# The capacity region of three user Gaussian inverse-compute-and-forward channels

Yanying Chen, Yiwei Song, Natasha Devroye
University of Illinois at Chicago, Chicago IL 60607, USA
Email: ychen90, ysong34, devroye @ uic.edu

***Abstract*—We consider a three user multiple access channel where transmitter $m$ has access to the linear equation $\mathbf{u}_m = \sum_{l=1}^{3} f_{ml}\mathbf{w}_l$ of independent messages $\mathbf{w}_1 \in \mathbb{F}_p^{k_1}, \mathbf{w}_2 \in \mathbb{F}_p^{k_2}, \mathbf{w}_3 \in \mathbb{F}_p^{k_3}$ (and $f_{ml} \in \mathbb{F}_p$), and the destination wishes to recover all three messages. This problem is motivated as the last hop in a network where relay nodes employ the Compute-and-Forward strategy and decode linear equations of messages; we seek to do the reverse and extract messages from sums over a multiple access channel. An achievable rate region for the two user problem was previously derived; here we extend and strengthen this work to show capacity for the two and three user Gaussian channel models subject to invertability conditions on the matrix of coefficients describing the given linear equations of messages. The optimal transmission scheme is not to independently send the three equations $\mathbf{u}_m$ over the MAC but rather to exploit their special correlation structure.**

## I. Introduction and outline

The recently proposed Compute-and-Forward (CF) framework [1] enables the decoding of linear combinations of messages at relays over Gaussian channels. That is, the decoding of integer combinations of lattice codewords is shown to correspond to decoding linear combinations of the underlying messages $\mathbf{w}$ assumed to be vectors of length $k$ of elements over a finite field of size $p$, $\mathbb{F}_p$, or $\mathbf{w} \in \mathbb{F}_p^k$. When decoding sums of messages suffices for one's communication needs, this may sometimes be done at higher rates using the CF rates than decoding individual messages as in a multiple access channel.

The CF model seeks to decode linear combinations of messages when individual messages are transmitted over a multiple access channel (MAC)[1]; the inverse compute-and-forward (ICF) channel model studied here seeks to do the reverse, and extract or decode individual messages over a MAC from relays which possess linear combinations of messages, previously obtained through CF. In a larger network one may envision source nodes having messages, destination nodes wanting to decode these messages, and intermediate relay nodes decoding individual or linear equations of messages according to the CF framework. In this paper we determine the rates at which we may extract individual messages from linear *message equations* known at relays over a Gaussian multiple access channel. This may then be combined with CF rates in deriving overall achievable rates in larger networks.

We focus on the three user ICF problem (contains the two user problem) where there are three relay nodes which possess linear combinations of three messages obtained using the CF framework. These relays then transmit over a Gaussian MAC to a single destination which seeks to decode the three individual messages. The relays have linear combinations, or equations, of independent messages. In order for the destination to even be able to obtain the individual messages, the matrix relating the messages to the equations must be invertible[2]. One might then initially consider sending the three equations to the destination using independent codebooks as in a classical MAC, and having the destination invert the message equations to obtain the original messages. However, the message equations are generally correlated when the rates of the messages are unequal, and a region larger than the MAC channel capacity region is in fact achievable.

Intuitively, there are two ways of improving upon the MAC region: 1) the equations are correlated, so if one equation is correct/wrong at the destination this will impact the number of possible choices of values the others may take, thereby reducing the rates on the left-hand side (LHS) of capacity expressions; and 2) Common messages may be extracted from the equations which may be coherently sent over the Gaussian MAC, intuitively lifting the right-hand side (RHS) of the capacity region expressions.

**Past Work.** This work, and in particular the problem statement builds upon the compute-and-forward framework [1]: it is assumed that message equations have been previously decoded at the relays, and that messages are length $k$ vectors of elements over a finite field $\mathbb{F}_p$. The ICF problem was first considered for the two user case in [2], where an achievable rate region was presented – we extend and improve upon this work by showing the capacity region for the three user case. One may show that the two user ICF problem is equivalent to sending one common message and two private messages over a MAC (as in the Slepian-Wolf MAC), whose capacity is known for both the discrete and Gaussian channels [3], [4], [5]. The three user problem is related to an extension of the two user problem (which may be mapped to the Slepian-Wolf MAC) which has a new ingredient – the equations may be shown to not only have common and independent components, but also pairwise independent components. To the best of our knowledge this cannot be mapped onto a solved problem.

[1] The CF framework may handle more general cases when combinations of messages are transmitted as well, but our statement was made for the sake of argument / intuitive definition of the ICF model.

[2] We impose additional constraints; see Section II and footnote 3.

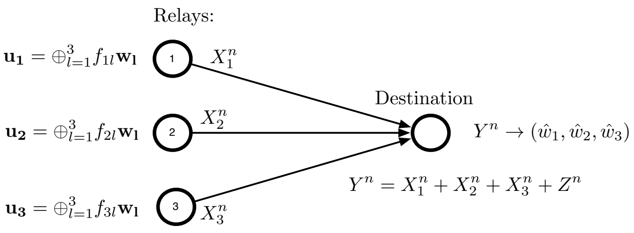
Relays:
$\mathbf{u_1} = \oplus_{l=1}^{3} f_{1l}\mathbf{w}_l$
1
$X_1^n$
$\mathbf{u_2} = \oplus_{l=1}^{3} f_{2l}\mathbf{w}_l$
2
$X_2^n$
$\mathbf{u_3} = \oplus_{l=1}^{3} f_{3l}\mathbf{w}_l$
3
$X_3^n$
Destination
$Y^n \rightarrow (\hat{w}_1, \hat{w}_2, \hat{w}_3)$
$Y^n = X_1^n + X_2^n + X_3^n + Z^n$

Fig. 1. Three user Gaussian ICF channel model. REPLACE

**Contribution and Outline.** Our main contribution is the derivation of the capacity region for decoding three messages over a Gaussian multiple access channel when the three transmitters each have a linear equation of the three messages, subject to invertibility conditions. We define this three user ICF channel in Section II, derive two key lemmas in Section III; present an achievable rate region for the discrete memoryless ICF in Section IV and the main capacity result in Section V.

**Notation.** Row vectors and matrices are written in bold font in lower and upper case, respectively, e.g., $\mathbf{u}$ and $\mathbf{U}$. In the context of coding and decoding, the length-$n$, $n \in \mathbb{N}$, vector codewords are represented by $X^n$ or $Y^n$. Define $C(x)$ as $\frac{1}{2}\log_2(1+x)$, $E[\cdot]$ as the expectation operator, and $\Pr\{A\}$ the probability of event $A$. Let $A \otimes B$ denote the Cartesian product of the sets $A$ and $B$, and $|A|$ denote the cardinality of set $A$. For $p$ prime, let $\mathbb{F}_p^k \cong \{0,1,\cdots,p-1\}^k$ ("$\cong$" denotes isomorphism) denote the field of length $k$ vectors of elements in the field $\mathbb{F}_p \cong \{0,1,\cdots,p-1\}$, under element-wise addition / multiplication modulo $p$. Let var$(X)$ denote the variance of $X$.

## II. Channel model

The three user ICF problem considers the decoding of three messages $\mathbf{w}_l$ which are uniformly distributed over $\mathbb{F}_p^{k_l}$ ($l = 1,2,3$) where we assume $k_1 \geq k_2 \geq k_3$; all messages are zero-padded at the head to a common length $k := \max_l k_l$. Let $\mathbf{W}$ denote the $3 \times k$ matrix whose $l$-th row is message $\mathbf{w}_l$. The *message rate* $R_l$ of message $\mathbf{w}_l$ at source node $l$ is $R_l := \frac{1}{n}\log_2(p^{k_l})$ where $n$ is the blocklength. Note that $R_1 \geq R_2 \geq R_3$. Because the messages may have different lengths, we may sub-divide the message into three portions: let $\mathbf{w}_{l,c}$ denote a *message section* of length $s_c := k_c - k_{c+1}$ (for $k_4 := 0$) which corresponds to the $c$-th segment of message $\mathbf{w}_l$ for $c \in \{1,2,3\}$ as shown in Fig. 2. Let $\mathbf{W}_{,c}$ be the matrix of dimension $3 \times s_c$ whose $l$-th row is $\mathbf{w}_{l,c}$.

Three transmitters (which in a larger network would be relay nodes) are assumed to have recovered a linear combination of the messages (as in the CF framework [1]):

$$\begin{aligned}\mathbf{u}_m &= \bigoplus_{l=1}^{3} f_{ml}\mathbf{w}_l \\ &= (\mathbf{u}_{m,1}, \mathbf{u}_{m,2}, \mathbf{u}_{m,3}) \quad m = 1,2,3\end{aligned}$$

which lies in $\mathbb{F}_p^k$ for some given $f_{ml} \in \mathbb{F}_p$. Let matrix $\mathbf{F}$ have $m$-th row $\mathbf{f}_m = (f_{m1}, f_{m2}, f_{m3})$. Then define $\mathbf{u}_m = \mathbf{f}_m \cdot \mathbf{W}$, and further define the *equation sections* $\mathbf{u}_{m,c} := \mathbf{f}_m \cdot \mathbf{W}_{,c}$. The *matrix of cth equation section* is $\mathbf{U}_{,c}$ whose $l$-th row is $\mathbf{u}_{l,c}$. In matrix form,

$$\begin{pmatrix}\mathbf{U}_{,1} & \mathbf{U}_{,2} & \mathbf{U}_{,3}\end{pmatrix} = \begin{pmatrix}\mathbf{F}\end{pmatrix} \cdot \begin{pmatrix}\mathbf{W}_{,1} & \mathbf{W}_{,2} & \mathbf{W}_{,3}\end{pmatrix}$$

or, breaking this into *message sections* and *equation sections*, as shown in Fig. 2

$$\begin{pmatrix}\mathbf{u}_{1,1} & \mathbf{u}_{1,2} & \mathbf{u}_{1,3} \\ \mathbf{u}_{2,1} & \mathbf{u}_{2,2} & \mathbf{u}_{2,3} \\ \mathbf{u}_{3,1} & \mathbf{u}_{3,2} & \mathbf{u}_{3,3}\end{pmatrix} = \begin{pmatrix} & & \\ & \mathbf{F} & \\ & & \end{pmatrix} \cdot \begin{pmatrix}\mathbf{w}_{1,1} & \mathbf{w}_{1,2} & \mathbf{w}_{1,3} \\ \mathbf{0} & \mathbf{w}_{2,2} & \mathbf{w}_{2,3} \\ \mathbf{0} & \mathbf{0} & \mathbf{w}_{3,3}\end{pmatrix}.$$

Each relay is equipped with an *encoder*, $\mathcal{E}_m : \mathbb{F}_p^k \rightarrow \mathcal{X}_m^n$, that maps the decoded equation $\mathbf{u}_m$, a length-$k$ vector, to a length-$n$ codeword in the input alphabet $\mathcal{X}_m$, i.e, $X_m^n(\mathbf{u}_m) = \mathcal{E}_m(\mathbf{u}_m)$. Codewords $X_m^n$ are inputs to a multiple access channel with input alphabets $\mathcal{X}_m$, output alphabet $\mathcal{Y}$ and memoryless transition probabilities $p(y|x_1,x_2,x_3)$. We will consider achievable rate regions for this general class of channels and will show capacity for the Gaussian multiple access channel where input and output alphabets are the real line, and channel outputs $Y^n$ are related to inputs as

$$Y^n = \sum_{m=1}^{3} X_m^n + Z^n, \tag{1}$$

where $Z^n$ is i.i.d. Gaussian noise, $Z^n \sim \mathcal{N}(\mathbf{0}_{n\times 1}, \mathbf{I}_{n\times n})$, and inputs are further subject to power constraints $E\left[\|X_m^n\|^2\right] \leq nP_m$ (which have absorbed any channel gains). The Gaussian channel model is shown in Fig. 1.

The single decoder wishes to decode $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$. Since the relays have $\mathbf{U} = \mathbf{F} \cdot \mathbf{W}$, this is only feasible if $\mathbf{F}$ is invertible, which we assume. *We furthermore assume that* $\mathbf{F}$ *and all of its square sub-matrices are non-singular.*[3] The decoder consists of a function $\mathcal{D}$ which takes received signals $Y^n$ and computes estimates $\{\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2, \hat{\mathbf{w}}_3\}$ of the messages. We note that since $\mathbf{F}$ is invertible, it is equivalent to obtain estimates of the equations $\{\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \hat{\mathbf{u}}_3\}$ and then invert to obtain the messages. We say that messages $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ are decoded with average probability of error $\epsilon$ if $\Pr\left(\bigcup_{l=1}^{3}\{\hat{\mathbf{w}}_l \neq \mathbf{w}_l\}\right) < \epsilon$.

Rates $(R_1, R_2, R_3)$, for $R_1 \geq R_2 \geq R_3$ are said to be *achievable* if for any $\epsilon > 0$ and $n$ large enough, there exist encoders $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$ and a decoder $\mathcal{D}$ such that the probability of error in decoding the messages of message rates $R_1, R_2, R_3$ is bounded by $\epsilon$. The capacity region for the ICF problem is defined as the closure of the set of all achievable rates.

*Remark 1:* Each equation $\mathbf{u}_m$ takes on values in $\mathbb{F}_p^k$, but the different equation values are not independent (over $m$). In particular, when the value(s) of one or several equations are given, the actual supports of the remaining equations are constrained. While each $\mathbf{u}_m$ may take on $2^{nR_1}$ values, the number of choices for the $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ is smaller than $(2^{nR_1})^3 = 2^{n3R_1}$ except for the special case when $R_1 =$

[3] While the invertibility constraint is necessary for feasibility, that all square sub-matrices are non-singular is for ease of exposition only; the extension to arbitrary invertible $\mathbf{F}$ poses no major technical issues, but the expressions are not succinct; deriving *succinct* expressions is current investigation [6].

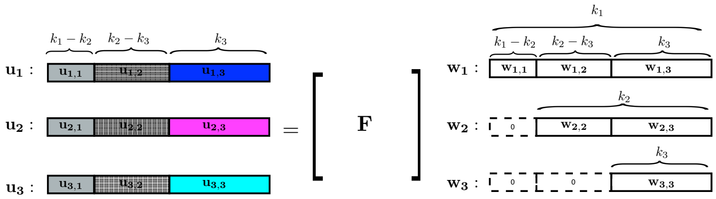

Fig. 2. Three user ICF message / equation structure. The grey color indicates that these equation sections ($\mathbf{u}_{*,1}$) are fully correlated; shading indicates that these three equation sections ($\mathbf{u}_{*,2}$) are pairwise independent, while different solid colors indicate that these three equation sections ($\mathbf{u}_{*,3}$) are mutually independent. All message sections $\mathbf{w}_{i,j}$ are mutually independent.

$R_2 = R_3$, i.e., $k_1 = k_2 = k_3$. This is exactly what we seek to exploit / characterize in deriving our achievable rate regions and capacity results, using the following two lemmas.

## III. Cardinality lemma and equation structure

The messages $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ are uniformly distributed over $\mathbb{F}_p^{k_1}, \mathbb{F}_p^{k_2}, \mathbb{F}_p^{k_3}$, respectively. We now seek to characterize the distributions of the equations $\mathbf{u}_m$ conditioned on other equations. This is useful for counting the error events in the achievability proof as well as for relating the entropies of messages to equations in the converse.

*Lemma 1 (Cardinality Lemma):* The triples $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_2)$ take on $2^{n(R_1+R_2+R_3)}$ values in $\mathbb{F}_p^k \bigotimes \mathbb{F}_p^k \bigotimes \mathbb{F}_p^k$ uniformly. The pairs $(\mathbf{u}_1, \mathbf{u}_2)$ take on $2^{n(R_2+R_3)}$ values in $\mathbb{F}_p^k \bigotimes \mathbb{F}_p^k$ uniformly for every fixed $\mathbf{u}_3$ (and likewise for the permutations $(\mathbf{u}_1, \mathbf{u}_3)$ given fixed $\mathbf{u}_2$, and $(\mathbf{u}_2, \mathbf{u}_3)$ given fixed $\mathbf{u}_1$). Finally, $\mathbf{u}_1$ takes on $2^{nR_3}$ values in $\mathbb{F}_p^k$ uniformly for each given pair $(\mathbf{u}_2, \mathbf{u}_3)$ (and likewise for the permutations $\mathbf{u}_2$ given $(\mathbf{u}_1, \mathbf{u}_3)$ and $\mathbf{u}_3$ given $(\mathbf{u}_1, \mathbf{u}_2)$.

*Proof:* $|(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)| = 2^{n(R_1+R_2+R_3)}$ since matrix $\mathbf{F}$ is of full rank and there is a one-to-one mapping between $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ and $(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3)$. Given one equation, say $\mathbf{u}_3$, we can eliminate the highest rate message, say $\mathbf{w}_1$, from both remaining equations $\mathbf{u}_1, \mathbf{u}_2$ to get two new equations that only depend on $\mathbf{w}_2$ and $\mathbf{w}_3$. The resulting two equations have $2^{n(R_2+R_3)}$ solutions for $(\mathbf{w}_2, \mathbf{w}_3)$, because every 2-by-2 sub-matrix of $\mathbf{F}$ is assumed to be non-singular. Thus, $|(\mathbf{u}_1, \mathbf{u}_2)|$ given a fixed $\mathbf{u}_3$ is $2^{n(R_2+R_3)}$. Similarly arguments follow by eliminating all but the lowest rate message to obtain $|\mathbf{u}_1| = 2^{nR_3}$ for fixed $(\mathbf{u}_2, \mathbf{u}_3)$. All values are taken on uniformly as $\mathbf{w}_l$ takes on all possible values in $\mathbb{F}_p^{k_l}$ uniformly.

From the relationship between the messages and the equations as shown in Fig. 2, we may state the following Lemma, which follows easily from the decomposition of the messages into message sections and the fact that $\mathbf{F}$ and all its square sub-matrices are full rank:

*Lemma 2 (Properties of equation sections):* The message and equation sections satisfy the following properties:

(I) $\mathbf{U}_{,1}$, or $\mathbf{u}_{1,1}, \mathbf{u}_{2,1}, \mathbf{u}_{3,1}$ are completely correlated, and may be used to reconstruct $\mathbf{w}_{1,1}$, a common message known to all relays.

(II) $\mathbf{u}_{1,2}, \mathbf{u}_{2,2}, \mathbf{u}_{3,2}$ are pairwise independent. That is, any two of these are pairwise independent and the third is a deterministic function of the other two. The three are not mutually independent.

(III) $\mathbf{u}_{1,3}, \mathbf{u}_{2,3}, \mathbf{u}_{3,3}$ are mutually independent.

This presence of pairwise independent (but not mutually independent) equation sections is the new ingredient in moving from the two [2] to the three user ICF problem.

## IV. Achievable rate region for the three user ICF

Using Lemmas 1 (count the error events) and 2 (dictate the form of the input distribution, and later on, used in converse), we may show the following achievable rate region.

*Theorem 3 (DM 3-user ICF achievability):* The messages $(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3)$ at rates $(R_1 \geq R_2 \geq R_3)$ may be recovered from $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ (where $\mathbf{U} = \mathbf{F} \cdot \mathbf{W}$) sent over a MAC, assuming $\mathbf{F}$ and all of its square sub-matrices are non-singular, if the rates lie in

$$\mathcal{R}_{IN} := \bigcup_{p(q)p(x_1|q)p(x_2|q)p(x_3|q)} \mathcal{R} \tag{2}$$

for $|Q| \leq \min\{|\mathcal{X}_1| \cdot |\mathcal{X}_2| \cdot |\mathcal{X}_3|, |\mathcal{Y}|\}$, where $\mathcal{R}$ is defined as the set of $(R_1, R_2, R_3)$ with $(R_1 \geq R_2 \geq R_3)$ and

$$\begin{aligned}
R_1 + R_2 + R_3 &\leq I(X_1, X_2, X_3; Y) && \text{(3a)}\\
2R_2 + R_3 &\leq I(X_1, X_2, X_3; Y|Q) && \text{(3b)}\\
R_2 + R_3 &\leq I(X_1, X_2; Y|X_3, Q) && \text{(3c)}\\
R_2 + R_3 &\leq I(X_1, X_3; Y|X_2, Q) && \text{(3d)}\\
R_2 + R_3 &\leq I(X_2, X_3; Y|X_1, Q) && \text{(3e)}\\
R_3 &\leq I(X_1; Y|X_2, X_3, Q) && \text{(3f)}\\
R_3 &\leq I(X_2; Y|X_1, X_3, Q) && \text{(3g)}\\
R_3 &\leq I(X_3; Y|X_1, X_2, Q) && \text{(3h)}
\end{aligned}$$

*Proof:* The proof follows from standard random coding arguments similar to the proof of the Slepian Wolf MAC [3], [5]. We extract the common message section $\mathbf{w}_{1,1}$ of rate $R_1 - R_2$ known to all relays by Lemma 2 (I) which we encode using a codebook i.i.d. distributed according to $p(q)$. For every $q^n$, we superpose codebooks for $x_1^n$, $x_2^n$ and $x_3^n$ of rates $R_2$ which are created i.i.d. according to $p(x_1|q), p(x_2|q)$ and $p(x_3|q)$ and indexed by $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$, resp. The decoder searches for unique indices of $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ such that $(Q^n(\mathbf{U}_{,1}), X_1^n(\mathbf{u}_1), X_2^n(\mathbf{u}_2), X_3^n(\mathbf{u}_3), Y^n)$ are jointly typical. The error analysis proceeds as usual, where the only difference is the number of error events, which may be obtained from Lemma 1. This affects the LHS of the expressions; the RHS follow as in a MAC with a common message. A full proof is omitted due to space constraints, will be in [6]. The cardinality bound on $Q$ follows from [7] and may likely be tightened. ■

*Remark 2:* To understand the form of the LHS, consider for example (3b). This results from the error event that all message sections except the common message ($\mathbf{w}_{1,1}$ or $\mathbf{U}_{,1}$) are incorrect. The rate of these incorrect message sections is $2(R_2 - R_3) + 3(R_3) = 2R_2 + R_3$. Similarly, (3c)–(3e) correspond to when the common message portion and one of the codewords is correct (or $[\mathbf{u}_1 = \mathbf{u}_1^0, \mathbf{u}_2 \neq \mathbf{u}_2^0, \mathbf{u}_3 \neq \mathbf{u}_3^0]$), and thus the rates of the incorrect message portions is

$1(R_2-R_3)+2(R_3) = R_2+R_3$ (or by Lemma 1). Finally, (3f)–(3h) correspond to when the common message and two entire codewords are correct, meaning that only the third independent message section of rate $R_3$ is wrong.

*Remark 3:* We note that the above theorem holds for $R_1 \geq R_2 \geq R_3$. One may obtain other regions for other relative sizes of $R_1, R_2, R_3$ analogously. The ICF problem is motivated as the final hop in a network where intermediate relays employ CF, leading to the special message structure at the relays. When deriving an achievable rate region for the entire network of CF and ICF links, one would take the intersection of the rate regions for the CF and ICF portions, and take the convex hull *after* taking this intersection. For this reason, we do not claim the convex hull of the rate regions for different orderings of $R_1, R_2, R_3$ to be achievable.

## V. Capacity region for three user Gaussian ICF

We now present our main result: the capacity region of the three user Gaussian ICF. Achievability follows by selecting Gaussian inputs in Theorem 3: we decompose the message equations into common, pairwise independent, and independent components,and coherently combine the common components by generating the same Gaussian codewords for the common parts. For both the pairwise independent and independent components, we use independent Gaussian codewords. The subtlety (and subtle difference with an extension of the Slepian-Wolf MAC) lies in the converse. There, besides the derivation of matching outer bound inequalities which follow by the properties in Lemma 1 and 2, the key idea is to exploit the linearity and second moment constraints of the AWGN channel and note that the conditional *pairwise* independence of the input distributions is sufficient to demonstrate Gaussian optimality of our new derived outer bound region.

*Theorem 4 (3-user Gaussian ICF capacity):* The capacity region for recovering messages $(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3)$ at rates $R_1 \geq R_2 \geq R_3$ from the equations $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ (where $\mathbf{U} = \mathbf{F} \cdot \mathbf{W}$) sent over an AWGN-MAC in (1), assuming $\mathbf{F}$ and all of its square sub-matrices are non-singular, is

$$\bigcup_{0 \leq b_1, b_2, b_3 \leq 1} \mathcal{R}^{\mathcal{G}}(b_1, b_2, b_3) \tag{4}$$

where

$$\begin{aligned} \mathcal{R}^{\mathcal{G}}(b_1, b_2, b_3) = \Big\{ &(R_1, R_2, R_3) : R_1 \geq R_2 \geq R_3 \\ &R_1 + R_2 + R_3 \leq C(P_1 + P_2 + P_3 + 2\sqrt{b_1 b_2 P_1 P_2} + \\ &\qquad 2\sqrt{b_1 b_3 P_1 P_3} + 2\sqrt{b_2 b_3 P_2 P_3}) \\ &2R_2 + R_3 \leq C((1-b_1)P_1 + (1-b_2)P_2 + (1-b_3)P_3) \\ &R_2 + R_3 \leq \min_{i \neq j \in \{1,2,3\}} C((1-b_i)P_i + (1-b_j)P_j) \\ &R_3 \leq \min_{i \in \{1,2,3\}} C((1-b_i)P_i) \Big\}. \end{aligned} \tag{5}$$

*Proof:* **Achievability.** Evaluating the region in Theorem 3 for the following choice of input distributions yields $\mathcal{R}^{\mathcal{G}}(b_1, b_2, b_3)$ for fixed $b_1, b_2, b_3 \in [0,1]$. Let $[T_0, T_1, T_2, T_3] \sim \mathcal{N}(\mathbf{0}_{4\times 1}, \mathbf{I}_{4\times 4})$, for $i = 1,2,3$ and select

$$Q = T_0, \; X_i = \sqrt{b_i}\sqrt{P_i} T_0 + \sqrt{1-b_i}\sqrt{P_i} T_i. \tag{6}$$

**Converse.** We first show that the form of the right hand sides of the outer bound follows those of (3) and discuss over which distributions this must be taken. Notice from the problem statement that the following Markov chain holds: $(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3) \to (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \to (X_1^n, X_2^n, X_3^n) \to Y^n \to (\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \hat{\mathbf{u}}_3)$. Eq. (3a) follows by MAC arguments. For (3b):

$$\begin{aligned} n(2R_2 + R_3) &= H(\mathbf{W}_{,3}, \mathbf{W}_{,2}) \overset{(a)}{=} H(\mathbf{U} | \mathbf{U}_{,1}) \\ &\leq I(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3; Y^n | \mathbf{U}_{,1}) + n \cdot \epsilon_n \\ &\overset{(b)}{\leq} \sum_{i=1}^{n} I(X_{1i}, X_{2i}, X_{3i}; Y_i | \mathbf{U}_{,1}) + n \cdot \epsilon_n \end{aligned} \tag{7}$$

For bounds (3c) (and (3d) and (3e)), the following holds:

$$\begin{aligned} n(R_2 + R_3) &= H(\mathbf{w}_{1,3}, \mathbf{w}_{2,3}, \mathbf{w}_{1,2}) \overset{(a)}{=} H(\mathbf{u}_1, \mathbf{u}_2 | \mathbf{u}_3) \\ &\leq I(\mathbf{u}_1, \mathbf{u}_2; Y^n | \mathbf{u}_3) + n \cdot \epsilon_n \\ &\overset{(b)}{=} \sum_{i=1}^{n} I(\mathbf{u}_1, \mathbf{u}_2, X_{1i}, X_{2i}; Y_i | Y^{i-1}, \mathbf{u}_3, X_{3i}, \mathbf{U}_{,1}) + n\, \epsilon_n \\ &\overset{(b)}{=} \sum_{i=1}^{n} I(X_{1i}, X_{2i}; Y_i | X_{3i}, \mathbf{U}_{,1}) + n \cdot \epsilon_n \end{aligned} \tag{8}$$

Finally, for the single rate bounds (3f) (and (3g) and (3h))

$$\begin{aligned} n(R_3) &= H(\mathbf{w}_{1,3}) \overset{(a)}{=} H(\mathbf{u}_1 | \mathbf{u}_2, \mathbf{u}_3) \leq I(\mathbf{u}_1; Y^n | \mathbf{u}_2, \mathbf{u}_3) + n\, \epsilon_n \\ &\overset{(b)}{\leq} \sum_{i=1}^{n} I(\mathbf{u}_1, X_{1i}; Y_i | Y^{i-1}, \mathbf{u}_2, \mathbf{u}_3, X_{2i}, X_{3i}, \mathbf{U}_{,1}) + n \cdot \epsilon_n \\ &\overset{(b)}{\leq} \sum_{i=1}^{n} I(X_{1i}; Y_i | X_{2i}, X_{3i}, \mathbf{U}_{,1}) + n \cdot \epsilon_n \end{aligned} \tag{9}$$

The equalities between message rates and equation entropies in (a) all follow by definitions, and Lemma 1 and 2. Steps (b) follow from the encoding functions, Markov chain and the memoryless channel properties. Now, in all bounds set $Q_i = \mathbf{U}_{,1}$ and further combine this into a time-sharing random variable $Q$; we then notice that since $\mathbf{u}_m$'s are pairwise conditionally independent given $\mathbf{U}_{,1}$ (Lemma 2) and since $X_m^n$ is a function of $\mathbf{u}_m$, then $X_m^n$ (and hence also $X_m$) are conditionally pairwise independent given $Q$, or the following Markov chains hold $X_1 \to Q \to X_2$, and $X_1 \to Q \to X_3$ and $X_2 \to Q \to X_3$. By further time-sharing arguments and Jensens' inequality we obtain the form in (3).

Let $\mathcal{P} := \{p(q, x_1, x_2, x_3) : X_1 \to Q \to X_2, X_2 \to Q \to X_3, X_2 \to Q \to X_3, E[X_m^2] \leq P_m, m = 1,2,3\}$. Our outer bound may thus be expressed as follows

$$\mathcal{R}_{\text{out}} = \bigcup_{p(x_1,x_2,x_3,q) \in \mathcal{P}} \mathcal{R} \tag{10}$$

We now essentially mimic the steps of [4, Proposition 3.14, Lemma 3.15, Appendix B.2] to show optimality of Gaussian distributions in $\mathcal{R}_{\text{out}}$. The key difference is that we have three inputs $X_1, X_2, X_3$ satisfying Markov relationships rather than two. First, for any fixed distribution in $\mathcal{P}$, the region $\mathcal{R}$ for this distribution may be further outer bounded as $\mathcal{R} \subseteq \mathcal{R}'$, for

$$\begin{aligned}
\mathcal{R}' := \Big\{ &(R_1, R_2, R_3) : R_1 \geq R_2 \geq R_3 \\
&R_1 + R_2 + R_3 \leq C(E[X_1^2] + E[X_2^2] + E[X_3^2] \\
&\qquad\qquad + 2E[X_1X_2] + 2E[X_1X_3] + 2E[X_2X_3]) \\
&2R_2 + R_3 \leq C(\text{var}(X_1|Q) + \text{var}(X_2|Q) + \text{var}(X_3|Q)) \\
&R_2 + R_3 \leq C(\text{var}(X_1|Q) + \text{var}(X_2|Q)) \\
&R_2 + R_3 \leq C(\text{var}(X_1|Q) + \text{var}(X_3|Q)) \\
&R_2 + R_3 \leq C(\text{var}(X_2|Q) + \text{var}(X_3|Q)) \\
&R_3 \leq \min_{i \in \{1,2,3\}} C(\text{var}(X_i|Q)) \Big\}.
\end{aligned} \tag{11}$$

This may be shown as in [4, Lemma B.4]. The key is to first apply the Max-Entropy Theorem [8] conditional on $Q = q$, $\forall q \in \mathbb{Q}$, to upper bound the conditional mutual information terms. Jensen's inequality is then applied to obtain (11). The proof of the third inequality is shown as an example.

$$\begin{aligned}
&I(X_1, X_2; Y|X_3, Q) = E_Q[I(X_1, X_2; Y|X_3, Q = q)] \\
&\overset{(a)}{=} E_Q[h(X_1 + X_2 + Z|Q = q) - h(Z)] \\
&\overset{(b)}{\leq} E_Q\left[\frac{1}{2}\log\left(\frac{\text{var}(X_1 + X_2 + Z|Q = q)}{\text{var}(Z)}\right)\right] \\
&\overset{(c)}{=} E_Q\left[\frac{1}{2}\log(1 + \text{var}(X_1|Q = q) + \text{var}(X_2|Q = q))\right] \\
&\overset{(d)}{\leq} \frac{1}{2}\log(1 + \text{var}(X_1|Q) + \text{var}(X_2|Q))
\end{aligned} \tag{12}$$

(a) follows since $X_1, X_3$ are independent given $Q$, as are $X_2$ and $X_3$, and hence $X_1+X_2$ is conditionally independent of $X_3$ given $Q$. The inequality in (b) is achieved when $X_1+X_2+Z$ is Gaussian conditioned on $Q = q$; (c) follows from the additive form of the channel and since $X_1$ and $X_2$ are conditionally independent given $Q$ (arguably the *key* step in this converse); finally (d) follows by Jensen's inequality.

Using [4, Lemma B.3] and the definitions for $i = 1, 2, 3$:

$$\rho_i = \frac{E[X_i^2] - \text{var}(X_i|Q)}{E[X_i^2]} \tag{13}$$

one can show that $\mathcal{R}' \subseteq \mathcal{R}''$ where

$$\begin{aligned}
\mathcal{R}'' := \Big\{ &(R_1, R_2, R_3) : R_1 \geq R_2 \geq R_3 \\
&R_1 + R_2 + R_3 \leq C(E[X_1^2] + E[X_2^2] + E[X_3^2] + \\
&\quad 2\sqrt{\rho_1\rho_2}\sqrt{E[X_1^2]E[X_2^2]} + \\
&\quad 2\sqrt{\rho_1\rho_3}\sqrt{E[X_1^2]E[X_3^2]} + 2\sqrt{\rho_2\rho_3}\sqrt{E[X_2^2]E[X_3^2]}) \\
&2R_2 + R_3 \leq C((1-\rho_1)E[X_1^2] + (1-\rho_2)E[X_2^2] + (1-\rho_3)E[X_3^2]) \\
&R_2 + R_3 \leq C((1-\rho_1)E[X_1^2] + (1-\rho_2)E[X_2^2]) \\
&R_2 + R_3 \leq C((1-\rho_1)E[X_1^2] + (1-\rho_3)E[X_3^2]) \\
&R_2 + R_3 \leq C((1-\rho_2)E[X_2^2] + (1-\rho_3)E[X_3^2])
\end{aligned} \tag{14}$$

$$R_3 \leq \min_{i \in \{1,2,3\}} C((1-\rho_i)E[X_i^2]) \Big\} \tag{15}$$

Consider the distributions in (6) with $b_i$ replaced by $\rho_i$. Substitution into the region $\mathcal{R}''$ yields the Theorem. ■

*Remark 4:* The key differences with the two-user Slepian-Wolf MAC are the (i) different message structure which led to different rate bounds (or rates on the LHS), and (ii) the key step in showing Gaussian optimality was in step (c) above (and the analogous step for other inequalities) – using the fact that our channel is linear and that the second moment or variance is the quantity of interest (due to the power constraints). All that is needed is *pairwise* independence between the inputs conditioned on $Q$ (and not mutual independence), which is exactly what we have.

## VI. Conclusion and Future work

We have obtained the capacity region for the three user ICF problem in which one destination node wishes to extract three messages $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3 \in \mathbb{F}_p^k$ over a MAC from three transmitting nodes which have linear combinations of these messages. This differs from a MAC in two ways: the equations may be correlated and hence knowledge of one equation affects the possible number of values of the others and also enables correlation between the transmitted signals when the message rates are not equal. While the three user case was presented, we conjecture that this proof may be extended to $L$ users [6], where the key to showing Gaussian optimality is that the equations will remain pairwise independent conditioned on the common message section. The capacity region here may be used as a building block for the "last hop" in relay networks where CF is employed at relay nodes, besides being of independent interest. Whether the achievable rate region presented for a discrete memoryless channel is capacity remains an interesting open question.

## Acknowledgment

The work of all authors was partially supported by NSF under award CCF-1216825. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NSF.

## References

[1] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, 2011.

[2] Y. Song, N. Devroye, and B. Nazer, "Inverse compute-and-forward: Extracting messages from simultaneously transmitted equations," in *Proc. IEEE Int. Symp. Inf. Theory*, Aug. 2011.

[3] D. Slepian and J. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Syst. Tech. Journal*, vol. 52, no. 7, pp. 1037–1076, 1973.

[4] M. A. Wigger, "Cooperation on the multiple-access channel," Ph.D. dissertation, 2008.

[5] F. M. Willems, "Information theoretical results for multiple access channels," Ph.D. dissertation, 1982.

[6] Y. Chen, Y. Song, and N. Devroye, "On the L user inverse compute and forward problem," *to be submitted to IEEE Trans. Inf. Theory*, 2013.

[7] M. Salehi, "Cardinality bounds on auxiliary variables in multiple-user theory via the method of Ahlswede and Körner," *Stanford Department of Statistics Technical Rrport*, no. 33, Aug. 1978.

[8] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York:Wiley, 2006.