New Sum-Rate Upper Bound for the Two-User Gaussian Interference Channel

Raul Etkin
Hewlett-Packard Laboratories
Palo Alto, CA 94304
Email: raul.etkin@hp.com

***Abstract*—A new upper bound for the sum-capacity of the two-user complex Gaussian interference channel is derived. The bound is based on genies that provide side information signals to the receivers. Various parameters are introduced to control the side information, which can be optimized to get the tightest possible bound. The new bound improves the performance of the best known bounds for sufficiently small cross-gains.**

I. Introduction

The 2-user Gaussian interference channel (GIFC) models a situation in which two transmitters communicate independent information to their corresponding receivers. Each receiver observes a component-wise linear combination of the transmitted signals plus additive memoryless Gaussian noise, and seeks to decode, with probability close to one, the message of its corresponding transmitter, in spite of the interfering signals and noise. The time averages of the squares of the transmitted signal values are required to not exceed certain power constraints. A rate pair (R_1, R_2) is said to be achievable for a GIFC if the transmitters can increase the sizes of their message sets as 2^{mR_i} with the signal length m, and signal in such a way that the power constraints are met and the receivers are able to correctly decode their corresponding messages with probability converging to 1, as m grows to infinity. The set of all achievable rate pairs is known as the capacity region of the GIFC. Determining it, as a function of the channel coefficients (specifying the linear combinations mentioned above), power constraints, and noise variances, has been an open problem in information theory for over 30 years.

The best known coding scheme is that presented in [2]. In a special range of channel parameters, the so called strong interference regime, the capacity region is completely known [1], [2]. In the remaining parameter ranges, i.e. weak and mixed interference regimes, [4] determines the capacity region to within a 1 bit margin using a carefully chosen version of the scheme of [2], and a new genie-aided outer bound.

While a tight characterization of the capacity region in the weak and mixed interference regimes remains an open problem, the maximum achievable sum-of-rates is known for some parameter ranges [3], [5], [6], [7]. In addition, [4] shows that in some asymptotic regimes the gap between some lower and upper bounds on the sum-rate shrinks to zero, thus obtaining an asymptotic characterization of the sum-capacity for some other regimes.

In this paper we present a new upper bound for the sum-capacity of the complex GIFC with weak interference, which strengthens the sum-rate bounds of [4] and [5], [6], [7]. All these bounds are based on a genie-aided approach, where genies provide side information signals to the receivers. The side information signals are carefully chosen to enable the computation of a sum-rate upper bound, and are defined in terms of some free parameters which can be optimized.

Regarding notation, we use lower and upper case letters for scalars, boldface letters for vectors, and upper case calligraphic letters for sets. For example, a is a scalar, $\mathbf{v}$ is a vector, and $\mathcal{R}$ is a set. The length of an m-vector is indicated by the superscript m, e.g. $\mathbf{v}^m = (v(1), \ldots, v(m))$. We write $\Re(x)$ to denote the real part of the complex number x and use $\mathbb{E}$ for the expectation operator. We use the standard notation $h(\cdot)$, $I(\cdot;\cdot)$ to denote differential entropy and mutual information. All logarithms are to the base 2, and rates are in bits per channel use.

The rest of the paper is organized as follows. In the next section we define the channel model and the side information signals that we will use in the bound. Section III presents our main result which is proved in Section IV. The new bound is simplified for the special case of a symmetric GIFC with real gains in Section V, where a numerical comparison of the new bound and previous bounds is presented. Finally, Section VI discusses whether the new bound can extend the range of parameters where treating interference as noise is known to be a sum-capacity achieving strategy.

II. Channel model and side information signals

In the complex case, the normalized[1] complex Gaussian interference channel is defined by

$$\begin{aligned} y_1(t) &= x_1(t) + h_{21}x_2(t) + z_1(t) \\ y_2(t) &= h_{12}x_1(t) + x_2(t) + z_2(t) \end{aligned} \quad (1)$$

where $x_i(t) \in \mathbb{C}$ is the channel input for user $i = 1, 2$ at time $t = 1, \ldots, m$, subject to an average power constraint P_i (i.e. $\sum_{t=1}^{m} |x_i(t)|^2 \le mP_i$), and $z_i(t) \sim \mathcal{CN}(0, 1)^2$ is white noise

[1]A normalized GIFC has direct gains equal to 1, and noise variances equal to 1. It can be easily shown that for every complex GIFC there exists a normalized interference channel with the same capacity region. Therefore, there is no loss of generality in restricting attention to normalized channels.

[2]$\mathcal{CN}(0, \Sigma)$ denotes a circularly symmetric complex Gaussian distribution with zero mean and covariance matrix Σ.

978-1-4244-4313-0/09/$25.00 ©2009 IEEE

$$
\begin{aligned}
K_1 &= \log\left[\frac{(P_1+|h_{21}|^2P_2+1)(P_1+|\alpha_1|^2P_2+\sigma_1^2)-|P_1+h_{21}\alpha_1^*P_2+\rho_1^*\sigma_1|^2}{(P_1+|\alpha_1|^2P_2)\gamma_2+(1-|\rho_2|^2)\sigma_2^2}\right]\\
&\quad+\log\left[\frac{(P_2+|h_{12}|^2P_1+1)(P_2+|\alpha_2|^2P_1+\sigma_2^2)-|P_2+h_{12}\alpha_2^*P_1+\rho_2^*\sigma_2|^2}{(P_2+|\alpha_2|^2P_1)\gamma_1+(1-|\rho_1|^2)\sigma_1^2}\right]\\
K_2 &= \log\left(P_1+|h_{21}|^2P_2+1-\frac{|P_1+h_{21}\alpha_1^*P_2+\rho_1^*\sigma_1|^2}{P_1+|\alpha_1|^2P_2+\sigma_1^2}\right)-\log(1-|\rho_1|^2)\\
&\quad+\log\left(P_2+|h_{12}|^2P_1+1-\frac{|P_2+h_{12}\alpha_2^*P_1+\rho_2^*\sigma_2|^2}{P_2+|\alpha_2|^2P_1+\sigma_2^2}\right)-\log(1-|\rho_2|^2)
\end{aligned}
$$

---

affecting receiver i, and is independent of the channel inputs. We restrict our attention to channels with weak interference, i.e. $|h_{ij}|<1$.

In our genie-aided bounds, side information signals s_1 and s_2 will be given by a genie to receivers 1 and 2 respectively. We define the side information signals as

$$
\begin{aligned}
s_1(t) &= x_1(t)+\alpha_1x_2(t)+n_1(t)\\
s_2(t) &= \alpha_2x_1(t)+x_2(t)+n_2(t)
\end{aligned} \tag{2}
$$

where $n_i(t)\sim\mathcal{CN}(0,\sigma_i^2)$ is i.i.d. over time, and independent of the channel inputs and $z_j(\tau)$, for all τ when $j\neq i$ and for all $\tau\neq t$ when $j=i$, but is correlated with $z_i(t)$,

$$\mathbb{E}[n_i(t)z_i^*(t)]=\rho_i\sigma_i$$

with $|\rho_i|\leq 1$. The parameters $\sigma_1,\sigma_2\in\mathbb{R}$, $\sigma_1,\sigma_2\geq 0$, $\alpha_1,\alpha_2,\rho_1,\rho_2\in\mathbb{C}$, are arbitrary within their corresponding valid ranges and will be later optimized.

The bounds of [4], and [5], [7], [6] also use side information signals, which can be obtained from (2) with specific parameter choices. The genies of [4], [5], [6], [7] restrict $\alpha_1=\alpha_2=0$, while in addition, the genies of [4] further restrict $\rho_i=0$ and $\sigma_i=1$.

III. Main result

Our new bound is presented in the following theorem.

Theorem 1: Any rate pair (R_1,R_2) in the capacity region of the 2-user complex GIFC in weak interference satisfies

$$R_1+R_2\leq\min\{B_{1a},B_{1b},B_{1c},B_{2a},B_{2b},B_{2c}\}$$

where

$$
\begin{aligned}
B_{1a} &= \min_{\boldsymbol{\eta}\in\mathcal{R}_{1a}}\log\left[1+\frac{|\alpha_1|^2P_2\gamma_2}{(1-|\rho_2|^2)\sigma_2^2}\right]\\
&\qquad+\log\left[1+\frac{|\alpha_2|^2P_1\gamma_1}{(1-|\rho_1|^2)\sigma_1^2}\right]+K_1\\
B_{1b} &= \min_{\boldsymbol{\eta}\in\mathcal{R}_{1b}}\log\left[1+\frac{(P_2+|\alpha_2|^2P_1)\gamma_1}{(1-|\rho_1|^2)\sigma_1^2}\right]+K_1\\
B_{1c} &= \min_{\boldsymbol{\eta}\in\mathcal{R}_{1c}}\log\left[1+\frac{(P_1+|\alpha_1|^2P_2)\gamma_2}{(1-|\rho_2|^2)\sigma_2^2}\right]+K_1\\
\gamma_1 &= |\alpha_1|^2-2\Re[h_{21}\rho_1\sigma_1\alpha_1^*]+|h_{21}|^2\sigma_1^2\\
\gamma_2 &= |\alpha_2|^2-2\Re[h_{12}\rho_2\sigma_2\alpha_2^*]+|h_{12}|^2\sigma_2^2\\
\mathcal{R}_{1a} &= \Big\{(\sigma_1,\sigma_2,\alpha_1,\alpha_2,\rho_1,\rho_2):(1-|\rho_1|^2)\sigma_1^2\geq\gamma_1\sigma_2^2,\\
&\qquad (1-|\rho_2|^2)\sigma_2^2\geq\gamma_2\sigma_1^2\Big\}\\
\mathcal{R}_{1b} &= \mathcal{R}_{1a}\cap\Big\{(\sigma_1,\sigma_2,\alpha_1,\alpha_2,\rho_1,\rho_2):\\
&\qquad (1-|\rho_1|^2)\sigma_1^2|\alpha_1|^2\gamma_2\leq(1-|\rho_2|^2)\sigma_2^2\gamma_1\Big\}\\
\mathcal{R}_{1c} &= \mathcal{R}_{1a}\cap\Big\{(\sigma_1,\sigma_2,\alpha_1,\alpha_2,\rho_1,\rho_2):\\
&\qquad (1-|\rho_2|^2)\sigma_2^2|\alpha_2|^2\gamma_1\leq(1-|\rho_1|^2)\sigma_1^2\gamma_2\Big\}\\
B_{2a} &= \min_{\boldsymbol{\eta}\in\mathcal{R}_{2a}}\log\left[1+\frac{|\alpha_1|^2P_2}{\sigma_1^2}\right]+\log\left[1+\frac{|\alpha_2|^2P_1}{\sigma_2^2}\right]\\
&\qquad+K_2\\
B_{2b} &= \min_{\boldsymbol{\eta}\in\mathcal{R}_{2b}}\log\left[1+\frac{|\alpha_2|^2P_1+P_2}{\sigma_2^2}\right]+K_2\\
B_{2c} &= \min_{\boldsymbol{\eta}\in\mathcal{R}_{2b}}\log\left[1+\frac{P_1+|\alpha_1|^2P_2}{\sigma_1^2}\right]+K_2\\
\mathcal{R}_{2a} &= \Big\{(\sigma_1,\sigma_2,\alpha_1,\alpha_2,\rho_1,\rho_2):(1-|\rho_1|^2)\sigma_1^2\leq\gamma_1\sigma_2^2,\\
&\qquad (1-|\rho_2|^2)\sigma_2^2\leq\gamma_2\sigma_1^2\Big\}\\
\mathcal{R}_{2b} &= \mathcal{R}_{2a}\cap\Big\{(\sigma_1,\sigma_2,\alpha_1,\alpha_2,\rho_1,\rho_2):\sigma_2^2|\alpha_1|^2\leq\sigma_1^2\Big\}\\
\mathcal{R}_{2c} &= \mathcal{R}_{2a}\cap\Big\{(\sigma_1,\sigma_2,\alpha_1,\alpha_2,\rho_1,\rho_2):\sigma_1^2|\alpha_2|^2\leq\sigma_2^2\Big\}
\end{aligned}
$$

with $\boldsymbol{\eta}\triangleq(\sigma_1,\sigma_2,\alpha_1,\alpha_2,\rho_1,\rho_2)$ satisfying the conditions stated in the previous section, and where K_1 and K_2 are defined in the display at the top of the page.

IV. Proof of Theorem 1

For any code of block-length m, the sum rate of the genie-aided channel can be upper bounded using Fano's inequality:

$$
\begin{aligned}
&m(R_1+R_2-\epsilon_m)\leq I(\mathbf{x}_1^m;\mathbf{y}_1^m,\mathbf{s}_1^m)+I(\mathbf{x}_2^m;\mathbf{y}_2^m,\mathbf{s}_2^m)\\
&=I(\mathbf{x}_1^m;\mathbf{s}_1^m)+I(\mathbf{x}_1^m;\mathbf{y}_1^m|\mathbf{s}_1^m)+I(\mathbf{x}_2^m;\mathbf{s}_2^m)+I(\mathbf{x}_2^m;\mathbf{y}_2^m|\mathbf{s}_2^m)\\
&=h(\mathbf{s}_1^m)-h(\alpha_1\mathbf{x}_2^m+\mathbf{n}_1^m)+h(\mathbf{y}_1^m|\mathbf{s}_1^m)\\
&\quad-h(h_{21}\mathbf{x}_2^m+\mathbf{z}_1^m|\alpha_1\mathbf{x}_2^m+\mathbf{n}_1^m)+h(\mathbf{s}_2^m)-h(\alpha_2\mathbf{x}_1^m+\mathbf{n}_2^m)\\
&\quad+h(\mathbf{y}_2^m|\mathbf{s}_2^m)-h(h_{12}\mathbf{x}_1^m+\mathbf{z}_2^m|\alpha_2\mathbf{x}_1^m+\mathbf{n}_2^m)\\
&=h(\mathbf{s}_1^m)-h(\alpha_1\mathbf{x}_2^m+\mathbf{n}_1^m)+h(\mathbf{y}_1^m|\mathbf{s}_1^m)
\end{aligned}
$$

$$
\begin{aligned}
&- h\big(\mathbf{z}_1^m - \frac{h_{21}}{\alpha_1}\mathbf{n}_1^m | \alpha_1\mathbf{x}_2^m + \mathbf{n}_1^m\big) + h(\mathbf{s}_2^m) - h(\alpha_2\mathbf{x}_1^m + \mathbf{n}_2^m) \\
&+ h(\mathbf{y}_2^m|\mathbf{s}_2^m) - h\big(\mathbf{z}_2^m - \frac{h_{12}}{\alpha_2}\mathbf{n}_2^m | \alpha_2\mathbf{x}_1^m + \mathbf{n}_2^m\big) \\
=& h(\mathbf{s}_1^m) + h(\mathbf{y}_1^m|\mathbf{s}_1^m) - h\big(\mathbf{z}_1^m - \frac{h_{21}}{\alpha_1}\mathbf{n}_1^m\big) \\
&- h\big(\alpha_1\mathbf{x}_2^m + \mathbf{n}_1^m | \mathbf{z}_1^m - \frac{h_{21}}{\alpha_1}\mathbf{n}_1^m\big) + h(\mathbf{s}_2^m) + h(\mathbf{y}_2^m|\mathbf{s}_2^m) \\
&- h\big(\mathbf{z}_2^m - \frac{h_{12}}{\alpha_2}\mathbf{n}_2^m\big) - h\big(\alpha_2\mathbf{x}_1^m + \mathbf{n}_2^m | \mathbf{z}_2^m - \frac{h_{12}}{\alpha_2}\mathbf{n}_2^m\big) \\
=& h(\mathbf{s}_1^m) + h(\mathbf{y}_1^m|\mathbf{s}_1^m) - h(\mathbf{v}_1^m) - h(\alpha_1\mathbf{x}_2^m + \mathbf{w}_1^m) \\
&+ h(\mathbf{s}_2^m) + h(\mathbf{y}_2^m|\mathbf{s}_2^m) - h(\mathbf{v}_2^m) - h(\alpha_2\mathbf{x}_1^m + \mathbf{w}_2^m) \qquad (3)
\end{aligned}
$$

where we defined

$$
\begin{aligned}
v_i(t) &\triangleq z_i(t) - \frac{h_{ji}}{\alpha_i}n_i(t) \\
w_i(t) &\triangleq n_i(t) - E[n_i(t)|v_i(t)]
\end{aligned}
$$

for $i = 1, 2,\ j \neq i$. Note that $v_i(t) \sim \mathcal{CN}(0, \sigma_{v_i}^2)$ with $\sigma_{v_i}^2 = 1 - 2\Re\{h_{ji}\rho_i\sigma_i/\alpha_i\} + \sigma_i^2|h_{ji}|^2/|\alpha_i|^2$, and $w_i(t) \sim \mathcal{CN}(0, \sigma_{w_i}^2)$ with $\sigma_{w_i}^2 = (1-|\rho_i|^2)\sigma_i^2/\sigma_{v_i}^2$.

Continuing with (3),

$$
\begin{aligned}
m(R_1 + R_2 - \epsilon_m) \le\ & h(\mathbf{x}_1^m + \alpha_1\mathbf{x}_2^m + \mathbf{n}_1^m) - h(\alpha_2\mathbf{x}_1^m + \mathbf{w}_2^m) \\
&+ h(\mathbf{y}_1^m|\mathbf{s}_1^m) - h(\mathbf{v}_1^m) + h(\mathbf{x}_2^m + \alpha_2\mathbf{x}_1^m + \mathbf{n}_2^m) \\
&- h(\alpha_1\mathbf{x}_2^m + \mathbf{w}_1^m) + h(\mathbf{y}_2^m|\mathbf{s}_2^m) - h(\mathbf{v}_2^m) \\
=\ & h(\mathbf{x}_1^m + \alpha_1\mathbf{x}_2^m + \mathbf{n}_1^m) - h\big(\mathbf{x}_1^m + \frac{\mathbf{w}_2^m}{\alpha_2}\big) + h(\mathbf{y}_1^m|\mathbf{s}_1^m) \\
&- h(\mathbf{v}_1^m) + h(\mathbf{x}_2^m + \alpha_2\mathbf{x}_1^m + \mathbf{n}_2^m) - h\big(\mathbf{x}_2^m + \frac{\mathbf{w}_1^m}{\alpha_1}\big) \\
&+ h(\mathbf{y}_2^m|\mathbf{s}_2^m) - h(\mathbf{v}_2^m) - m\log(|\alpha_1\alpha_2|^2). \qquad (4)
\end{aligned}
$$

We will bound (4) differently depending on whether $\sigma_{w_i}^2/|\alpha_i|^2 \ge \sigma_j^2,\ (i,j) \in \{(1,2),(2,1)\}$, or $\sigma_{w_i}^2/|\alpha_i|^2 \le \sigma_j^2,\ (i,j) \in \{(1,2),(2,1)\}$. These cases correspond to $(\sigma_1, \sigma_2, \alpha_1, \alpha_2, \rho_1, \rho_2) \in \mathcal{R}_{1a}$ and $(\sigma_1, \sigma_2, \alpha_1, \alpha_2, \rho_1, \rho_2) \in \mathcal{R}_{2a}$ respectively.

A. Case $\boldsymbol{\eta} \in \mathcal{R}_{1a}$: $\sigma_{w_1}^2/|\alpha_1|^2 \ge \sigma_2^2$ and $\sigma_{w_2}^2/|\alpha_2|^2 \ge \sigma_1^2$

We can write:

$$
\begin{aligned}
h(\mathbf{x}_1^m + \alpha_1\mathbf{x}_2^m + \mathbf{n}_1^m) =& - I(\tilde{\mathbf{n}}_1^m; \mathbf{x}_1^m + \alpha_1\mathbf{x}_2^m + \mathbf{n}_1^m + \tilde{\mathbf{n}}_1^m) \\
&+ h(\mathbf{x}_1^m + \alpha_1\mathbf{x}_2^m + \mathbf{n}_1^m + \tilde{\mathbf{n}}_1^m) \quad (5) \\
h(\mathbf{x}_2^m + \alpha_2\mathbf{x}_1^m + \mathbf{n}_2^m) =& - I(\tilde{\mathbf{n}}_2^m; \mathbf{x}_2^m + \alpha_2\mathbf{x}_1^m + \mathbf{n}_2^m + \tilde{\mathbf{n}}_2^m) \\
&+ h(\mathbf{x}_2^m + \alpha_2\mathbf{x}_1^m + \mathbf{n}_2^m + \tilde{\mathbf{n}}_2^m) \quad (6)
\end{aligned}
$$

where $\tilde{\mathbf{n}}_i^m$ $(i = 1, 2)$ has i.i.d. circularly symmetric complex Gaussian components of variance $\sigma_{\tilde{n}_i}^2$, and is independent of everything else. Choosing $\sigma_{\tilde{n}_i}^2$ such that $\sigma_{\tilde{n}_1}^2 + \sigma_1^2 = \sigma_{w_2}^2/|\alpha_2|^2$ and $\sigma_{\tilde{n}_2}^2 + \sigma_2^2 = \sigma_{w_1}^2/|\alpha_1|^2$ we can upper bound (5) and (6) using the fact that for Gaussian inputs the worst noise distribution for given covariance is white circularly symmetric complex Gaussian:

$$
\begin{aligned}
h(\mathbf{x}_1^m + \alpha_1\mathbf{x}_2^m + \mathbf{n}_1^m) \le& -m\log\left[1 + \frac{(\sigma_{w_2}^2/|\alpha_2|^2) - \sigma_1^2}{P_1 + |\alpha_1|^2P_2 + \sigma_1^2}\right] \\
&+ h\left(\mathbf{x}_1^m + \alpha_1\mathbf{x}_2^m + \frac{\mathbf{w}_2^m}{\alpha_2}\right) \qquad (7) \\
h(\mathbf{x}_2^m + \alpha_2\mathbf{x}_1^m + \mathbf{n}_2^m) \le& -m\log\left[1 + \frac{(\sigma_{w_1}^2/|\alpha_1|^2) - \sigma_2^2}{P_2 + |\alpha_2|^2P_1 + \sigma_2^2}\right] \\
&+ h\left(\mathbf{x}_2^m + \alpha_2\mathbf{x}_1^m + \frac{\mathbf{w}_1^m}{\alpha_1}\right). \qquad (8)
\end{aligned}
$$

Using (7) and (8) in (4) we obtain

$$
\begin{aligned}
m(R_1 + R_2 - \epsilon_m) \le\ & I\left(\mathbf{x}_2^m; \mathbf{x}_1^m + \alpha_1\mathbf{x}_2^m + \frac{\mathbf{w}_2^m}{\alpha_2}\right) + h(\mathbf{y}_1^m|\mathbf{s}_1^m) \\
&- h(\mathbf{v}_1^m) - m\log\left(1 + \frac{(\sigma_{w_2}^2/|\alpha_2|^2) - \sigma_1^2}{P_1 + |\alpha_1|^2P_2 + \sigma_1^2}\right) \\
&+ I\left(\mathbf{x}_1^m; \mathbf{x}_2^m + \alpha_2\mathbf{x}_1^m + \frac{\mathbf{w}_1^m}{\alpha_1}\right) + h(\mathbf{y}_2^m|\mathbf{s}_2^m) \\
&- h(\mathbf{v}_2^m) - m\log\left(1 + \frac{(\sigma_{w_1}^2/|\alpha_1|^2) - \sigma_2^2}{P_2 + |\alpha_2|^2P_1 + \sigma_2^2}\right) \\
&- m\log(|\alpha_1\alpha_2|^2). \qquad (9)
\end{aligned}
$$

We can think of the sum

$$
T_1 \triangleq I\big(\mathbf{x}_2^m; \mathbf{x}_1^m + \alpha_1\mathbf{x}_2^m + \frac{\mathbf{w}_2^m}{\alpha_2}\big) + I\big(\mathbf{x}_1^m; \mathbf{x}_2^m + \alpha_2\mathbf{x}_1^m + \frac{\mathbf{w}_1^m}{\alpha_1}\big)
$$

as the achievable sum-rate in the interference channel

$$
\begin{aligned}
\tilde{y}_1 &= \alpha_2 x_1 + x_2 + w_1/\alpha_1 \\
\tilde{y}_2 &= \alpha_1 x_2 + x_1 + w_2/\alpha_2.
\end{aligned}
$$

In this virtual channel, giving the interfering signals as side information to the receivers we obtain the bound

$$
T_1 \le m\log\left(1 + \frac{|\alpha_1|^2P_2}{\sigma_{w_2}^2/|\alpha_2|^2}\right) + m\log\left(1 + \frac{|\alpha_2|^2P_1}{\sigma_{w_1}^2/|\alpha_1|^2}\right). \qquad (10)
$$

If the condition $\sigma_{w_1}^2 \le \sigma_{w_2}^2/|\alpha_2|^2$ holds, which corresponds to $\boldsymbol{\eta} \in \mathcal{R}_{1b}$, the first receiver in the virtual channel experiences strong interference, which implies that it can decode the signals of both users. Therefore, using a bound for the sum rate in the Gaussian multiple access channel formed by the two transmitters and virtual receiver 1, we have in this case

$$
T_1 \le m\log\left(1 + \frac{P_2 + |\alpha_2|^2P_1}{\sigma_{w_1}^2/|\alpha_1|^2}\right). \qquad (11)
$$

Similarly, if the condition $\sigma_{w_2}^2 \le \sigma_{w_1}^2/|\alpha_1|^2$ holds, which corresponds to $\boldsymbol{\eta} \in \mathcal{R}_{1c}$, we obtain

$$
T_1 \le m\log\left(1 + \frac{P_1 + |\alpha_1|^2P_2}{\sigma_{w_2}^2/|\alpha_2|^2}\right). \qquad (12)
$$

It remains to upper bound $h(\mathbf{y}_i^m|\mathbf{s}_i^m),\ i = 1, 2$. Using the chain rule for differential entropy, dropping conditioning, using the fact that for given covariance matrix jointly Gaussian random variables maximize conditional differential entropy, and Jensen's inequality, it can be shown that

$$
\begin{aligned}
h(\mathbf{y}_1^m|\mathbf{s}_1^m) \le\ & m\log\Bigg[\pi e\Bigg(P_1 + |h_{21}|^2P_2 + 1 \\
&- \frac{|P_1 + h_{21}\alpha_1^*P_2 + \rho_1^*\sigma_1|^2}{P_1 + |\alpha_1|^2P_2 + \sigma_1^2}\Bigg)\Bigg] \qquad (13)
\end{aligned}
$$

$$R_1+R_2 \leq \min\left\{\log\left[1+\frac{P(1+\alpha^2)\gamma}{(1-\rho^2)\sigma^2}\right];2\log\left(1+\frac{\alpha^2 P\gamma}{(1-\rho^2)\sigma^2}\right)\right\} + 2\log\left[\frac{(P(1+h^2)+1)(P(1+\alpha^2)+\sigma^2)-(P(1+h\alpha)+\rho\sigma)^2}{P(1+\alpha^2)\gamma+(1-\rho^2)\sigma^2}\right] \quad (14)$$

with a similar bound holding for $h(\mathbf{y}_2^m|\mathbf{s}_2^m)$.

Using each of the bounds (10), (11), (12) with (13) in (9), and letting $m\to\infty$ so that $\epsilon_m\to 0$ we obtain the bounds B_{1a}, B_{1b}, and B_{1c}.

B. Case $\boldsymbol{\eta}\in\mathcal{R}_{2a}$: $\sigma_{w_1}^2/|\alpha_1|^2\leq\sigma_2^2$ and $\sigma_{w_2}^2/|\alpha_2|^2\leq|\sigma_1|^2$

Let $\beta_1,\beta_2\in\mathbb{R}$ be such that $\beta_1^2=\sigma_1^2-\sigma_{w_2}^2/|\alpha_2|^2$, $\beta_2^2=\sigma_2^2-\sigma_{w_1}^2/|\alpha_1|^2$, and $\tilde{\mathbf{z}}^m\sim\mathcal{CN}(0,I_m)$ be independent of everything else.

We bound the following 4 terms of (4) as follows:

$$\begin{aligned}
&h(\mathbf{x}_1^m+\alpha_1\mathbf{x}_2^m+\mathbf{n}_1^m)-h\big(\mathbf{x}_1^m+\frac{\mathbf{w}_2^m}{\alpha_2}\big) + h(\mathbf{x}_2^m+\alpha_2\mathbf{x}_1^m+\mathbf{n}_2^m)-h\big(\mathbf{x}_2^m+\frac{\mathbf{w}_1^m}{\alpha_1}\big)\\
&=h\big(\mathbf{x}_1^m+\alpha_1\mathbf{x}_2^m+\beta_1\tilde{\mathbf{z}}^m+\frac{\mathbf{w}_2^m}{\alpha_2}\big)-h\big(\mathbf{x}_1^m+\frac{\mathbf{w}_2^m}{\alpha_2}\big)\\
&\quad+h\big(\mathbf{x}_2^m+\alpha_2\mathbf{x}_1^m+\beta_2\tilde{\mathbf{z}}^m+\frac{\mathbf{w}_1^m}{\alpha_1}\big)-h\left(\mathbf{x}_2^m+\frac{\mathbf{w}_1^m}{\alpha_1}\right)\\
&=I\big(\mathbf{x}_2^m,\tilde{\mathbf{z}}^m;\mathbf{x}_1^m+\alpha_1\mathbf{x}_2^m+\beta_1\tilde{\mathbf{z}}^m+\frac{\mathbf{w}_2^m}{\alpha_2}\big)\\
&\quad+I\big(\mathbf{x}_1^m,\tilde{\mathbf{z}}^m;\mathbf{x}_2^m+\alpha_2\mathbf{x}_1^m+\beta_2\tilde{\mathbf{z}}^m+\frac{\mathbf{w}_1^m}{\alpha_1}\big)\\
&=T_2+I\big(\tilde{\mathbf{z}}^m;\mathbf{x}_1^m+\alpha_1\mathbf{x}_2^m+\beta_1\tilde{\mathbf{z}}^m+\frac{\mathbf{w}_2^m}{\alpha_2}\big|\mathbf{x}_2^m\big)\\
&\quad+I\big(\tilde{\mathbf{z}}^m;\mathbf{x}_2^m+\alpha_2\mathbf{x}_1^m+\beta_2\tilde{\mathbf{z}}^m+\frac{\mathbf{w}_1^m}{\alpha_1}\big|\mathbf{x}_1^m\big)\\
&\leq T_2+I\big(\tilde{\mathbf{z}}^m;\beta_1\tilde{\mathbf{z}}^m+\frac{\mathbf{w}_2^m}{\alpha_2}\big)+I\big(\tilde{\mathbf{z}}^m;\beta_2\tilde{\mathbf{z}}^m+\frac{\mathbf{w}_1^m}{\alpha_1}\big)\\
&=T_2+m\log\big(1+\frac{\beta_1^2}{\sigma_{w_2}^2/|\alpha_2|^2}\big)+\log\big(1+\frac{\beta_2^2}{\sigma_{w_1}^2/|\alpha_1|^2}\big)\\
&=T_2+m\log\big(\frac{|\alpha_2|^2\sigma_1^2}{\sigma_{w_2}^2}\big)+\log\big(\frac{|\alpha_1|^2\sigma_2^2}{\sigma_{w_1}^2}\big)
\end{aligned}$$

where we defined $T_2\triangleq I(\mathbf{x}_2^m;\mathbf{x}_1^m+\alpha_1\mathbf{x}_2^m+\mathbf{n}_1^m)+I(\mathbf{x}_1^m;\mathbf{x}_2^m+\alpha_2\mathbf{x}_1^m+\mathbf{n}_2^m)$. As was done in subsection IV-A, we can bound T_2 by considering the virtual interference channel

$$\begin{aligned}\bar{y}_1 &= \alpha_2x_1+x_2+n_2\\ \bar{y}_2 &= \alpha_1x_2+x_1+n_1\end{aligned}$$

and bounding the achievable sum-rate in three different ways depending on the parameter ranges. In fact, the same bounds derived in subsection IV-A can be used here if we replace $\sigma_{w_2}^2/|\alpha_2|^2\to\sigma_1^2$ and $\sigma_{w_1}^2/|\alpha_1|^2\to\sigma_2^2$. The upper bounds B_{2a}, B_{2b} and B_{2c} follow by replacing these three bounds together with the bounds for $h(\mathbf{y}_i^m|\mathbf{s}_i^m)$ (cf. eq. (13)) in (4) and letting $m\to\infty$ so that $\epsilon_m\to 0$.

V. Bound for the symmetric GIFC

We can specialize Theorem 1 for the case of a symmetric GIFC where $P_1=P_2=P$ and $h_{12}=h_{21}=h$. Furthermore, we can choose the auxiliary parameters in a symmetric way: $\alpha_1=\alpha_2=\alpha$, $\sigma_1=\sigma_2=\sigma$ and $\rho_1=\rho_2=\rho$. With these restrictions on the parameters choices, we have $\sigma_{v_1}=\sigma_{v_2}=\sigma_v$ and $\sigma_{w_1}=\sigma_{w_2}=\sigma_w$. For simplicity, we only consider real parameters, i.e. $h,\alpha,\rho\in\mathbb{R}$.

When $\boldsymbol{\eta}\in\mathcal{R}_{1a}$, due to the symmetric parameter choices, the condition $\sigma_w^2/\alpha^2\geq\sigma^2$ reduces to $(1-\rho^2)\geq\alpha^2-2h\rho\sigma\alpha+h^2\sigma^2$. Furthermore, choosing $|\alpha|\leq 1$, the three bounds (10), (11), and (12) apply simultaneously. As a result, we obtain for the symmetric channel with symmetric parameter choices the bound (14) given at the top of the page, where $\gamma=\alpha^2-2h\rho\sigma\alpha+h^2\sigma^2$, valid for $(1-\rho^2)\geq\alpha^2-2h\rho\sigma\alpha+h^2\sigma^2$ and $|\alpha|\leq 1$.

Similarly, when $\boldsymbol{\eta}\in\mathcal{R}_{2a}$, we obtain:

$$R_1+R_2\leq\min\left\{\log\left[1+\frac{P(1+\alpha^2)}{\sigma^2}\right];2\log\left(1+\frac{\alpha^2P}{\sigma^2}\right)\right\} + 2\log\left[P(1+h^2)+1-\frac{(P(1+h\alpha)+\rho\sigma)^2}{P(1+\alpha^2)+\sigma^2}\right] - 2\log(1-\rho^2) \quad (15)$$

valid for $(1-\rho^2)\leq\alpha^2-2h\rho\sigma\alpha+h^2\sigma^2$ and $|\alpha|\leq 1$.

In order to evaluate the performance of the new bound we consider a symmetric GIFC, with power constraint $P=6$, and vary the cross-gain h between 0 and 1. Figure 1 shows different bounds on the symmetric rate $R=(R_1+R_2)/2$. Since evaluating the general lower bound of [2] is prohibitively complex, in Fig. 1 we use a simple lower bound obtained by considering two achievable strategies:

- treating interference as noise: $R=\log\big(1+\frac{P}{1+h^2P}\big)$,
- time-sharing: $R=\frac{1}{2}\log(1+P)$,

and choosing the best of the two. (See [3] for more sophisticated lower bounds). The curves for the new upper bound are obtained by minimizing (14) and (15) within the valid parameter ranges. For reference, we also plot the upper bound of [8, Theorem 2] and the new bounds[3] of [5], [6], [7].

We see in Fig. 1 that the new upper bound significantly improves over the bounds of [5], [6], [7] and [8] for small values of h. For h large enough the bound of [8] eventually becomes better.

When minimizing the bounds (14) and (15) for various choices of P and h we observed that the optimal parameters satisfy $\sigma_w/\alpha=\sigma$. In this case, the bounding steps of

[3] In [5], [6], [7] the bound of [8, Theorem 2] is included to further improve the new bounds proposed therein.

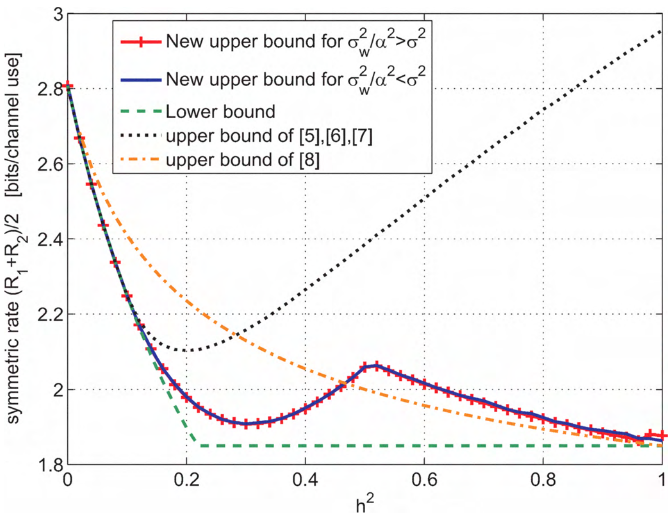

Fig. 1. Upper and lower bounds on the symmetric rate of a symmetric complex GIFC as a function of the channel (power) cross gain h^2, for $P = 6$. The new upper bound is obtained by optimizing (14) and (15) within the valid parameter ranges. Also included are the upper bounds of [5], [6], [7] (which coincide) and the upper bound of [8, Theorem 2].

subsections IV-A and IV-B simultaneously apply and result in the same bound. Therefore, for $\sigma_w/\alpha = \sigma$, (14) and (15) coincide, as seen in Fig. 1.

For $|\alpha| \leq 1$, assuming real parameters, $\sigma_w/\alpha = \sigma$ requires:

$$\sigma^2 \leq \frac{1}{h^2}\ ; \quad \rho = h\alpha\sigma \pm \sqrt{(1-\alpha^2)(1-h^2\sigma^2)}. \qquad (16)$$

It is easy to check that the two possible values of ρ resulting from (16) are in $[-1, 1]$, and therefore are valid. As a result, both values of ρ need to be checked when optimizing the parameters.

VI. Sum-Capacity of GIFC with Very Weak Interference

In this section we will restrict attention to the symmetric Gaussian interference channel with real gains, and consider only symmetric parameter choices (i.e. $\alpha_1 = \alpha_2 = \alpha$, $\sigma_1 = \sigma_2 = \sigma$, $\rho_1 = \rho_2 = \rho$) for our bound.

Let $h_0(P)$ be the positive solution to $|h_0 + h_0^3 P| = 0.5$. It was shown in [5], [6], [7] that if $|h| \leq h_0(P)$, treating interference as noise is sum-rate optimal. The condition $|h| \leq h_0$ is sufficient but may not be necessary. Does our new bound extend the range of values of h such that treating interference as noise can be shown to be sum-rate optimal?

To have an upper bound that matches the lower bound obtained by treating interference as noise with Gaussian signals we need: (a) that i.i.d. Gaussian inputs achieve the capacity of the genie-aided channel, *and* (b) that for i.i.d. Gaussian inputs the side information provided by the genies does not provide gains in rate.

Assuming that conditions (a) and (b) hold, we will sketch an argument that shows that our bound can only be tight if $\alpha = 0$. To get a contradiction, assume that (a) and (b) hold for some $|\alpha| > 0$.

First consider the case $\boldsymbol{\eta} \in \mathcal{R}_{1a}$. For condition (a) we need equality in either (10), (11) or (12) when the inputs are i.i.d. circularly symmetric complex Gaussian.

With symmetric parameters, (11) and (12) are equivalent, and we need that

$$2\log\left[1 + \frac{|\alpha|^2 P}{\sigma_w^2/|\alpha|^2 + P}\right] = \log\left[1 + \frac{P(1+|\alpha|^2)}{\sigma_w^2/|\alpha|^2}\right].$$

After some algebraic manipulations, it can be shown that this equality can only occur when $|\alpha| > 1$. But the bounds (11) and (12) are only valid when $|\alpha| \leq 1$.

To have equality in (10) we need

$$2\log\left[1 + \frac{|\alpha|^2 P}{(\sigma_w^2/|\alpha|^2) + P}\right] = 2\log\left[1 + \frac{|\alpha|^2 P}{(\sigma_w^2/|\alpha|^2)}\right]. \quad (17)$$

Since we assumed $|\alpha| > 0$, as long as $|h| > 0$, we have:

$$\frac{\sigma_w^2}{\alpha^2} = \frac{(1-\rho^2)\sigma^2}{(\alpha - h\sigma)^2 + 2h\sigma\alpha(1-\rho)} \leq \frac{(1+\rho)\sigma}{2h\alpha} < \infty$$

implying that (17) cannot occur. It follows that when $\boldsymbol{\eta} \in \mathcal{R}_{1a}$, condition (a) can only hold for $\alpha = 0$.

Now we consider the case $\boldsymbol{\eta} \in \mathcal{R}_{2a}$. Using a similar argument as was used for the case $\boldsymbol{\eta} \in \mathcal{R}_{1a}$, we can show that the bounding steps leading to the first argument of the $\min\{\cdot,\cdot\}$ in (15) can only be tight with Gaussian inputs if $\alpha = 0$. The bounding steps leading to the second argument of the $\min\{\cdot,\cdot\}$ in (15) can only be tight with Gaussian inputs if either $\alpha = 0$ or $\sigma \to \infty$. In the case $\sigma \to \infty$ we can obtain lower bounds for (15) which are strictly greater than the sum-rate obtained by treating interference as noise, leaving $\alpha = 0$ as the only option to obtain a tight bound with Gaussian inputs.

When $\alpha = 0$, the side information signals s_1 and s_2 reduce to the ones used in [5], [6], [7] and our bound does not allow to obtain any weaker condition than $|h| \leq h_0(P)$ for the sum-rate optimality of treating interference as noise.

References

[1] A. B. Carleial, "A case where interference does not reduce capacity," (Corresp.), *IEEE Trans. Inf. Theory*, vol. IT-21, pp. 569–570, Sep. 1975.

[2] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49-60, Jan. 1981.

[3] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1345–1356, Jun. 2004.

[4] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534-5562, Dec. 2008.

[5] X. Shang, G. Kramer, and B. Chen, "A new outer bound and noisy-interference sum-rate capacity for Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 689-699, Feb. 2009.

[6] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: sum capacity in the low interference regime and new outer bounds on the capacity region," submitted to *IEEE Trans. Inf. Theory*, Feb. 2008.

[7] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 620-643, Feb. 2009.

[8] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 50, pp. 581-586, Mar. 2004.