Coded SumComp: Linear Coding over Integer Rings for Improved SumComp Performance

Vaishnavi Anekar and Anjana A. Mahesh
Department of Electrical Engineering
IIT Hyderabad, Hyderabad, India
Email: ee24mtech14009@iith.ac.in, anjana.am@ee.iith.ac.in,

***Abstract*—Over-the-air computation (OAC) enables efficient aggregation of distributed data by exploiting the superposition property of wireless multiple-access channels. Digital OAC frameworks such as SumComp leverage integer-ring structures for reliable computation of nomographic functions, but their performance is limited by the minimum distance of the effective constellation. In this paper, we propose a *Coded SumComp* scheme that incorporates linear coding over integer rings to improve reliability. The proposed design preserves the algebraic structure required for computation while enhancing distance properties through coding, with the resulting modulo-folded channel naturally inducing Lee-type errors. We derive an upper bound on the mean squared error (MSE) in terms of the decoder failure probability under Lee-metric decoding. Simulation results show that the proposed scheme achieves about 2–3 dB improvement in MSE over uncoded SumComp in the moderate-to-high SNR regime.**

I. Introduction

Over-the-air computation (OAC) has emerged as a promising paradigm for enabling distributed computation over wireless networks. By leveraging the superposition property of the wireless multiple access channel (MAC), multiple devices can simultaneously transmit signals such that a desired function of their data is computed directly at the receiver [1]. Traditional OAC schemes [2], [3] rely on analog transmission, which is sensitive to channel noise and hardware imperfections. Recently, digital over-the-air computation methods have been proposed to address these limitations. The ChannelComp [4] framework enables the computation of functions using digital modulations but requires solving a high-complexity optimization problem for modulation design.

To reduce the complexity associated with modulation design, the SumComp [5] coding scheme was proposed, which leverages the algebraic structure of integer rings to enable efficient digital over-the-air computation for nomographic functions. By exploiting the ring structure of Gaussian integers, SumComp provides a low-complexity framework that supports standard digital modulation schemes while enabling computation over the wireless channel. However, its performance is limited by the minimum distance of the effective constellation formed at the receiver, which degrades as the number of transmitting nodes increases.

In this paper, we propose a Coded SumComp framework that incorporates linear coding over integer rings to improve reliability of function computation. Specifically, we modify the encoder and decoder structures of the SumComp scheme to integrate linear coding, which increases the effective minimum distance of the aggregated constellation at the receiver and improves the symbol error rate and mean squared error performance. We analyze the performance of the proposed scheme in terms of the mean squared error (MSE) and derive an upper bound in terms of the decoder failure probability.

The main contributions of this paper are as follows:

- We propose a coded extension of SumComp using linear codes over $\mathbb{Z}_q$, improving reliability while preserving the computation structure.
- We design encoder and decoder architectures that integrate ring-based coding with SumComp and enable recovery via syndrome decoding.
- We characterize the modulo-q folded channel and show that the resulting errors are naturally captured by the Lee metric over $\mathbb{Z}_q$.
- We derive an upper bound on mean-squared error (MSE) in terms of decoder failure probability and demonstrate 2–3 dB gains over uncoded SumComp through simulations.

The remainder of this paper is organized as follows. Section II presents the SumComp system model. Section III describes the proposed Coded SumComp scheme. Section IV derives analytical expressions for the mean squared error of the sum function. Section V presents simulation results validating the theoretical analysis. Finally, Section VI concludes the paper.

Notation: Scalars are denoted by lowercase letters and vectors by bold lowercase letters. The sets of integers, real numbers, and complex numbers are denoted by $\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{C}$, respectively. $\mathbb{Z}_q$ denotes the ring of integers modulo q and $\mathbb{Z}_q^n$ denotes the n dimensional vector space over the ring $\mathbb{Z}_q$. The Gaussian integer ring is denoted by $\mathbb{Z}[i]$. The probability and expectation operators are denoted by $\mathbb{P}(.)$ and $\mathbb{E}[\cdot]$ respectively, and $\|\cdot\|$ denotes the Euclidean norm. The Gaussian Q-function is defined as $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt$.

II. SumComp System Model

A large class of functions that can be computed over the MAC belong to the class of *nomographic functions*. In this section, we define nomographic functions and present the SumComp system model in [5] used for computing such functions over a multiple-access channel.

Definition 1 (Nomographic Function [6], [7])**.** *Let $\mathcal{S}$ be a compact set and $K \geq 2$. A function $f : \mathcal{S}^K \to \mathbb{R}$ is called* *a* nomographic function *if there exist pre-processing functions* $\phi_k : \mathcal{S} \to \mathbb{R}$ *for* $k = 1, \ldots, K$, *and a post-processing function* $\psi : \mathbb{R} \to \mathbb{R}$ *such that*

$$f(s_1, s_2, \ldots, s_K) = \psi\left(\sum_{k=1}^{K} \phi_k(s_k)\right). \quad (1)$$

Thus, computing a nomographic function reduces to computing the summation $\sum_{k=1}^{K} \phi_k(s_k)$ over the channel followed by a postprocessing function $\psi(\cdot)$ at the receiver.

A. SumComp System Model

Consider a network consisting of K distributed nodes that communicate with a computation point (CP) over a shared MAC. Each node k observes a local measurement $s_k \in \mathcal{S}$, where $\mathcal{S}$ is a finite alphabet and the goal of the system is to compute a nomographic function $f(s_1, s_2, \ldots, s_K)$ at the CP.

1) Encoder Structure: The transmitter architecture at each node consists of three stages: source pre-processing, channel encoding, and modulation mapping, as illustrated in Fig. 1.

- Source Encoding: Each node first applies a pre-processing function $\phi_k(.)$ on the observed source symbol s_k to give $c_k = \phi_k(s_k)$, where c_k is an element in the ring $\mathbb{Z}_q$.
- Channel Encoding: The pre-processed symbol c_k is then encoded using a linear channel encoder $\mathcal{E}_k(.)$, which maps c_k to the Gaussian integer ring $\mathbb{Z}[i]$ as
$$g_k := \mathcal{E}_k(c_k) = m\mu_1 + kq_2 + (m\mu_2 - kq_1)i, \quad (2)$$
where, q_1, q_2, μ_1, μ_2 satisfy the Bézout's identity [8] for $q_1, q_2,\ k \in \mathbb{Z}$. The values of q_1, q_2, μ_1, μ_2 and k are design parameters which affect the spacing in between the constellation points and hence the error probability.
- Modulation Mapping: The encoded symbol $g_k = a + ib$ is then mapped to a constellation point via $\mathcal{G}_\rho(\cdot)$ as $x_k = \mathcal{G}_\rho(g_k) = a + b\rho i, \quad a, b \in \mathbb{Z}$, where $\mathcal{G}_\rho(\cdot)$ maps g_k into the ring of integers $\mathbb{Z}[\rho]$.

Finally, each node selects a subset $\Lambda_q \subset \mathbb{Z}[\rho]$ as the constellation set and transmits the symbol x_k over the MAC.

2) Multiple-Access Channel: All nodes transmit simultaneously over a shared MAC. The received signal at the CP is given by $r = \sum_{k=1}^{K} x_k + z$, where x_k denotes the transmitted symbol from node k and $z \sim \mathcal{CN}(0, \sigma^2)$ represents circularly symmetric complex Gaussian noise.

3) Decoder Structure: The receiver performs a sequence of operations to recover the desired function value from the superimposed signal, as illustrated in Fig. 1.

- Quantization: The received signal r is first quantized using $x = Q_\rho(r)$, which maps the received signal to the nearest constellation point in $\mathbb{Z}[\rho]$.
- Inverse Modulation Mapping: The quantized symbol is then mapped back to $\mathbb{Z}[i]$ using $g = \mathcal{G}_\rho^{-1}(x)$.
- Channel Decoding: A channel decoder $\mathcal{D}$ is applied to obtain an estimate of the sum of the encoded symbols.
$$c = \mathcal{D}(g) = \mathrm{Re}(g) \cdot q_1 + \mathrm{Im}(g) \cdot q_2 \quad (3)$$
Due to the linear structure of the code, the decoder recovers the sum of the pre-processed symbols.
- Source Decoding: Finally, the post-processing function is applied to compute the estimate of the desired function $\hat{f} = \psi(c)$. Thus, the receiver directly recovers the desired function value without individually decoding the messages of all the nodes.

Example 1. *We illustrate the operation of the SumComp scheme through a simple two–node* ($K = 2$) *example, where the objective is to compute the sum function* $f(s_1, s_2) = s_1 + s_2$. *Since the desired function is the sum, the pre-processing functions at the transmitters and the post-processing function at the receiver are chosen as identity maps, i.e.,* $\phi_1(s) = \phi_2(s) = s$ *and* $\psi(c) = c$. *Let* $q = 4$, *with constellation parameters* $q_1 = 1$, $q_2 = 4$, $\mu_1 = -3$, $\mu_2 = 1$, *and* $\rho = 1$. *Each node transmits one symbol from the alphabet* $\{0, 1, 2, 3\}$. *Suppose the two nodes observe* $s_1 = 3$ *and* $s_2 = 2$, *which after pre-processing gives* $c_1 = 3$ *and* $c_2 = 2$. *The desired sum at the computation point is* $c_1 + c_2 = 5$.

Each node then applies the encoder $\mathcal{E}_k(\cdot)$ *to map its symbol into the Gaussian integer ring* $\mathbb{Z}[i]$. *Using the given parameters (which satisfy Bézout's identity), node 1 maps* $c_1 = 3$ *to* $g_1 = -5 + 2i$, *while node 2 maps* $c_2 = 2$ *to* $g_2 = 2 + 0i$. *With* $\rho = 1$, *the modulation mapping is identity, so the transmitted symbols are* $x_1 = -5 + 2i$ *and* $x_2 = 2$.

These are transmitted simultaneously over the MAC, and the received signal is given by $r = x_1 + x_2 + z$, *where* $z \sim \mathcal{CN}(0, \sigma^2)$ *is complex Gaussian noise. Assume that after quantization, the receiver recovers the lattice point* $-3 + 2i$. *The receiver then applies the inverse mapping and the decoder* $\mathcal{D}(\cdot)$, *which computes a linear combination of the real and imaginary parts:* $c = \mathrm{Re}(g)\, q_1 + \mathrm{Im}(g)\, q_2 = (-3)\cdot 1 + (2)\cdot 4 = -3 + 8 = 5$. *This equals* $c_1 + c_2$, *and since* ψ *is the identity function, the final output matches the desired function value.*

The next section presents the proposed *Coded SumComp* framework, which uses linear coding over integer rings to improve the reliability of function computation over the MAC.

III. Coded SumComp: Encoder and Decoder

As in the SumComp system model, we consider a wireless network with K nodes communicating with a computation point over a MAC. Each node observes a symbol $s_k \in \mathcal{S},\ k = 1, \ldots, K$, where $\mathcal{S}$ is a finite alphabet, and the receiver aims to compute a nomographic function $f(s_1, \ldots, s_K)$. Instead of symbol-by-symbol processing, each node applies the Coded SumComp encoder $\mathscr{E}_k^c$ to a block $\mathbf{c}_k$ of m pre-processed symbols $[c_{k,1}, c_{k,2}, \cdots, c_{k,m}]$, corresponding to m source symbols $s_{k,1}, s_{k,2}, \ldots, s_{k,m} \in \mathcal{S}$, generating a transmit vector $\mathbf{x}_k^c \in \mathbb{C}^n$. All nodes transmit simultaneously, and the received signal is

$$\mathbf{r}^c = \sum_{k=1}^{K} h_k p_k \mathbf{x}_k^c + \mathbf{z}, \quad (4)$$

where h_k is the channel coefficient (known at the transmitter). As the final transmit step, each node applies precoding $p_k = h_k^*/|h_k|^2$ to invert the channel, resulting in a unit effective gain and coherent signal addition at the receiver. The noise vector

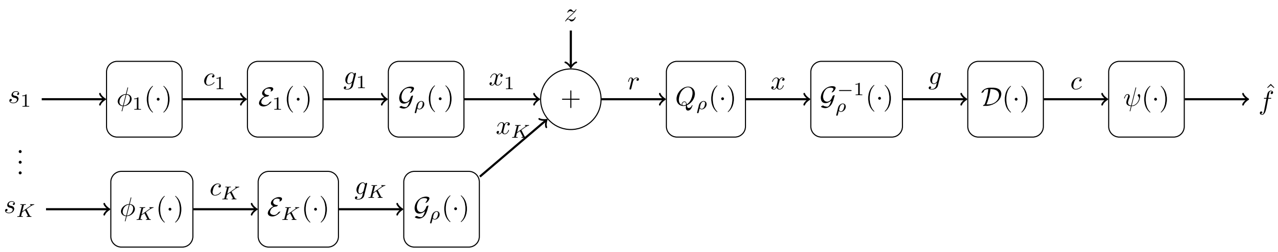

Fig. 1. Block diagram of the SumComp system for computing a nomographic function over a multiple-access channel.

$\mathbf{z} \sim \mathcal{CN}(0, \sigma^2)$ is AWGN. The receiver applies the decoder $\mathscr{D}^c$ to estimate the desired function,

$$\hat{\mathbf{f}}^c = \mathscr{D}^c(\mathbf{r}^c). \tag{5}$$

A. Coded SumComp Encoder Structure

The encoder $\mathscr{E}_k^c$ architecture of the proposed Coded SumComp scheme is shown in Fig. 2. Let node k observe a sequence of m source symbols $s_{k,1}, s_{k,2}, \ldots, s_{k,m} \in \mathcal{S}$. Each symbol is first processed using the pre-processing function $\varphi_k(\cdot)$. The resulting sequence is given by $c_{k,i} = \varphi_k(s_{k,i}) \in \mathbb{Z}_q, \quad i = 1, \ldots, m$. We denote the vector of processed symbols by $\mathbf{c}_k = (c_{k,1}, c_{k,2}, \ldots, c_{k,m}) \in \mathbb{Z}_q^m$. In the proposed scheme, the vector $\mathbf{c}_k$ is encoded using a linear code defined over the integer ring $\mathbb{Z}_q$. Let G denote the generator matrix of the linear code. The encoded codeword is given by

$$\mathbf{w}_k = \mathcal{E}_k^c(\mathbf{c}_k) = \big(\mathbf{c}_k G\big) \ \mathit{mod}\ q \qquad \in \mathbb{Z}_q^n \tag{6}$$

The codeword symbols are further processed using the SumComp encoder $\mathcal{E}_k(\cdot)$ as mentioned in (2) and subsequently mapped to constellation points via the modulation mapper $\mathcal{G}_\rho(\cdot)$ to obtain the transmit vector $\mathbf{x}_k^c$, which is finally precoded for channel inversion prior to transmission.

B. Coded SumComp Decoder Structure

At the receiver, the signals transmitted by all nodes superpose over the MAC yielding the received signal $\mathbf{r}^c = \sum_{k=1}^{K} \mathbf{x}_k^c + \mathbf{z}$. Thus, $\mathbf{r}^c \in \mathbb{C}^n$ is the aggregate of all n-length transmitted vectors corrupted by a circularly symmetric AWGN noise vector $\mathbf{z} \sim \mathcal{CN}(0, \sigma^2)$.

The decoder architecture $\mathscr{D}^c$ of the proposed Coded SumComp scheme is illustrated in Fig. 3. The received vector $\mathbf{r}^c$ is first passed through the quantizer $\mathcal{Q}_\rho^c(\cdot)$, which maps it to the nearest valid sum of K transmitted codewords to obtain the estimate $\mathbf{x}^c$. The estimated vector $\mathbf{x}^c$ is then processed by the inverse modulation mapper $\mathcal{G}_\rho^{-1}(\cdot)$ followed by the SumComp decoder $\mathcal{D}(\cdot)$ to obtain the vectors $\mathbf{g}^c$ and $\hat{\mathbf{c}}_s^c$, respectively. The vector $\hat{\mathbf{c}}_s^c$ is subsequently reduced modulo q to obtain $\hat{\mathbf{u}}_s$. This vector is then passed to the syndrome decoder $\mathscr{S}(\cdot)$ to recover the estimated codeword $\mathbf{v}_s$. The first m symbols of $\mathbf{v}_s$ are extracted to form $\mathbf{v}'_s$.

Finally, $\mathbf{v}'_s$ is processed by the mapping function $\mathcal{M}(\cdot)$, which performs *position-wise lifting* followed by maximum-likelihood (ML) decoding. Since the modulo-q reduction introduces ambiguity, each component of $\mathbf{v}'_s \in \mathbb{Z}_q^m$ can correspond to multiple integer values. In particular, as the sum of K symbols each from $\{0, \ldots, q-1\}$ lies in $\{0, \ldots, K(q-1)\}$, the true sum can differ from its modulo-q version by integer multiples of q up to $(K-1)q$. Accordingly, for each coordinate, all possible lifted values $v'_s,\ v'_s + q,\ v'_s + 2q,\ \ldots,\ v'_s + (K-1)q$ are considered. By independently lifting each coordinate, a set of K^m candidate vectors is generated.

Assuming a *systematic encoding structure*, the first m symbols of $\hat{\mathbf{c}}_s^c$ correspond directly to the (noisy) sum of the original pre-processed symbols. The ML decoding is therefore performed by computing the Euclidean distance between these first m symbols of $\hat{\mathbf{c}}_s^c$ and each of the K^m candidate lifted vectors, and selecting the one with the minimum distance. The selected vector is declared as the estimate $\hat{\mathrm{s}}^c$, after which the post-processing function $\psi(\cdot)$ is applied to obtain the desired function value $\hat{\mathbf{f}}^c$.

C. Why Linear Coding over Integer Ring?

Linear coding over the integer ring $\mathbb{Z}_q$ [9], [10] is adopted in the proposed scheme due to the following reasons:

- Compatibility with MAC superposition and closure under addition: The MAC naturally produces a sum of transmitted signals, and linear codes preserve this structure. In particular, $(\mathbf{c}_1 G \bmod q) + (\mathbf{c}_2 G \bmod q) = (\mathbf{c}_1 + \mathbf{c}_2) G \bmod q$, ensuring that the superposition of codewords remains a valid codeword.
- Consistency with ring-based modulation: Linear codes over $\mathbb{Z}_q$ naturally align with the algebraic structure used in the SumComp modulation mapping.
- Robustness to modulo-q noise via Lee metric: The combined effect of quantization and modulo-q reduction induces a wrapped additive noise over $\mathbb{Z}_q$, where integer-valued perturbations are folded modulo q. Such errors are naturally measured by the Lee metric, where the Lee weight of $a \in \mathbb{Z}_q$ is $w_L(a) = \min(a, q-a)$, capturing circular deviations. This makes Lee-metric decoding well-matched to the effective channel.

Nonlinear codes, in general, do not preserve the additive structure of the transmitted signals and therefore cannot reliably leverage the superposition property of the MAC for function computation.

D. Illustrative Example

Example 2. *We illustrate the proposed* Coded SumComp *scheme using the example in Section II. Consider a* $(5, 3)$

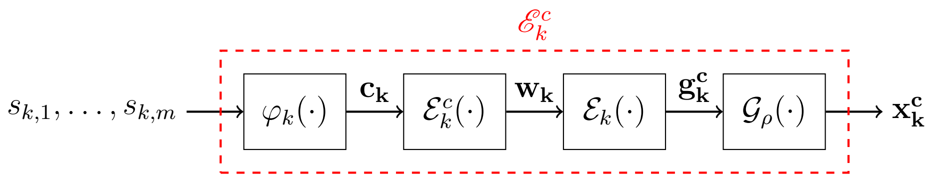

Fig. 2. Proposed Coded SumComp Encoder $\mathscr{E}_k^c$ structure at node k.

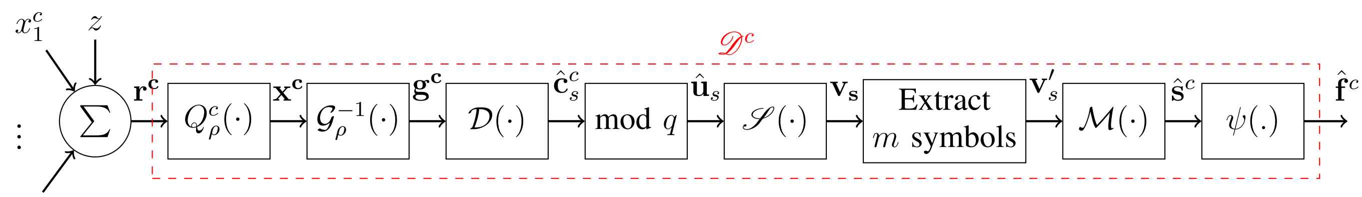

Fig. 3. Proposed coded SumComp Decoder $\mathscr{D}^c$ structure at the computation point.

linear code over $\mathbb{Z}_4$ with minimum Lee distance $d_L = 3$, yielding error-correction capability $t = \lfloor \frac{d_L-1}{2} \rfloor = 1$.

The generator and parity-check matrices are

$$G = \begin{bmatrix} 1 & 0 & 0 & 3 & 3 \\ 0 & 1 & 0 & 3 & 1 \\ 0 & 0 & 1 & 2 & 3 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 1 & 2 & 1 & 0 \\ 1 & 3 & 1 & 0 & 1 \end{bmatrix}.$$

The inputs to encoder $\mathcal{E}_k^c(.)$ at nodes 1 and 2 are $\mathbf{c}_1 = [3,3,0]$ and $\mathbf{c}_2 = [3,2,0]$ respectively, with desired sum $\mathbf{c}_1 + \mathbf{c}_2 = [6,5,0]$. Encoding over $\mathbb{Z}_4$ gives $\mathbf{w}_1 = [3,3,0,2,0]$ and $\mathbf{w}_2 = [3,2,0,3,3]$. After SumComp encoding $\mathcal{E}_k(\cdot)$ and modulation $\mathcal{G}_\rho(\cdot)$, the transmitted vectors are

$$\mathbf{x}_1^c = \begin{bmatrix} -5+2i \\ -5+2i \\ 4-1i \\ -2+1i \\ 4-1i \end{bmatrix}, \quad \mathbf{x}_2^c = \begin{bmatrix} -1+1i \\ 2+0i \\ 8-2i \\ -1+1i \\ -1+1i \end{bmatrix}.$$

The received signal $\mathbf{r}^c = \mathbf{x}_1^c + \mathbf{x}_2^c + \mathbf{z}$ is quantized, inverse-mapped, and decoded to $\hat{\mathbf{c}}_s^c = [6,6,0,5,3]$, with an error in the second symbol. Modulo-4 reduction gives $\hat{\mathbf{u}}_s = [2,2,0,1,3]$. The syndrome decoder $\mathscr{S}(\cdot)$ corrects a $+1$ error in the second symbol, yielding $\mathbf{v}_s = [2,1,0,1,3]$. Extracting the first $m = 3$ symbols gives $\mathbf{v}_s' = [2,1,0]$.

Since the computation is performed over $\mathbb{Z}_4$, the possible sum values can go up to 6. Therefore integer lifting is applied in $\mathcal{M}(\cdot)$ by adding patterns formed using elements $\{0,4\}$ to each symbol position. All $K^m = 2^3 = 8$ possible lifting patterns are evaluated using maximum-likelihood (ML) decoding with the first three symbols of $\hat{\mathbf{c}}_s^c$ which are our message sum symbols, namely $[6,5,0]$. For example, adding the pattern $[4,0,4]$ to $[2,1,0]$ produces $[6,1,4]$. The lifting pattern $[4,4,0]$ yields zero ML distance and therefore the final estimated sum $\hat{\mathbf{s}}^c = [6,5,0]$, is given as the output which matches the original sum of the transmitted messages.

IV. MSE Analysis of the Coded SumComp Scheme

Let s denote the true sum and $\hat{\mathbf{s}}^c$ the estimated sum obtained after decoding. Since the target function is the sum, the post-processing function simplifies to $\psi(x) = x$. The mean squared error (MSE) is defined as

$$\text{MSE} = \mathbb{E}\left[\|\hat{\mathbf{s}}^c - \mathbf{s}\|^2 \right]. \tag{7}$$

Theorem 1 (MSE for Coded SumComp)**.** *Consider the Coded SumComp system over $\mathbb{Z}_q$ with K transmitting nodes and an (n, m, d_L) linear code over $\mathbb{Z}_q$ used at the receiver for syndrome decoding. Let $t = \lfloor (d_L - 1)/2 \rfloor$ denote the maximum number of correctable symbol errors under Lee metric decoding.*

Let $\tilde{\boldsymbol{\epsilon}} = (\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_n)$ denote the effective noise vector after modulo-q folding, where each $\tilde{\epsilon}_i \in \mathbb{Z}_q$. Define the decoder failure probability as

$$P_f = \mathbb{P}\big(w_L(\tilde{\boldsymbol{\epsilon}}) > t\big) = \sum_{w=t+1}^{n} \sum_{\substack{\tilde{\boldsymbol{\epsilon}}: \\ w_L(\tilde{\boldsymbol{\epsilon}})=w}} \prod_{i=i}^{n} \mathbb{P}(\tilde{\epsilon}_i)$$

where $w_L(\tilde{\boldsymbol{\epsilon}})$ denotes the Lee weight of the error vector $\tilde{\boldsymbol{\epsilon}}$.

Then the mean squared error (MSE) of the recovered sum satisfies the bound

$$\text{MSE} \le \sum_{w=t+1}^{n} w^2 \sum_{\substack{\tilde{\boldsymbol{\epsilon}}: \\ w_L(\tilde{\boldsymbol{\epsilon}})=w}} \prod_{i=1}^{n} \mathbb{P}(\tilde{\epsilon}_i). \tag{8}$$

Proof: Let $\mathbf{r} = (r_1, \ldots, r_n)$ denote the received vector, where $r_i = \sum_{k=1}^K x_{k,i}^c + z_i$, $z_i \sim \mathcal{CN}(0, \sigma^2)$. Writing $z_i = z_{i,R} + j z_{i,I}$, we have $z_{i,R}, z_{i,I} \sim \mathcal{N}(0, \sigma^2/2)$, independent.

Quantization: Applying nearest Gaussian integer quantization $x_i^c = Q_\rho^c(r_i)$, we obtain $x_i^c = \sum_{k=1}^K x_{k,i}^c + e_i$, where $e_i = e_{i,R} + j e_{i,I}$ with $e_{i,R} = \mathcal{Q}_\rho^c(z_{i,R}) \in \mathbb{Z}$, $e_{i,I} = \mathcal{Q}_\rho^c(z_{i,I}) \in \mathbb{Z}$.

Sum Computation: Substituting $e_i = e_{i,R} + j e_{i,I}$, we obtain $x_i^c = \sum_{k=1}^K x_{k,i}^c + e_{i,R} + j e_{i,I}$. The desired sum is $s_i = \sum_{k=1}^K c_{k,i}$, and the decoded sum (after the SumComp decoder in (3)) is $\hat{c}_{s,i}^c = s_i + q_1 e_{i,R} + q_2 e_{i,I}$. Defining $\epsilon_i = q_1 e_{i,R} + q_2 e_{i,I}$, we get $\hat{c}_{s,i}^c = s_i + \epsilon_i$.

Effective Noise Distribution: The components $e_{i,R}, e_{i,I}$ are independent with ([5])

$$\mathbb{P}(e_{i,R} = l_1) = Q\left(\frac{(2|l_1|-1)d_1}{\sigma_1}\right) - Q\left(\frac{(2|l_1|+1)d_1}{\sigma_1}\right), \tag{9}$$

$$\mathbb{P}(e_{i,I} = l_2) = Q\left(\frac{(2|l_2|-1)d_2}{\sigma_2}\right) - Q\left(\frac{(2|l_2|+1)d_2}{\sigma_2}\right). \tag{10}$$

where d_1, d_2 and σ_1, σ_2 denote the minimum decision distances and noise standard deviations in the real and imaginary dimensions, respectively.

Thus, $\mathbb{P}(\epsilon_i = y) = \sum_{q_1 a + q_2 b = y} \mathbb{P}(e_{i,R} = a)\mathbb{P}(e_{i,I} = b)$.

Modulo-q Folding: After modulo-q reduction, the decoded codeword becomes $\hat{\mathbf{u}}_s = (\hat{\mathbf{c}}_s^c \bmod q) = \tilde{\mathbf{c}}_s + \tilde{\boldsymbol{\epsilon}}$, where $\tilde{\epsilon}_i = \epsilon_i \bmod q$ and Thus the effective channel observed by the syndrome decoder at the i^{th} instant is $\hat{u}_{s,i} = \tilde{c}_{s,i} + \tilde{\epsilon}_i$. The distribution of $\tilde{\epsilon}_i$ is obtained by folding the integer error distribution of ϵ_i. For $j \in \{0, 1, \ldots, q-1\}$, $\mathbb{P}(\tilde{\epsilon}_i = j) = \sum_{b=-\infty}^{\infty} \mathbb{P}(\epsilon_i = j + bq)$. Assuming independence across coordinates, $\mathbb{P}(\tilde{\boldsymbol{\epsilon}}) = \prod_{i=1}^{n} \mathbb{P}(\tilde{\epsilon}_i)$.

After reduction, $\hat{\mathbf{u}}_s = (\hat{\mathbf{c}}_s^c \bmod q) = \tilde{\mathbf{c}}_s + \tilde{\boldsymbol{\epsilon}}$, where $\tilde{\epsilon}_i = \epsilon_i \bmod q$ and

$$\mathbb{P}(\tilde{\epsilon}_i = j) = \sum_{b=-\infty}^{\infty} \mathbb{P}(\epsilon_i = j + bq), \quad j \in \mathbb{Z}_q. \tag{11}$$

Assuming independence, $\mathbb{P}(\tilde{\boldsymbol{\epsilon}}) = \prod_{i=1}^{n} \mathbb{P}(\tilde{\epsilon}_i)$.

Decoding Error: For an $(n, m, d_{\min})$ linear code over $\mathbb{Z}_q$, decoding succeeds if $w_L(\tilde{\boldsymbol{\epsilon}}) \le t = \lfloor \frac{d_{\min}-1}{2} \rfloor$. Thus,

$$P_f = \sum_{w=t+1}^{n} \sum_{\substack{\tilde{\boldsymbol{\epsilon}}: \\ w_L(\tilde{\boldsymbol{\epsilon}})=w}} \prod_{i=1}^{n} \mathbb{P}(\tilde{\epsilon}_i). \tag{12}$$

MSE Analysis: Let $\Delta\mathbf{s} = \hat{\mathbf{s}}^c - \mathbf{s}$. Then

$$MSE = \mathbb{E}[\|\Delta\mathbf{s}\|^2] = \mathbb{E}[\|\Delta\mathbf{s}\|^2 \mid \text{failure}] P_f, \tag{13}$$

since the error is zero upon success. When $w_L(\tilde{\boldsymbol{\epsilon}}) = w$, an error of Lee weight w is introduced into the original sum $\mathbf{s}$, therefore $\|\Delta\mathbf{s}\|^2 \le w_L(\tilde{\boldsymbol{\epsilon}})^2 = w^2$, hence

$$MSE \le \sum_{w=t+1}^{n} w^2 \sum_{\substack{\tilde{\boldsymbol{\epsilon}}: \\ w_L(\tilde{\boldsymbol{\epsilon}})=w}} \prod_{i=1}^{n} \mathbb{P}(\tilde{\epsilon}_i). \tag{14}$$

■

V. Simulation Results

In this section, we evaluate the performance of the proposed Coded SumComp scheme through Monte Carlo simulations. We consider a system with 2 transmitting nodes performing over-the-air sum computation over the ring $\mathbb{Z}_4$. For the SumComp system $q_1 = 101, q_2 = 64, \mu_1 = -19, \mu_2 = 30$ and $k = 1, 2$ are taken for simulation purposes. A $(5, 3)$ linear code over $\mathbb{Z}_4$ with minimum Lee distance $d_L = 3$ is employed at the transmitter and receiver with G and H as given in III-D. For each SNR point, results are averaged over 10^6 independent Monte Carlo frames, each containing 3000 transmitted symbols. Fig. 4 compares the MSE performance of uncoded and Coded SumComp. For the uncoded scheme, the analytical curve overlaps the empirical results, while for

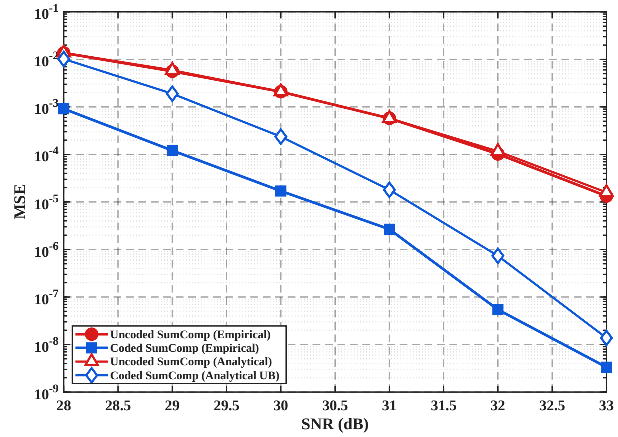

Fig. 4. Mean squared error (MSE) versus SNR for uncoded and Coded SumComp over $\mathbb{Z}_4$.

the coded scheme the analytical upper bound lies above the empirical curve validating the theoretical analysis. The coded scheme provides an MSE gain of approximately 2–3 dB compared to the uncoded system.

VI. Conclusion

In this paper, we analyzed the mean squared error (MSE) performance of coded digital over-the-air computation (SumComp) system. An analytical upper bound on the MSE was derived in terms of the decoder failure probability under Lee-metric decoding for linear codes over $\mathbb{Z}_q$. Simulation results using a $(5, 3)$ code over $\mathbb{Z}_4$ confirm the analysis and demonstrate that coding significantly reduces the MSE compared to the uncoded SumComp scheme, particularly in the moderate-to-high SNR regime.

References

[1] M. Goldenbaum and S. Stańczak, "Harnessing interference for analog function computation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 61, no. 20, pp. 4893–4906, 2013.

[2] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3498–3516, Oct. 2007.

[3] M. Goldenbaum and S. Stańczak, "Robust analog function computation via wireless multiple-access channels," *IEEE Transactions on Communications*, vol. 61, no. 9, pp. 3863–3877, Sep. 2013.

[4] J. M. B. d. S. S. Razavikia and C. Fischione, "Channelcomp: A general method for computation by communications," *IEEE Transactions on Communications*, vol. 72, no. 2, pp. 692–706, 2024.

[5] S. Razavikia *et al.*, "Sumcomp: Coding for digital over-the-air computation via the ring of integers," *IEEE Transactions on Communications*, vol. 73, no. 2, pp. 752–767, 2025.

[6] M. Goldenbaum and S. Stanczak, "On the channel estimation effort for analog computation over wireless multiple-access channels," *IEEE Wireless Communications Letters*, vol. 3, no. 3, pp. 261–264, 2014.

[7] R. Buck, "Approximate complexity and functional representation," *Journal of Mathematical Analysis and Applications*, vol. 70, no. 1, pp. 280–298, 1979.

[8] D. M. Burton, *Elementary Number Theory*, 7th ed. McGraw-Hill, 2011.

[9] J. A. Wood, "Duality for modules over finite rings and applications to coding theory," *American Journal of Mathematics*, vol. 121, no. 3, pp. 555–575, 1999.

[10] J. A. R. Hammons, P. V. Kumar, A. R. Calderbank, N. J. A. Sloane, and P. Solé, "The $\mathbb{Z}_4$-linearity of Kerdock, Preparata, Goethals, and related codes," *IEEE Transactions on Information Theory*, vol. 40, no. 2, pp. 301–319, Mar. 1994.