Explicit Correction Probabilities for Montgomery Multiplication via Arithmetic Analysis

Daichi Aoki
NEC Corporation
Kanagawa, Japan
daichi_aoki@nec.com

Tsuyoshi Takagi
The University of Tokyo
Tokyo, Japan
takagi@g.ecc.u-tokyo.ac.jp

***Abstract*—Montgomery multiplication is a core primitive in public-key and post-quantum cryptography, replacing division operations with multiplications and shifts. Its final conditional subtraction, called the correction step, can be implemented in constant time, but its occurrence remains input-dependent and may contribute to leak secret information through timing or micro-architectural effects. Thus, correction-probability analysis is important beyond implementation-level countermeasures. Existing analyses mainly consider moduli close to the Montgomery radix and often rely on heuristic distributional assumptions, leaving the small-modulus regime largely unexplored.**

We give a simple number-theoretic characterization of the Montgomery correction condition for any odd modulus and non-negative inputs, and analyze two input regimes. In the modulus-size setting $[0, m)^2$, we prove that for $m^2 < R$ the correction probability is exactly $(P(m) - 2m + 1)/m^2$, where P is the gcd-sum function, making the dependence on the multiplicative structure of m explicit. In the full-size arithmetic counting model $0 < xy \leq mR$, we derive an exact divisor-sum representation and a heuristic main-term approximation, based on elementary averaging and Dirichlet's divisor-sum estimate, explaining why the probability is close to $1/2$ when m is close to R. Numerical experiments support the analysis and illustrate how the arithmetic structure of the modulus shapes Montgomery correction events.

***Index Terms*—Montgomery multiplication, correction probability, analytic number theory, gcd-sum function, divisor function.**

I. Introduction

Efficient modular reduction is a core primitive across public-key and post-quantum cryptography: RSA relies on modular exponentiation, while lattice-based schemes such as Kyber [1] and Dilithium [2] rely heavily on modular multiplications in their polynomial arithmetic. Among the available modular reduction techniques, Montgomery's method [3], [4], [5] and Barrett's method [6] remain dominant due to their low-cost use of multiplication and logical shifts instead of division. In contrast, Plantard-style reductions [7], [8], [9] are attractive alternatives in word-size arithmetic but differ in the structure of their final correction steps.

A distinct aspect of Montgomery and Barrett reduction is the final range-correction, a conditional subtraction of the modulus. While this subtraction can be implemented in constant time, the need for the correction itself depends on the input. Hence, a quantitative analysis of the correction probability is useful for assessing potential leakage beyond implementation-level countermeasures. As demonstrated by prior timing and micro-architectural attacks [10], [11], the distribution of this correction event can leak information about secret values. Although branchless implementations can avoid direct timing leakage from the conditional subtraction itself, the correction event remains an input-dependent arithmetic condition and is relevant to implementations where it affects timing, micro-architectural behavior, or lazy-reduction strategies.

Earlier works [12], [13] analyze the correction probability primarily in the regime where the modulus m is close to the Montgomery radix $R = 2^n$, leading to estimated probabilities of approximately $1/3$ for squaring and $1/4$ for general multiplication. However, these works leave open the question of how the correction probability behaves when m is small relative to R, a common scenario in word-size arithmetic and CRT sub-moduli.

This work fills that gap by reformulating Montgomery's correction condition in a simple number-theoretic form and analyzing its behavior in two regimes:

1) **Modulus-size setting** $(x, y) \in [0, m)^2$: We express the correction condition for input pairs (x, y) by congruence and order constraints and derive an exact formula for the correction probability as a sum involving the greatest common divisors. In particular, we show that for small moduli $m < \sqrt{R}$, the correction probability is expressed using the gcd-sum function $P(m)$. This makes the dependence on the multiplicative structure of m explicit, allowing precise comparison across different choices of moduli such as primes, prime powers, and semiprimes.
2) **Full-size setting** $xy \in (0, mR]$: We express the correction count as a divisor-sum over residue classes determined by the Montgomery correction condition. We use elementary divisor-sum decompositions together with Dirichlet's classical estimate for $\sum_{n \leq X} \tau(n)$. This yields a transparent main-term approximation for the correction probability and explains the numerical observation that the probability is close to $1/2$ in the tested full-size regime. The full-size analysis should be interpreted as an exact counting identity followed by a heuristic main-term evaluation based on averaging over primitive residue classes.

Together, these results provide a unified arithmetic de-

scription of Montgomery's correction probability across both modulus-size and full-size input models. In addition to clarifying the theoretical behavior of the correction event, our analysis may lead to a more detailed understanding of side-channel leakage patterns in actual implementations.

II. Preliminaries

In this section, we outline some fundamental topics in analytic number theory, following Apostol's textbook [14].

Let $\mathbb{Z}_{>0}$ denote the set of positive integers. An arithmetic function is a map $f : \mathbb{Z}_{>0} \to \mathbb{C}$. We use standard arithmetic functions: Euler's totient function $\varphi(n) = |(\mathbb{Z}/n\mathbb{Z})^\times|$, the divisor function $\tau(n) = \sum_{d|n} 1$, and the Möbius function μ. We also use the gcd-sum function

$$P(n) = \sum_{k=1}^{n} \gcd(k, n).$$

Definition 1. The Möbius function μ is defined as follows:

$$\mu(1) = 1;$$

If $n = p_1^{e_1} \cdots p_k^{e_k}$, where $p_1, \ldots, p_k$ are distinct primes and each $e_i \geq 1$, then

$$\mu(n) = \begin{cases} (-1)^k & \text{if each } e_i = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Definition 2 (Dirichlet convolution)**.** If f and g are arithmetic functions, then their Dirichlet convolution $f * g$ is defined as follows:

$$(f * g)(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right).$$

It is well known that

$$P = N * \varphi = N * N * \mu,$$

where $N(n) = n$. Hence P is multiplicative, and for any prime power p^a,

$$P(p^a) = (a+1)p^a - ap^{a-1}.$$

III. Montgomery Multiplication

Modular multiplication typically involves two steps: computing integer products and computing modular reductions. The modular reduction step of $xy \bmod m$ can naively be computed using division as follows:

$$xy \bmod m = xy - \left\lfloor \frac{xy}{m} \right\rfloor m.$$

However, in general, performing divisions on a CPU is a costly process. Montgomery multiplication [3] is an efficient method for modular multiplication. As shown in Algorithm 1, Montgomery multiplication uses the unsigned low product and logical right shift operations. Both operations are efficient on modern CPUs.

In some implementations, the final subtraction is omitted or delayed by lazy reduction, allowing intermediate values to remain in a larger range. Nevertheless, the correction condition remains relevant for analyzing timing, micro-architectural behavior, and the validity of output-range assumptions.

Algorithm 1 Montgomery multiplication [3]

Input: integers $R = 2^n$ and $m < R$ with $\gcd(m, R) = 1$
Input: an integer $m' \coloneqq -m^{-1} \bmod R$
Input: integers $0 \leq x, y < m$
Output: $A \equiv xyR^{-1} \bmod m$ with $0 \leq A < m$

1: $u \leftarrow xym' \bmod R$ ▷ unsigned low product
2: $A \leftarrow (xy + um)/R$ ▷ logical right shift
3: **if** $A \geq m$ **then** $A \leftarrow A - m$ ▷ correction step
4: **return** A

For inputs $x, y \geq 0$, $\tilde{A} = \frac{xy + (xym' \bmod R)m}{R} < \frac{xy}{R} + m$ holds. Then the correction step is performed if and only if $\tilde{A} \geq m$. It is often assumed that either $xy \leq mR$ or $x, y \in [0, m)$ in order to guarantee that after at most one correction the output satisfies $A \in [0, m)$.

In summary, the correction step is performed if and only if

$$xy + (xym' \bmod R)m \geq mR. \tag{1}$$

We show that this condition can be expressed more simply.

Theorem 3. *Let R be a power-of-two integer and m be an odd modulus $m < R$. For integers $x, y \geq 0$, the correction step of Montgomery multiplication is performed if and only if there exists an integer $k \geq 0$ such that*

$$\begin{cases} xy \equiv kR \bmod m, \\ xy > kR. \end{cases} \tag{2}$$

Proof. First, we show (2) $\Rightarrow$ (1). Assuming (2), there exists an integer $\ell > 0$ satisfying $xy = kR + \ell m$. Let $\ell' = \lceil \ell / R \rceil$. The integer $\ell' \geq 1$ satisfies $(\ell' - 1)R < \ell \leq \ell' R$. Then we have

$$\begin{aligned} xy + (xym' \bmod R)m &= xy + (xy(-m^{-1}) \bmod R)m \\ &= kR + \ell m + (-\ell \bmod R)m \\ &= kR + \ell m + (\ell' R - \ell)m \\ &= (k + \ell' m)R \\ &\geq mR. \end{aligned}$$

Next, we prove the converse: (1) $\Rightarrow$ (2). As the left-hand side of (1) is divisible by R, there exists an integer $k \geq 0$ satisfying $xy + (xym' \bmod R)m = (m + k)R$. Therefore,

$$\begin{aligned} & xy - kR = m(R - (xym' \bmod R)) > 0 \\ \Rightarrow \quad & xy \equiv kR \bmod m \text{ and } xy > kR. \end{aligned}$$

□

Although Algorithm 1 is stated for the standard case $0 \leq x, y < m$, the correction condition itself is meaningful for arbitrary non-negative inputs. In the subsequent analysis, we consider two finite input models: the modulus-size model $x, y \in [0, m)$, and the full-size arithmetic counting model $0 < xy \leq mR$.

IV. Analysis in Modulus-Size Setting

In this section, we examine the probability of the correction step for the input $(x, y) \in [0, m) \times [0, m)$. By using Theorem 3, we can determine the number of inputs that cause a correction step for each k. We further investigate the probabilities under the additional condition that the modulus m is small, namely $m < \sqrt{R}$. The asymptotic behavior in this modulus-size input model when m approaches R remains a future challenge.

A. Exact Count of Inputs Triggering a Correction

First, we define the subset of $[0, m)^2$ that satisfies the correction condition (2).

Definition 4. Let $S_m(k)$ be the set of pairs $(x, y) \in [0, m) \times [0, m)$ that satisfy (2) for k. Let $S(m)$ be the union of $S_m(k)$ for all possible values of k, i.e.,

$$S_m(k) = \left\{ (x, y) \in \mathbb{Z}^2_{\geq 0} \;\middle|\; \begin{array}{l} xy \equiv kR \bmod m, \\ xy > kR, \\ x, y < m \end{array} \right\},$$

$$S(m) = \bigcup_{k=0}^{m-1} S_m(k).$$

We note that $0 \leq k < m$ in this setting. For any distinct $0 \leq k, k' < m$, the sets $S_m(k)$ and $S_m(k')$ are disjoint. Hence $|S(m)| = \sum_k |S_m(k)|$.

To make the correction condition in Theorem 3 more intuitive, Fig. 1 visualizes the correction sets $S(m)$ with $R = 256$ for two small moduli, $m = 37$ and $m = 39$. Each plotted point corresponds to an input pair $(x, y) \in [0, m)^2$ that causes a correction step. For a fixed k, the inequality $xy > kR$ restricts the points to the region above the hyperbola $xy = kR$, while the congruence $xy \equiv kR \pmod m$ selects arithmetic progressions across that region. The comparison between $m = 37$ (prime) and $m = 39$ (composite) highlights how the multiplicative structure of the modulus affects the resulting distribution of correction events. As a result, corrections are not uniformly scattered in the input square but follow arithmetic patterns induced by the modulus.

The following theorem gives the size of $S_m(k)$.

Theorem 5. *For an integer $R = 2^n$ and an odd integer $m < R$, we have*

$$|S_m(k)| = \sum_{d \mid \gcd(k,m)} \sum_{\substack{\frac{kR}{m} < x < m \\ \gcd(x,m) = d}} \left(d - 1 - \left\lfloor \frac{\frac{kR}{x} - y_0(x,k)}{m/d} \right\rfloor \right)$$

if $k < m^2/R$; otherwise $|S_m(k)| = 0$, where $y_0(x, k) \in \{0, \ldots, m/d - 1\}$ is the unique solution modulo m/d of $(x/d)y \equiv (k/d)R \bmod (m/d)$. In particular, $|S_m(0)| = P(m) - 2m + 1$, where P is the gcd-sum function.

Proof. Since $xy > kR$ and $0 \leq x, y < m$, there is no pair $(x, y) \in S_m(k)$ if $kR \geq m^2$. Otherwise, if $kR < m^2$, any pair $(x, y) \in S_m(k)$ satisfies $kR/m < x < m$ and $kR/x < y < m$. Let $d = \gcd(x, m)$ for a fixed $x \in (kR/m, m)$. Then the congruence equation $xy \equiv kR \bmod m$ as a congruence

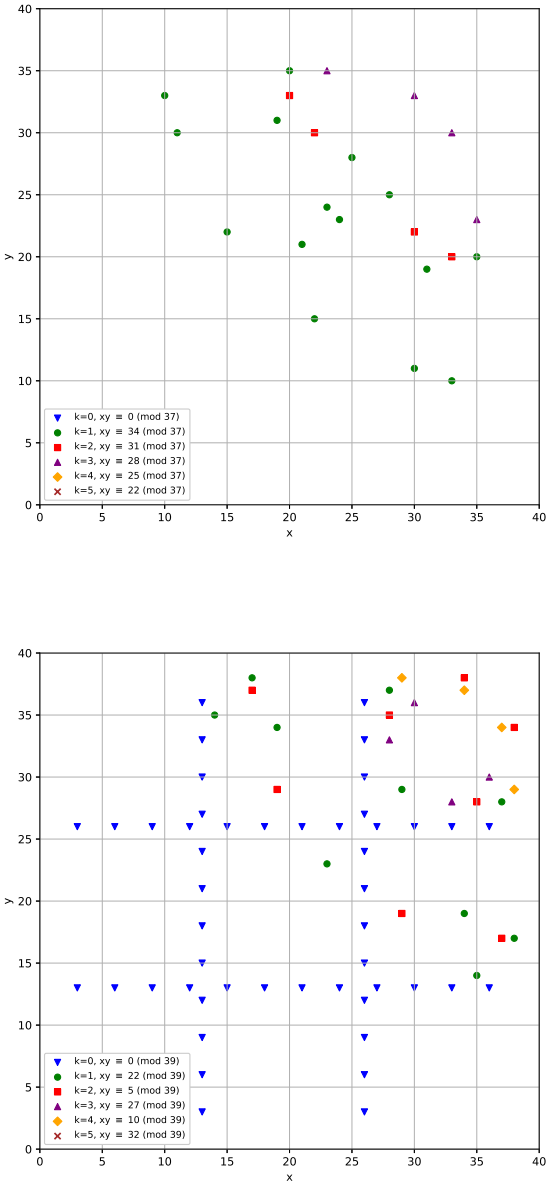

Fig. 1. Examples of the correction set $S(m)$ with $R = 256$ for two small moduli: (above) $m = 37$ and (below) $m = 39$. Each point represents an input pair $(x, y) \in [0, m)^2$ that triggers the correction step in Montgomery multiplication. For each $k < m^2/R$, i.e., $k = 0, \ldots, 5$ for both cases $m = 37$ and $m = 39$, the nonempty plotted subsets satisfy the conditions of Theorem 3, namely $xy \equiv kR \pmod m$ and $xy > kR$. Different markers correspond to different values of k. The figures illustrate that correction events form structured arithmetic patterns depending on the modulus, rather than being uniformly distributed over the input domain.

in y has exactly d solutions modulo m if and only if $d|kR$, equivalently $d|k$. Let $y_0 \in \{0, 1, \ldots, m/d - 1\}$ be a solution. Then the general solutions are represented as $y = y_0 + j(m/d)$ with integers $j \in \{0, 1, \ldots, d-1\}$. Since $kR/x < y < m$, we obtain

$$\frac{kR/x - y_0}{m/d} < j < d - \frac{y_0 d}{m}.$$

Hence, $j = \lfloor \alpha_x \rfloor + 1, \lfloor \alpha_x \rfloor + 2, \ldots, d-1$, where

$$\alpha_x \coloneqq \frac{kR/x - y_0}{m/d}.$$

Therefore, we obtain

$$\begin{aligned}
|S_m(k)| &= \sum_{\substack{\frac{kR}{m} < x < m \\ \gcd(x,m)|k}} (\gcd(x,m) - 1 - \lfloor \alpha_x \rfloor) \\
&= \sum_{\substack{d|k \\ d|m}} \sum_{\substack{\frac{kR}{m} < x < m \\ \gcd(x,m) = d}} (d - 1 - \lfloor \alpha_x \rfloor) \\
&= \sum_{d | \gcd(k,m)} \sum_{\substack{\frac{kR}{m} < x < m \\ \gcd(x,m) = d}} (d - 1 - \lfloor \alpha_x \rfloor).
\end{aligned}$$

In particular, for $k = 0$, the number of y is exactly $\gcd(x, m) - 1$ for a fixed $x \in (0, m)$. Therefore, we have

$$\begin{aligned}
|S_m(0)| &= \sum_{x=1}^{m-1} (\gcd(x,m) - 1) \\
&= \sum_{x=1}^{m} \gcd(x,m) - (m-1) - \gcd(m,m) \\
&= P(m) - 2m + 1.
\end{aligned}$$

□

B. Correction Probability in the Small-Modulus Regime

Let $f(m)$ be the probability that the correction step is performed for a uniformly randomly sampled pair $(x, y) \xleftarrow{\$} \{0, 1, \ldots, m-1\}^2$, namely $f(m) = |S(m)|/m^2$.

In this section, we focus on relatively small moduli m such that $m^2 < R$. In this case, $|S_m(k)|$ is non-zero only when $k = 0$. Then we have

$$f(m) = \frac{|S_m(0)|}{m^2} = \frac{P(m) - 2m + 1}{m^2} =: f_0(m).$$

Note that $f_0(p) = 0$ for any prime number p. Intuitively, this is because $xy \equiv 0 \bmod p$ is satisfied only when $x = 0$ or $y = 0$, and in this case, the correction step does not occur.

Since the correction probability in the small-modulus regime is given explicitly by $f_0(m) = (P(m) - 2m + 1)/m^2$, it is natural to ask how large this value can be over all admissible moduli. Fig. 2 plots $f_0(m)$ for all odd moduli $3 \le m < 256$. The maximum is attained at $m = 15$, and no larger value is observed for any tested modulus. This motivates the following empirical conjecture.

Conjecture 6. *For any odd integer $m \ge 3$, where $f_0(m) = (P(m) - 2m + 1)/m^2$, $f_0(m) \le f_0(15)$.*

Theorem 7 and the subsequent analysis for products of two distinct primes provide theoretical support for this conjecture, since neither prime powers nor products of two distinct primes exceed the value at $m = 15$.

Theorem 7. *Let m be a prime power p^a for $p \ge 3$ and $a \ge 1$. Then we have*

$$f_0(m) = f_0(p^a) \le f_0(p^2) \le f_0(3^2) = \frac{4}{81} \approx 0.0494$$

Proof. For a fixed prime $p \ge 3$, we define

$$\begin{aligned}
g(a) &\coloneqq f_0(p^a) = \frac{a-1}{p^a} - \frac{a}{p^{a+1}} + \frac{1}{p^{2a}}, \\
\Delta_a &\coloneqq g(a+1) - g(a).
\end{aligned}$$

We have $\Delta_1 > 0$ and $\Delta_a < 0$ for $a \ge 2$ (see Lemma 8). Therefore $g(a) \le g(2)$, namely $f_0(p^a) \le f_0(p^2)$. Since $f_0(p^2) = (p-1)^2/p^4$ is monotonically decreasing as a function of p, we have $f_0(p^2) \le f_0(3^2)$. □

Lemma 8. *Let Δ_a be as defined in the proof of Theorem 7. Then it holds that $\Delta_1 > 0$ and $\Delta_a < 0$ for $a \ge 2$.*

Proof. For $a = 1$, we have

$$\Delta_1 = g(2) - g(1) = g(2) = \frac{(p-1)^2}{p^4} > 0.$$

For $a \ge 2$, we have

$$\begin{aligned}
\Delta_a &= -\frac{a-1}{p^a} + \frac{2a}{p^{a+1}} - \frac{a+1}{p^{a+2}} - \frac{1}{p^{2a}} + \frac{1}{p^{2(a+1)}} \\
&= -\frac{(a-1)p^2 - 2ap + (a+1)}{p^{a+2}} - \frac{p^2 - 1}{p^{2a+2}} \\
&= -\frac{(p-1)\{(a-1)(p-1) - 2\}}{p^{a+2}} - \frac{p^2 - 1}{p^{2a+2}} < 0
\end{aligned}$$

because $(a-1)(p-1) - 2 \ge 0$ for $a \ge 2$ and $p \ge 3$. □

Next, we investigate the case where m is a product of two distinct primes. Let $m = pq$, where p and q are two distinct primes with $3 \le p < q$. Then we have

$$\begin{aligned}
f_0(m) &= \frac{P(m) - 2m + 1}{m^2} \\
&= \frac{(2p-1)(2q-1) - 2pq + 1}{p^2 q^2} \\
&= \frac{2(p-1)(q-1)}{p^2 q^2}.
\end{aligned}$$

$f_0(pq)$ decreases in each of p and q. Therefore,

$$f_0(pq) \le f_0(3 \cdot 5) = \frac{16}{225} \approx 0.0711.$$

V. Analysis in Full-Size Setting

As mentioned in Section III, inputs $(x, y) \in \mathbb{Z}_{\ge 0}^2$ to Montgomery multiplication (Algorithm 1) are often assumed to satisfy either $xy \le mR$ or $x, y \in [0, m)$. We analyzed the probability of the correction step in the latter case in Section IV. In this section, we examine the probability in the former case.

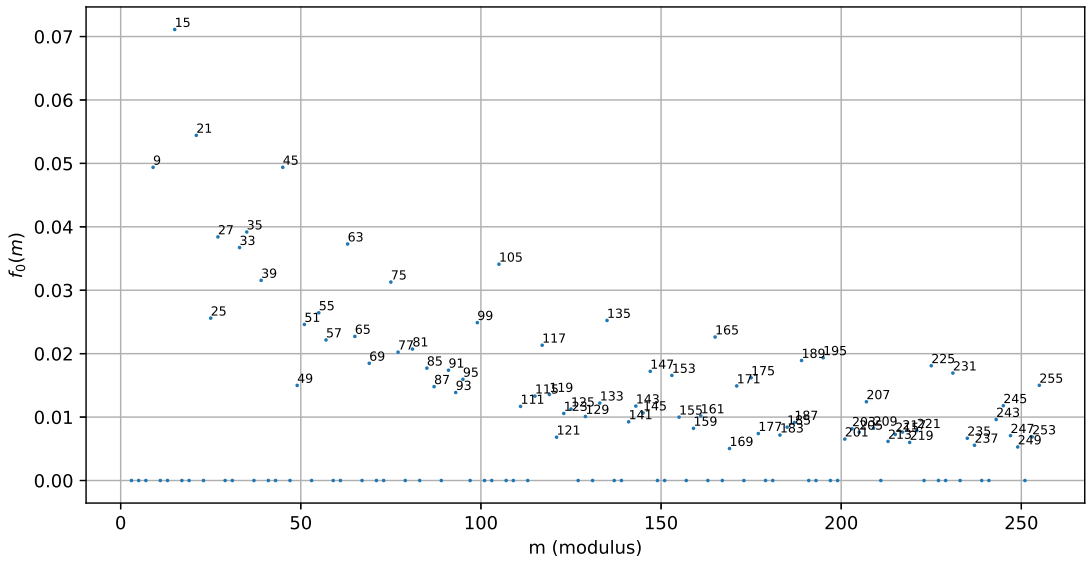

Fig. 2. Values of $f_0(m) = \frac{P(m)-2m+1}{m^2}$ for every odd modulus $m \in \{3, 5, \ldots, 255\}$. This quantity equals the correction probability in the small-modulus regime $m^2 < R$. The number next to each plotted point is the value of m. Unlabeled zero-valued points correspond to prime moduli.

A. Total Number of Input Pairs

Let $F_m = \{(x, y) \in \mathbb{Z}^2_{>0} \mid xy \leq mR\}$. Note that we exclude the case where $xy = 0$, since otherwise infinitely many pairs with $xy = 0$ would be included. This full-size model is an arithmetic counting model over a finite hyperbolic region; it is distinct from sampling x and y independently from a fixed interval. The size of F_m can be represented by the sum of the divisor function τ as follows:

$$|F_m| = \sum_{\substack{x,y\geq 1 \\ xy\leq mR}} 1 = \sum_{n\leq mR} \tau(n).$$

Dirichlet's asymptotic formula for the partial sum of the divisor function is well known (see [14]).

Lemma 9. *For every real number $X \geq 1$ we have*

$$\sum_{n\leq X} \tau(n) = X \log X + (2\gamma - 1)X + O(\sqrt{X}),$$

where γ is Euler's constant.

From Lemma 9, we obtain that

$$|F_m| \approx mR \log mR + (2\gamma - 1)mR. \tag{3}$$

B. Evaluation of the Correction Count

Next, we define the set of pairs that satisfy the correction condition (2). Let $T(m)$ denote the set of such pairs. We examine the ratio of the sizes of $T(m)$ and F_m, i.e., the probability that the correction step is performed for an input $(x, y) \xleftarrow{\$} F_m$.

Definition 10. For a fixed k, let $T_m(k)$ be the set of pairs of positive integers (x, y) with $xy \leq mR$ that satisfy the condition (2). Let $T(m)$ be the union of $T_m(k)$, namely,

$$T_m(k) = \left\{ (x, y) \in \mathbb{Z}^2_{>0} \;\middle|\; \begin{array}{l} xy \equiv kR \bmod m, \\ kR < xy \leq mR \end{array} \right\},$$

$$T(m) = \bigcup_{k=0}^{m-1} T_m(k).$$

For distinct $0 \leq k, k' < m$, $T_m(k)$ and $T_m(k')$ are disjoint. Therefore $|T(m)| = \sum_{k=0}^{m-1} |T_m(k)|$. Since the number of pairs $(x, y) \in \mathbb{Z}^2_{>0}$ with $xy = n$ is $\tau(n)$, we have the exact representation

$$|T_m(k)| = \sum_{\substack{kR<n\leq mR \\ n\equiv kR \bmod m}} \tau(n).$$

For each k, put

$$g = \gcd(k, m), \quad k = gh, \quad m = gq.$$

Then $0 \leq h < q$. If $k > 0$, then $\gcd(h, q) = 1$. Since R is a power of two and m is odd, we have $\gcd(R, q) = 1$. Thus, for $k > 0$, $\gcd(hR, q) = 1$. For the case $k = 0$, we have $q = 1$ and $h = 0$. In applying the main-term expression below, we interpret the term $h \log(hR)$ as zero, corresponding to the exact identity $D_m(0; 1, 0) = 0$.

The congruence condition in $T_m(k)$ is $n \equiv kR \bmod m$. Since $g = \gcd(k, m)$, we have $g|n$ and write $n = gN$. Therefore,

$$|T_m(k)| = \sum_{\substack{hR<N\leq qR \\ N\equiv hR \bmod q}} \tau(gN).$$

Define

$$D_g(X;q,a)=\sum_{\substack{N\le X\\ N\equiv a \bmod q}}\tau(gN).$$

Then

$$|T_m(k)|=D_g(qR;q,hR)-D_g(hR;q,hR).$$

We now replace $D_g(X;q,a)$ by its average over primitive residue classes modulo q. This replacement is motivated by the standard equidistribution principle for the divisor function in primitive arithmetic progressions. Fouvry, Iwaniec, and Katz [15] note that, for primitive residue classes, elementary estimates already approximate individual residue-class sums by their average over primitive classes in a suitable range. We use this principle here as a main-term approximation for the weighted divisor sum $D_g(X;q,a)$.

For $\gcd(a,q)=1$, define

$$D_g(X;q)=\frac{1}{\varphi(q)}\sum_{\substack{N\le X\\ \gcd(N,q)=1}}\tau(gN). \tag{4}$$

The difference between $D_g(X;q,a)$ and $D_g(X;q)$ is absorbed into the error term below. Since the number of divisors of the fixed integer g is finite, the following elementary decomposition reduces the mean value to ordinary divisor sums:

$$\tau(gN)=\sum_{d\mid \gcd(g,N)}\mu(d)\tau(g/d)\tau(N/d). \tag{5}$$

Lemma 11. *Let $D_g(X;q)$ be given by* (4)*. For $g,q\ge 1$,*

$$D_g(X;q)=\frac{A_{g,q}}{\varphi(q)}X\log X+O_{g,q}\left(X\right),$$

where

$$A_{g,q}=\left(\frac{\varphi(q)}{q}\right)^2\sum_{\substack{d\mid \mathrm{rad}(g)\\ \gcd(d,q)=1}}\frac{\mu(d)\tau(g/d)}{d}.$$

The error term is not intended to be uniform in g and q.

Proof. Using (5), we obtain

$$\sum_{\substack{N\le X\\ \gcd(N,q)=1}}\tau(gN)=\sum_{\substack{d\mid \mathrm{rad}(g)\\ \gcd(d,q)=1}}\mu(d)\tau(g/d)\sum_{\substack{M\le X/d\\ \gcd(M,q)=1}}\tau(M).$$

The elementary estimate gives

$$\begin{aligned}\sum_{\substack{M\le X/d\\ \gcd(M,q)=1}}\tau(M)&=\sum_{\substack{M\le X/d\\ \gcd(M,q)=1}}\sum_{ab=M}1=\sum_{\substack{ab\le X/d\\ \gcd(a,q)=1\\ \gcd(b,q)=1}}1\\&=\sum_{\substack{a\le X/d\\ \gcd(a,q)=1}}\#\{b\le X/da\mid \gcd(b,q)=1\}\\&=\sum_{\substack{a\le X/d\\ \gcd(a,q)=1}}\left(\frac{\varphi(q)}{q}\frac{X}{da}+O_q(1)\right)\\&=\frac{\varphi(q)}{q}\frac{X}{d}\sum_{\substack{a\le X/d\\ \gcd(a,q)=1}}\frac{1}{a}+O_q(X/d)\\&=\frac{\varphi(q)}{q}\frac{X}{d}\left(\frac{\varphi(q)}{q}\log\frac{X}{d}+O_q(1)\right)+O_q(X/d)\\&=\left(\frac{\varphi(q)}{q}\right)^2\frac{X}{d}\log X+O_q(X/d).\end{aligned}$$

Then we obtain

$$\sum_{\substack{N\le X\\ \gcd(N,q)=1}}\tau(gN)=A_{g,q}X\log X+O_{g,q}(X).$$

Dividing by $\varphi(q)$ proves the lemma. □

Replacing $D_g(X;q,a)$ by the primitive-class average and applying Lemma 11, we obtain

$$|T_m(k)|=\frac{A_{g,q}}{\varphi(q)}R(q\log qR-h\log hR)+\mathcal{E}_k. \tag{6}$$

Here $\mathcal{E}_k$ denotes the total error coming from the $O_{g,q}(X)$ terms in Lemma 11 and from replacing individual residue classes by their primitive average. We do not require a sharp bound for $\mathcal{E}_k$ in the following main-term approximation.

We now sum over $0\le k<m$. The parametrization $g=\gcd(k,m)$, $k=gh$, $m=gq$ is equivalent to summing over $q|m$ and $0\le h<q$, $\gcd(h,q)=1$, with $g=m/q$. For $h=0$, we use the convention $h\log hR=0$, consistent with $D_m(0;1,0)=0$. Using (6), we obtain

$$|T(m)|=R\sum_{q\mid m}\frac{A_{m/q,q}}{\varphi(q)}\sum_{\substack{0\le h<q\\ \gcd(h,q)=1}}(q\log qR-h\log hR)+\mathcal{E}(m,R),$$

where $\mathcal{E}(m,R)=\sum_k\mathcal{E}_k$. The main term can be decomposed into a $\log R$-part and an m-dependent logarithmic part. Since

$$q\log qR-h\log hR=(q-h)\log R+q\log q-h\log h,$$

we have

$$\begin{aligned}|T(m)|=R\log R\sum_{q\mid m}\frac{A_{m/q,q}}{\varphi(q)}&\sum_{\substack{0\le h<q\\ \gcd(h,q)=1}}(q-h)\\&+R\sum_{q\mid m}\frac{A_{m/q,q}}{\varphi(q)}\sum_{\substack{0\le h<q\\ \gcd(h,q)=1}}(q\log q-h\log h)\\&+\mathcal{E}(m,R),\end{aligned} \tag{7}$$

where $0 \log 0$ is interpreted as 0. For $q = 1$, the inner sum of the $\log R$-part equals 1. For $q > 1$,

$$\sum_{\substack{1 \le h < q \\ \gcd(h,q)=1}} (q - h) = q\varphi(q) - \frac{q\varphi(q)}{2} = \frac{q\varphi(q)}{2}.$$

Therefore, the $\log R$-part equals

$$\frac{1}{2}\left(A_{m,1} + \sum_{q|m} qA_{m/q,q}\right) R \log R.$$

The following lemma gives the sum of $qA_{m/q,q}$.

Lemma 12. *For every positive integer m,*

$$\sum_{q|m} qA_{m/q,q} = m.$$

Proof. Both sides are multiplicative in m. It is therefore enough to prove the claim for $m = p^\alpha$. Let $q = p^\beta$, where $0 \le \beta \le \alpha$. Then $g = m/q = p^{\alpha-\beta}$. If $\beta = 0$, then $A_{p^\alpha,1} = \alpha + 1 - \alpha/p$. If $1 \le \beta \le \alpha$, then

$$A_{p^{\alpha-\beta},p^\beta} = (\alpha - \beta + 1)\left(1 - \frac{1}{p}\right)^2.$$

Thus

$$\sum_{q|p^\alpha} qA_{p^\alpha/q,q} = \left(\alpha + 1 - \frac{\alpha}{p}\right) + \sum_{\beta=1}^{\alpha} p^\beta(\alpha-\beta+1)\left(1 - \frac{1}{p}\right)^2.$$

A direct calculation gives $\sum_{q|p^\alpha} qA_{p^\alpha/q,q} = p^\alpha$. Multiplying over all prime-power factors of m proves the claim. □

By Lemma 12, the $\log R$-part becomes

$$\frac{1}{2}(m + A_{m,1})R \log R,$$

where $A_{m,1} = \prod_{p^\alpha \| m}\left(\alpha + 1 - \frac{\alpha}{p}\right)$.

It remains to describe the second term in (7). Set

$$\mathcal{L}(m) = \sum_{q|m} \frac{A_{m/q,q}}{\varphi(q)} \sum_{\substack{0 \le h < q \\ \gcd(h,q)=1}} (q \log q - h \log h).$$

In the term $q = 1$, the inner sum is interpreted using the convention $0 \log 0 = 0$, and hence this term contributes zero to $\mathcal{L}(m)$. Then

$$|T(m)| = \frac{1}{2}(m + A_{m,1})R \log R + R\mathcal{L}(m) + \mathcal{E}(m, R). \quad (8)$$

For a simple approximation, we use the standard heuristic that reduced residues modulo q are equidistributed in $[1, q]$. Then, for $q > 1$,

$$\sum_{\substack{1 \le h < q \\ \gcd(h,q)=1}} h \log h \approx \frac{\varphi(q)}{q} \int_1^q t \log t dt.$$

Since

$$\int_1^q t \log t dt = \frac{q^2}{2} \log q - \frac{q^2}{4} + O(1),$$

we obtain

$$\sum_{\substack{0 \le h < q \\ \gcd(h,q)=1}} (q \log q - h \log h) = \frac{q\varphi(q)}{2} \log q + O(q\varphi(q)).$$

Consequently,

$$\mathcal{L}(m) \approx \frac{1}{2} \sum_{q|m} qA_{m/q,q} \log q.$$

Lemma 12 shows that the normalized quantities $qA_{m/q,q}/m$ may be viewed as a weighted average of $\log q$. In typical cases, especially when the weights are not concentrated on small divisors, this weighted average is expected to be of order $\log m$. Therefore, we use

$$\sum_{q|m} qA_{m/q,q} \log q \approx m \log m$$

as a heuristic main-term estimate. A sharper divisor-structure-dependent analysis of this logarithmic average would refine the error term in the approximation. Substituting this into (8), we obtain the main approximation

$$|T(m)| \approx \frac{1}{2}(m + A_{m,1})R \log R + \frac{1}{2}mR \log m. \quad (9)$$

C. Approximation of the Correction Probability in the Full-Size Setting

In this section, we obtain an approximation of the probability that the correction step occurs from the results obtained in Sections V-A and V-B. Furthermore, we evaluate the probability when m is large and close to R. We also show the numerical results for specific parameter values.

The probability that the correction step is performed for an input $(x, y) \xleftarrow{\$} F_m$ is $|T(m)|/|F_m|$. Combining (3) with the main approximation (9), we obtain

$$\begin{aligned}\frac{|T(m)|}{|F_m|} &\approx \frac{\frac{1}{2}(m + A_{m,1})R \log R + \frac{1}{2}mR \log m}{mR \log mR + (2\gamma - 1)mR} \\ &= \frac{1}{2} \frac{\log mR + \frac{A_{m,1}}{m} \log R}{\log mR + (2\gamma - 1)} \\ &\approx \frac{1}{2} \quad \text{when } m \text{ is close to } R \text{ and both are large.}\end{aligned}$$

$A_{m,1}/m \to 0$ as $m \to \infty$, since

$$A_{m,1} = \prod_{p^\alpha \| m}\left(\alpha + 1 - \frac{\alpha}{p}\right) \le \prod_{p^\alpha \| m} (\alpha+1) = \tau(m) = o(m^\varepsilon)$$

for every $\varepsilon > 0$. Hence, $A_{m,1}/m = o(m^{\varepsilon-1})$, in particular $A_{m,1}/m \to 0$.

The values in Fig. 3 were computed by evaluating the exact divisor-sum representation for $|T(m)|$ and the exact divisor-sum formula for $|F_m|$. Fig. 3 shows the numerical results for the probability $|T(m)|/|F_m|$ with $R = 4096$. The large-m region is consistent with the approximation $|T(m)|/|F_m| \approx 1/2$.

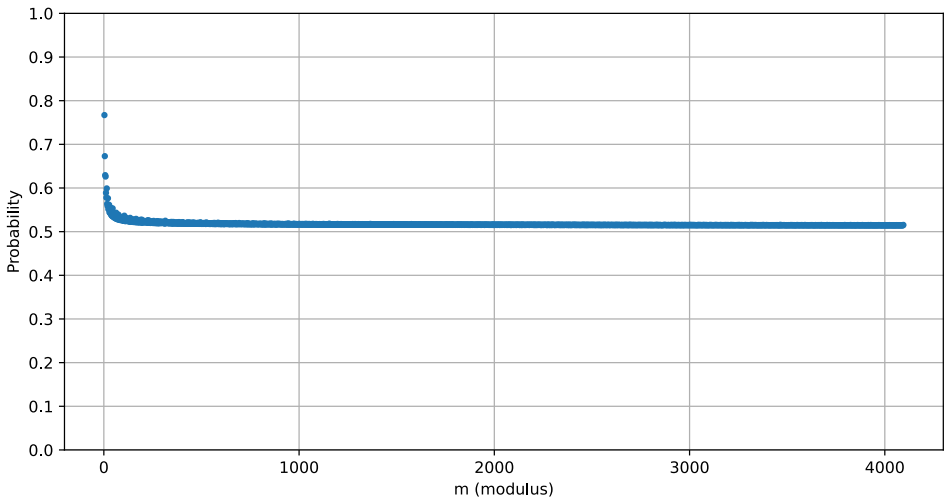

Fig. 3. Plot of the probability $|T(m)|/|F_m|$ for odd moduli $3 \leq m < R$, with $R = 4096$.

VI. Conclusion

This work analyzed input-dependent correction events in Montgomery multiplication in both the modulus-size and full-size input regimes. For any odd modulus and any non-negative input, we reformulated the correction condition into a simple number-theoretic representation. In the modulus-size setting $(x, y) \in [0, m)^2$, we obtained an exact correction probability when $m^2 < R$:

$$f_0(m) = \frac{P(m) - 2m + 1}{m^2}.$$

Thus, the correction probability is governed by the multiplicative structure of the modulus. We also established precise bounds for prime powers and for products of two distinct primes. In the full-size setting, where $xy \in (0, mR]$, we derived an exact divisor-sum representation of the correction count. Starting from this representation, we obtained a heuristic main-term approximation by elementary averaging and by applying Dirichlet's classical estimate for $\sum_{n \leq X} \tau(n)$. The resulting approximation explains the numerical observation that the correction probability is close to $1/2$. These results provide arithmetic information on Montgomery correction events that may be relevant to future leakage analysis. Future work includes deriving a full asymptotic expansion in the modulus-size model for general m, and obtaining uniform error terms for the full-size main-term approximation.

References

[1] J. W. Bos, L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, J. M. Schanck, P. Schwabe, G. Seiler, and D. Stehlé, "CRYSTALS - Kyber: A CCA-secure module-lattice-based KEM," in *Proc. 2018 IEEE European Symposium on Security and Privacy*, London, United Kingdom, Apr. 2018, pp. 353–367.

[2] L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, P. Schwabe, G. Seiler, and D. Stehlé, "CRYSTALS-Dilithium: A lattice-based digital signature scheme," *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, vol. 2018, no. 1, pp. 238–268, 2018.

[3] P. L. Montgomery, "Modular multiplication without trial division," *Math. Comput.*, vol. 44, no. 170, pp. 519–521, 1985.

[4] G. Seiler, "Faster AVX2 optimized NTT multiplication for Ring-LWE lattice cryptography," Cryptology ePrint Archive, Report 2018/039, 2018.

[5] H. Becker, V. Hwang, M. J. Kannwischer, B.-Y. Yang, and S.-Y. Yang, "Neon NTT: Faster Dilithium, Kyber, and Saber on Cortex-A72 and Apple M1," *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, vol. 2022, no. 1, pp. 221–244, 2022.

[6] P. Barrett, "Implementing the Rivest Shamir and Adleman public key encryption algorithm on a standard digital signal processor," in *Proc. Advances in Cryptology — CRYPTO '86*, ser. LNCS, vol. 263, Santa Barbara, CA, USA, 1987, pp. 311–323.

[7] T. Plantard, "Efficient word size modular arithmetic," *IEEE Trans. Emerg. Top. Comput.*, vol. 9, no. 3, pp. 1506–1518, 2021.

[8] D. Aoki, K. Minematsu, T. Okamura, and T. Takagi, "Efficient word size modular multiplication over signed integers," in *Proc. 29th IEEE Symposium on Computer Arithmetic*, Lyon, France, Sep. 2022, pp. 94–101.

[9] J. Huang, J. Zhang, H. Zhao, Z. Liu, R. C. C. Cheung, Ç. K. Koç, and D. Chen, "Improved Plantard arithmetic for lattice-based cryptography," *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, vol. 2022, no. 4, pp. 614–636, 2022.

[10] W. Schindler, "A timing attack against RSA with the Chinese Remainder Theorem," in *Proc. Cryptographic Hardware and Embedded Systems — CHES 2000*, ser. LNCS, vol. 1965, Worcester, MA, USA, Aug. 2000, pp. 109–124.

[11] J.-F. Dhem, F. Koeune, P.-A. Leroux, P. Mestré, J.-J. Quisquater, and J.-L. Willems, "A practical implementation of the timing attack," in *Proc. Smart Card Research and Applications, CARDIS '98*, ser. LNCS, vol. 1820, Louvain-la-Neuve, Belgium, Sep. 1998, pp. 167–182.

[12] C. D. Walter and S. Thompson, "Distinguishing exponent digits by observing modular subtractions," in *Proc. Topics in Cryptology — CT-RSA 2001*, ser. LNCS, vol. 2020, San Francisco, CA, USA, Apr. 2001, pp. 192–207.

[13] H. Sato, D. Schepers, and T. Takagi, "Exact analysis of Montgomery multiplication," in *Proc. Progress in Cryptology — INDOCRYPT 2004*, ser. LNCS, vol. 3348, Chennai, India, Dec. 2004, pp. 290–304.

[14] T. M. Apostol, *Introduction to Analytic Number Theory*. New York, NY: Springer, 1976.

[15] E. Fouvry, H. Iwaniec, and N. Katz, "The divisor function over arithmetic progressions," *Acta Arith.*, vol. 61, no. 3, pp. 271–287, 1992.