Maximum Likelihood Estimation Based Robust Chinese Remainder Theorem for Real Numbers and Its Fast Algorithm

Wenjie Wang, *Member, IEEE*, Xiaoping Li, Wei Wang, and Xiang-Gen Xia, *Fellow, IEEE*

***Abstract*—Robust Chinese remainder theorem (CRT) has been recently investigated for both integers and real numbers, where the folding integers are accurately recovered from erroneous remainders. In this paper, we consider the CRT problem for real numbers with noisy remainders that follow wrapped Gaussian distributions. We propose the maximum-likelihood estimation (MLE) based CRT when the remainder noises may not necessarily have the same variances. Furthermore, we present a fast algorithm for the MLE based CRT algorithm that only needs to search for the solution among L elements, where L is the number of remainders. Then, a necessary and sufficient condition on the remainder errors for the MLE CRT to be robust is obtained, which is weaker than the existing result. Finally, we compare the performances of the newly proposed algorithm and the existing algorithm in terms of both theoretical analysis and numerical simulations. The results demonstrate that the proposed algorithm not only has a better performance especially when the remainders have different error levels/variances, but also has a much lower computational complexity.**

***Index Terms*—Chinese remainder theorem (CRT), phase unwrapping, residue number system, robustness.**

I. Introduction

THE traditional Chinese remainder theorem (CRT) is to reconstruct a positive integer from its remainders modulo a series of integer moduli, which has tremendous applications [4], [7], [17], [21]. In some applications, such as, phase unwrapping in radar imaging in [28], real numbers need to be reconstructed from their remainders. One obvious way to connect the real number reconstruction problem with the integer reconstruction problem from their remainders is to reconstruct its integer part from the remainders using the CRT and then the fractional part can be any fractional number between -0.5 and 0.5. However, it may not be optimal, in particular when the remainders have errors in practice.

For the traditional CRT when all the moduli are pair-wisely co-prime, it is not robust in the sense that a small error in its remainders may cause a large reconstruction error. Recently robust CRT has been studied in [5], [12]–[14], [25], [28]–[30], [32], where it basically says that when all the moduli have a common divisor M, if the remainder errors are within the range of $M/4$, a robust CRT for integers is possible and it is also generalized to real numbers in [25], and an improved version for integers called multi-stage robust CRT is found in [29]. The basic idea of the robust CRT for integers and reals is to accurately determine the unknown folding integers from the erroneous remainders. A different probabilistic approach to deal with noises in CRT is proposed in [23], where the prime moduli are required. A lattice based method is proposed in [15] to estimate a real unknown distance using the phase measurements taken at multiple co-prime wavelengths. There are many applications of robust CRT, see, for example, [3]–[6], [8]–[15], [20], [22]–[34].

In this paper, we use the approach recently studied in [5], [12]–[14], [25], [28]–[30], [32], i.e., uniquely recover the folding integers from noisy remainders. Although real number reconstruction has been considered in [25] from noisy real-valued remainders, it may not be optimal in general, i.e., it is not the maximum-likelihood estimation (MLE) when the remainder noises have different variances, which is usually the case since their remainder noise variances may be proportional to the moduli [8], [15], [24]. In this paper, the MLE is proposed, where the remainder noises may not necessarily have the same variance. We prove that the MLE can be obtained by only searching for the optimal among L elements, where L is the number of remainders and therefore it has a fast algorithm. As it is the MLE, compared with [25], it has a better performance, due to its fast algorithm, it also has a much lower computational complexity. Another contribution of this paper is that we obtain a necessary and sufficient condition on the remainder errors for the MLE to be robust, which is weaker than what we have previously obtained in [25].

The remaining of this paper is organized as follows. In Section II, we first recall the basics of robust CRT for both integers and reals. In Section III, we present the MLE and its fast algorithm. In Section IV, we present a necessary and sufficient condition for the MLE to be robust. We then calculate the probability of the robust MLE CRT. Lastly, in Section V, we present some simulation results to verify the obtained theory.

Manuscript received October 18, 2014; revised January 04, 2015 and February 27, 2015; accepted February 28, 2015. Date of publication April 21, 2015; date of current version May 29, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Antonio Napolitano. This work was supported in part by the NSFC (Nos. 61172092, 61302069), the Research Fund for the Doctoral Programs of Higher Education of China (No. 20130201110014), and the Air Force Office of Scientific Research (AFOSR) under Grant FA9550-12-1-0055.

Wenjie Wang and X. Li are with the MOE Key Lab for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China (e-mail: wjwang@xjtu.edu.cn; lixiaoping@stu.xjtu.edu.cn).

Wei Wang is with the College of Information Engineering, Tarim University, Alar, Xinjiang 843300, China (e-mail: wangwei.math@gmail.com).

X.-G. Xia is with Xi'dian University, Xi'an, China, and also with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716 USA (e-mail: xxia@ee.udel.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2015.2413378

II. Maximum Likelihood Estimation Based Robust CRT for Real Numbers

In this section, we first briefly recall the basics of robust CRT for both integers and reals. Then we propose the MLE based robust CRT for real numbers, where the remainder errors are assumed to follow wrapped normal distributions with zero mean and possibly different variances.

A. Problem of Interest, CRT, and Robust CRT

Let N be a real number, $M_1, \ldots, M_L$ be L positive integers called moduli with $0 < M_1 < \cdots < M_L$, and $r_1, \ldots, r_L$ be L remainders of N modulo M_i as follows:

$$r_i \equiv N \mod M_i \text{ or } N = n_i M_i + r_i, \tag{1}$$

where $0 \le r_i < M_i$, denoted by $r_i = \langle N \rangle_{M_i}$, n_i are unknown integers called folding integers for $i = 1, \ldots, L$. This remaindering problem for a real number N only makes sense when the folding numbers n_i are integers in (1). Clearly, the traditional remaindering problem for an integer N is a special case here. After saying so, when N and M_i are not small, reals and integers N and r_i can approximate each other well and it may be more convenient to study one of the integer and real remaindering problems than to study the other. Furthermore, the above real number reconstruction problem may occur in some applications, such as phase unwrapping in radar imaging [28] as mentioned in Introduction.

If all the moduli M_i are co-prime and N is a positive integer less than the product of the moduli, then integer N can be uniquely reconstructed by CRT. If the moduli are non-pairwise co-prime, integer N can also be uniquely reconstructed by an extended CRT if and only if N is less than the least common multiple (lcm) of all the moduli [18].

The problem we are interested in this paper is to robustly and optimally recover a real number N from its erroneous remainders, where remainders are real numbers with errors, denoted as $\hat{r}_i$, i.e.,

$$\hat{r}_i = r_i + \Delta r_i,\ i = 1, \ldots, L, \tag{2}$$

Δr_i denote errors and are independent each other. In order to resist errors, we consider a special remainder redundancy, where the gcd of all the moduli M_i is larger than 1 and the remaining integers factorized by the gcd of all M_i are co-prime. The robust remaindering problem is how to robustly estimate N from the erroneous remainders $\hat{r}_i$ modulo M_i, which has many applications in engineering, see, for example [10], [11], [27], [33].

The basic idea of the robust CRT for integers and reals in the recent studies [12]–[14], [25], [28], [29], [32] is to accurately determine the unknown folding integers n_i in (1), which may cause large errors in the reconstruction if they are not correctly determined. Hence, the problem is transformed to determine the folding integers from these noisy remainders. Once n_i are determined, the unknown real number N can be estimated as [25]:

$$\hat{N} = \frac{1}{L}\sum_{i=1}^{L}(n_i M_i + \hat{r}_i) = N + \frac{1}{L}\sum_{i=1}^{L}\Delta r_i. \tag{3}$$

If N is an integer, its estimate is then a rounded integer of $\hat{N}$ in (3).

In order to determine the unknown folding integers n_i, a searching based robust algorithm is proposed in [28], followed by [12]–[14]. It is proved in [14] that if the remainder error bound τ is smaller than a quarter of M, i.e., $\tau < M/4$, where M is the greatest common divisor (gcd) of all the moduli, then n_i can be accurately determined. Therefore, we know from (3) that the estimation error of N is bounded by τ, i.e.,

$$\left|\hat{N} - N\right| \le \tau. \tag{4}$$

In addition, a fast searching algorithm was proposed in [14], where the number of searches is sharply reduced. However, the computational complexity is still high when the number of moduli is large. A closed-form robust CRT and its improved version for both integers and reals are proposed to estimate the folding integers n_i in [25], where the remainder difference operation is used and no searching is needed. It is noted that the reference remainder is arbitrarily selected for the closed-form robust CRT. By using the statistically optimal selection process of the reference remainder, the probability of the successful robust estimation is improved greatly [25]. Most recently, a more simpler form of the robust CRT for integers is proposed in [30], where the estimate of a positive integer N is obtained from the erroneous remainders $\hat{r}_i$ directly.

The above algorithms and conclusions are based on the assumption that the remainder errors have the same level/variance, where the error absolute values are within τ, i.e.,

$$|\hat{r}_i - r_i| \le \tau \text{ and } 0 \le \hat{r}_i \le M_i - 1. \tag{5}$$

This may not be practical in applications since different quantizations (different moduli) may contain different noise levels. Take the distance measurement (ranging) system as an example. For the phase measurement, which is interpreted as a distance measurement, the associated noise is proportional to the related wavelength [8], [15], [24]. In this paper, we assume that for each i, the error Δr_i follows a wrapped normal distribution with mean zero and variance σ_i^2, and the variances may be different from each other and may be related to the moduli. In the robust CRT for reals in [25], the remainder noises are assumed to follow Gaussian distribution with the same variance. It is noted that the optimal reference remainder that is subtracted from the other remainders is based on the reference common remainder, which can only be appropriately determined in the case when the noise variances are the same. When the remainder noises have different variances, the method of determining the reference common remainder in [25] may be ineffective, which may lead to a wrong selection of the optimal reference remainder. More details can be seen later. Motivated by this, we propose the MLE based robust CRT and its fast algorithm, where the estimate of the common remainder [18] is optimal from the erroneous remainders with possibly different error variances. To describe it, we begin with the congruence problem proposed in (1).

Let M be the gcd of all the moduli M_i, and let $M_i = M\Gamma_i$, where $\Gamma_1 < \cdots < \Gamma_L$, are assumed relatively co-prime, i.e., for $i \ne j$, $\gcd(\Gamma_i, \Gamma_j) = 1$. Define $\Gamma \triangleq \Gamma_1 \cdots \Gamma_L$. Let $\gamma_i = \Gamma_1 \cdots \Gamma_{i-1}\Gamma_{i+1} \cdots \Gamma_L = \Gamma/\Gamma_i$ for $i = 1, \ldots, L$, and let the modular multiplicative inverse of γ_i modulo Γ_i be $\bar{\gamma}_i$, i.e.,

$$\bar{\gamma}_i \gamma_i \equiv 1 \mod \Gamma_i \text{ or } \bar{\gamma}_i \gamma_i + k\Gamma_i = 1 \text{ for some } k \in \mathbb{Z}, \tag{6}$$

where $\mathbb{Z}$ denotes the set of integers. From (1), we have

$$r_i \equiv N \mod M,\ i = 1, \ldots, L.$$

That is, all remainders r_i modulo M have the same value, named common remainder [18], denoted as r^c. Let

$$q_i = (r_i - r^c)/M,\ i = 1, \ldots, L, \tag{7}$$

and let $N_0 = (N - r^c)/M$. Subtracting r^c and then dividing M from congruence (1), we have

$$q_i \equiv N_0 \mod \Gamma_i,\ i = 1, \ldots, L. \tag{8}$$

According to the classical CRT formula, N_0 can be uniquely reconstructed as

$$N_0 \equiv \sum_{i=1}^{L} \bar{\gamma}_i \gamma_i q_i \mod \Gamma \tag{9}$$

if and only if $0 \leq N_0 < \Gamma$. Therefore, N can be uniquely reconstructed by

$$N = MN_0 + r^c. \tag{10}$$

B. Maximum Likelihood Estimation Based Robust CRT for Reals

Before considering the MLE problem, we first introduce a circular distance function that is useful for the following derivation. For real numbers x and y, the circular distance of x to y for a non-zero positive number C is defined as

$$d_C(x, y) \triangleq x - y - k_0 C \tag{11}$$

where

$$k_0 = \left[\frac{x - y}{C}\right], \tag{12}$$

and $[\cdot]$ stands for the rounding integer, i.e., for any $x \in \mathbb{R}$, where $\mathbb{R}$ denotes the set of all reals, $[x]$ is an integer and subject to

$$-\frac{1}{2} \leq x - [x] < \frac{1}{2}. \tag{13}$$

It is not hard to see that for any real x and integer k, we have

$$[x + k] = [x] + k. \tag{14}$$

Note that in the above, the non-absolute-valued $d_C(\cdot, \cdot)$ is used for convenience but $|d_C(\cdot, \cdot)|$ is in fact the distance effectively involved in the following optimizations and is the real version of the Lee distance [2]. For any real numbers x, y, z and C, where $C > 0$, we have the following properties, which are not hard to see.

Property 1

$$d_C(x, y + kC) = d_C(x, y) \tag{15}$$

Property 2

$$d_C(x + z, y + z) = d_C(x, y) \tag{16}$$

Property 3

$$d_C(x, y) = d_C(d_C(x, z), d_C(y, z)) \tag{17}$$

Property 4

$$-C/2 \leq d_C(x, y) < C/2. \tag{18}$$

Property 5 If $-C/2 \leq x - y < C/2$, then

$$d_C(x, y) = x - y \tag{19}$$

Property 6

$$|d_C(x, y)| \leq |x - y|. \tag{20}$$

We now consider the erroneous remainders described by (2). According to [16], the probability density function (pdf) of the wrapped normal distribution with mean r_i and variance σ_i^2 for $\hat{r}_i$ for a given N is

$$f_{\hat{r}_i}(x|N) = \frac{1}{\sqrt{2\pi}\sigma_i} \sum_{k=-\infty}^{+\infty} \exp\left\{-\frac{(x - r_i + kM\Gamma_i)^2}{2\sigma_i^2}\right\},\ 0 \leq x < M\Gamma_i.$$

That is,

$$f_{\hat{r}_i}(x|N) = \frac{1}{\sqrt{2\pi}\sigma_i} \sum_{k=-\infty}^{+\infty} \exp\left\{-\frac{(x - r_i - k_i M\Gamma_i + (k + k_i)M\Gamma_i)^2}{2\sigma_i^2}\right\},$$

where $k_i = \left[\frac{x - r_i}{M\Gamma_i}\right]$, $0 \leq x < M\Gamma_i$. By the definition of the circular distance in (11), we have

$$\begin{aligned} f_{\hat{r}_i}(x|N) &= \frac{1}{\sqrt{2\pi}\sigma_i} \sum_{k=-\infty}^{+\infty} \exp\left\{-\frac{(d_{M\Gamma_i}(x, r_i) + kM\Gamma_i)^2}{2\sigma_i^2}\right\} \\ &= \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left\{-\frac{d_{M\Gamma_i}^2(x, r_i)}{2\sigma_i^2}\right\} + \\ &\quad \frac{1}{\sqrt{2\pi}\sigma_i} \sum_{k \neq 0} \exp\left\{-\frac{(d_{M\Gamma_i}(x, r_i) + kM\Gamma_i)^2}{2\sigma_i^2}\right\}, \end{aligned} \tag{21}$$

where $0 \leq x < M\Gamma_i$. For convenience, we denote

$$x_i = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left\{-\frac{d_{M\Gamma_i}^2(x, r_i)}{2\sigma_i^2}\right\},$$

$$e_i = \frac{1}{\sqrt{2\pi}\sigma_i} \sum_{k \neq 0} \exp\left\{-\frac{(d_{M\Gamma_i}(x, r_i) + kM\Gamma_i)^2}{2\sigma_i^2}\right\},$$

for $i = 1, \ldots, L$. Since $\hat{r}_1, \ldots, \hat{r}_L$ are independent for a given N, their joint pdf for a given N is

$$\begin{aligned} f_{\hat{r}_1, \ldots, \hat{r}_L}(x|N) &= \prod_{i=1}^{L}(x_i + e_i) = \prod_{j=1}^{L} e_j + \sum_{i=1}^{L} \prod_{j \neq i} x_i e_j + \\ &\quad \sum_{i_1=1}^{L} \sum_{i_1 < i_2} \prod_{j \neq i_1, i_2} x_{i_1} x_{i_2} e_j + \cdots + \prod_{i=1}^{L} x_i. \end{aligned} \tag{22}$$

As $M\Gamma_i$ is usually much larger than σ_i, the terms of $k \neq 0$ in (21) are much smaller than the term of $k = 0$. That is, e_i are

much smaller than x_i for $i = 1, \ldots, L$. Hence, we can approximate (22) as

$$f_{\hat{r}_1,\ldots,\hat{r}_L}(x|N) \approx \prod_{i=1}^{L} x_i$$
$$= (2\pi)^{-L/2}(\sigma_1 \cdots \sigma_L)^{-1} \exp\left\{-\sum_{i=1}^{L} \frac{1}{2\sigma_i^2} d^2_{M\Gamma_i}(r_i, x)\right\}. \quad (23)$$

Clearly, the error of the joint pdf for a given N is

$$\sum_{i=1}^{L} \prod_{j\neq i} x_i e_j + \sum_{i_1=1}^{L} \sum_{i_2\neq i_1} \prod_{j\neq i_1,i_2} x_{i_1} x_{i_2} e_j + \cdots + \prod_{j=1}^{L} e_j.$$

Given L erroneous remainders $\hat{r}_1, \ldots, \hat{r}_L$, parameter M, and moduli $\Gamma_1, \ldots, \Gamma_L$, we now show how to robustly reconstruct N by the MLE. From (23), we obtain the approximation of the log likelihood function

$$\mathcal{L}\left(\bar{N}\right) \approx -\frac{L}{2}\ln 2\pi - \sum_{i=1}^{L} \ln \sigma_i - \sum_{i=1}^{L} \frac{1}{2\sigma_i^2} d^2_{M\Gamma_i}\left(\hat{r}_i, \bar{N}\right). \quad (24)$$

The MLE maximizes $\mathcal{L}(\bar{N})$ with respect to the unknown real number $\bar{N} \in [0, M\Gamma)$, which yields the following minimization problem

$$\hat{N}_{\mathrm{MLE}} = \arg \max_{0\leq \bar{N} < M\Gamma} \mathcal{L}\left(\bar{N}\right)$$
$$= \arg \min_{0\leq \bar{N} < M\Gamma} \sum_{i=1}^{L} \frac{1}{\sigma_i^2} d^2_{M\Gamma_i}\left(\hat{r}_i, \bar{N}\right), \quad (25)$$

where $d_{M\Gamma_i}\left(\hat{r}_i, \langle \bar{N}\rangle_{M\Gamma_i}\right)$ is simplified as $d_{M\Gamma_i}\left(\hat{r}_i, \bar{N}\right)$ according to Property 1 in (15). Then, $\hat{N}_{\mathrm{MLE}}$ is the MLE of N. In Fig. 1, we show the right-hand side of the log likelihood function in (23), where $N = 569$, $M = 10$, $\Gamma_1 = 3, \Gamma_2 = 5, \Gamma_3 = 7$, and σ_1 to σ_3 are 0.5, 0.8 and 1, respectively. By (25), we have $\hat{N}_{\mathrm{MLE}} = 576$.

Notice that the argument variable $\bar{N}$ in the minimization problem in (25) is real and may take any real value in the interval $[0, M\Gamma)$. Thus, in general, solving the minimization problem (25) may have a high computational complexity. In the next section, we will propose a fast algorithm that has a much lower computational complexity.

III. Fast MLE Algorithm

From (7) in Section II, one can see that the common remainder r^c is significant to the estimation of q_i and consequently N. In the noise free case, r^c can be determined from any remainder r_i modulo M. But, for noisy remainders $\hat{r}_i$ of N modulo M_i, their remainders modulo M, i.e.,

$$\hat{r}_i^c = \langle \hat{r}_i \rangle_M, \ i = 1, \ldots, L, \quad (26)$$

may be different from each other due to the errors. In order to obtain the optimal estimate of N, intuitively the common remainder r^c should be optimally determined. As $\hat{r}_1^c, \ldots, \hat{r}_L^c$ are folded real numbers, we can not estimate r^c by simply averaging them. Instead, we define a special averaging operation of $\hat{r}_i^c$ as

$$\hat{r}^c \triangleq \arg \min_{0\leq x < M} \sum_{i=1}^{L} \frac{1}{\sigma_i^2} d_M^2\left(\hat{r}_i^c, x\right), \quad (27)$$

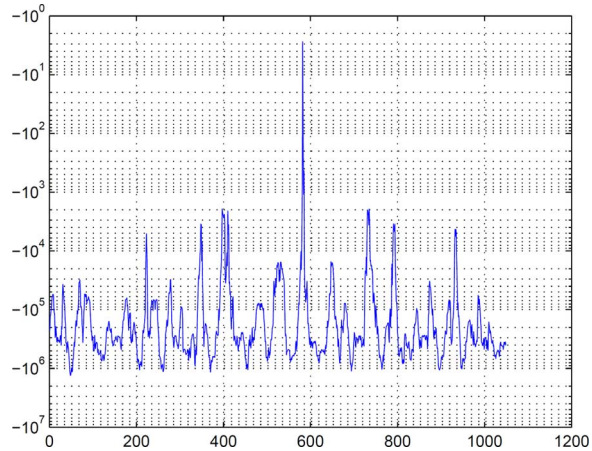

Fig. 1. The log likelihood function (24).

where again x takes real values in the interval $[0, M)$. After the common remainder r^c is estimated above, we can use $[(\hat{r}_i - \hat{r}^c)/M]$ as an estimate of q_i, i.e.,

$$\hat{q}_i = \left[\frac{\hat{r}_i - \hat{r}^c}{M}\right], \ i = 1, \ldots, L. \quad (28)$$

Recall that $[\cdot]$ stands for the rounding integer defined in (13). Then, N_0 can be reconstructed as

$$\hat{N}_0 \equiv \sum_{i=1}^{L} \bar{\gamma}_i \gamma_i \hat{q}_i \mod \Gamma. \quad (29)$$

Therefore, N can be reconstructed by

$$\hat{N} = M\hat{N}_0 + \hat{r}^c. \quad (30)$$

The following result says that $\hat{N}$ obtained from the above algorithm is indeed the MLE when the estimate $\hat{r}^c$ of common remainder r^c is (27).

Theorem 1: Assume that all moduli $\Gamma_1, \ldots, \Gamma_L$ are pair-wisely co-prime. If $0 \leq N < M\Gamma$, then $\hat{N}$ in (30) is the MLE of N, that is, $\hat{N} = \hat{N}_{\mathrm{MLE}}$.

Proof: Clearly, congruence (29) means

$$\hat{N}_0 \equiv \hat{q}_i \mod \Gamma_i, \ i = 1, \ldots, L.$$

Hence, there exist integers k_i such that

$$\hat{N}_0 = k_i\Gamma_i + \hat{q}_i, \ i = 1, \ldots, L.$$

According to (30) and Property 1, the circular distance of $\hat{r}_i$ to $\hat{N}$ for $M\Gamma_i$ can be written as

$$d_{M\Gamma_i}(\hat{r}_i, \hat{N}) = d_{M\Gamma_i}\left(\hat{r}_i, M\hat{q}_i + \hat{r}^c\right).$$

From Property 2 and (28), we have

$$d_{M\Gamma_i}\left(\hat{r}_i, \hat{N}\right) = d_{M\Gamma_i}\left(\hat{r}_i - M\hat{q}_i - \hat{r}^c, 0\right)$$
$$= d_{M\Gamma_i}\left(\hat{r}_i - \hat{r}^c - M\left[\frac{\hat{r}_i - \hat{r}^c}{M}\right], 0\right).$$

From the definition of the circular distance in (11), we have

$$d_{M\Gamma_i}\left(\hat{r}_i, \hat{N}\right) = d_{M\Gamma_i}\left(d_M(\hat{r}_i, \hat{r}^c), 0\right).$$

As $-M/2 \le d_M(\hat{r}_i, \hat{r}^c) < M/2$, from Properties 1, 5 and the definition of $\hat{r}_i^c$ in (26), we obtain

$$d_{M\Gamma_i}(\hat{r}_i, \hat{N}) = d_M(\hat{r}_i, \hat{r}^c) = d_M(\hat{r}_i^c, \hat{r}^c). \tag{31}$$

On the other hand, for a real number $0 \le \tilde{N} < M\Gamma$, $\tilde{N} \ne \hat{N}$, let

$$\tilde{N} = M\tilde{N}_0 + \tilde{r}^c, \tag{32}$$

where $0 \le \tilde{N}_0 < \Gamma$ and $0 \le \tilde{r}^c < M$. From Properties 1, 3 and (32), we have

$$\begin{aligned} d_M\left(d_{M\Gamma_i}(\hat{r}_i, \tilde{N}), 0\right) &= d_M\left(\hat{r}_i - \tilde{N} - M\Gamma_i\left[\frac{\hat{r}_i - \tilde{N}}{M\Gamma_i}\right], 0\right) \\ &= d_M(\hat{r}_i - \tilde{r}^c, 0) \\ &= d_M(\hat{r}_i^c, \tilde{r}^c). \end{aligned} \tag{33}$$

According to the definition of the circular distance, we obtain

$$\begin{aligned} d^2_{M\Gamma_i}\left(\hat{r}_i, \tilde{N}\right) &= d^2_{M\Gamma_i}\left(d_{M\Gamma_i}(\hat{r}_i, \tilde{N}), 0\right) \\ &\ge d^2_M\left(d_{M\Gamma_i}(\hat{r}_i, \tilde{N}), 0\right). \end{aligned} \tag{34}$$

It follows from (33) and (34) that

$$\begin{aligned} \sum_{i=1}^{L} \frac{1}{\sigma_i^2} d^2_{M\Gamma_i}\left(\hat{r}_i, \tilde{N}\right) &\ge \sum_{i=1}^{L} \frac{1}{\sigma_i^2} d^2_M\left(d_{M\Gamma_i}(\hat{r}_i, \tilde{N}), 0\right) \\ &= \sum_{i=1}^{L} \frac{1}{\sigma_i^2} d^2_M(\hat{r}_i^c, \tilde{r}^c). \end{aligned} \tag{35}$$

According to the definition of $\hat{r}^c$ in (27), we have

$$\sum_{i=1}^{L} \frac{1}{\sigma_i^2} d^2_M(\hat{r}_i^c, \tilde{r}^c) > \sum_{i=1}^{L} \frac{1}{\sigma_i^2} d^2_M(\hat{r}_i^c, \hat{r}^c). \tag{36}$$

Combining (31), (35) and (36), we obtain

$$\sum_{i=1}^{L} \frac{1}{\sigma_i^2} d^2_{M\Gamma_i}\left(\hat{r}_i, \tilde{N}\right) > \sum_{i=1}^{L} \frac{1}{\sigma_i^2} d^2_{M\Gamma_i}\left(\hat{r}_i, \hat{N}\right).$$

It follows from (24) that

$$\mathcal{L}(\tilde{N}) < \mathcal{L}(\hat{N}).$$

This proves the theorem. ∎

From Theorem 1, we know that the MLE of N depends on the estimate of the common remainder. Note that the estimate of the reference common remainder of the improved robust CRT proposed in [25] is only the special case in the above when all the variances σ_i^2 in (27) are equal. Thus, when the variances σ_i^2 in (27) are not equal, which is the case when the remainder noises do not have the same variance, the method in [25] is no longer optimal. A detailed comparison will be given later.

For the computational complexity, the optimal searching of the unknown real number N among all the reals in the range $[0, M\Gamma)$ is reduced to the searching of the common remainder r^c among all the reals in the range of $[0, M)$. Although the searching region is significantly reduced, since Γ is usually much larger than M, it still has infinite possibilities, i.e., infinitely many reals in the range of $[0, M)$. Next, we propose a novel fast algorithm to obtain the optimal estimate, where the optimal one is in a finite set with L candidates that is independent of M.

Let

$$w_i = \frac{\frac{1}{\sigma_i^2}}{\sum_{i=1}^{L} \frac{1}{\sigma_i^2}},\ i = 1, \ldots, L. \tag{37}$$

Then, $0 < w_i \le 1$ and $\sum_{i=1}^{L} w_i = 1$. So, we can view w_i as the weights of the remainders later. Note that $\sum_{i=1}^{L} \frac{1}{\sigma_i^2}$ is a constant when all the variances of the noise are known. Hence, the optimal estimate in (27) can be rewritten as

$$\hat{r}^c = \arg \min_{x \in [0, M)} \sum_{i=1}^{L} w_i d^2_M(\hat{r}_i^c, x). \tag{38}$$

Then, we have the following result.

Theorem 2: The optimal estimate $\hat{r}^c$ in (38) belongs to the following set:

$$\Omega = \left\{ \left\langle \sum_{i=1}^{L} w_i \hat{r}_i^c + M \sum_{i=1}^{t} w_{\rho(i)} \right\rangle_M, t = 1, \ldots, L \right\}, \tag{39}$$

where ρ is a permutation of the set $\{1, \ldots, L\}$ such that

$$\hat{r}^c_{\rho(1)} \le \cdots \le \hat{r}^c_{\rho(L)}, \tag{40}$$

and $w_{\rho(i)}$ is the weight of $\hat{r}^c_{\rho(i)}$.

Proof: Let $f(x) = \sum_{i=1}^{L} w_i d^2_M(\hat{r}_i^c, x)$, and

$$X = \left\{ x \in \mathbb{R} : d_M(\hat{r}_j^c, x) = -M/2 \text{ for some } j \right\}.$$

Then, we have

$$f(x) = \sum_{i=1}^{L} w_i \left(\hat{r}_i^c - x - M\left[\frac{\hat{r}_i^c - x}{M}\right] \right)^2.$$

For $x_0 \in X$, we obtain

$$\hat{r}_j^c = x_0 + M\left[\frac{\hat{r}_j^c - x_0}{M}\right] - M/2 \text{ for some } j.$$

Hence,

$$\begin{aligned} \lim_{x \to x_0^+} f(x) &= \lim_{x \to x_0^+} w_j \left(\hat{r}_j^c - x - M\left[\frac{\hat{r}_j^c - x}{M}\right] \right)^2 + \\ &\quad \lim_{x \to x_0^+} \sum_{i \ne j} w_i \left(\hat{r}_i^c - x - M\left[\frac{\hat{r}_i^c - x}{M}\right] \right)^2 \\ &= \lim_{x \to x_0^+} w_j (x_0 - x + M/2)^2 + \\ &\quad \lim_{x \to x_0^+} \sum_{i \ne j} w_i \left(\hat{r}_i^c - x - M\left[\frac{\hat{r}_i^c - x}{M}\right] \right)^2. \end{aligned}$$

Since

$$f(x_0) = w_j M^2/4 + \sum_{i \ne j} w_i \left(\hat{r}_i^c - x_0 - M\left[\frac{\hat{r}_i^c - x_0}{M}\right] \right)^2$$

and

$$\begin{aligned} &\lim_{x \to x_0^+} \sum_{i \ne j} w_i \left(\hat{r}_i^c - x - M\left[\frac{\hat{r}_i^c - x}{M}\right] \right)^2 \\ &= \sum_{i \ne j} w_i \left(\hat{r}_i^c - x_0 - M\left[\frac{\hat{r}_i^c - x_0}{M}\right] \right)^2, \end{aligned}$$

we have

$$\begin{aligned} f'_+(x_0) &= \lim_{x\to x_0+} \frac{f(x)-f(x_0)}{x-x_0} \\ &= \lim_{x\to x_0+} \frac{w_j\left(x_0-x+M/2\right)^2 - w_j M^2/4}{x-x_0} \\ &= -M. \end{aligned}$$

Similarly, $f'_-(x_0) = M$. Clearly, $f(x)$ is not differentiable at x_0. Since $f'_+(x_0) < 0$ and $f'_-(x_0) > 0$, x_0 is not the local minimum point. This means that the optimal estimate $\hat{r}^c$ is not in X, i.e.,

$$\hat{r}^c \notin X. \tag{41}$$

Since $f(x)$ is differentiable in $(-\infty, \infty)$ except X, we have

$$f'(\hat{r}^c) = 0.$$

This leads to

$$\sum_{i=1}^{L} w_i \left(\hat{r}_i^c - \hat{r}^c - M\left[\frac{\hat{r}_i^c - \hat{r}^c}{M}\right]\right) = 0. \tag{42}$$

We have two cases below.

Case 1 $\hat{r}^c \in [0, M/2)$.

Since $\hat{r}_i^c \in [0, M)$, we have $\hat{r}_i^c - \hat{r}^c \in (-M/2, M)$. From (41), we have $\hat{r}_i^c - \hat{r}^c \neq M/2$, implying either $\hat{r}_i^c - \hat{r}^c \in (-M/2, M/2)$ or $\hat{r}_i^c - \hat{r}^c \in (M/2, M)$. Define

$$k = \begin{cases} 0, & \text{if } \hat{r}^c_{\rho_{(1)}} - \hat{r}^c > \frac{M}{2}, \\ \max\left\{ i : \hat{r}^c_{\rho_{(i)}} - \hat{r}^c < \frac{M}{2} \right\}, & \text{otherwise.} \end{cases}$$

Then, (42) can be rewritten as

$$\sum_{i=1}^{k} w_{\rho_{(i)}} \left(\hat{r}^c_{\rho_{(i)}} - \hat{r}^c \right) + \sum_{i=k+1}^{L} w_{\rho_{(i)}} \left(\hat{r}^c_{\rho_{(i)}} - \hat{r}^c - M \right) = 0.$$

Consequently,

$$\hat{r}^c = \sum_{i=1}^{L} w_i \hat{r}_i^c - M \sum_{i=k+1}^{L} w_{\rho_{(i)}},$$

and hence

$$\begin{aligned} \hat{r}^c &= \left\langle \sum_{i=1}^{L} w_i \hat{r}_i^c - M \sum_{i=k+1}^{L} w_{\rho_{(i)}} \right\rangle_M \\ &= \left\langle \sum_{i=1}^{L} w_i \hat{r}_i^c + M \sum_{i=1}^{k} w_{\rho_{(i)}} \right\rangle_M . \end{aligned} \tag{43}$$

Note that $\sum_{i=1}^{L} w_i = 1$. Then, for $k = 0$ and $k = L$ the right-hand side expression in (43) has the same value, i.e.,

$$\begin{aligned} \left\langle \sum_{i=1}^{L} w_i \hat{r}_i^c - M \sum_{i=1}^{L} w_{\rho_{(i)}} \right\rangle_M &= \left\langle \sum_{i=1}^{L} w_i \hat{r}_i^c + M \sum_{i=1}^{L} w_{\rho_{(i)}} \right\rangle_M \\ &= \left\langle \sum_{i=1}^{L} w_i \hat{r}_i^c \right\rangle_M . \end{aligned}$$

Thus, $\hat{r}^c \in \Omega$.

Case 2 $\hat{r}^c \in [M/2, M)$.

In this case, either $\hat{r}_i^c - \hat{r}^c \in (-M, -M/2)$ or $\hat{r}_i^c - \hat{r}^c \in (-M/2, M/2)$. Similarly, we define

$$k' = \begin{cases} 0, & \text{if } \hat{r}^c_{\rho_{(1)}} - \hat{r}^c > -\frac{M}{2}, \\ \max\left\{ i : \hat{r}^c_{\rho_{(i)}} - \hat{r}^c < -\frac{M}{2} \right\}, & \text{otherwise.} \end{cases}$$

Then (42) can be rewritten as

$$\sum_{i=1}^{k'} w_{\rho_{(i)}} \left(\hat{r}^c_{\rho_{(i)}} - \hat{r}^c + M \right) + \sum_{i=k'+1}^{L} w_{\rho_{(i)}} \left(\hat{r}^c_{\rho_{(i)}} - \hat{r}^c \right) = 0.$$

That is,

$$\hat{r}^c = \sum_{i=1}^{L} w_i \hat{r}_i^c + M \sum_{i=1}^{k'} w_{\rho_{(i)}}.$$

Hence,

$$\hat{r}^c = \left\langle \sum_{i=1}^{L} w_i \hat{r}_i^c + M \sum_{i=1}^{k'} w_{\rho_{(i)}} \right\rangle_M . \tag{44}$$

Similarly, for $k' = 0$ and $k' = L$ the right-hand side expression in (44) has the same value. Thus, $\hat{r}^c \in \Omega$. This proves the theorem. ■

Theorem 2 tells us that for a given erroneous remainder sequence $\hat{r}_1, \ldots, \hat{r}_L$, the optimal estimate $\hat{r}^c$ of the common remainder r^c belongs to the finite set Ω of L elements in (39). Hence, the optimal estimate in (38) can be simplified as

$$\hat{r}^c = \arg\min_{x\in\Omega} \sum_{i=1}^{L} w_i d_M^2\left(\hat{r}_i^c, x\right), \tag{45}$$

where Ω has only L elements.

Comparison With the Improved Robust CRT in [25]: In the fast MLE algorithm, the key processes to have the robustness and better performance are the common remainder r^c estimation and the remainder rounding after canceling r^c described in (27) and (28), respectively. The improved closed-form robust CRT in [25] also involves with r^c estimation and canceling the optimal reference remainder from the other remainders, where the optimal reference remainder has the minimum circular distance for M between the estimate of r^c and the remainders. Recall that the estimate of r^c proposed in [25] is based on the assumption that all the remainder errors have the same variance and it is determined by

$$\hat{r}^c \triangleq \arg\min_{0\le x<M} \sum_{i=1}^{L} d_M^2\left(\hat{r}_i^c, x\right). \tag{46}$$

When the remainder errors have the same error variance, i.e., $\sigma_1 = \cdots = \sigma_L$, the two estimates, (27) and (46) are the same. When the remainder errors have different variances, the two estimates are different. Hence, the estimate of r^c obtained by (46) is not optimal when the remainder error variances are different, which may lead to a wrong selection of the optimal reference remainder. For the obtained fast MLE algorithm here, the estimate of r^c in (27) is always the optimal one when either the remainder errors have the same variance or different variances. So, the proposed fast MLE algorithm has a better performance than the improved closed-form robust CRT in [25]. Another advantage of the proposed algorithm is that it has a lower complexity, since the algorithm in [25] estimates r^c by searching all

the reals in the range of $[0, M)$, while the proposed algorithm in this paper determines the optimal estimate of r^c from a finite set of L elements that is independent of M.

IV. Robust MLE

In this section, we present a necessary and sufficient condition for the MLE to be robust. Then, we calculate the probability of the robust MLE CRT.

A. Robust Estimation

We first consider the condition of the remainder errors that leads the MLE to be a robust estimation, i.e., if $|\Delta r_i| \leq \tau$ for all $i = 1, \ldots, L$, then $\left|\hat{N} - N\right| \leq \tau$. Suppose that the estimation error of N is ϵ:

$$\hat{N} - N = M\hat{N}_0 + \hat{r}^c - MN_0 - r^c = \epsilon, \tag{47}$$

where $-M\Gamma_i < \epsilon < M\Gamma_i$. Note that (29) is equivalent to

$$\hat{N}_0 \equiv \hat{q}_i \mod \Gamma_i.$$

That is,

$$\begin{aligned} M\hat{N}_0 &\equiv M\hat{q}_i \mod M\Gamma_i \\ &\equiv M\left[\frac{\hat{r}_i - \hat{r}^c}{M}\right] \mod M\Gamma_i. \end{aligned}$$

Similarly,

$$MN_0 \equiv Mq_i \mod M\Gamma_i.$$

Hence, (47) can be rewritten as

$$\epsilon \equiv M\left[\frac{\hat{r}_i - \hat{r}^c}{M}\right] - Mq_i + \hat{r}^c - r^c \mod M\Gamma_i. \tag{48}$$

According to (2) and (7), we have

$$\hat{r}_i = Mq_i + r^c + \Delta r_i. \tag{49}$$

Let the error of the common remainder be

$$\Delta r^c \triangleq \hat{r}^c - r^c. \tag{50}$$

Combining (48), (49) and (50), we obtain

$$\epsilon \equiv M\left[\frac{\Delta r_i - \Delta r^c}{M}\right] + \Delta r^c \mod M\Gamma_i. \tag{51}$$

Since $-M\Gamma_i < \epsilon < M\Gamma_i$, we have

$$\Delta r^c = kM + \epsilon \text{ for some } k \in \mathbb{Z}. \tag{52}$$

Plugging (52) into (51), using (14), we have

$$M\left[\frac{\Delta r_i - \epsilon}{M}\right] \equiv 0 \mod M\Gamma_i.$$

That is,

$$\left[\frac{\Delta r_i - \epsilon}{M}\right] \equiv 0 \mod \Gamma_i. \tag{53}$$

If

$$-\Gamma_i < \left[\frac{\Delta r_i - \epsilon}{M}\right] < \Gamma_i, \tag{54}$$

then we obtain from (53) that

$$\left[\frac{\Delta r_i - \epsilon}{M}\right] = 0.$$

By the definition of the rounding operation in (13), we obtain

$$-M/2 \leq \Delta r_i - \epsilon < M/2. \tag{55}$$

From $|\Delta r_i| \leq \tau$ and $\left|\hat{N} - N\right| = |\epsilon| \leq \tau$, we can obtain that the error bound τ for the robust estimation of N satisfies

$$\tau < M/4, \tag{56}$$

which coincides with the result obtained in [14].

From the above analysis, we see that (55) is a necessary condition for a robust estimation. Thus, in what follows, in order to discuss the robustness of the MLE based robust CRT, we suppose that (55) is always satisfied. Let ν be a permutation of the set $\{1, \ldots, L\}$ satisfying

$$\Delta r_{\nu_{(1)}} \leq \cdots \leq \Delta r_{\nu_{(L)}}, \tag{57}$$

and the weights in the corresponding order be $w_{\nu_{(1)}}, \ldots, w_{\nu_{(L)}}$. Then, we obtain from (55) that

$$\Delta r_{\nu_{(L)}} - \Delta r_{\nu_{(1)}} = \left(\Delta r_{\nu_{(L)}} - \epsilon\right) - \left(\Delta r_{\nu_{(1)}} - \epsilon\right) < M, \tag{58}$$

and we have the following result.

Lemma 1: If $\Delta r_{\nu_{(L)}} - \Delta r_{\nu_{(1)}} < M$, then the set Ω in (39) can be rewritten as

$$\Omega = \left\{ \left\langle r^c + \sum_{i=1}^{L} w_i \Delta r_i + M \sum_{i=1}^{t} w_{\nu_{(i)}} \right\rangle_M, t = 1, \ldots, L \right\}. \tag{59}$$

Proof: Note that from (26),

$$\begin{aligned} \hat{r}_i^c &= \langle \hat{r}_i \rangle_M = \langle r^c + \Delta r_i \rangle_M \\ &= r^c + \Delta r_i - M\left\lfloor \frac{r^c + \Delta r_i}{M} \right\rfloor, \end{aligned} \tag{60}$$

where $\lfloor \cdot \rfloor$ denotes the floor operation. Let $\ell = \left\lfloor \frac{r^c + \Delta r_{\nu_{(1)}}}{M} \right\rfloor$. Then, we obtain from (57) that

$$\left\lfloor \frac{r^c + \Delta r_{\nu_{(1)}}}{M} \right\rfloor \leq \cdots \leq \left\lfloor \frac{r^c + \Delta r_{\nu_{(L)}}}{M} \right\rfloor.$$

From (58), we have $\Delta r_{\nu_{(L)}} < M + \Delta r_{\nu_{(1)}}$. Hence,

$$\left\lfloor \frac{r^c + \Delta r_{\nu_{(L)}}}{M} \right\rfloor \leq \left\lfloor \frac{r^c + \Delta r_{\nu_{(1)}}}{M} \right\rfloor + 1 = \ell + 1. \tag{61}$$

We have two cases below.

Case 1 The inequality in (61) is strict.

In this case, we have

$$\left\lfloor \frac{r^c + \Delta r_{\nu_{(1)}}}{M} \right\rfloor = \cdots = \left\lfloor \frac{r^c + \Delta r_{\nu_{(L)}}}{M} \right\rfloor.$$

It follows from (60) that

$$\hat{r}^c_{\nu_{(1)}} \leq \cdots \leq \hat{r}^c_{\nu_{(L)}},$$

which means that ν in (57) and ρ in (40) are the same permutation. Hence,

$$\begin{aligned}&\left\langle \sum_{i=1}^{L} w_i \hat{r}_i^c + M \sum_{i=1}^{t} w_{\rho_{(i)}} \right\rangle_M \\ &= \left\langle \sum_{i=1}^{L} w_i (r^c + \Delta r_i - \ell M) + M \sum_{i=1}^{t} w_{\nu_{(i)}} \right\rangle_M \\ &= \left\langle r^c + \sum_{i=1}^{L} w_i \Delta r_i + M \sum_{i=1}^{t} w_{\nu_{(i)}} \right\rangle_M .\end{aligned}$$

Case 2 The inequality in (61) is an equality.

Let $k \in \{1, \ldots, L-1\}$ be the subscript of the errors satisfying

$$\left\lfloor \frac{r^c + \Delta r_{\nu_{(k)}}}{M} \right\rfloor < \left\lfloor \frac{r^c + \Delta r_{\nu_{(k+1)}}}{M} \right\rfloor .$$

Due to (61), there only exists one such k in the above inequality. Then,

$$\left\lfloor \frac{r^c + \Delta r_{\nu_{(i)}}}{M} \right\rfloor = \begin{cases} \ell, & \text{if } i \in \{1, \ldots, k\} \\ \ell + 1, & \text{if } i \in \{k+1, \ldots, L\}. \end{cases}$$

It follows from (60) that

$$\begin{aligned} \hat{r}^c_{\nu_{(1)}} &= r^c + \Delta r_{\nu_{(1)}} - \ell M, \\ \hat{r}^c_{\nu_{(L)}} &= r^c + \Delta r_{\nu_{(L)}} - (\ell + 1) M. \end{aligned} \tag{62}$$

Since $\Delta r_{\nu_{(L)}} - \Delta r_{\nu_{(1)}} < M$, we obtain from (62) that

$$\hat{r}^c_{\nu_{(L)}} < \hat{r}^c_{\nu_{(1)}}.$$

According to (57), we have

$$\hat{r}^c_{\nu_{(k+1)}} \le \cdots \le \hat{r}^c_{\nu_{(L)}} < \hat{r}^c_{\nu_{(1)}} \le \cdots \le \hat{r}^c_{\nu_{(k)}},$$

which means that

$$\left(\rho_{(1)}, \ldots, \rho_{(L)}\right) = \left(\nu_{(k+1)}, \ldots, \nu_{(L)}, \nu_{(1)}, \ldots, \nu_{(k)}\right). \tag{63}$$

Thus,

$$\begin{aligned}&\left\langle \sum_{i=1}^{L} w_i \hat{r}_i^c + M \sum_{i=1}^{t} w_{\rho_{(i)}} \right\rangle_M = \left\langle \sum_{i=1}^{k} w_{\nu_{(i)}} \left(r^c + \Delta r_{\nu_{(i)}} - \ell M\right) \right. \\ &\left. + \sum_{i=k+1}^{L} w_{\nu_{(i)}} \left(r^c + \Delta r_{\nu_{(i)}} - (\ell+1) M\right) + M \sum_{i=1}^{t} w_{\rho_{(i)}} \right\rangle_M \\ &= \left\langle r^c + \sum_{i=1}^{L} w_i \Delta r_i + M \sum_{i=1}^{k} w_{\nu_{(i)}} + M \sum_{i=1}^{t} w_{\rho_{(i)}} \right\rangle_M . \end{aligned} \tag{64}$$

If $t \le L - k$, then we obtain from (63) that

$$\sum_{i=1}^{t} w_{\rho_{(i)}} = \sum_{i=k+1}^{k+t} w_{\nu_{(i)}}.$$

Hence, (64) can be simplified as

$$\left\langle \sum_{i=1}^{L} w_i \hat{r}_i^c + M \sum_{i=1}^{t} w_{\rho_{(i)}} \right\rangle_M = \left\langle r^c + \sum_{i=1}^{L} w_i \Delta r_i + M \sum_{i=1}^{k+t} w_{\nu_{(i)}} \right\rangle_M .$$

If $t > L - k$, then we obtain from (63) that

$$\begin{aligned} \sum_{i=1}^{t} w_{\rho_{(i)}} &= \sum_{i=1}^{L-k} w_{\rho_{(i)}} + \sum_{i=L-k+1}^{t} w_{\rho_{(i)}} \\ &= \sum_{i=k+1}^{L} w_{\nu_{(i)}} + \sum_{i=1}^{t-(L-k)} w_{\nu_{(i)}}. \end{aligned}$$

Hence, (64) can be simplified as

$$\left\langle \sum_{i=1}^{L} w_i \hat{r}_i^c + M \sum_{i=1}^{t} w_{\rho_{(i)}} \right\rangle_M = \left\langle r^c + \sum_{i=1}^{L} w_i \Delta r_i + M \sum_{i=1}^{t-(L-k)} w_{\nu_{(i)}} \right\rangle_M .$$

This proves the lemma. ∎

Lemma 1 gives another expression of the candidate set of the optimal estimate $\hat{r}^c$ of the common remainder r^c. It is noted that the two weights, $w_{\rho_{(i)}}$ and $w_{\nu_{(i)}}$, may be not equal because of the different sorting ways. For a given erroneous remainder sequence $\hat{r}_i^c$ of $\hat{r}_i$ modulo M and the variances, we can determine Ω only by (39) but not (59). This is because the common remainder r^c is unknown and needs to be estimated. However, (59) is helpful to analyze the estimation error of N, which will be seen in the following theorems. For convenience, we denote the remainder error set as

$$U = \{\Delta r_1, \ldots, \Delta r_L\}, \tag{65}$$

and the weighted average of the remainder errors as

$$\overline{\Delta r} = \sum_{i=1}^{L} w_i \Delta r_i. \tag{66}$$

Then, we have the following results.

Theorem 3: If the weighted average error in (66) satisfies $|\overline{\Delta r}| < M/2$ and

$$-\frac{M}{2} \le \sum_{\Delta r_i \in S} \frac{w_i}{\sum_{\Delta r_j \in S} w_j} \Delta r_i - \sum_{\Delta r_i \in \overline{S}} \frac{w_i}{\sum_{\Delta r_j \in \overline{S}} w_j} \Delta r_i < \frac{M}{2} \tag{67}$$

holds for any subset S of U and $\overline{S} = U \setminus S$ is the complement of S in U, then the optimal $\hat{r}^c$ in (38) has the form

$$\hat{r}^c = \langle r^c + \overline{\Delta r} \rangle_M. \tag{68}$$

Moreover,

$$\Delta r^c = \hat{r}^c - r^c = \overline{\Delta r} + \begin{cases} M, & \text{if } r^c + \overline{\Delta r} < 0 \\ 0, & \text{if } 0 \le r^c + \overline{\Delta r} < M \\ -M, & \text{if } r^c + \overline{\Delta r} \ge M. \end{cases} \tag{69}$$

The proof of this theorem is in Appendix A.

Theorem 3 gives a condition of the remainder errors and their weights when the optimal estimate $\hat{r}^c$ of the common remainder is $\langle r^c + \overline{\Delta r} \rangle_M$. As discussed above, the estimate of the common remainder is in the set Ω with L elements described in (39) or (59). Clearly, $\langle r^c + \overline{\Delta r} \rangle_M$ belongs to the set Ω. Next theorem shows that $\hat{r}^c = \langle r^c + \overline{\Delta r} \rangle_M$ is a necessary condition for the robust estimation of N. It also shows that the condition in (67) is the necessary and sufficient condition for the robust estimation of N.

Theorem 4: Let the weighted average remainder in (66) satisfy $|\overline{\Delta r}| < M/2$, and N be a real number in the range

$[M, M(\Gamma-1))$. Then, for the above MLE $\hat{N}$ of N, we have that

$$\hat{N} - N = \overline{\Delta r} \tag{70}$$

holds if and only if

$$-\frac{M}{2} \le \sum_{\Delta r_i \in S} \frac{w_i}{\sum_{\Delta r_j \in S} w_j} \Delta r_i - \sum_{\Delta r_i \in \overline{S}} \frac{w_i}{\sum_{\Delta r_j \in \overline{S}} w_j} \Delta r_i < \frac{M}{2} \tag{71}$$

holds for any subset S of U, where $\overline{S}$ is its complement.

The proof of this theorem is in Appendix B.

The condition in (71) or (67) looks complicated and strong. In fact, when

$$|\Delta r_i| < M/4,\ i = 1, \ldots, L, \tag{72}$$

it is not hard to see

$$\begin{aligned} &\sum_{\Delta r_i \in S} \frac{w_i}{\sum_{\Delta r_j \in S} w_j} \Delta r_i - \sum_{\Delta r_i \in \overline{S}} \frac{w_i}{\sum_{\Delta r_j \in \overline{S}} w_j} \Delta r_i \\ &\ge -\sum_{\Delta r_i \in S} \frac{w_i}{\sum_{\Delta r_j \in S} w_j} |\Delta r_i| - \sum_{\Delta r_i \in \overline{S}} \frac{w_i}{\sum_{\Delta r_j \in \overline{S}} w_j} |\Delta r_i| \\ &> -M/2 \end{aligned}$$

and

$$\begin{aligned} &\sum_{\Delta r_i \in S} \frac{w_i}{\sum_{\Delta r_j \in S} w_j} \Delta r_i - \sum_{\Delta r_i \in \overline{S}} \frac{w_i}{\sum_{\Delta r_j \in \overline{S}} w_j} \Delta r_i \\ &\le \sum_{\Delta r_i \in S} \frac{w_i}{\sum_{\Delta r_j \in S} w_j} |\Delta r_i| + \sum_{\Delta r_i \in \overline{S}} \frac{w_i}{\sum_{\Delta r_j \in \overline{S}} w_j} |\Delta r_i| \\ &< M/2 \end{aligned}$$

hold for any subset S of U. This tells that the above simpler condition (72) implies the condition (71) and therefore leads to the following robustness of the MLE.

Corollary 1: If $|\Delta r_i| \le \tau$ for all $i = 1, \ldots, L$, and $\tau < M/4$, then $\hat{N} - N = \overline{\Delta r}$.

Clearly, when $|\Delta r_i| \le \tau$ for $i = 1, \ldots, L$, $|\overline{\Delta r}| \le \tau$. Thus, from Corollary 1, $|\hat{N} - N| = |\overline{\Delta r}| \le \tau$. This coincides with the robustness we have obtained previously in [25] but the result obtained in Corollary 1 is stronger than the robustness.

B. Probability of the Robust Estimation

Now, we calculate the probability of the robust MLE estimation of N from erroneous remainders for a given real number N, where the remainder errors Δr_i are assumed to follow wrapped normal distribution with mean zero and variances σ_i^2. As $M\Gamma_i$ are usually much larger than σ_i^2, we approximate the distributions of Δr_i as normal distributions in the following.

Theorem 4 shows that N can be robustly recovered if and only if its remainder errors satisfy (71). Note that the equal sign of the left-hand side in (71) has no effect on the probability for continuous random variables. Then, (71) can be substituted by

$$\left| \sum_{\Delta r_i \in S} \frac{w_i}{\sum_{\Delta r_j \in S} w_j} \Delta r_i - \sum_{\Delta r_i \in \overline{S}} \frac{w_i}{\sum_{\Delta r_j \in \overline{S}} w_j} \Delta r_i \right| < M/2. \tag{73}$$

Next, we consider two cases depending on the number of remainders, i.e., $L = 3$ and $L \ge 4$. Note that the case of $L = 2$ is trivial and the probability is

$$\begin{aligned} &p((\Delta r_1, \Delta r_2) \in \Sigma) \\ &\quad = \frac{1}{2\pi\sigma_1\sigma_2} \iint_{\Sigma} \exp\left\{ -\frac{x_1^2}{2\sigma_1^2} - \frac{x_2^2}{2\sigma_2^2} \right\} dx_1 dx_2, \end{aligned} \tag{74}$$

where $\Sigma = \{(x_1, x_2) : |x_1 - x_2| < M/2\}$.

1) $L = 3$: In this case, (73) is equivalent to

$$\left| \Delta r_i - \sum_{j \ne i} \frac{w_j}{1 - w_i} \Delta r_j \right| < M/2,\ i = 1, 2, 3. \tag{75}$$

Let $X = A\Delta r$, where $X = [x_1, x_2, x_3]^T$, $\Delta r = [\Delta r_1, \Delta r_2, \Delta r_3]^T$, and $A = \begin{bmatrix} 1-w_1 & -w_2 & -w_3 \\ -w_1 & 1-w_2 & -w_3 \\ -w_1 & -w_2 & 1-w_3 \end{bmatrix}$. Then, (75) can be simplified as

$$|x_i| < M(1 - w_i)/2,\ i = 1, 2, 3. \tag{76}$$

As the noises for different remainders, Δr_1, Δr_2 and Δr_3 are independently and identically normal distributed random variables for a given N, X is a random vector with normal distribution. Let $\Lambda = \mathrm{diag}(\sigma_1, \sigma_2, \sigma_3)$. Then the covariance matrix of X, $A\Lambda^2 A^T$, is a singular matrix since the determinant of the matrix A is 0, which is because the sum of all the column vectors of A is the zero vector. Hence, the joint pdf of X does not exist [1]. But the characteristic function of X exists, i.e., $\varphi(\mathbf{t}) = \exp\left\{-\frac{1}{2}\mathbf{t}^T A\Lambda^2 A^T \mathbf{t}\right\}$, where $\mathbf{t} = [t_1, t_2, t_3]^T$. According to the inversion formula [19], we have

$$\begin{aligned} &p((\Delta r_1, \Delta r_2, \Delta r_3) \in \Sigma) \\ &= \frac{1}{\pi^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{i=1}^{3} \frac{\sin\left(M(1-w_i)t_i/2\right)}{t_i} \times \\ &\quad \exp\left\{ -\frac{1}{2}\mathbf{t}^T A\Lambda^2 A^T \mathbf{t} \right\} dt_1 dt_2 dt_3, \end{aligned} \tag{77}$$

where Σ is specified by (73).

For $\sigma_1 = \sigma_2 = \sigma_3 = \sigma$, i.e., $w_1 = w_2 = w_3 = 1/3$, (77) can be simplified as

$$\begin{aligned} &p((\Delta r_1, \Delta r_2, \Delta r_3) \in \Sigma) \\ &= \frac{1}{\pi^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{i=1}^{3} \frac{\sin(Mt_i/3)}{t_i} \times \\ &\quad \exp\left\{ -\frac{\sigma^2}{6} \sum_{i=1}^{3} \sum_{i<j} (t_i - t_j)^2 \right\} dt_1 dt_2 dt_3. \end{aligned} \tag{78}$$

2) $L \ge 4$: In this case, (73) is equivalent to the following equation shown at the bottom of the next page, where $K \le \lfloor \frac{L}{2} \rfloor$.

As $\Delta r_1, \ldots, \Delta r_L$ are independent of each other for a given N, the joint pdf of $\Delta r = [\Delta r_1, \ldots, \Delta r_L]^T$ is

$$f(x_1, \ldots, x_L) = |2\pi V|^{-1/2} \exp\left\{ -\frac{1}{2}\mathbf{x}^T V^{-1} \mathbf{x} \right\}, \tag{80}$$

where the covariance matrix $V = \Lambda\Lambda^T = \Lambda^2$, $\Lambda = \mathrm{diag}(\sigma_1, \ldots, \sigma_L)$, and $\mathbf{x} = [x_1, \ldots, x_L]^T$. Then, the probability can be expressed as

$$p((\Delta r_1, \ldots, \Delta r_L) \in \Sigma) = \int_{\Sigma} \cdots \int |2\pi V|^{-1/2} \exp\left\{-\frac{1}{2}\mathbf{x}^T V^{-1}\mathbf{x}\right\} dx_1 \cdots dx_L, \quad (81)$$

where the integral region Σ is specified by (73) or (79).

When the remainder errors have the same variance, i.e., $\sigma_1 = \cdots = \sigma_L = \sigma$ or $w_1 = \cdots = w_L = 1/L$, we obtain the simpler form of (81) for $L = 4$ and $L > 4$ as follows.

① $L = 4$

In this case, condition (73) or (79) is equivalent to

$$\begin{cases} \left|3\Delta r_i - \sum_{j\neq i} \Delta r_j\right| < 3M/2,\ i = 1, \ldots, 4 \\ |\Delta r_1 + \Delta r_2 - \Delta r_3 - \Delta r_4| < M \\ |\Delta r_1 - \Delta r_2 + \Delta r_3 - \Delta r_4| < M \\ |\Delta r_1 - \Delta r_2 - \Delta r_3 + \Delta r_4| < M. \end{cases} \quad (82)$$

Let

$$\begin{cases} x_1 = \Delta r_1 + \Delta r_2 - \Delta r_3 - \Delta r_4 \\ x_2 = \Delta r_1 - \Delta r_2 + \Delta r_3 - \Delta r_4 \\ x_3 = \Delta r_1 - \Delta r_2 - \Delta r_3 + \Delta r_4. \end{cases} \quad (83)$$

Then, (82) can be simplified as

$$\begin{cases} |\ x_1 + x_2 + x_3\ | < 3M/2 \\ |\ x_1 + x_2 - x_3\ | < 3M/2 \\ |\ x_1 - x_2 + x_3\ | < 3M/2 \\ |\ -x_1 + x_2 + x_3\ | < 3M/2 \\ |\ x_i\ | < M, i = 1, 2, 3. \end{cases} \quad (84)$$

As $\Delta r_1, \ldots, \Delta r_4$ are independent of each other for a given N, x_i follow normal distributions with mean zero and variance $4\sigma^2$. Then, the joint pdf of $X = [x_1, x_2, x_3]^T$ is

$$f(x_1, x_2, x_3) = (2\sigma\sqrt{2\pi})^{-3} \exp\left\{-\frac{1}{8\sigma^2}(x_1^2 + x_2^2 + x_3^2)\right\}. \quad (85)$$

Thus,

$$p((\Delta r_1, \ldots, \Delta r_4) \in \Sigma) = p((x_1, x_2, x_3) \in \Delta) = \int\int\int_{\Delta} f(x_1, x_2, x_3) dx_1 dx_2 dx_3, \quad (86)$$

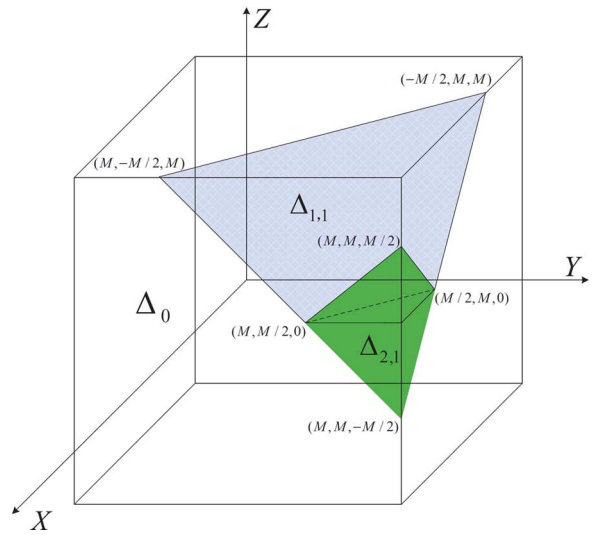

Fig. 2. Integral region Δ of (84).

where the integral region Δ is specified by (84). By carefully examining the region Δ, it can be obtained by Δ_0 subtracting $\Delta_{1,1}$ eight times and adding $\Delta_{2,1}$ twelve times as shown in Fig. 2, where Δ_0, $\Delta_{1,1}$ and $\Delta_{2,1}$ are

$$\Delta_0 = \{(x_1, x_2, x_3) : -M < x_1 < M, -M < x_2 < M, -M < x_3 < M\},$$
$$\Delta_{1,1} = \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 > 3M/2, x_1 < M, x_2 < M, x_3 < M\}$$

and

$$\Delta_{2,1} = \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 > 3M/2, x_1 < M, x_1 + x_2 - x_3 > 3M/2, x_2 < M\},$$

respectively. Then, (86) becomes

$$\begin{aligned} p((x_1, x_2, x_3) \in \Delta) &= 8\int_0^M \int_0^M \int_0^M f(x_1, x_2, x_3) dx_3 dx_2 dx_1 - \\ &\quad 8\int_{-\frac{M}{2}}^M \int_{\frac{M}{2}-x_1}^M \int_{\frac{3M}{2}-x_1-x_2}^M f(x_1, x_2, x_3) dx_3 dx_2 dx_1 + \\ &\quad 24\int_{\frac{M}{2}}^M \int_{\frac{M}{2}-x_1}^M \int_{\frac{3M}{2}-x_1-x_2}^0 f(x_1, x_2, x_3) dx_3 dx_2 dx_1. \end{aligned} \quad (87)$$

② $L > 4$

$$\begin{cases} \left|\Delta r_i - \sum_{j\neq i} \frac{w_j}{1-w_i}\Delta r_j\right| < \frac{M}{2},\ i = 1, \ldots, L \\ \left|\frac{1}{\sum_{k=1}^2 w_{i_k}} \sum_{k=1}^2 w_{i_k}\Delta r_{i_k} - \frac{1}{1-\sum_{k=1}^2 w_{i_k}} \sum_{j\neq i_1, i_2} w_j \Delta r_j\right| < \frac{M}{2} \\ \cdots \\ \left|\frac{1}{\sum_{k=1}^K w_{i_k}} \sum_{k=1}^K w_{i_k}\Delta r_{i_k} - \frac{1}{1-\sum_{k=1}^K w_{i_k}} \sum_{j\neq i_1, \ldots, i_K} w_j \Delta r_j\right| < \frac{M}{2}, \end{cases} \quad (79)$$

In this case, (81) can be simplified as

$$p((\Delta r_1,\ldots,\Delta r_L)\in\Sigma) = (2\pi\sigma^2)^{-L/2}\int_{\Sigma}\cdots\int\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{L}x_i^2\right\}dx_1\cdots dx_L. \quad (88)$$

V. Simulation Results

In the simulations, $M = 10$, moduli from Γ_1 to Γ_{10} are 2, 3, 5, 7, 11, 13, 17, 19, 23, 29. Real number N is uniformly distributed between M and $(M-1)\Gamma$. The remainder errors follow wrapped normal distributions with zero mean and variances σ_i^2, and the standard deviation σ_i is assumed to be proportional to a fraction of the modulus M_i [8], [15], i.e., $\sigma_i = \mu M_i$, where μ is a small positive factor. We call the process of determining N as a trial. In each trial, if N is robustly determined, the trial is passed, otherwise the trial is failed. By Theorem 4, we can express the trial fail rate (TFR) as

$$p_{\mathrm{TFR}}^{L} = 1 - p((\Delta r_1,\ldots,\Delta r_L)\in\Sigma). \quad (89)$$

The root mean square error (RMSE) of N is

$$N_{\mathrm{RMSE}} = \sqrt{\mathrm{E}\left\{\left|\hat{N}-N\right|^2\right\}}, \quad (90)$$

where $\mathrm{E}\{\cdot\}$ denotes the expectation. We obtain from Theorem 4 that the RMSE for the fast MLE algorithm is

$$N_{\mathrm{RMSE}} = \sqrt{\mathrm{E}\left\{\left|\overline{\Delta r}\right|^2\right\}} = \sqrt{\sum_{i=1}^{L} w_i^2\sigma_i^2}. \quad (91)$$

We first consider the TFR and RMSE performances of the fast MLE algorithm and the improved closed-form robust CRT [25], where the moduli Γ_i are 2, 3, 5, 7, 11, and the total number of trials is 10000 for each algorithm. We say that a test succeeds if N is robustly determined, i.e., its estimate $\hat{N}$ satisfies

$$\left|\hat{N}-N\right| \le \tau$$

where τ is a small nonnegative number. Otherwise, the test is failed and the TFR is given by (89). In Figs. 3(a) and 3(b), we show the TFR performances of the two algorithms versus the factor μ when $\tau = 2.5$ and 1.2, respectively. The figures show that the fast MLE algorithm has a better performance than the improved closed-form robust CRT in [25] when the remainders have different error levels, which is in agreement with theoretical analysis in Section IV. The improvement becomes more significant when the error level τ becomes smaller.

In Fig. 4, we compare the performances of the two algorithms by investigating their RMSE versus the factor μ. The theory curve for RMSE is based on (91). Fig. 4 shows that the RMSE of the two algorithms decrease linearly when $-20\log_{10}\mu$ is larger than 40, i.e., $\mu < 0.01$. It also shows that if $\mu < 0.01$, the RMSE of the proposed fast MLE algorithm matches the theoretical values very well as predicted by (91), but the improved closed-form robust CRT in [25] has a gap with the theoretical values, no matter how small μ is.

In Fig. 5, we show the TFR performance versus the number of remainders L, where all the variances are set to be equal. In the simulation, we set $\sigma = 1.25$ and 2.5, respectively. The theory curves are based on (78), (87), (88) in Section IV and (92) in [25]. As analyzed in Section III, the two algorithms, fast MLE algorithm and the improved closed-form robust CRT, have the same performance because the remainders have the same error level. The results also demonstrate that the TFR of the improved closed-form robust CRT in [25] does not match the theory obtained in [25] for a small L, which is due to the assumption of large L used in [25] in the derivations of the theoretical performance analysis. For the proposed fast MLE algorithm, the simulation and the theory are always matched very well no matter L is small or large.

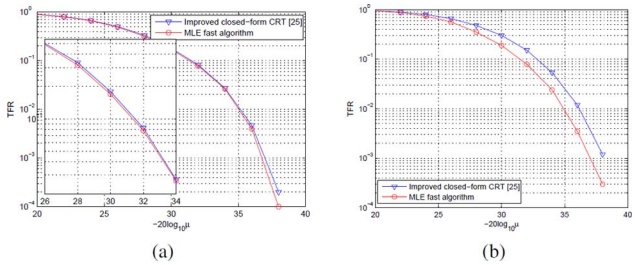

Fig. 3. TFR versus μ for different methods: (a) $\tau = 2.5$; (b) $\tau = 1.2$.

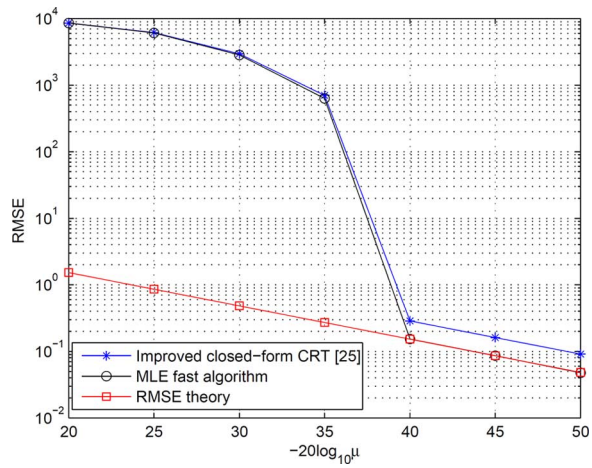

Fig. 4. RMSE versus μ for different methods.

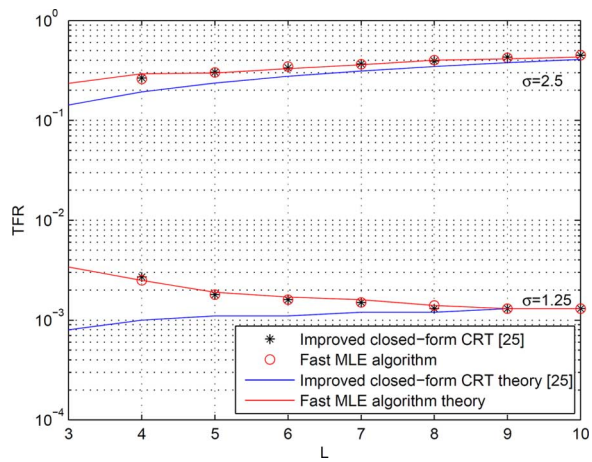

Fig. 5. TFR versus number of remainders L.

VI. Conclusion

In this paper, we have proposed an MLE based reconstruction of a real number from its smaller erroneous remainders modulo several moduli. We have proved that the MLE solution can be obtained when the estimation of the common remainder is optimal, where the remainder errors may not necessarily have the same variances. Instead of searching the common remainder from the infinite possibilities, we have obtained a novel method that finds the optimal common remainder by searching a finite set of L elements, where L is the number of remainders. This provides a fast algorithm for the MLE based reconstruction that only needs to search for the optimal among L many elements. Based on the proposed algorithm, we have obtained a necessary and sufficient condition on the remainder errors for the MLE to be robust, which is weaker than the existing known result. Simulations are provided to verify the efficiency of the proposed algorithm and the correctness of the theoretical analysis. Compared with the improved closed-form robust CRT in [25], it not only has a better performance especially when the remainders have different error levels/variances, but also has a much lower computational complexity.

Appendix

A. Proof of Theorem 3

Proof: Otherwise, by Lemma 1, the optimal estimate is $\hat{r}^c = \left\langle r^c + \overline{\Delta r} + M\sum_{i=1}^{t} w_{\nu_{(i)}} \right\rangle_M$ for some t with $1 \le t \le L-1$. Let

$$\delta = \sum_{i=1}^{L} w_i d_M^2\left(\hat{r}_i^c, \hat{r}^c\right) - \sum_{i=1}^{L} w_i d_M^2\left(\hat{r}_i^c, \left\langle r^c + \overline{\Delta r}\right\rangle_M\right).$$

Then, we obtain from the definition of the optimal estimate $\hat{r}^c$ in (38) that

$$\delta < 0. \tag{92}$$

Let $S = \{\Delta r_i\}$ for any $1 \le i \le L$. Then, we obtain from (67) that

$$-M/2 \le \Delta r_i - \sum_{j\ne i} \frac{w_j}{\sum_{j\ne i} w_j}\Delta r_j < M/2.$$

Since $\sum_{j\ne i} w_j = 1 - w_i$, we have

$$-M/2 \le \frac{1}{1-w_i}\Delta r_i - \frac{1}{1-w_i}\overline{\Delta r} < M/2.$$

Consequently,

$$-M(1-w_i)/2 \le \Delta r_i - \overline{\Delta r} < M(1-w_i)/2. \tag{93}$$

Recall that $0 \le \sum_{i=1}^{t} w_{\nu_{(i)}} \le 1$ for $1 \le t \le L$. Hence,

$$-3M/2 \le \Delta r_i - \overline{\Delta r} - M\sum_{i=1}^{t} w_{\nu_{(i)}} < M/2. \tag{94}$$

Then, all the remainder errors can be categorized by two sets, S_1 and its complement set $\overline{S}_1$, with

$$S_1 = \left\{\Delta r_i : -\frac{3M}{2} \le \Delta r_i - \overline{\Delta r} - M\sum_{i=1}^{t} w_{\nu_{(i)}} < -\frac{M}{2}\right\}$$

and

$$\overline{S}_1 = \left\{\Delta r_i : -\frac{M}{2} \le \Delta r_i - \overline{\Delta r} - M\sum_{i=1}^{t} w_{\nu_{(i)}} < \frac{M}{2}\right\}.$$

Note that the optimal estimate $\hat{r}^c$ satisfies $f'(\hat{r}^c) = 0$ as previously shown in the proof of Theorem 2, where $f(x)$ is the objective function in the right-hand side of (38) as defined in the beginning of the proof of Theorem 2. That is,

$$\frac{d}{dx}\sum_{i=1}^{L} w_i d_M^2\left(\hat{r}_i^c, x\right)\Big|_{x=\hat{r}^c} = 0.$$

Then,

$$\sum_{i=1}^{L} w_i d_M(\hat{r}_i^c, \hat{r}^c) = 0.$$

Note that $d_M(\hat{r}_i^c, \hat{r}^c) = d_M(\hat{r}_i, \hat{r}^c)$. By (2), (7), and (26) and Property 1 in (15), we have

$$\begin{aligned}
0 &= \sum_{i=1}^{L} w_i d_M\left(\hat{r}_i, \left\langle r^c + \overline{\Delta r} + M\sum_{i=1}^{t} w_{\nu_{(i)}}\right\rangle_M\right)\\
&= \sum_{i=1}^{L} w_i d_M\left(Mq_i + r^c + \Delta r_i, r^c + \overline{\Delta r} + M\sum_{i=1}^{t} w_{\nu_{(i)}}\right)\\
&= \sum_{i=1}^{L} w_i d_M\left(\Delta r_i, \overline{\Delta r} + M\sum_{i=1}^{t} w_{\nu_{(i)}}\right)\\
&= \sum_{\Delta r_i \in S_1} w_i\left(\Delta r_i - \overline{\Delta r} - M\sum_{i=1}^{t} w_{\nu_{(i)}} + M\right) +\\
&\quad \sum_{\Delta r_i \in \overline{S}_1} w_i\left(\Delta r_i - \overline{\Delta r} - M\sum_{i=1}^{t} w_{\nu_{(i)}}\right).
\end{aligned}$$

By re-organizing the right-hand side of the last equation above, we have

$$\sum_{i=1}^{t} w_{\nu_{(i)}} = \sum_{\Delta r_i \in S_1} w_i. \tag{95}$$

Note that

$$\begin{aligned}
\delta &= \sum_{i=1}^{L} w_i d_M^2(\hat{r}_i, \hat{r}^c) - \sum_{i=1}^{L} w_i d_M^2(\hat{r}_i, \langle r^c + \overline{\Delta r}\rangle_M)\\
&= \sum_{i=1}^{L} w_i d_M^2\left(Mq_i + r^c + \Delta r_i, r^c + \overline{\Delta r} + M\sum_{i=1}^{t} w_{\nu_{(i)}}\right) -\\
&\quad \sum_{i=1}^{L} w_i d_M^2\left(Mq_i + r^c + \Delta r_i, r^c + \overline{\Delta r}\right)\\
&= \sum_{i=1}^{L} w_i d_M^2\left(\Delta r_i, \overline{\Delta r} + M\sum_{i=1}^{t} w_{\nu_{(i)}}\right) - \sum_{i=1}^{L} w_i d_M^2\left(\Delta r_i, \overline{\Delta r}\right).
\end{aligned}$$

From (93), we have

$$-M/2 \le \Delta r_i - \overline{\Delta r} < M/2,$$

which leads to

$$d_M^2(\Delta r_i, \overline{\Delta r}) = (\Delta r_i - \overline{\Delta r})^2.$$

Hence,

$$\delta = \sum_{\Delta r_i \in S_1} w_i \left(\Delta r_i - \overline{\Delta r} - M \sum_{i=1}^{t} w_{\nu_{(i)}} + M \right)^2 + \sum_{\Delta r_i \in \overline{S}_1} w_i \left(\Delta r_i - \overline{\Delta r} - M \sum_{i=1}^{t} w_{\nu_{(i)}} \right)^2 - \sum_{i=1}^{L} w_i \left(\Delta r_i - \overline{\Delta r} \right)^2. \tag{96}$$

By (95), we can simplify (96) as

$$\delta = 2M \sum_{\Delta r_i \in S_1} w_i \left(\Delta r_i - \overline{\Delta r} \right) + M^2 \left(\sum_{\Delta r_i \in S_1} w_i \right) \left(1 - \sum_{\Delta r_i \in S_1} w_i \right).$$

According to (92), $\delta < 0$, that is,

$$\sum_{\Delta r_i \in S_1} w_i \left(\Delta r_i - \overline{\Delta r} \right) < -\frac{M}{2} \left(\sum_{\Delta r_i \in S_1} w_i \right) \left(1 - \sum_{\Delta r_i \in S_1} w_i \right),$$

then we have

$$\sum_{\Delta r_i \in S_1} w_i \Delta r_i - \left(\sum_{\Delta r_i \in S_1} w_i \Delta r_i + \sum_{\Delta r_i \in \overline{S}_1} w_i \Delta r_i \right) \sum_{\Delta r_i \in S_1} w_i < -\frac{M}{2} \left(\sum_{\Delta r_i \in S_1} w_i \right) \left(\sum_{\Delta r_i \in \overline{S}_1} w_i \right).$$

Here, we use the equations

$$\overline{\Delta r} = \sum_{\Delta r_i \in S_1} w_i \Delta r_i + \sum_{\Delta r_i \in \overline{S}_1} w_i \Delta r_i$$

and

$$1 - \sum_{\Delta r_i \in S_1} w_i = \sum_{\Delta r_i \in \overline{S}_1} w_i.$$

Thus,

$$\left(\frac{1}{\left(\sum_{\Delta r_i \in S_1} w_i \right) \left(\sum_{\Delta r_i \in \overline{S}_1} w_i \right)} - \frac{1}{\sum_{\Delta r_i \in \overline{S}_1} w_i} \right) \sum_{\Delta r_i \in S_1} w_i \Delta r_i - \frac{1}{\sum_{\Delta r_i \in \overline{S}_1} w_i} \sum_{\Delta r_i \in \overline{S}_1} w_i \Delta r_i < -\frac{M}{2},$$

and then

$$\frac{1}{\sum_{\Delta r_i \in S_1} w_i} \sum_{\Delta r_i \in S_1} w_i \Delta r_i - \frac{1}{\sum_{\Delta r_i \in \overline{S}_1} w_i} \sum_{\Delta r_i \in \overline{S}_1} w_i \Delta r_i < -\frac{M}{2},$$

which contradicts with (67). Thus, $\hat{r}^c = \left\langle r^c + \overline{\Delta r} \right\rangle_M$.

By the definition of Δr^c in (50), we can get (69) directly. ■

B. Proof of Theorem 4

Proof: We now show the sufficiency. Note that in this case the conditions of Theorem 3 are satisfied. According to the three cases of Δr^c in Theorem 3, we have $\Delta r^c = \overline{\Delta r} + M$, $\Delta r^c = \overline{\Delta r}$ and $\Delta r^c = \overline{\Delta r} - M$, respectively.

In the case when $\Delta r^c = \overline{\Delta r} + M$, we obtain from (2), (7), (13) and (28) that

$$\hat{q}_i = q_i + \left[\frac{\Delta r_i - \Delta r^c}{M} \right] = q_i - 1 + \left[\frac{\Delta r_i - \overline{\Delta r}}{M} \right].$$

As shown in the proof of Theorem 3, (71) can lead to

$$-M/2 \le \Delta r_i - \overline{\Delta r} < M/2,\ i = 1, \ldots, L.$$

Hence,

$$\hat{q}_i = q_i - 1,\ i = 1, \ldots, L,$$

and then

$$\hat{N}_0 = \left\langle \sum_{i=1}^{L} \bar{\gamma}_i \gamma_i (q_i - 1) \right\rangle_\Gamma = N_0 - 1.$$

Therefore,

$$\hat{N} = M(N_0 - 1) + r^c + \Delta r^c = N + \overline{\Delta r}.$$

Similarly, in the cases when $\Delta r^c = \overline{\Delta r}$ and $\Delta r^c = \overline{\Delta r} - M$, we can get the same conclusion (70). This completes the proof of the sufficiency.

We next show the necessity. Let $\epsilon = \overline{\Delta r}$ in (47), then (55) is equivalent to

$$-M/2 \le \Delta r_i - \overline{\Delta r} < M/2. \tag{97}$$

Suppose that there exists a nonempty set $S_1 \subseteq \{\Delta r_1, \ldots, \Delta r_L\}$ satisfying

$$\sum_{\Delta r_i \in S_1} \frac{w_i}{\sum_{\Delta r_j \in S_1} w_j} \Delta r_i - \sum_{\Delta r_i \in \overline{S}_1} \frac{w_i}{\sum_{\Delta r_j \in \overline{S}_1} w_j} \Delta r_i < -\frac{M}{2}.$$

That is,

$$\frac{1}{\sum_{\Delta r_i \in S_1} w_i} \sum_{\Delta r_i \in S_1} w_i \Delta r_i - \frac{1}{1 - \sum_{\Delta r_i \in S_1} w_i} \sum_{\Delta r_i \in \overline{S}_1} w_i \Delta r_i < -\frac{M}{2},$$

and then

$$\sum_{\Delta r_i \in S_1} w_i \left(\Delta r_i - \overline{\Delta r} \right) < -\frac{M}{2} \left(\sum_{\Delta r_i \in S_1} w_i \right) \left(1 - \sum_{\Delta r_i \in S_1} w_i \right). \tag{98}$$

Let $\tilde{r}^c = \left\langle r^c + \overline{\Delta r} + M\sum_{\Delta r_i \in S_1} w_i \right\rangle_M$, and let

$$\delta = \sum_{i=1}^{L} w_i d_M^2\left(\hat{r}_i^c, \tilde{r}^c\right) - \sum_{i=1}^{L} w_i d_M^2\left(\hat{r}_i^c, \left\langle r^c + \overline{\Delta r}\right\rangle_M\right).$$

Then, we have

$$\begin{aligned}
\delta &= \sum_{i=1}^{L} w_i d_M^2\left(\hat{r}_i, r^c + \overline{\Delta r} + M\sum_{\Delta r_i \in S_1} w_i\right) - \\
&\quad \sum_{i=1}^{L} w_i d_M^2\left(\hat{r}_i, r^c + \overline{\Delta r}\right) \\
&= \sum_{i=1}^{L} w_i d_M^2\left(Mq_i + r^c + \Delta r_i, r^c + \overline{\Delta r} + M\sum_{\Delta r_i \in S_1} w_i\right) - \\
&\quad \sum_{i=1}^{L} w_i d_M^2\left(Mq_i + r^c + \Delta r_i, r^c + \overline{\Delta r}\right) \\
&= \sum_{i=1}^{L} w_i d_M^2\left(\Delta r_i, \overline{\Delta r} + M\sum_{\Delta r_i \in S_1} w_i\right) - \\
&\quad \sum_{i=1}^{L} w_i d_M^2\left(\Delta r_i, \overline{\Delta r}\right) \\
&= \sum_{\Delta r_i \in S_1} w_i d_M^2\left(\Delta r_i + M, \overline{\Delta r} + M\sum_{\Delta r_i \in S_1} w_i\right) + \\
&\quad \sum_{\Delta r_i \in \overline{S}_1} w_i d_M^2\left(\Delta r_i, \overline{\Delta r} + M\sum_{\Delta r_i \in S_1} w_i\right) - \\
&\quad \sum_{i=1}^{L} w_i\left(\Delta r_i - \overline{\Delta r}\right)^2.
\end{aligned}$$

According to Property 6 in (20), we obtain

$$\begin{aligned}
\delta &\le \sum_{\Delta r_i \in S_1} w_i\left(\Delta r_i - \overline{\Delta r} - M\sum_{\Delta r_i \in S_1} w_i + M\right)^2 + \\
&\quad \sum_{\Delta r_i \in \overline{S}_1} w_i\left(\Delta r_i - \overline{\Delta r} - M\sum_{\Delta r_i \in S_1} w_i\right)^2 - \\
&\quad \sum_{i=1}^{L} w_i\left(\Delta r_i - \overline{\Delta r}\right)^2 \\
&= M^2\left(\sum_{\Delta r_i \in S_1} w_i\right)\left(1 - \sum_{\Delta r_i \in S_1} w_i\right) + \\
&\quad 2M\sum_{\Delta r_i \in S_1} w_i\left(\Delta r_i - \overline{\Delta r}\right).
\end{aligned}$$

It follows from (98) that

$$\delta < 0. \tag{99}$$

According to the definition of $\hat{r}^c$ in (27), we obtain that $\left\langle r^c + \overline{\Delta r}\right\rangle_M$ is not the optimal estimate $\hat{r}^c$, i.e.,

$$\hat{r}^c \neq \left\langle r^c + \overline{\Delta r}\right\rangle_M. \tag{100}$$

On the other hand, we obtain from (30) and (70) that

$$M\hat{N}_0 + \hat{r}^c - MN_0 - r^c = \overline{\Delta r}.$$

Hence,

$$\hat{r}^c = \left\langle r^c + \overline{\Delta r}\right\rangle_M. \tag{101}$$

Clearly, (100) contradicts with (101).

In the case when

$$\sum_{\Delta r_i \in S_1} \frac{w_i}{\sum_{\Delta r_j \in S_1} w_j}\Delta r_i - \sum_{\Delta r_i \in \overline{S}_1} \frac{w_i}{\sum_{\Delta r_j \in \overline{S}_1} w_j}\Delta r_i \ge \frac{M}{2},$$

we let

$$\tilde{r}^c = \left\langle r^c + \overline{\Delta r} - M\sum_{\Delta r_i \in S_1} w_i \right\rangle_M,$$

then the same contradiction occurs. Therefore, the remainder errors satisfy (71). This completes the proof of the necessity. ■

ACKNOWLEDGMENT

The authors would like to thank the reviewers for their useful comments that have helped to improve the presentation of this paper.

REFERENCES

[1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd ed. New York, NY, USA: Wiley, 2003.
[2] E. Berlekamp, *Algebraic Coding Theory*. New York, NY, USA: McGraw-Hill, 1968.
[3] J. Bioucas-Dias, V. Katkovnik, J. Astola, and K. Egiazarian, "Multifrequency phase unwrapping from noisy data: adaptive local maximum likelihood approach," *Image Analysis, Lecture Notes in Computer Science*, vol. 5575/2009, pp. 310–320, Jul. 2009.
[4] C. Ding, D. Pei, and A. Salomaa, *Chinese Remainder Theorem: Applications in Computing, Coding, Cryptography*. Singapore: World Scientific, 1999.
[5] Z.-X. Huang and Z. Wan, "Range ambiguity resolution in multiple PRF pulse Doppler radars," in *Proc. ICASSP*, Dallas, TX, USA, Apr. 1987, pp. 1786–1789.
[6] D. P. Jorgensen, T. R. Shepherd, and A. S. Goldstein, "A dual-pulse repetition frequency scheme for mitigating velocity ambiguities of the NOAA P-3 airborne doppler radar," *J. Atmos. Ocean. Technol.*, vol. 17, no. 5, pp. 585–594, May 2000.
[7] N. Koblitz, *A Course in Number Theory and Cryptography*. New York, NY, USA: Springer-Verlag, 1994.
[8] B. Kusy, A. Ledeczi, M. Maroti, and L. Meertens, "Node-density independent localization," presented at the 5th Int. Conf. Inf. Process. Sens. Netw., Nashville, TN, USA, Apr. 2006.
[9] G. Li, J. Xu, Y.-N. Peng, and X.-G. Xia, "Location and imaging of moving targets using non-uniform linear antenna array," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 3, pp. 1214–1220, Jul. 2007.
[10] G. Li, H. Meng, X.-G. Xia, and Y.-N. Peng, "Range and velocity estimation of moving targets using multiple stepped-frequency pulse trains," *Sens.*, vol. 8, pp. 1343–1350, 2008.
[11] X. W. Li and X.-G. Xia, "Location and imaging of elevated moving target using multi-frequency velocity SAR with cross-track interferometry," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 2, pp. 1203–1212, Apr. 2011.
[12] G. Li, J. Xu, Y.-N. Peng, and X.-G. Xia, "An efficient implementation of a robust phase-unwrapping algorithm," *IEEE Signal Process. Lett.*, vol. 14, no. 6, pp. 393–396, Jun. 2007.
[13] X. W. Li and X.-G. Xia, "A fast robust Chinese remainder theorem based phase unwrapping algorithm," *IEEE Signal Process. Lett.*, vol. 15, pp. 665–668, Oct. 2008.
[14] X. W. Li, H. Liang, and X.-G. Xia, "A robust Chinese remainder theorem with its applications in frequency estimation from undersampled waveforms," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4314–4322, Nov. 2009.
[15] W. C. Li, X. Z. Wang, X. M. Wang, and B. Moran, "Distance estimation using wrapped phase measurements in noise," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1676–1688, Apr. 2013.
[16] K. V. Mardia and P. E. Jupp, *Directional Statistics*, 2nd ed. New York, NY, USA: Wiley, 2000.
[17] J. H. McClellan and C. M. Rader, *Number Theory in Signal Digital Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1979.
[18] O. Ore, "The general Chinese remainder theorem," *Amer. Math. Month.*, vol. 59, no. 6, pp. 365–370, Jun.-Jul. 1952.
[19] E. Parzen, *Modern Probability Theory and its Applications*. New York, NY, USA: Wiley, 1992.
[20] W.-K. Qi, Y.-W. Dang, and W.-D. Yu, "Deblurring velocity ambiguity of distributed space-borne SAR based on Chinese remainder theorem," *J. Electron. Inf. Technol.*, vol. 31, no. 10, pp. 2493–2496, Oct. 2009.

[21] K. H. Rosen, *Elementary Number Theory and its Applications*, 5th ed. Dedham, MA, USA: Addison-Wesley, 2010.
[22] M. Ruegg, E. Meier, and D. Nuesch, "Capabilities of dual-frequency millimeter wave SAR with monopulse processing for ground moving target indication," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 539–553, Mar. 2007.
[23] I. E. Shparlinski and R. Steinfeld, "Noisy Chinese remaindering in the Lee norm," *J. Complex.*, vol. 20, pp. 423–437, 2004.
[24] I. Vrana, "Optimal ststistical estimates in conditions of ambiguity," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 1023–1030, May 1993.
[25] W. J. Wang and X.-G. Xia, "A closed-form robust Chinese remainder theorem and its performance analysis," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5655–5666, Nov. 2010.
[26] G. Wang, X.-G. Xia, V. C. Chen, and R. L. Fiedler, "Detection, location, and imaging of fast moving targets using multifrequency antenna array SAR," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 1, pp. 345–355, Jan. 2004.
[27] C. Wang, Q. Y. Yin, and W. J. Wang, "An efficient ranging method for wireless sensor networks," in *Proc. ICASSP*, Dallas, TX, USA, Mar. 2010, pp. 2846–2849.
[28] X.-G. Xia and G. Wang, "Phase unwrapping and a robust Chinese remainder theorem," *IEEE Signal Process. Lett.*, vol. 14, no. 4, pp. 247–250, Apr. 2007.
[29] L. Xiao, X.-G. Xia, and W. J. Wang, "Multi-stage robust Chinese remainder theorem," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4772–4785, Sep. 2014.
[30] G. W. Xu, "On solving a generalized chinese remainder theorem in the presence of remainder errors," Aug. 2014 [Online]. Available: http://arxiv-web3.library.cornell.edu/pdf/1409.0121.pdf
[31] W. Xu, E. C. Chang, L. K. Kwoh, H. Lim, and W. C. A. Heng, "Phase unwrapping of SAR interferogram with multi-frequency or multi-baseline," in *Proc. IGARSS*, 1994, pp. 730–732.
[32] B. Yang, W. J. Wang, Q. Y. Yin, and X.-G. Xia, "Phase detection based range estimation with a dual-band robust Chinese remainder theorem," *Sci. China-Inf. Sci.*, vol. 57, no. 2, pp. 1–9, Feb. 2014.
[33] Z. H. Yuan, Y. K. Deng, F. Li, R. Wang, G. Liu, and X. L. Han, "Multichannel InSAR DEM reconstruction through improved closed-form robust Chinese remainder theorem," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1314–1318, Nov. 2013.
[34] Y. M. Zhang and M. Amin, "MIMO radar exploiting narrowband frequency-hopping waveforms," presented at the 16th Eur. Signal Process. Conf. (EUSIPCO 2008), Lausanne, Switzerland, Aug. 25–29, 2008.

Wenjie Wang (M'10) received the B.S., M.S., and Ph.D. degrees in information and communication engineering from Xi'an Jiaotong University, Xi'an, China, in 1993, 1998, and 2001, respectively.

From 2009 to 2010, he was a visiting scholar with the Department of Electrical and Computer Engineering, University of Delaware, Newark, USA. Currently, he is a Professor at Xi'an Jiaotong University. His main research interests include information theory, broadband wireless communications, signal processing with applications in communication systems, array signal processing, and cooperative communications in distributed networks.

Xiaoping Li received the B.S. degree in mathematics from Sichuan Normal University, Chengdu, China, in 2006.

From 2006 to 2010, he was an Assistant Professor in the College of Information Engineering, Tarim University, Alar, China. From December 2013 to January 2015, he was a visiting scholar with the Department of Electrical and Computer Engineering, University of Delaware, Newark, USA. Currently, he is working toward the Ph.D. degree with the School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an, China. His main research interests include signal processing with applications in communication systems and coding theory.

Wei Wang received the B.S. and M.S. degrees from Central South University, Changsha, China, in 2004 and 2006, respectively, both in mathematics.

Currently, he is an Associate Professor in the College of Information Engineering, Tarim University, Alar, China. His main research interests are digital signal processing and graph theory especially graph coloring.

Prof. Wang has been a reviewer for *Mathematical Reviews* since 2013.

Xiang-Gen Xia (M'97–S'00–F'09) received the B.S. degree in mathematics from Nanjing Normal University, Nanjing, China, and the M.S. degree in mathematics from Nankai University, Tianjin, China, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, USA, in 1983, 1986, and 1992, respectively.

He was a Senior/Research Staff Member at Hughes Research Laboratories, Malibu, CA, during 1995–1996. In September 1996, he joined the Department of Electrical and Computer Engineering, University of Delaware, Newark, USA, where he is the Charles Black Evans Professor. His current research interests include space-time coding, MIMO and OFDM systems, digital signal processing, and SAR and ISAR imaging. He is the author of the book *Modulated Coding for Intersymbol Interference Channels* (New York: Marcel Dekker, 2000).

Dr. Xia received the National Science Foundation (NSF) Faculty Early Career Development (CAREER) Program Award in 1997, the Office of Naval Research (ONR) Young Investigator Award in 1998, and the Outstanding Overseas Young Investigator Award from the National Nature Science Foundation of China in 2001. He also received the Outstanding Junior Faculty Award of the Engineering School of the University of Delaware in 2001. He is currently serving and has served as an Associate Editor for numerous international journals including the IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON MOBILE COMPUTING, and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He is Technical Program Chair of the Signal Processing Symp., Globecom 2007 in Washington DC, and the General Co-Chair of ICASSP 2005 in Philadelphia.