^2]\right)-2x_{1:n}\right)\right)$$

$$C_u=\sum_{t=1}^{v}\left(D_t\left(x_{1:n}^2+\left(\overline{x}x_{1:n}-\left(\overline{x^2}\right)\right)\left(E_{\xi_{1:n}|c=Cc_t,Min[r_s],n}[\xi_{1:n}^2]\right)\right)\right))$$

Chapter 6 The Asymptotic Minimax Optimization for Mean Estimation of Combined Signals

We now follow the robust statistical method "asymptotic minimax principle" to realize the mean estimation of combined signals. It is referred to as QMLE. We derive the QMLE via solving the problem of maximizing the objective function $QMLE(\mu,\sigma)$ defined by:

$$\begin{cases} QMLE(\mu,\sigma) = (-\frac{n}{2}\log 2\pi - n\log\sigma) - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2} \quad , \\ \mu = x_{p:n} - \sigma\xi_{p:n}. \qquad \text{for } p=1 \text{ or } n \end{cases} \tag{6-1}$$

where $x_{p:n}$ is the minimum order (for p=1) or maximum order (for p=n) of samples x_i, for $1 \le i \le n$; $\xi_{p:n}$ is a standard normal random variable normalized from $x_{p:n}$; and n is the sample size. The solution derived in detail in the Appendix is given below:

$$\mu_p^* = x_{p:n} - \sigma_p^*\xi_{p:n} \tag{6-2}$$

where

$$\sigma_p^* = \frac{\xi_{p:n}\left(n(\overline{x}-x_{p:n})\right)}{2n} \pm \frac{\sqrt{\left(\xi_{p:n}\left(n(\overline{x}-x_{p:n})\right)\right)^2 + 4n(\sum_{i=1}^{n}(x_i-x_{p:n})^2)}}{2n} \tag{6-3}$$

with the constraint $\sigma^* > 0$, and $\overline{x}$ is the sample mean.

If we emulate the *pdf* of combined quantities as a quasi-normal distribution (see an example shown in Fig. 4), one of its two extreme shapes looks like a rectangular *pdf* for large UR. Fig. 19 demonstrates the first order and last order random variables (i.e. QSQ) of the rectangular and normal *pdf*s with the same standard uncertainty. From the figure, we find that the dispersion-areas of QSQ for the rectangular *pdf* are more concentrated than these for the normal *pdf*.

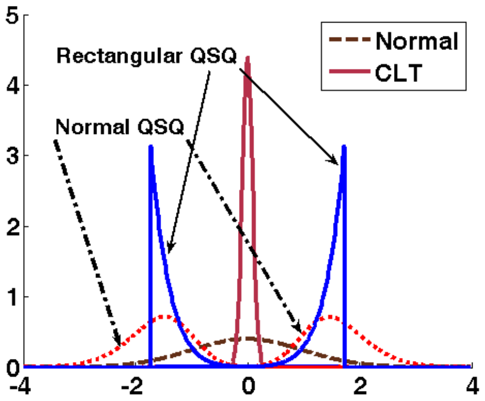

Fig. 19: Standard normal *pdf* combined with its CLT *pdf* and QSQ *pdf* for sample size=11. We plot the QSQ of equal variance rectangular *pdf* $[-\sqrt{3},\sqrt{3}]$ as the blue solid line.

6.1 Establish the Minimax Structure

Huber [60] addressed the robust statistical method via the least possible variance searching algorithm given below:

Asymptotic minimax results [60]: Let κ be a convex compact set of distribution F on the real line. To find a sequence T_n of estimators of location which have a small asymptotic variance over the whole of κ; more precisely, the supremum over κ of the asymptotic variance should be least possible.

According to the above theorem, we need three components to establish a minimax searching algorithm. They are the convex set, least variance and a minimax optimization objective function. We describe them in detail as follows.

6.1.1 Convex Set

Eq.(6-1) is a quadratic equation so that its global extreme does exist. According to this property, we construct the convex set comprising the candidates of population mean. Using three normal *pdf*s, $N(10,1^2)$, $N(2.3,0.8^2)$ and $N(3.7,1.2^2)$ as examples, we form their convex sets by using Eqs.(6-2) and (6-3). There are 1,000 trials with 15 samples in each trail. For each trial, the 15 samples are firstly sorted in the ascending order to find the two endpoints $x_{1:n}$ and $x_{n:n}$. They are then transformed into the standard normal distributed versions, $\xi_{1:n}$ and $\xi_{n:n}$, by using a pre-assumed pseudo mean μ_{ps} and the true standard deviation σ if it is known (or the samples' standard deviation $\sigma_s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$). Then, the estimate σ^* is calculated by Eq.(6-3). We denote it as σ_p^*. The final mean estimate is obtained substituting σ_p^* into Eq.(6-2), to obtain $u_p^* = x_{p:n} - \sigma_p^* \xi_{p:n}$ for $p = 1$ or n.

To evaluate the performance of the QMLE estimator, an averaged mean square error (MSE) defined by:

$$MSE = \frac{1}{1000}\sum_{i=1}^{1000}\frac{\left((\mu_1^*(i) - \mu_{ps})^2 + (\mu_n^*(i) - \mu_{ps})^2\right)}{2} \tag{6-4}$$

is calculated for each test. We take the error between the pseudo mean and real mean, $\left(\mu_{ps} - \mu\right)$, as the reference. We set the inspection interval of μ_{ps} to be $[\mu - 2\sigma/\sqrt{n}, \mu + 2\sigma/\sqrt{n}]$ and take 50 pseudo means distributed uniformly over the interval as the candidates of population mean. Fig. 20 displays the average MSEs of QMLE versus $\left(\mu_{ps} - \mu\right)$. It can be clearly found from the figure that, for all the three test cases using different normal distributions, the average MSEs of QMLE are characterized as convergence curves to become smaller as the absolute value of the difference between μ_{ps} and μ decreases.

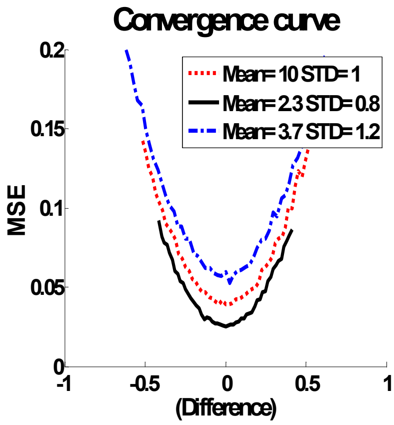

Fig. 20: MSE of QMLE versus difference=$\left(\mu_{ps} - \mu\right)$ for three normal distributions. Note that $\xi_{1:n}$ is calculated using true standard deviation σ.

6.1.2 Asymptotic Efficient near the Minimal Average of MSEs

Fig. 20 shows that the three average MSE curves are convex functions of $\left(\mu_{ps}-\mu\right)$ with their minima located at the zero of $\left(\mu_{ps}-\mu\right)$. Based on the observation, we therefore suggest letting the selection criterion of the pseudo mean, μ_{ps}, correspond to the minimal average MSE, and expect that the resulting QMLE has higher efficiency than the sample mean.

6.1.3 Minimax Structure for the Objective Function

Now, we add a punishment term to form a new objective function and find the optimal pseudo mean estimate by:

$$
\begin{aligned}
&\arg\max_{\mu_{ps}}\left\{QMLE(\sigma,\mu_{ps})-\frac{1}{2}\left((\mu_1^*-\mu)^2+(\mu_n^*-\mu)^2\right)\right\}\\
&=\arg\max_{\mu_{ps}}\left\{(-\frac{n}{2}\log 2\pi-n\log\sigma)-\frac{1}{2}\sum_{i=1}^{n}(\frac{x_i-x_{p:n}}{\sigma})^2\right.\\
&\quad-\sum_{i=1}^{n}\frac{x_i-x_{p:n}}{\sigma}\cdot\frac{x_{p:n}-\mu_{ps}}{\sigma_s}-\frac{n}{2}(\frac{x_{p:n}-\mu_{ps}}{\sigma_s})^2\\
&\quad\left.-\frac{1}{2}\{((x_{1:n}-\sigma_1^*\cdot\frac{x_{1:n}-\mu_{ps}}{\sigma_s})-\mu)^2+((x_{n:n}-\sigma_n^*\cdot\frac{x_{n:n}-\mu_{ps}}{\sigma_s})-\mu)^2\}\right\}
\end{aligned}
\qquad (6\text{-}5)
$$

The minimax operation is thus constructed completely. The corresponding criterion of optimization is a combination of maximum QMLE and minimum MSE (MMSE) on QSQ.

Table 9 lists four possible conditions that we will encounter in setting the inspection area for searching the optimal μ_{ps}. They specify the conditions whether the population's mean and population's standard uncertainty are given or not. Basically, the inspection area is set as $[\mu-2\sigma/\sqrt{n},\mu+2\sigma/\sqrt{n}]$. If the combined (population's) mean is unknown, the best searching interval for determining the candidates is also unknown. In this case, we use the sample mean to determine the searching interval. Similarly, if the combined (population's) standard uncertainty, σ, is unknown, we use the samples' standard uncertainty, σ_S, for its substitution. Table 9 shows the test conditions for the four combinations.

Table 9: Table of confusion for the conditions of combined mean and combined standard uncertainty

		Combined Mean (CLT searching interval)	
		Known	Unknown
Standard uncertainty of combined quantities	Known	A	B
	Unknown	C	D

6.2 QMLE optimization on MMSE of the Two Endpoints of Range, (QSQ)

In the proposed QMLE mean estimator, the quantiles are determined by the maximum percentage of its original population, i.e. coverage. Since the coverage-constrained quantiles obey the properties of symmetric quantiles, the QMLE mean estimator may be efficient and robust with variance asymptotically approaching the Cramer-Rao lower bound. It is worthy noting that since the QSQ usually covers a significant portion of the population, it is therefore popular to apply the double censoring scheme for the observations of small sample size, especially in the sport contest. We know that adopting such a strategy can avoid the large variation occurring in the mean estimation. Based on above discussions, we apply the above QMLE+MMSE optimization search only on QSQ, and call it the Q2MMSE-CLT scheme.

We now examine the performance of Q2MMSE-CLT by simulations. Suppose that the combined quantity is composed of four independent random input quantities with two normal distributions, $x = z_1 + z_2 + z_3 + z_4$, $z_1 \sim N(0.1, 1^2)$ and $z_2 \sim N(2.15, 1.5^2)$, and two rectangular distributions, $z_3 \sim rect[-2\sqrt{3} - 1.05, 2\sqrt{3} - 1.05]$ and $z_4 \sim rect[-10\sqrt{3} + 1.45, 10\sqrt{3} + 1.45]$. We perform 10,000 trials to test Q2MMSE-CLT for each of the four conditions listed in Table 9. The testing sample size ranges from 11 to 40 for each trial. Fig. 21 and Fig. 22 display the experimental results. It can be found from these two figures that Q2MMSE-CLT significantly outperforms the sample mean for Conditions A and B, and is slightly better for Conditions C and D. In other words, Q2MMSE-CLT has much lower MSEs when the standard uncertainty is known.

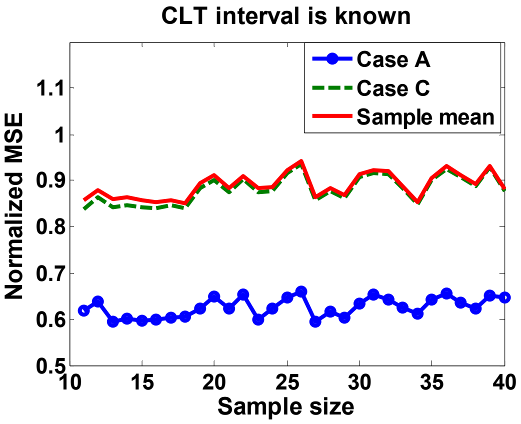

Fig. 21: Conditions A and C. y axis is normalized to $u_c^2(x)/n$

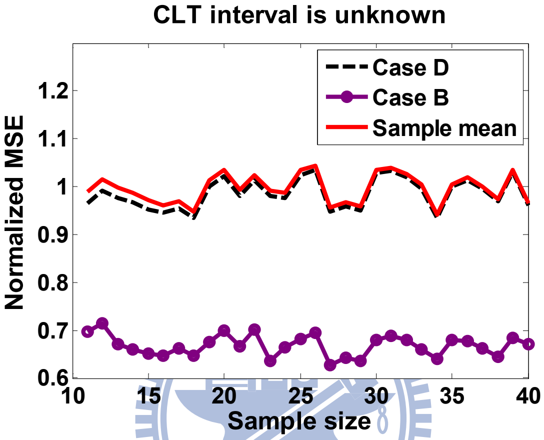

Fig. 22: Conditions B and D. y axis normalized to $u_c^2(x)/n$

6.2.1 Test the Robustness of Q2MMSE-CLT for Different Uncertainty Ratio

Here we test Q2MMSE-CLT for two different values of UR. As demonstrated in Fig. 4, the R*N distribution is more flat in its central part as UR increases. It is a general issue to study whether Q2MMSE-CLT performs better for larger UR. We perform 10,000 trials for two cases of combined quantities composing of four different distributions. One has $z_1 \sim N(0.1,1^2)$, $z_2 \sim N(0.2,1.5^2)$, $z_3 \sim rec[-2\sqrt{3}+0.15,2\sqrt{3}+0.15]$, and $z_4 \sim rec[-10\sqrt{3}-0.1,10\sqrt{3}-0.1]$. Its UR is equal to 3.7 evaluated according to Eq.(2-25). Another is the same as the first case except that $z_4 \sim rec[-10\sqrt{3}-0.1,10\sqrt{3}-0.1]$ is changed to $z_4 \sim rec[-28\sqrt{3}-0.1,28\sqrt{3}-0.1]$. The UR is accordingly changed to 10.4. Fig. 23 displays the histograms of 50,000 outputs of combined quantities for the two cases. It shows the property of quasi-normal distribution for the output of combined quantities.

To compare the two cases of Q2MMSE-CLT, a robustness function of gain relative to sample mean is defined as

$$G = 1 - \frac{Average\ MSEs\ of\ (Q2MMSE)}{Average\ MSEs\ of\ (sample\ \ mean)}, \quad (unit: \frac{u_c^2(x)}{n}) \tag{6-6}$$

Fig. 24 displays the experimental results. It can be found from the figure that Q2MMSE-CLT outperforms sample mean for both cases of UR=3.7 and UR=10.4. Moreover, the performance is better for larger UR.

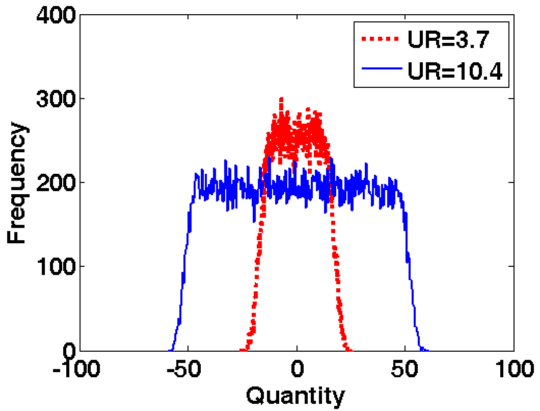

Fig. 23: Histogram of 50,000 combined quantities for different URs. x-axis is the output of combined quantities and y-axis is the frequency count

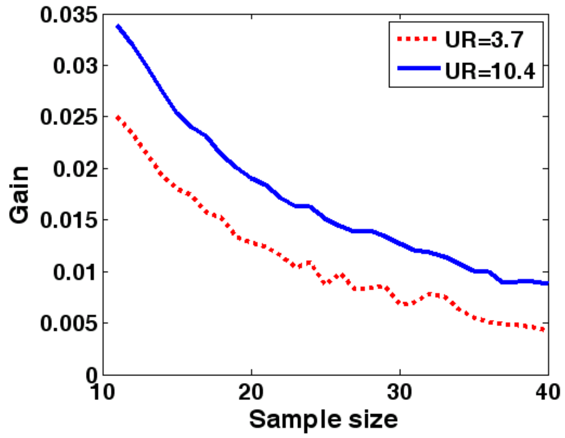

Fig. 24: Gain performance for the different URs. The unit is $u_c^2(x)/n$

6.2.2 An Advanced Refinement of the QMLE

Although Q2MMSE-CLT follows the paradigm of asymptotic minimax principle, there are only about 2%~3% gains, for Conditions C and D, over the sample mean in the mean estimation for the output of combined quantities. By considering the practical applications, we only further discuss Condition D. As was noted previously, the testing data of combined quantities are formed in the same manner and we execute

1,000 trials with 15 observations in each trial. We select 60 candidates of population mean and arrange them to be symmetric to the sample mean within the interval of $[-2\sigma_s / \sqrt{n} + \bar{x}, 2\sigma_s / \sqrt{n} + \bar{x}]$. Then we evaluate the QMLE via the Q2MMSE-CLT scheme. In our maneuver, we first plot the convex curves according to the three different clusters of Z score (i.e., quantile of the signal transformed to standard normal *pdf*) of sample mean: $Z < -2$, $-0.5 \le Z \le 0.5$, and $Z > 2$. We then define the cluster $-0.5 \le Z \le 0.5$ as good sample mean and the other two clusters, $Z < -2$ and $Z > 2$, as the bad sample means. Fig. 25 is the convex sets conditioned on the good sample mean. Here, the dot line is the convex set for the original signal of combined quantities and the green solid line represents the convex set due to enlarging standard uncertainty (ESU) to 4 times of the original signal with the same reference candidates of population mean. We find from the figure that for the good sample mean case QMLE converges near the symmetric location, (i.e., the 30-th candidate) for both the original and ESU signals. So, in the good sample mean case the convergence of QMLE to population mean on heavy observations will be guaranteed.

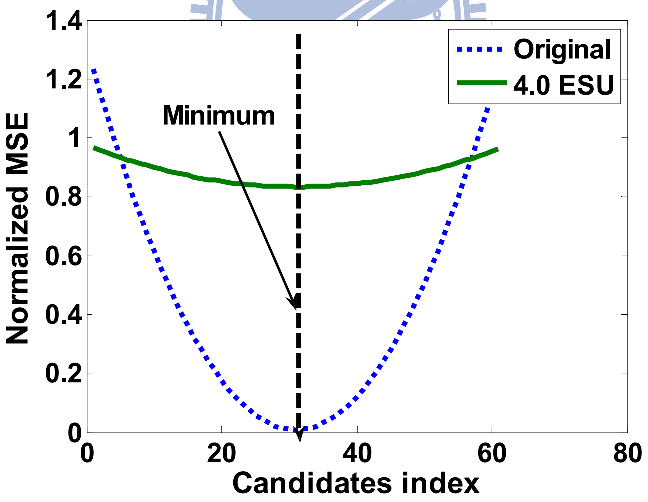

Fig. 25: Good sample mean tested with the convex sets, normalized by $u_c^2(x)/n$, sample size is 15, 4 combined quantities

Fig. 26 and Fig. 27 show respectively the results for the two bad cases of biased Z score to be less than -2 and greater than 2 when applying the Q2MMSE-CLT and enlarging standard uncertainty Q2MMSE-CLT (ESQ2MMSE-CLT). We plot the details shown as the double y-axes representation in which the dash line represents the original signal evaluated by Q2MMSE-CLT and the solid line represents the signal evaluated by ESQ2MMSE-CLT with 4 times of combined standard uncertainty. An important fact is found from these two figures that the original signal will be affected by the sample mean if it only takes the Q2MMSE-CLT operations. The resulting MSE curves converge to the near symmetric location which is the sample mean, but we know it is a bad sample mean. We also found from these two figures that, as we apply the ESQ2MMSE-CLT algorithm with 4 times of combined standard uncertainty, the MSE curves converge to locations deviated away from the bad sample mean and toward the true population mean. Why does it act like this as the action? The reason is that the ESQ2MMSE-CLT enlarges the combined standard uncertainty to 4 times of the original signal. Thus the Z score of the general maximum bias sample mean will be reduced to 25% of that of the original signal. It means that the Z score of bias is constrained to $-0.5 \le Z \le 0.5$. This in turn will guarantee the convergence to the good sample mean (also the population mean) as shown in Fig. 25,

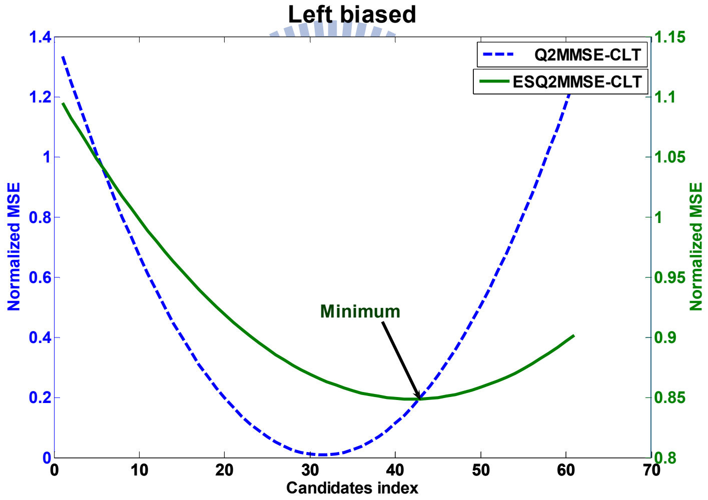

Fig. 26: Originally left biased of bad sample mean tested with the convex sets, double y-axes, normalized by $u_c^2(x)/n$, sample size is 15, 4 combined quantities

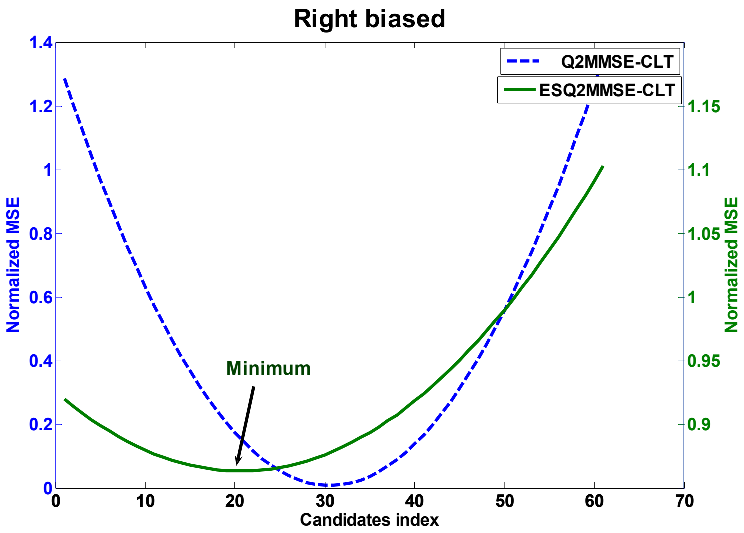

Fig. 27: Originally right biased of bad sample mean tested with the convex sets, double y-axes, normalized by $u_c^2(x)/n$, sample size is 15, 4 combined quantities

Fig. 28 displays the refined results of ESQ2MMSE-CLT for sample size from 11~40. We find from the figure that ESQ2MMSE-CLT significantly outperforms the sample mean by 40% MSE reduction. So it is a promising mean estimator.

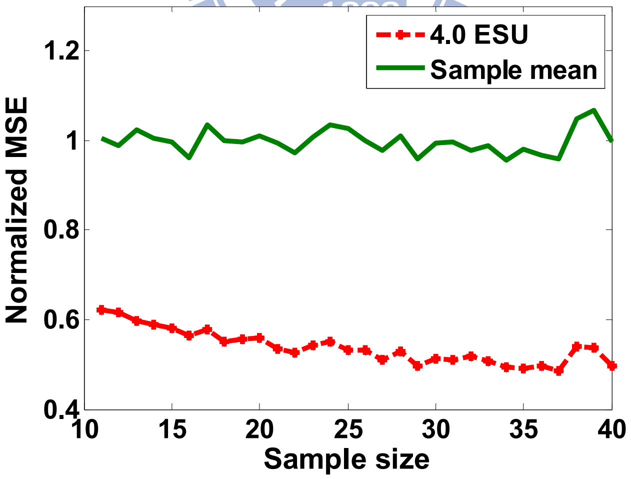

Fig. 28: Refined Q2MMSE-CLT with the enlarging standard uncertainty, y-axis is normalized by $u_c^2(x)/n$, 4 combined quantities

6.3 Change the Variable to Obtain a Nonlinear Estimator for Mean Estimation

We now derive a new nonlinear equations for variable u from Eq.(4-8). By letting $\overline{x} = u + \Delta h$ and $\overline{x^2} = S_n^2 + \overline{x}^2$. Then Eq.(5-15) becomes

$$\begin{aligned}&\sum_{t=1}^{v}(D_t)u^2+\left\{\sum_{t=1}^{v}\left\{D_t\left((-\Delta h)\left(E_{\xi_{p:n}|c=Cc_t,\mathrm{E}\{r_s\},n}\left\{\xi_{p:n}^2\right\}\right)-2x_{p:n}\right)\right\}\right\}u\\&\quad+\sum_{t=1}^{v}\ (D_t(x_{p:n}^2+\left(\Delta h\cdot x_{p:n}-S_n^2-\Delta h^2\right)\times\left(E_{\xi_{p:n}|c=Cc_t,\mathrm{E}\{r_s\},n}\left\{\xi_{p:n}^2\right\}\right)))=0\end{aligned}\tag{6-7}$$

where Δh is the sample mean bias relative to the population mean μ.

Eq.(6-7) is a quadratic equations for variable μ. So, we can easily obtain the estimated mean u^*. Because the variance, S_n^2, is a function of μ, we can regard Eq.(6-7) as the nonlinear equation. If we set Δh equal zero for Eq.(6-7), and the new equation is changed to:

$$\begin{aligned}&\mu_p^2-2x_{p:n}\cdot\mu_p+\xi_{p:n}^2\cdot(x_{p:n}^2-S_n^2)=0,\ \ p=1\ \ or\ \ p=n\\&\mu^*=\frac{1}{2}(\mu_1+\mu_n)\end{aligned}\tag{6-8}$$

$$\xi_{1:n}=(\Phi^{-1}(\frac{1}{n+1})),\qquad \xi_{n:n}=(\Phi^{-1}(\frac{n}{n+1}))\tag{6-9}$$

where $\Phi(.)$ is the *cdf* of $N(0,1^2)$, and S_n^2 is the variance of input sequence. Let us test the combined quantities of four input signals expressed by

$$x=z_1+z_2+z_3+z_4\tag{6-10}$$

We perform 10,000 trials for two cases of combined quantities composing of four different distributions. One has $z_1\sim N(0.1,1^2)$, $z_2\sim N(0.2,1.5^2)$, $z_3\sim rec[-2\sqrt{3}+0.15,2\sqrt{3}+0.15]$, and $z_4\sim rec[-10\sqrt{3}-0.1,10\sqrt{3}-0.1]$. Its UR is equal to 3.7 evaluated according to Eq.(2-25). Another is the same as the first case except that $rec[-10\sqrt{3}-0.1,10\sqrt{3}-0.1]$ is changed to $rec[-28\sqrt{3}-0.1,28\sqrt{3}-0.1]$ and the new UR is changed to 10.4. The experimental results are shown in Fig. 29. We find from the figure that the nonlinear mean estimator outperforms the sample-mean estimator for both cases. Moreover, it performs better for the case of large UR(=10.4). MSE reductions of about 30% and 70% were achieved for the two cases of UR=3.7 and 10.4, respectively.

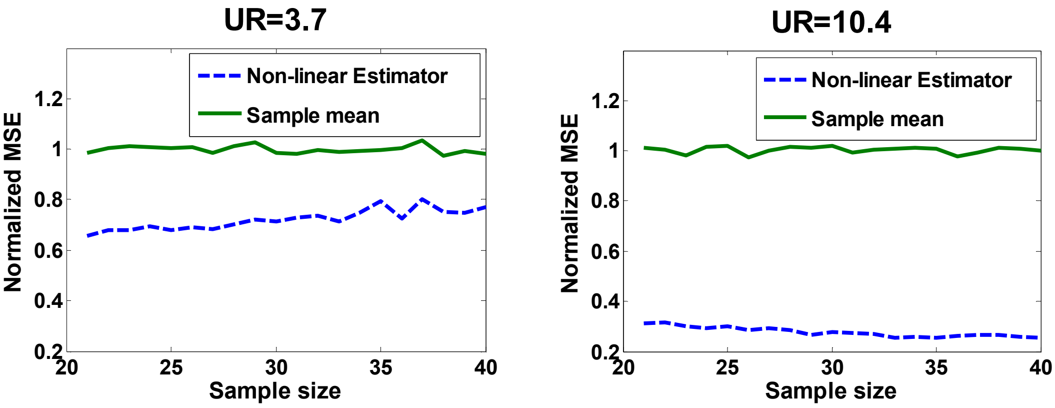

Fig. 29: Nonlinear estimators compared to the sample mean estimator for different uncertainty ratio (UR), y axis is normalized to the combined standard uncertainty, $u_c^2(x)/n$

6.4 Conclusions

In this chapter, the issue of applying quantile-based maximum likelihood estimation (QMLE) to mean value estimation of normally-distributed signal in sparse data condition is addressed. It proposes to incorporate order statistics into QMLE to take the maximum coverage as quantiles so as to conform to the requirement of symmetric quantiles. Simulation results confirm that the new Q2MMSE-CLT performs very well to outperform the conventional sample mean estimator. The proposed Q2MMSE-CLT reaches the highest gain when the combined mean is known and obtains the least benefit if we take the sample mean to substitute for the combined mean. In spite of the fact, ESQ2MMSE-CLT can compensate this shortcoming. The robustness of ESQ2MMSE-CLT to its usage of sample mean makes it a promising mean estimator for practical applications.

It is worthy to note that Q2MMSE-CLT is free to the standard uncertainty of population. The standard uncertainty of combined quantities can therefore be ignored and replaced with the samples' standard uncertainty in the estimation process. We also find that the nonlinear mean estimator solved from Eq.(6-8) outperforms all other estimators when UR is high.

6.5 Appendix: Derivation of the Quantile-based Mean Estimator

By substituting $\mu = x_{p:n} - \sigma\xi_{p:n}$, for $p = 1$ or n, into $QMLE(\mu,\sigma)$ defined in Eq.(6-1), we obtain

$$QMLE(\mu,\sigma)=(-\frac{n}{2}\log 2\pi-n\log\sigma)-\frac{1}{2}\sum_{i=1}^{n}(\frac{x_i-x_{p:n}}{\sigma})^2 \\ -\sum_{i=1}^{n}\frac{x_i-x_{p:n}}{\sigma}\cdot\xi_{p:n}-\frac{n}{2}\xi_{p:n}^2 \tag{6-11}$$

Taking the partial derivative of Eq.(6-11) with respect to σ and setting it to zero, we obtain

$$n\sigma^2-\xi_{p:n}\sum_{i=1}^{n}(x_i-x_{p:n})\sigma-\sum_{i=1}^{n}(x_i-x_{p:n})^2=0 \tag{6-12}$$

Solving Eq.(6-12) to obtain an estimate of the standard deviation of population:

$$\sigma^*=\frac{B_\sigma\pm\sqrt{\left(B_\sigma\right)^2+4nC_\sigma}}{2n} \tag{6-13}$$

where $B_\sigma=\xi_{p:n}\sum_{i=1}^{n}(x_i-x_{p:n})$, $C_\sigma=\sum_{i=1}^{n}(x_i-x_{p:n})^2$, and $\sigma^*>0$.

Chapter 7 An Efficient Representation for Combined Signal Activity Detection in Sparse Data Condition

7.1 The Simplified Quantile-based Mean Estimator

We have shown that the proposed VTNJE is an efficient mean estimator in the previous chapter. But, its computational cost is too high. Now we proposed a new mean estimator for the combined signals. The idea is to keep using the same QMI principle and to inspect its efficiency in terms of the representation of the maximum eigenvalue which has been reviewed in Sub-section 2.6. We continue the work of [41] to employ the simplified representation of UBE and developed a new algorithm to against the uncertainty increasing on the sparse data condition such as Eqs. (2-27) - (2-28). The study considers that the correlation matrix, $R_{mxm} = \left[r_{ij} \right]$, is estimated from n observed samples of m-dimensional random vector by

$$r_{ij} = \frac{c_{ij}}{\sigma_i \sigma_j} \quad \text{for } 1 \le i, j \le m, \tag{7-1}$$

where $C = \left[c_{ij} \right] = \frac{1}{n-1}(X_{m\times n} - U_{m\times n})(X_{m\times n} - U_{m\times n})^T$ is the sample covariance of the observed *i.i.d.* random vectors $\mathbf{x}_i$, for $1 \le i \le n$; $X_{m\times n} = [\mathbf{x}_1\ \mathbf{x}_2\ \cdots\ \mathbf{x}_n]$; $U_{m\times n} = [\mathbf{u}\ \mathbf{u}\ \cdots\ \mathbf{u}]$ is the mean matrix with identical column vector $\mathbf{u} = E[\mathbf{x}_i]$; and σ_j is the standard deviation of the j-th component of $\mathbf{x}_i$. The conventional approach to mean matrix estimation is by the sample mean method. Since the variance of sample mean is known to increase as the sample size decreases, the resulting mean matrix estimate is hence unreliable in the sparse data condition. This will make the uncertainty of the estimated correlation matrix increase accordingly; which in turn affects the UBE finding.

In accordance with the discussions in Section 6.5, we have suggested a new approach for the mean estimation. Due to the fact that symmetric quantiles are efficient [36], we use QSQ to form a QMLE-based mean estimator by

$$QMLE\text{-}QSQ = \frac{1}{2}(\mu_1^* + \mu_n^*) \tag{7-2}$$

To use Eq.(7-2), two problems are still needed to be solved. One is that Eq.(6-2) is derived based on the assumption of ideal additive mixture signal with normal distribution, while the realistic signal is QSAW. Another is that the transform-domain QSQ,i.e. , $\xi_{1:n}$ and $\xi_{n:n}$, are unknown.

To solve the first problem, we define the match pair (MP) for the QMLE analysis. From Fig. 30, we find that the two extreme *pdf*s of the QSAW signal are normal and rectangular distributions. We hence define MP as the following two *pdf*s with the same mean μ and standard deviation σ by

$$N(\mu, \sigma^2) \leftrightarrow \text{Rect}[-\sqrt{3}\sigma + \mu, \sqrt{3}\sigma + \mu] \tag{7-3}$$

Now we use order statistics to express the *pdf* of QSQ [61].

$$f_{x_{k:n}}(x_{k:n}) = \frac{n!}{(k-1)!(n-k)!}(F_x(x_{k:n}))^{k-1}(1 - F_x(x_{k:n}))^{n-k} \cdot f_x(x_{k:n}) \tag{7-4}$$

where $f_x(x)$ and $F_x(x)$ are the *pdf* and *cdf* of *x*, and *k* is the order index restricted to *k* =1 or *n*. We plot the *pdf*s of QSQ of the MP in Fig. 30 for the case of μ=0, σ=1, and *n*=11. The two red dash curves represent the *pdf*s of the standard normal QSQ and the two green solid curves represent those with rectangular QSQ. Fig. 30 reveals that the rectangular QSQ are spanned in two smaller areas covered completely by their corresponding standard normal counterparts. Since the *pdf*s of MP are the two extremes of the *pdf* of QSAW, the QSQ of QSAW will also be dispersed in two areas covered by those of the normal QSQ. So, applying the transform-domain QSQ of QSAW to Eq.(6-2), derived based on the assumption of normal distribution, will cause no troubles at all.

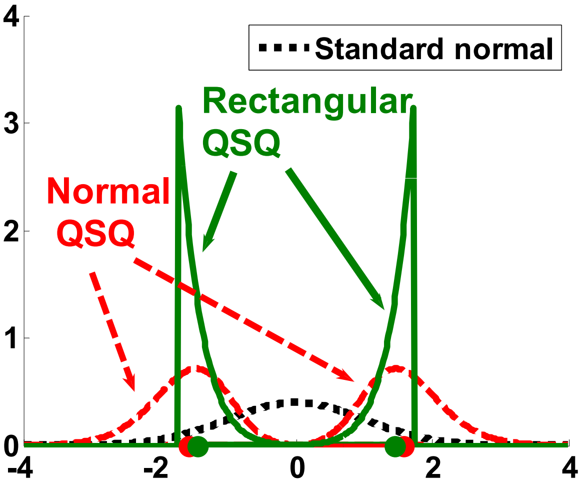

Fig. 30: The *pdf*s of QSQ of the MP pair (i.e., normal and rectangular distributions) for the case of μ=0, σ=1, and n=11. Dots on x-axis are expectation values of QSQ.

The second problem is solved by replacing $\xi_{1:n}$ and $\xi_{n:n}$ with their expectation values. But, calculating the expectation of Eq.(7-4) is still difficult to implement for QSAW signal with *pdf* given in Eq.(2-24). Thus an alternative approach is adopted. Since the expectation values of QSQ for a QSAW signal are located between those of its two extreme MP *pdf*s which are very close to each other (see Fig. 30: green and red dots on the *x*-axis), we can therefore use the expectation values of either rectangular QSQ or normal QSQ to approximate them. Besides, we use an indirect way to calculate the expectation values of rectangular QSQ and normal QSQ. For any *pdf* $f_x(x)$, if we consider to transform the quantile $x_{k:n}$ to its cumulative probability by Eq.(3-8), the distribution of the cumulative probability is subject to $rect[0,1]$. The expectation of the cumulative probability of minimum-order quantile, $p_{1:n}$, can then be easily obtained from Eq.(7-4) by:

$$E_{p_{1:n}|n}[p_{1:n}] = \int_0^1 p_{1:n}\cdot f_{p_{1:n}}(p_{1:n})dp_{1:n} = \int_0^1 p_{1:n}\cdot\frac{n!}{0!(n-1)!}\cdot 1\cdot(1-p_{1:n})^{n-1}\cdot 1\cdot dp_{1:n} = \frac{1}{n+1} \qquad (7\text{-}5)$$

Similarly, the expectation of cumulative probability for the maximum quantile is $E_{p_{n:n}|n}[p_{n:n}] = n/(n+1)$. Since the formulations for these two expectations are distribution-free, we can therefore calculate the expectations of normal QSQ by

$$\bar{\xi}^N_{1:n} = (\Phi^{-1}(\frac{1}{n+1})), \qquad \bar{\xi}^N_{n:n} = (\Phi^{-1}(\frac{n}{n+1})) \qquad (7\text{-}6)$$

and those of rectangular QSQ by

$$\bar{\xi}^{R}_{1:n} = (\Psi^{-1}(\frac{1}{n+1})), \qquad \bar{\xi}^{R}_{n:n} = (\Psi^{-1}(\frac{n}{n+1})) \tag{7-7}$$

where $\Phi(.)$ is the *cdf* of $N(0,1^2)$ and $\Psi(.)$ is the *cdf* of $rect[-\sqrt{3},\sqrt{3}]$.

To justify the feasibility of the scheme of replacing $\xi_{1:n}$ and $\xi_{n:n}$ with their expectation values, the following experiment is conducted. Consider the signal x formed by four independently combined quantities:

$$x = z_1 + z_2 + z_3 + z_4 \tag{7-8}$$

where two input quantities are normally distributed, $z_1 \sim N(0.1,1^2)$ and $z_2 \sim N(2.15,1.5^2)$, and the other two are rectangular distributed, $z_3 \sim \mathrm{rect}[-2\sqrt{3}+0.15, 2\sqrt{3}+0.15]$ and $z_4 \sim \mathrm{rect}[-10\sqrt{3}-0.1, 10\sqrt{3}-0.1]$. We generate 50,000 samples of x to calculate its mean μ and standard deviation σ as the true parameters. We then perform 10,000 trials for each sample size n in the range of 11~40. In each trial, we generate a set of samples and transform the minimal and maximal samples, $x_{1:n}$ and $x_{n:n}$, to produce the true transform-domain quantiles and denote them as $\xi^{T}_{1:n}$ and $\xi^{T}_{n:n}$. We then simulate a pair of $\xi_{1:n}$ and $\xi_{n:n}$ by

$$\xi_{k:n} \sim rect[-0.5VR + \xi^{T}_{k:n}, 0.5VR + \xi^{T}_{k:n}] \tag{7-9}$$

for k=1 or n, and use them in Eqs.(7-5)~(7-9) to generate a *QMLE-QSQ* estimate. Here, VR is the dynamic range with unit of standard deviation of a single quantile of QSQ (denoted as SQSQ STD). We test several cases of VR which shows different degree of deviation of the transform-domain quantiles corresponding to the true ones. Fig. 31 shows the experimental results.

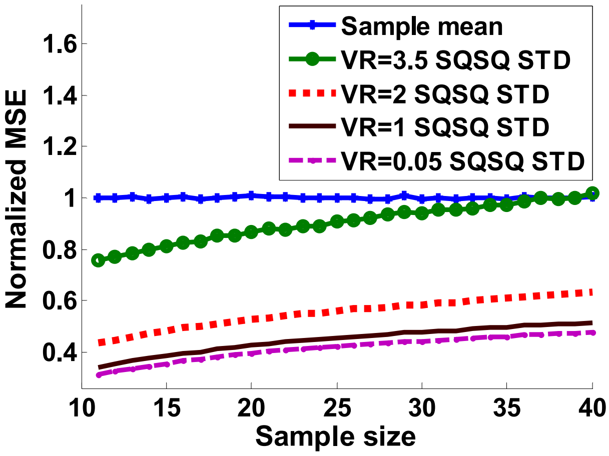

Fig. 31: The performance of *QMLE-QSQ* mean estimation for different degree of deviation of the transform-domain quantile used to the true one. Note that y axis is normalized by $u_c^2(x)/n$.

As shown in Fig. 31, *QMLE-QSQ* outperforms the sample mean estimator if VR is smaller than 3.5 SQSQ STD. The average MSE decreases by about 50-60% if the transform-domain quantile is approximately known, e.g. VR=0.05 SQSQ STD. The improvement gradually degrades as VR increases. These results show that using approximate quantiles in the *QMLE-QSQ* mean estimator will not cause big trouble if they deviate not too far away from the true values. So, the proposed replacement scheme is appropriate.

We then examine the effect of UR variation of the QSAW signal on the performance of *QMLE-QSQ* by simulations. Consider the previous combined signal x shown in Eq.(7-8). Its UR is 3.7 evaluated according to Eq.(2-25). We then simulate another combined signal x' formed by z_1, z_2, z_3, and $z'_4 \sim \text{rect}[-28\sqrt{3}-0.1, 28\sqrt{3}-0.1]$. The UR of x' is 10.4. To test *QMLE-QSQ*, both cases of replacing $(\xi_{1:n}, \xi_{n:n})$ by their expectation values, calculated using Eq.(7-6) based on the normal assumption and Eq.(7-7) based on the assumption of rectangular *pdf* are examined. For these two cases, the *QMLE-QSQ* are calculated, respectively, by

$$\frac{1}{2}(\mu_1^* + \mu_n^*) = \frac{1}{2}\left\{\left(x_{1:n} - \sigma_1^* \overline{\xi}_{1:n}^N\right) + \left(x_{n:n} - \sigma_n^* \overline{\xi}_{n:n}^N\right)\right\} \tag{7-10}$$

$$\frac{1}{2}(\mu_1^* + \mu_n^*) = \frac{1}{2}\left\{\left(x_{1:n} - \sigma_1^* \overline{\xi}_{1:n}^R\right) + \left(x_{n:n} - \sigma_n^* \overline{\xi}_{n:n}^R\right)\right\}. \tag{7-11}$$

We perform 10,000 trials for each sample size *n* ranging from 11 to 40. The normalized MSEs for UR=3.7 and UR=10.4 are displayed in Fig. 32. It can be found from these two figures that *QMLE-QSQ* significantly outperforms sample mean for both cases of normal and rectangular *pdf* assumptions. Moreover, *QMLE-QSQ* performs better for large UR. This can be explained as follows. As shown in Fig. 30, rectangular QSQ have much smaller variances than their counterparts of normal QSQ. The error caused by the expectation replacement scheme will be smaller for the rectangular *pdf* to make it outperform the normal *pdf* on mean estimation. Since the *pdf* of QSAW signal with larger UR looks more like the rectangular *pdf* (see Fig. 4), it is hence expected to perform better on mean estimation. Moreover, we find that the performances of *QMLE-QSQ* using the two quantile-expectation (QE) replacement schemes, based respectively on standard normal *pdf* and standard rectangular *pdf* assumptions, are almost the same. The reason is obvious because their locations are closed to each other.

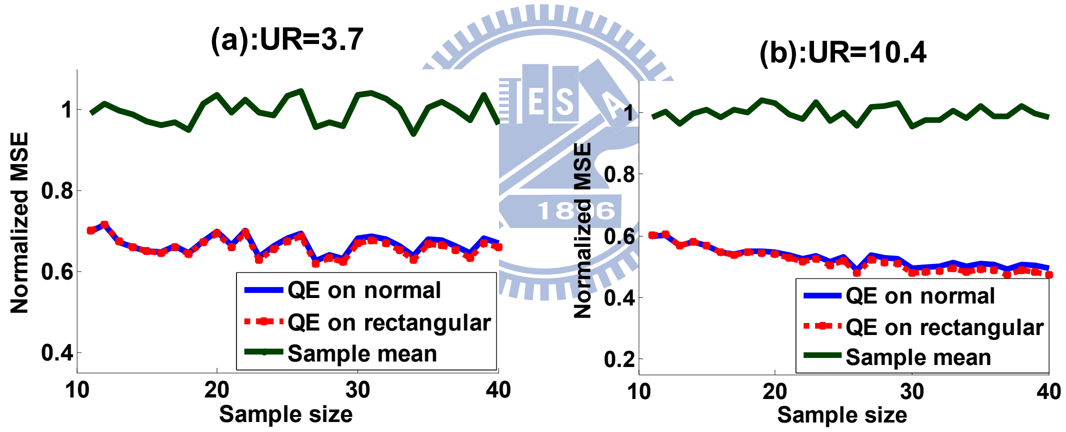

Fig. 32: The mean estimation performance of *QMLE-QSQ* using two different quantile-expectation (QE) replacement schemes for two QSAW signals: (a) UR=3.7 and (b) UR=10.4. The y axis is normalized by $u_c^2(x)/n$. Note that the MSEs of *QMLE-QSQ* and sample mean are $\frac{n}{10000\cdot u_c^2(x)}\sum_{i=1}^{10000}(\frac{\mu_1^*(i)+\mu_n^*(i)}{2}-\mu)^2$ and $\frac{n}{10000\cdot u_c^2(x)}\sum_{i=1}^{10000}(\bar{\mu}-\mu)^2$, respectively.

7.2 Simulation Results for UBE Finding with QMLE-QSQ Mean Estimation

Now, we examine the effect of applying the *QMLE-QSQ* mean estimator on UBE finding. We set *r*=1 in Eq.(2-28) for the first-order UBE evaluation so that it is tighter than the Dembo’s bound which is the case of r=0 [41]. The first-order UBE is the maximal real root of Eq.(2-28). We denote $\eta_m^Q(t)$ as the UBE estimate using the *QMLE-QSQ* mean estimate to perform the mean matrix $U_{m\times n}$, and $\eta_m^s(t)$ as that of using sample mean. Here t denotes the trial index. Define the following five factors to measure the performance of UBE finding:

$$ET = \text{count of } \left\{ t \,\middle|\, \eta_m^s(t) > \eta_m^Q(t) > \max_k\{\varepsilon_k(t)\} \right\} \tag{7-12}$$

$$Yield = ET / T \tag{7-13}$$

$$OD = \frac{1}{ET}\sum_{t=1}^{ET}(\eta_m^s(t) - \max_k\{\varepsilon_k(t)\}) \tag{7-14}$$

$$CD = \frac{1}{ET}\sum_{t=1}^{ET}(\eta_m^Q(t) - \max_k\{\varepsilon_k(t)\}) \tag{7-15}$$

$$IR = (OD - CD) / OD \tag{7-16}$$

where ET denotes the number of effective (or successful) trails; $\varepsilon_k(t), 1 \le k \le m$, are eigenvalues of trial t; “*Yield*” measures the percentage of effective trails; T is the total number of trials; OD and CD denote the average distances from the upper bound to the maximal eigenvalue for effective trails using sample mean and *QMLE-QSQ*, respectively; and IR denotes the improvement factor of the proposed UBE finding method over that using sample mean. We take 100 trials (T=100) for each sample size n ranging from 11 to 40. In each trial, m is set to equal to n, and each row vector of $X_{m\times n}$ is formed by i.i.d. random variables generated by Eq.(7-8) using the same mean and the same UR (in the range of 8~10.4). The values of mean and UR for different row vector are different. The testing procedure spent 8 days on a PC with Intel Pentium 4 CPU run at a clock of 2.84 GHz. The experimental results are displayed in Fig. 33. It can be found from the figure that the *yield* is over 85% for n in the range of 11~40. The corresponding IR is 25% for n=11 and decreases gradually to 13% when n=40. These results show that the proposed *QMLE-QSQ* mean estimator can improve the UBE finding.

1896

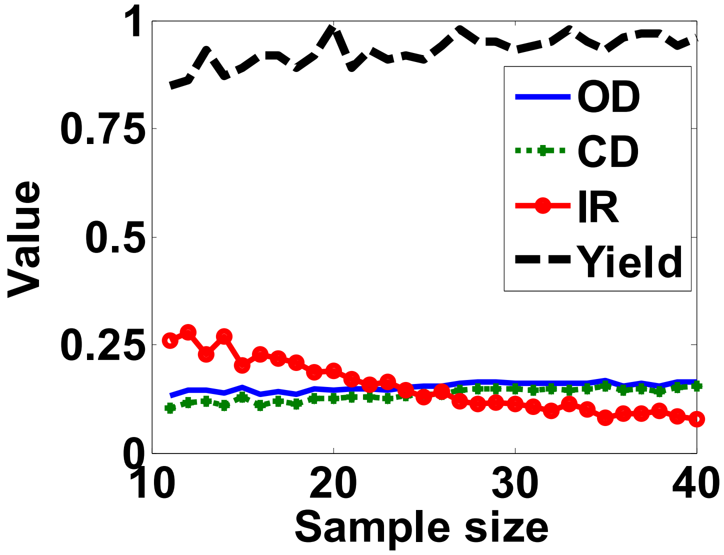

Fig. 33: The performance of UBE finding using the proposed QMLE-QSQ mean estimator

7.3 Conclusions

We measured the combined signal as a result of linear combination of input quantities and have proved that its shape looks like the QSAW in the previous chapter. In this chapter, we propose the QMLE-QSQ to estimate the mean value of QSAW signals and apply it to find the upper bound of eigenvalues for signal activity detection (SAD) in combined signals. The propagation of additive model is appropriate for assuming the QSAW to be quasi-normally distributed whenever its mean value is needed to be estimated. Either Eq.(7-6) or Eq.(7-7) is simple and unique to obtain based on the QMI transform domain working where Fig. 31 demonstrates the sensitivity analysis and Fig. 30 shows the dispersion of QSAW being smaller than 3.5 SQSQ STD. Lastly, the result in Fig. 32 admits the above facts.

This topic related to the signal activity detection brings the issue of UBE approximation from deterministic to stochastic analysis via considering the case that the correlation matrix of signal is estimated from sparse observed sample vectors. A tighter upper bound of eigenvalues can be obtained as the correlation matrix is calculated by using the mean matrix formed by the proposed QMLE-QSQ mean estimates. More reliable correlation matrix can be explained in the system obtained by the *QMLE-QSQ* mean estimation to result in better UBE finding than that by the conventional sample mean estimation.

Chapter 7 Conclusions and Further Works

In this dissertation, we devote to the robust representation of combined signals in sparse data condition in terms of the formulation of JCGM expression and take advantage to the unified *pdf* of coverage interval for uncertainty measurement. However, this unified *pdf* expression of coverage interval shows that the shortest coverage interval is good enough to represent the whole distribution of coverage interval when the *pdf* of population is asymptotically symmetric. Due to the fact that the two endpoints of coverage interval are decided in one step, we reverse the traditional procedure, which finds the endpoints after the mean estimation, to estimate the mean value of population after finding the endpoints of coverage interval. We find that given with an accurate coverage interval is capable of improving the mean estimation by the way of regarding the coverage interval as a result of variably truncated normal distribution. Besides the improvement on the mean estimation, a robust estimation for truncated normal *pdf* is also reached when we take the quantile-based estimation combined with the output of unified *pdf* of coverage interval. The result is better as compared with the model derived by Cohen.

We also use quantile to derive a nonlinear equation for mean estimation. Simulation results demonstrate that it performs well. We last try a novel algorithm, named "The robust statistical principle of minimax optimization", to use the unified *pdf* of coverage interval in mean estimation. It is a convex optimization method for the general mean estimation. The optimization process converges exactly to the true mean direction so that it may be considered as a new search algorithm as well as the steepest gradient descent algorithm without the quadratic object function. Finally, we apply the new mean estimator, QMLE-QSQ, to the application of signal activity detection in terms of finding the upper bound of eigenvalues. We find that the QMLE-QSQ can replace the classical sample mean to obtain a more accurate correlation matrix estimate, which in turn leads to a more efficient representation of the maximum eigenvalue. Thus, our study extend the previous UBE finding studies, which use deterministic correlation matrix, to employ stochastic correlation matrix via introducing the uncertainty of mean estimation on spare data condition. And our solution can obtain better UBE for improving signal activity detection..

Our work still leaves several warm topics about CI which are worthy of studying in the future. For instance, “the shortest CI” should be replaced with “the probably shortest CI” whenever the *pdf* is skew. But we don’t know how the skewness of signal *pdf* affects the *pdf* shape of CI. Secondly, we have proved some properties of endpoints of quantiles based on the QMI principle. They include the structure of left endpoint mapping to the quantile of $1/(n+1)$, the right endpoint mapping to the quantile of $n/(n+1)$, and coverage being equal to $(n-1)/(n+1)$ which is the expectation $E_{c|n}[c]$ shown in Eq.(3-11). Hence, the endpoint-decision with the QMI principle is deterministic. So, this criterion can not support the exploration of the random effects of endpoints. We suspect that the quantile-based mean estimation ought to be suffered from the random effects of the endpoints expression.

E S A 1896

Publication List

Journal Papers

[1] **Wen-Hui Lo** and Sin-Horng Chen, “The analytical estimator for sparse data”, IAENG International journal of applied Mathematics, vol. 39, iss. 1, 2009, pp. 71-81.

[2] **Wen-Hui Lo** and Sin-Horng Chen, “The uncertainty reduction for the refined sample mean of combined quantities,” IAENG International journal of applied Mathematics, vol. 39, iss. 3, 2009, pp. 192-197.

Submitted Papers

[1] **Wen-Hui Lo** and Sin-Horng Chen, “The Probability Distribution Function of Coverage Interval and Its Optimal Expression for Sparse Data Condition,” IEEE Transaction on Instrumentation and Measurement, Revised, 2009.

[2] **Wen-Hui Lo** and Sin-Horng Chen, “An efficient representation for combined signal activity detection in sparse data condition,” IEEE Transaction on Instrumentation and Measurement, Revised, 2010.

Conference Papers

Wen-Hui Lo and Sin-Horng Chen, “The wide-sense parametric coverage estimator against the distribution mismatch problem for sparse data”, World Congress on Engineering, **(The best student paper award)**, London, U.K., 2008, pp.1005-1010.

Wen-Hui Lo and Sin-Horng Chen, “The coverage-based estimator for sparse data”, The 2008 Asian International Workshop on Advanced Reliability Modeling, Taichung, Taiwan., 2008, pp. 382-389.

Wen-Hui Lo and Sin-Horng Chen, “Robust estimation for sparse data”, IEEE 19th International Conference on Pattern Recognition, Tampa Bay, Florida, USA., 2008, pp.1-5.

Wen-Hui Lo and Sin-Horng Chen, “Theoretical and practical realization for the uncertainty measurement by coverage interval,” IEEE International Conference on Instrumentation and Measurement Technology **(The best paper award)**, Singapore, 2009, pp. 1562-1567.

Wen-Hui Lo and Sin-Horng Chen, “The mean estimation of the combined quantities by the asymptotic minimax optimization,” IEEE International Workshop on

Advanced Methods for Uncertainty Estimation in Measurement, Bucharest, Romania, 2009, pp. 63-68.

博士候選人資料

姓名：羅文輝

性別：男

出生年月日：民國58年3月30日

出生地：台灣新竹

學歷：

台灣省立新竹師範專科學校普通科數學組畢業(73年9月～79年6月）

私立逢甲大學交通工程與管理學系暨電子工程輔學系畢業(79年9月～82年6月）

國立台灣大學土木工程研究所交通工程組碩士畢業(82年9月～85年1月）

國立交通大學電機資訊學院碩士在職專班電信學程畢業(88年9月～95年7月）

國立交通大學電信工程研究所博士畢業(90年9月～99年7月）

論文題目：

稀少取樣下之組合式訊號測不準表示法研究與其在訊號平均值估計之應用

An Efficient Representation of Uncertainty Measurement for Combined Signals on Small Sampling Size Condition and its Application to Signal Mean Estimation

References

---

[1] CLSI, Defining, establishing, and verifying reference intervals in the clinical laboratory: approved guideline - third edition. CLSI Document C28-A3. Wayne, PA: Clinical and Laboratory Standards Institute, 2008.

[2] P. Fotowicz, “An analytical method for calculating a coverage interval,” Metrologia, vol.43, 2006, pp.42-45.

[3] S. Nadarajah, “Exact calculation of the coverage interval for the convolution of two Student’s t-distributions,” Metrologia, vol.43, 2006, pp.L21-L22.

[4] S.S. Wilks, “Determination of sample sizes for setting tolerance limits,” Ann. Math. Stat., 1941, pp.1291-1296.

[5] A. Wald, “An extension of Wilks’ method for setting tolerance limits,” Ann. Math. Stat., 1943, pp.1445–1455.

[6] J. K. Patel, “Tolerance limits — a review,” Commun. Stat.—Theory Methods, vol.15, 1986, pp.2719–2762.

[7] Wilks, S. S., “Statistical prediction with special reference to the problem of tolerance limits,” Ann. Math. Statist. vol. 13, 1948, pp. 400-409.

[8] Wilks, S. S., Mathematical Statistics, Wiley, New York, 1962.

[9] S.-H. Lin, W. Chan and L.-A. Chen, “A non-parametric coverage interval,” Metrologia, vol.45, 2008, pp.L1-L4.

[10] H.E. Solberg, “Approved recommendation on the theory of reference values. Part 1. The concept of reference values,” Journal of Clinical Chemistry and Clinical Biochemistry, vol.25, 1987, pp.337-342.

[11] R. Dybkar and H.E. Solberg, “Approved recommendation on the theory of reference values. Part 6. Presentation of observed values related to reference values,” Journal of Clinical Chemistry and Clinical Biochemistry, vol.25, 1987, pp.657-662.

[12] A. Heathcote, S. Brown, and D.J.K. Mewhort, “Quantile maximum likelihood estimation of response time distribution,” Psychonomic Bulletin and Review, vol.9, 2002, pp.394-401.

[13] N. Balarkrishnan and A. Clifford Cohen, Order statistics and inference estimation methods, Academic Press, Inc., 1991.

[14] E. H. Lloyd, “Least-square estimation of location and scale parameters using order statistics,” Biometrika, vol. 39, 1952, pp.88-95.

[15] D. Teichroew, “Tables of expected values of order statistics for samples of size twenty and less from the normal distribution,” The Ann. of Math. Stat., vol. 27, 1956, pp.410-426.

[16] Ahmed E. Sarhan and Bernard G. Greenberg, “Estimation of location and scale parameters by order statistics from singly and doubly censored samples, part one.

The normal distribution up to size 10," The Ann. of Math. Stat., vol. 27, 1956, pp.427-451, (correction , vol. 40, p.325)

[17] Ahmed E. Sarhan and Bernard G. Greenberg. eds., Contributions to order statistics, Wiley, New York, 1962.

[18] A. C. Cohen, "On the solution of estimating equations for truncated and censored samples from normal populations", Biometrika, Vol. 44, No. 1/2, Jun. 1957, pp. 225-236.

[19] V.K. Srivastava, "A note on the estimation of mean in normal population," Metrika, vol.27, 1980, pp.99-102.

[20] B. A. Bowen, "An alternate proof of the central limit theorem for sums of independent processes," Proceedings of the IEEE, vol. 54, iss. 6, 1966, pp. 878-879.

[21] D. T. Searls, "The utilization of a known coefficient of variation in the estimation procedure," Journal of American Statistical Association, vol. 59, 1964, pp. 1225-1226.

[22] L. J. Gleser and J. D. Healy, "Estimating the mean of a normal distribution with known coefficient of variation," Journal of American Statistical Association, vol. 71, 1976, pp, 977-981.

[23] Ashok Sahai, M. Raghunadh and Hydar Ali, "Efficient estimation of normal population mean," Journal of Applied Science, vol. 6, iss. 9, 2006, pp. 1966-1968.

[24] ISO, "Guide to the expression of uncertainty in measurement (GUM)—Supplement1: Numerical methods for the propagation of distributions," International Organization for Standardization, 2004.

[25] R. Koenker and G.J. Bassett, "Regression quantiles," Econometrica, vol. 46, 1978, pp. 33–50.

[26] W. G. Gilchrist, Statistical modeling with quantile function, London: Chapman and Hall/CRC, 2002.

[27] G. GIORGI and C. NARDUZZI "Uncertainty of quantile estimates in the measurement of self-similar processes," IEEE International workshop on advanced methods for uncertainty estimation in measurement, AMUEM 2008, pp.78-83.

[28] L.-A. Chen and Y.-C. Chiang, "Symmetric type quantile and trimmed means for location and linear regression model," Journal of Nonparametric Statistics, vol. 7, 1996, pp. 171–185.

[29] W.-H. Lo and S.-H. Chen, "The analytical estimator for sparse data," International Association of Engineers (IAENG) Journal of Applied Mathematics, vol. 39, iss. 1, 2009, pp. 71-81.

[30] W.-H Lo and S.-H. Chen, "Robust estimation for sparse data," The 19-th international conference on pattern recognition, 2008, pp.1-5.

[31] Y. Zeng, C.-L. Koh and Y.-C. Liang, "Maximum eigenvalue detection: theory and application," IEEE inter. conf. comm., ICC 2008, pp. 4160-4164.

[32] E. Candès and M. Wakin, "An introduction to compressive sampling [A sensing/sampling paradigm that goes against the common knowledge in data acquisition]," IEEE Signal Processing Mag. , vol. 25, iss. 2, 2008, pp. 21-30.

[33] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, Evaluation of measurement data-supplement 1 to the 'guide to the expression of uncertainty in measurement'-propagation of distributions using a Monte Carlo method in preparation (Joint Committee for Guides in Metrology (JCGM), 2008.

[34] J. Letessier, B. Vrigneau, P. Rostaing and G. Burel, “New closed-form of the largest eigenvalue PDF for Max-SNR MIMO system performances,” IEICE Trans. vol. E91-A, no. 7, 2008, pp. 1791-1796.

[35] E. M. Ma. and C.J. Zarowski, “On lower bounds for the smallest eigenvalue of a Hermitian matrix,” IEEE Trans. Inform. Theory, vol. 41, no. 22, 1995, pp. 539-540.

[36] Y.-C. Chiang, L.-A. Chen and H.-C., P. Yang, “Symmetric quantiles and its application,” Journal of applied Statistics, vol. 33, 2006, pp.807-817.

[37] T. Taniguchi, S. Sha, and Y. Karasawa, “Largest eigenvalue analysis in correlated MIMO rayleigh channels,” 2nd Eurepean conf. on antennas and propagation (EuCAP), 2007, pp.1-5.

[38] A. Taparugssanagorn and J. Ylitalo , “Characteristics of short-term phase noise of MIMO channel sounding and its effect on capacity estimation,“ IEEE Trans. Instrum. Meas., vol. 58, no. 1, 2009, pp. 196-201.

[39] J.Q. Zhang and S.J. Ovaska, “ADC characterization based on singular value decomposition,” IEEE Trans. Instrum. Meas., vol. 51, no. 1, 2002, pp. 138-143.

[40] H. Wu, M. Siegel, and P. Khosla, “Vehicle sound signature recognition by frequency vector principal component analysis,” IEEE Trans. Instrum. Meas., vol. 48. no. 5, 1999, pp. 1005-1009.

[41] D. Park and B. G. Lee, “On determining upper bounds of maximal eigenvalue of Hermitian positive-define matrix ,” IEEE Signal process. Lett., vol.10, no. 9, 2003, pp. 267-269.

[42] A. M. Zoubir and B. Boashash, “The bootstrap and its application in signal processing,” IEEE Signal Processing Mag. , vol. 15, iss. 1, 1998, pp. 56-76.

[43] A. Denguir-Rekik, G. Mauris, and J. Montmain, “Propagation of uncertainty by the possibility theory in Choquet integral-based decision making: application to an e-commerce website choice support,” IEEE Trans. Instrum. Meas., vol. 55, no. 3, 2006, pp. 721-728.

[44] A. Ferrero and S. Salicone, “The random-fuzzy variables: A new approach to the expression of uncertainty,” IEEE Trans. Instrum. Meas., vol. 53, no. 5, Oct. 2004, pp. 1370–1377.

[45] L.-A. Chen and H.-N. Hung, “Extending the discussion on coverage intervals and statistical coverage intervals,” Metrologia, vol.43, 2006, pp.L43-L44.

[46] JCGM, Evaluation of measurement data —supplement 1 to the guide to the expression of uncertainty in measurement— propagation of distributions using a Monte Carlo method, Joint Committee for Guides in Metrology, final draft, 2006.

[47] L.-A. Chen, J.-Y. Huang and H.-C. Chen, “Parametric coverage interval,” Metrologia, vol.44, 2007, pp.L7-L9.

[48] H. A. David, *Order statistics*, 2nd ed., Iowa State University, 1981, p.9.

[49] R. Willink, “Coverage intervals and statistical coverage intervals,” Metrologia, vol.41, 2004, pp.L5-L6.

[50] ISO 3534-1, Statistics—vocabulary and symbols—part 1: probability and general statistical terms, Geneva: International Organization for Standardization, 1993.

[51] R. B. Murphy, “Non-parametric tolerance limits,” The Ann. Math. Stat., vol.19, no.4, 1948, pp.581-589.

[52] G. D. Faulkenberry and D. L. Weeks, “Sample size determination for tolerance limits,” Technometrics, vol.10, 1968, pp.343-348.

[53] P. Abbott, “Tricks of the Trade: Legendre-Gauss Quadrature”, Mathematica Journal, vol.9, 2005, pp.689-691.

[54] A. Clifford Cohen, Truncated and censored samples — theory and applications, Marcel Dekker, New York, 1991, pp.31-43.

[55] Lin-An Chen and Hui-Nien Hung, "Extending the discussion on coverage intervals and statistical coverage intervals," Metrologia, vol. 43, 2006, pp.L43-L44.

[56] John W. Pratt and Jean D. Gibbsons, Concepts of nonparametric theory, Springer series in Statistics, Spring Verlag, 1981.

[57] "The NIST reference on constants, units, and uncertainty," available: http://physics.nist.gov/cuu/Uncertainty/index.html.

[58] P. Fotowicz, "A method approximation of the coverage factor in calibration," Measurement, vol.35, 2004, pp.251-256.

[59] E. Parzen, "Nonparametric statistical data modeling," Journal of the American Statistical Association, vol.74, 1979, pp.105-121.

[60] Peter J. Huber, "Robust statistics: a review," Ann. Math. Stat. vol. 43, No. 4, 1972, pp.1041-1067.

[61] Ahmed E. Sarhan and Bernard G. Greenberg., Contributions to order statistics, Wiley, New York, 1962.