Sum Rates of Random Beamforming MISO Downlink Systems with Other Cell Interference

Sung-Hyun Moon, Student Member, IEEE, Sang-Rim Lee, Student Member, IEEE, and Inkyu Lee, Senior Member, IEEE
School of Electrical Eng., Korea University, Seoul, Korea
Email: {shmoon, sangrim78, inkyu}@korea.ac.kr

Abstract—Random beamforming (RBF) is a simple yet effective technique for multiuser multiple-input multiple-output systems with limited feedback. In this paper, we analyze the performance of the RBF in the presence of other cell interference (OCI), where a base station with $M$ antennas supports $M_s$ users ($M_s \leq M$) selected by scheduling among $K$ single antenna users. Employing extreme value theory, we derive a closed-form approximation on the asymptotic ergodic sum rate by examining the limiting distribution of the largest signal-to-interference-plus-noise ratio (SINR) among $K$ independent users. Also, we prove that even if the OCI exists, we have the same sum rate scaling law of $M_s \log_2 \log_2 K$ as the system without OCI. Simulation results verify that our analysis provides an accurate estimation for the average sum rate performance even when $K$ is not so large.

I. INTRODUCTION

Until recently, significant research effort has been devoted to multiple-input multiple-output (MIMO) wireless communication systems, which can attain substantial capacity gains. In particular, MIMO broadcast channels have been widely studied due to their role in emerging cellular standards. In order to improve the sum rate, a variety of precoding schemes have been proposed, including optimal dirty paper coding (DPC) [1] and simple linear beamforming techniques [2]–[4].

Although theoretically attractive, the actual performance promised by the MIMO techniques can be severely degraded in a realistic cellular system. Especially, users located in the cell edge area suffer from poor channel conditions due to severe signal attenuation and the effect of other-cell interference (OCI). As one way of reducing the OCI, cooperative processing among base stations (BSs), called network MIMO, has been attracting a lot of attention [5]. However, the network MIMO leads to an increase of the backhaul capacity and a synchronization issue. Alternatively, the OCI aware multiuser precoding schemes have been proposed [6] [7], which utilize the exact estimation of the OCI plus noise covariance.

Another fundamental limit in the real-world system design is the feedback overhead. Currently, several related works addressing limited feedback strategies are in progress [8]. Random beamforming (RBF) is one of effective techniques for MIMO downlink channels which takes advantage of its simple structure and small feedback load [9]. It was shown in [9] that the RBF achieves the same optimal sum rate growth as the DPC. Also in practice, a research in [10] indicates that a simple RBF approach, known as per user unitary and rate control (PU2RC), can outperform zero-forcing beamforming when the number of users is large. Recently, there have been efforts to evaluate the performance of the RBF [11]–[13]. However, the effect of the OCI was not taken into account in the prior works.

In this paper, we analyze the sum rate performance of the RBF in the presence of the OCI when the BS with $M$ antennas supports $M_s$ users ($M_s \leq M$) who are selected from scheduling among $K$ single antenna users. Deriving an exact sum rate is difficult, since the distribution of the signal-to-interference-plus-noise ratio (SINR) is quite complicated to deal with. Thus, we seek a tight closed-form approximation on the asymptotic ergodic sum rate in the limit of large $K$ by employing extreme value theory [14]. Furthermore, based on our derivation, we show that the sum rate scales like $M_s \log_2 \log_2 K$ as $K$ goes to infinity, which is the same growth rate as the conventional result in [9]. From numerical simulations, we verify the validity of our approximation for a wide range of the number of users.

Throughout this paper, we use the following notations. Normal letters represent scalar quantities, bold face letters indicate vectors, and boldface uppercase letters designate matrices. The superscripts $(\cdot)^T$ and $(\cdot)^H$ stand for the transpose and Hermitian transpose, respectively. The expectation of a random variable is given by $\mathbb{E}(\cdot)$.

II. SYSTEM DESCRIPTION

A. System model

Consider a multiuser multiple-input single-output (MISO) downlink channel where a BS communicates to $K$ mobile users. The BS is equipped with $M$ transmit antennas and each user has a single antenna. Among $K$ users, $M_s$ users ($M_s \leq M$) are selected via multiuser scheduling in each transmission. We assume that there exist $L$ cochannel interferers for each user from neighboring cells. First, we define the precoded signal vector $s \in \mathbb{C}^{M \times 1}$ as $s = \sum_{j=1}^{M_s} \mathbf{w}_j u_j$ where $u_j$ and $\mathbf{w}_j \in \mathbb{C}^{M \times 1}$ denote the transmitted data symbol and the beamforming vector, respectively. In the same way, the $l$-th OCI signal vector for user $k$ is given by $\tilde{\mathbf{s}}_{k,l} = \sum_{j=1}^{M_s} \mathbf{w}_{k,l,j} \tilde{u}_j$ where $\tilde{u}_j$ and $\mathbf{w}_{k,l,j} \in \mathbb{C}^{M \times 1}$ are the data symbol and the beamformer transmitted at the $l$-th interfering cell, respectively. Throughout this paper, we use the bar notation to represent the terms related to the OCI. We assume that each BS follows the sum power constraint $P$ such that $\text{Tr}(\mathbb{E}[ss^H]) \leq P$ and $\text{Tr}(\mathbb{E}[\tilde{\mathbf{s}}_{k,l}\tilde{\mathbf{s}}_{k,l}^H]) \leq P$ for all $k$ and $l$.

Then, the received signal $y_k$ of user $k$ is written as

$$y_k = \sqrt{\alpha} \mathbf{h}_k^H s + \sum_{l=1}^{L} \sqrt{\beta_l} \tilde{\mathbf{h}}_{k,l}^H \tilde{\mathbf{s}}_{k,l} + n_k$$

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2010-0017099).
where \( \mathbf{h}_k \in \mathbb{C}^{M \times 1} \) and \( \mathbf{H}_{k,l} \in \mathbb{C}^{M \times 1} \) indicate the desired channel vector and the \( l \)-th OCI channel vector for user \( k \), respectively, whose entries are independent and identically distributed (i.i.d.) complex Gaussian \( \mathcal{CN}(0, 1) \). \( \alpha \) and \( \pi_l \) denote the signal attenuation from the serving BS and the \( l \)-th neighboring BS, and \( \eta_k \) stands for the additive white Gaussian noise (AWGN) with \( \mathcal{CN}(0, 1) \). We assume homogeneous OCI signal intensity between users so that the set \( \{ \eta_1, \cdots, \eta_L \} \) is identical for all users. It is also assumed that the local channel state information (CSI) \( \sqrt{\eta}_k \) and \( \{ \sqrt{\eta}[k,1] \}_{l=1}^L \) are perfectly known at the \( k \)-th user. At every BS, uniform power \( P/M \) is allocated to all data streams.

**B. Review of RBF technique**

The RBF utilizes \( M \) orthonormal vectors \( \Phi_1, \cdots, \Phi_M \in \mathbb{C}^{M \times 1} \) so that \( \mathbf{w} \) is set to \( \mathbf{w}_m = \phi_m, 1 \leq m \leq M \), which is generated independently at each cell in a pseudo-random fashion. After every \( T \) channel uses, we choose another independent set \( \Phi_1, \cdots, \Phi_M \) to obtain multiuser diversity [15], where \( T \) is determined according to the channel coherence time. The beam vectors for the \( l \)-th neighboring cell of user \( k \) are denoted by \( \mathbf{H}_{k,l,m} \in \mathbb{C}^{M \times 1} \) for \( 1 \leq m \leq M \).

At the receiver, each user computes its signal-to-interference-plus-noise ratio (SINR) values for each of \( M \) beams, and feeds back the highest SINR along with its index \( m \). Considering both the intra- and inter-cell interference signals from the model (1), the \( k \)-th user calculates its SINR for the \( m \)-th beam vector \( \phi_m \) as

\[
\text{SINR}_{k,m} = \frac{a|h_k^H \phi_m|^2}{M_s/P + a \sum_{j \neq m} |h_k^H \phi_j|^2 + \sum_{l=1}^L \pi_l \sum_{j=1}^L |h_{k,l,j}^H \phi_{k,l,j}|^2}.
\]

Then, based on the SINR feedback (2) from every user, the BS performs scheduling by assigning its \( m \)-th data stream to user \( \hat{k} \) who reports the maximum SINR such as \( \hat{k} = \arg \max_{k=1, \cdots, K} \text{SINR}_{k,m} \).

Under the above scheduling policy at the BS, the achievable sum rate \( R_{\text{sum}} \) can be written as

\[
R_{\text{sum}} = \sum_{m=1}^M R_m \approx \sum_{m=1}^M \mathbb{E} \left[ \log_2 \left( 1 + \max_{k=1, \cdots, K} \text{SINR}_{k,m} \right) \right]
\]

\[
= M_s \mathbb{E} \left[ \max_{k=1, \cdots, K} \log_2 \left( 1 + \text{SINR}_{k,m} \right) \right]
\]

where \( R_m \) is defined as the rate for the \( m \)-th data stream. Here, the approximation is due to a small probability that the SINR of a certain user may be the highest for more than one data stream [9]. However, since this probability is negligible if \( K \) is not very small, we can consider (3) as the exact expression for \( R_{\text{sum}} \).

**III. ASYMPTOTIC SUM RATE ANALYSIS**

In this section, we evaluate the sum rate of the RBF system. Since it is hard to obtain an exact solution, we first seek an approximate distribution for the received SINR. After that, by applying extreme value theory with the large \( K \) assumption, we derive an asymptotic expression for the sum rate in (4).

**A. Approximate SINR distribution**

As the first step, we start with finding the distribution of \( \text{SINR}_{k,m} \) which quantifies \( R_m \). For convenience, (2) can be rewritten as

\[
\text{SINR}_{k,m} = \frac{aX}{M_s/P + aW + \sum_{l=1}^L \pi_l W_l}
\]

where \( X = |h_k^H \phi_m|^2 \), \( W = \sum_{j \neq m} |h_k^H \phi_j|^2 \) and \( \pi_l = \sum_{l=1}^M |h_{k,l,j}^H \phi_{k,l,j}|^2 \). Note that \( |h_k^H \phi_m|^2 \) is i.i.d. over both \( k \) and \( m \) with \( \chi^2(2) \) distribution since \( \phi_1, \cdots, \phi_M \) are orthonormal [9]. Hence, \( X, W \) and \( \pi_l \) follow \( \chi^2(2) \) and \( \chi^2(2M_s - 2) \) and \( \chi^2(2M_s) \) distribution, respectively. We can also check that \( \text{SINR}_{k,1}, \cdots, \text{SINR}_{k,M} \) are identically distributed (obviously not independent).

In order to treat the interference term, we introduce \( V \) as

\[
V \triangleq aW + \sum_{l=1}^L \pi_l W_l.
\]

Since \( W \) and all \( \pi_l \)'s are independent, we find that \( V \) is a weighted sum of independent Chi-square random variables, whose probability density function (PDF) is complicated to compute. Instead of finding the exact distribution of (6), we make use of an approximation provided in [16], which is shown to be very close to the actual PDF. The argument in [16] is that the PDF of \( \sum_{k=1}^n d_k N_i^2 \) for real positive weights \( d_k \) and i.i.d. standard normal random variables \( N_i \) is well approximated by the gamma distribution as

\[
f(v; \alpha, \beta) = v^{\alpha-1} e^{-v/\beta} / \beta^\alpha \Gamma(\alpha)
\]

where \( \Gamma(\alpha) \) is the gamma function \( \Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt \) and the parameters \( \alpha \) and \( \beta \) follow the relation

\[
\alpha = \frac{(\sum_{k=1}^n d_k)^2}{2 \sum_{k=1}^n d_k^2}, \quad \beta = 2 \sum_{k=1}^n d_k^2 / \sum_{k=1}^n d_k.
\]

By appropriately applying this approach to our case in (6), the PDF of \( V \) is given by (7) with parameters

\[
\alpha = \frac{\Phi_1^2}{\Phi_2}, \quad \beta = \frac{\Phi_2}{\Phi_1},
\]

where \( \Phi_1 = (M - 1)\alpha + M \sum_{i=1}^L \pi_i \) and \( \Phi_2 = (M - 1)\alpha^2 + M \sum_{i=1}^L \pi_i^2 \). In Figure 1, we compare the CDF of the approximated gamma distribution of (7) with the actual CDF of \( V \) obtained from numerical simulations with \( M = 4 \), where 100,000 Monte Carlo runs were carried out. Simply, the weight coefficients \( \alpha, \pi_1, \cdots, \pi_L \) are uniformly generated between \(-10\) dB and \( 0 \) dB and are fixed during the simulation. From this plot, we verify that the approximation on the distribution of \( V \) is highly accurate for various configurations.

Now, using the gamma PDF (7) and (8), we can calculate the approximate PDF \( f_S(x) \) of \( \text{SINR}_{k,m} \) for all \( m \) as

\[
f_S(x) = \int_0^\infty f_{S|V}(x|v) f_V(v; \alpha, \beta) dv
\]

\[
\approx \int_0^\infty \frac{M_s/P + v}{a} e^{-(M_s/P + v)x/a} v^{\alpha-1} e^{-v/\beta} / \beta^\alpha \Gamma(\alpha) dv
\]

\[
= \frac{\beta e^{M_s x / \beta} + \alpha (\beta e^{M_s x / \beta} + \alpha) (\beta e^{M_s x / \beta} + \alpha)^{-\alpha-1}}{\alpha}
\]
using integration by parts, the exact sum rate expression is

Also by substituting (10) into (11), we have an approximation

B. Asymptotic sum rate approximation

Since users are homogeneous, the parameters $\alpha$, $\beta$ and $\sigma$ in

are i.i.d. over all $k = 1, \ldots, K$, which gives the PDF of its maximum

If $s = K$, which gives the PDF of its maximum

and its sample maximum as

The corresponding cumulative distribution function (CDF)

is a direct consequence of the LTD Theorem.

Unfortunately, it appears that the integration

is not solvable. In [18], the authors analyzed this problem under the

SNR assumption when the OCI does not exist. However, it is not easy to directly obtain $F_{\Omega_k}^*(x)$ since the PDF of $\Omega_k$ has a complicated form. Therefore, before considering the property of $\Omega_k$, we first investigate the limiting behavior of $\text{SINR}_{k,m}$ based on (9) and (10). We state the following two lemmas.

Lemma 1: Denote $\text{SINR}_{k,m} = \max_{k=1,\ldots,K} \text{SINR}_{k,m}$. Then, as $K \to \infty$, $(\text{SINR}_{k,m} - \mu_k)/\sigma_k$ converges to a standard Gumbel random variable whose distribution function is $\exp(-e^{-x})$ for $-\infty < x < \infty$. The location and scale parameters $\mu_k$ and $\sigma_k$ can be selected as

where $F_S^{-1}(x) = \inf\{y : x \leq F_S(y)\}$ represents the quantile function of the distribution of $\text{SINR}_{k,m}$.

Proof: See Appendix A.

Lemma 2: For the parent distribution $F_S(x)$ given in (10), its quantile function $F_S^{-1}(x)$ can be calculated as

where $W(x)$ is the Lambert W function [19].

Proof: See Appendix B.

Lemma 1 and 2 show that as $K \to \infty$, the CDF of $\text{SINR}_{k,m}$ becomes $F_{\Omega_k}^*(x) = \exp(-e^{-(x-\mu_k)/\sigma_k})$, and its mean $\mu_k$ and variance $\sigma_k^2$ are determined by substituting (16) into (14) and (15). For computing the Lambert W function, a couple of simple Newton’s iterations are enough to approach an exact solution [19].

Next, we identify the asymptotic distribution of $\Omega_k$. In this step, we utilize a theorem provided in [20], named as limiting throughput distribution (LTD) Theorem, which explains the limiting behavior of $\Omega_k$ without checking the condition (22) in Appendix A for $\Omega_k$. The following lemma is a direct consequence of the LTD Theorem.

Lemma 3: The distribution of $\Omega_k$ belongs to the domain of the attraction of the Gumbel distribution. The normalizing constants $\mu'_k$ and $\sigma'_k$ are given by

where $\mu'_k = M_s \log_2(1 + \mu_k)$ and $\sigma'_k = M_s \log_2(1 + (\mu_k + \sigma_k)) - \mu'_k$.

Proof: See [20].

This statement leads to our main result. Lemma 3 implies that $(\Omega_k - \mu'_k)/\sigma'_k$ also follows the standard Gumbel distribution in the asymptotic regime. Noting that the standard Gumbel distribution has a mean $\gamma_0 = 0.5772 \cdots$ (Euler’s constant), we have $\mathbb{E}[\Omega_k - \mu'_k]/\sigma'_k = \gamma_0$. Then, from (13), we finally arrive at a closed-form approximation on the ergodic sum rate $R_{\text{sum}}$ as

max_{k=1,\ldots,K} \Omega_k$. Then, the rate expression (4) can be simply rewritten as

Now, clearly, our goal is to find the limiting distribution of $\Omega_k$. However, it is not easy to directly obtain $F_{\Omega_k}^*(x)$ since the PDF of $\Omega_k$ has a complicated form. Therefore, before considering the property of $\Omega_k$, we first investigate the limiting behavior of $\text{SINR}_{k,m}$ based on (9) and (10). We state the following two lemmas.

Lemma 1: Denote $\text{SINR}_{k,m} = \max_{k=1,\ldots,K} \text{SINR}_{k,m}$. Then, as $K \to \infty$, $(\text{SINR}_{k,m} - \mu_k)/\sigma_k$ converges to a standard Gumbel random variable whose distribution function is $\exp(-e^{-x})$ for $-\infty < x < \infty$. The location and scale parameters $\mu_k$ and $\sigma_k$ can be selected as

where $F_S^{-1}(x) = \inf\{y : x \leq F_S(y)\}$ represents the quantile function of the distribution of $\text{SINR}_{k,m}$.

Proof: See Appendix A.

Lemma 2: For the parent distribution $F_S(x)$ given in (10), its quantile function $F_S^{-1}(x)$ can be calculated as

where $W(x)$ is the Lambert W function [19].

Proof: See Appendix B.

Lemma 1 and 2 show that as $K \to \infty$, the CDF of $\text{SINR}_{k,m}$ becomes $F_{\Omega_k}^*(x) = \exp(-e^{-(x-\mu_k)/\sigma_k})$, and its mean $\mu_k$ and variance $\sigma_k^2$ are determined by substituting (16) into (14) and (15). For computing the Lambert W function, a couple of simple Newton’s iterations are enough to approach an exact solution [19].

Next, we identify the asymptotic distribution of $\Omega_k$. In this step, we utilize a theorem provided in [20], named as limiting throughput distribution (LTD) Theorem, which explains the limiting behavior of $\Omega_k$ without checking the condition (22) in Appendix A for $\Omega_k$. The following lemma is a direct consequence of the LTD Theorem.

Lemma 3: The distribution of $\Omega_k$ belongs to the domain of the attraction of the Gumbel distribution. The normalizing constants $\mu'_k$ and $\sigma'_k$ are given by

where $\mu'_k = M_s \log_2(1 + \mu_k)$ and $\sigma'_k = M_s \log_2(1 + (\mu_k + \sigma_k)) - \mu'_k$.

Proof: See [20].

This statement leads to our main result. Lemma 3 implies that $(\Omega_k - \mu'_k)/\sigma'_k$ also follows the standard Gumbel distribution in the asymptotic regime. Noting that the standard Gumbel distribution has a mean $\gamma_0 = 0.5772 \cdots$ (Euler’s constant), we have $\mathbb{E}[\Omega_k - \mu'_k]/\sigma'_k = \gamma_0$. Then, from (13), we finally arrive at a closed-form approximation on the ergodic sum rate $R_{\text{sum}}$ as

$$R_{\text{sum}} = \mathbb{E}[\Omega_k] \approx \mu'_k + \gamma_0 \sigma'_k$$

(17)
where $\mu'_K$ and $\sigma'_K$ are obtained from Lemma 3 and (16) as

$$
\begin{align*}
\mu'_K &= M_s \log_2 \left( 1 + \frac{P_{\alpha}}{M_s} W \left( \frac{M_s}{P_{\alpha} \beta} K^{\frac{1}{2}} e^{\frac{\sigma^2}{2}} \right) - \frac{a}{\beta} \right) \\
\sigma'_K &= M_s \log_2 \left( 1 + \frac{P_{\alpha}}{M_s} W \left( \frac{M_s}{P_{\alpha} \beta} (K e)^{\frac{1}{2}} e^{\frac{\sigma^2}{2}} \right) - \frac{a}{\beta} \right) - \mu_K.
\end{align*}
$$

(18) and (19)

The approximation follows from the gamma approximation in (7) and the large $K$ assumption. Nevertheless, we note that (17) is also accurate for small number of users, which will be verified in the simulation section. For the convergence proof of (17), one may refer to [20, Lemma 2].

To further simplify (18) and (19), we can employ the approximation for $W(x)$ in [11] as

$$
W(x) \approx c_1 \log_2(x + c_2) + c_3
$$

(20)

where $c_1$, $c_2$, and $c_3$ are fixed coefficients. This approximation is tight for $x > 0$, which is always met in our case. We determine the constants as $c_1 = 0.4264$, $c_2 = 0.6083$ and $c_3 = 0.2547$ using a usual curve-fitting method. By adopting the approximation (20), we can reach the following conclusion.

**Theorem 1:** For fixed $M_s$ and $P$, $R_{\text{sum}}$ obeys the asymptotic growth rate as

$$
\lim_{K \rightarrow \infty} \frac{R_{\text{sum}}}{M_s \log_2 \log_2 K} \approx 1.
$$

(21)

**Proof:** See Appendix C.

The result of Theorem 1 guarantees that even when the OCI exists, the optimal sum rate scaling behavior of $M_s \log_2 \log_2 K$ is preserved. Note that this property is true regardless of the strength of the total OCI signal. This is because that although the users’ channel condition may become poor on average, the randomness existing in the small scale fading still can offer multiuser diversity.

**IV. SIMULATION RESULTS**

In this section, we verify our sum rate analysis with numerical simulations for MISO downlink channels employing the RBF technique. For all simulations, we use spatially uncorrelated channels which are randomly and independently generated for each transmission. As explained before, each user feeds back the maximum SINR value and $\lfloor \log_2 M_s \rfloor$ bits for its index. As for the attenuation parameters, $a$ is set to be 0 dB and $\{\bar{\alpha}_1, \cdots, \bar{\alpha}_L\}$ are uniformly generated between $-10$ dB and 0 dB and fixed during the simulation. This setting roughly reflects the cell edge environment where the OCI is very strong such that the users’ signal-to-interference ratio (SIR) might be considerably lower than 0 dB [21].

Figure 2 shows the average sum rate performance of the RBF with respect to $K$ for $M = M_s = 4$ and $P = 10$ dB. Both the simulations and the analytical results based on (17) are plotted together in a wide range of $K \in [5, 100]$. We observe that the effect of the OCI severely degrades the performance. By comparing the simulation and the analysis, we emphasize that our analysis is well fitted with the empirical curves for various $L$ and $K$, even with small numbers of $K$.

In Figure 3, we present the average sum rate curves with respect to $P$. This figure shows that our asymptotic approximation is tight over the whole range of signal-to-noise ratios (SNRs). From the two figures, we confirm that our result provides an accurate approximation for the sum rate of the RBF systems with arbitrary level of the OCI.

**V. CONCLUSIONS**

In this paper, we have analyzed the sum rate of the RBF technique for MISO downlink channels in the presence of the OCI. By utilizing extreme value theory, we have derived a closed-form approximation on the ergodic sum rate when the number of users is asymptotically large. In addition, we have shown that even when the OCI exists, the sum rate scales like $M_s \log_2 \log_2 K$ as the conventional no OCI case. From numerical simulations, we have confirmed that our asymptotic analysis is valid for practical system configurations.
A. Proof of Lemma 1

We need to prove that the sample maximum of $\text{SINR}_{k,m}, \ldots, \text{SINR}_{K,m}$ follows the Type I extreme value distribution (Gumbel distribution). The condition for attraction to the Gumbel distribution is given by [14]

$$
\lim_{K \to \infty} \frac{d}{dx} \left[ \frac{1 - F_{S}(x)}{f_{S}(x)} \right] = 0.
$$

(22)

From (9) and (10), the distribution of $\text{SINR}_{k,m}$ leads to

$$
\frac{d}{dx} \left[ \frac{1 - F_{S}(x)}{f_{S}(x)} \right] = \frac{d}{dx} \left[ \frac{x + \alpha}{\beta} \right] = \frac{d}{dx} \left[ \frac{P_{a}}{M} + O(x^{-1}) \right] = O(x^{-2}).
$$

(23)

Thus, (22) is satisfied and then $(\text{SINR}_{(K)} - \mu_{K})/\sigma_{K}$ converges to a standard Gumbel random variable.

B. Proof of Lemma 2

By substituting $z = \beta x/a$ into (10), we have

$$
y = 1 - e^{-\frac{M_{s}}{\beta_{s}}(z + 1)^{a}}.
$$

(24)

Similar to the approach in [11], we can reformulate the above equation as the form of $v = w e^{w}$ with

$$
v = \frac{M_{s}}{P_{a} \alpha} \exp \left( \frac{M_{s}}{P_{a} \alpha} - \frac{1}{\alpha} \log(1 - y) \right) = \frac{M_{s}}{P_{a} \alpha} (z + 1) \exp \left( \frac{M_{s}}{P_{a} \alpha} (z + 1) \right)
$$

and $w = \frac{M_{s}}{P_{a} \alpha} (z + 1)$. Noting that the solution of $v = w e^{w}$ is $w = W(v)$, which is the definition of the Lambert W function, it follows

$$
x = \frac{\alpha}{\beta} z = \frac{\alpha}{\beta} \left( \frac{P_{a} \alpha}{M_{s}} W(v(y)) - 1 \right) = F_{S}^{-1}(y).
$$

C. Proof of Theorem 1

Due to space limitation, we only give a simple and intuitive proof here. By inserting (20) into (18) and neglecting $c_{2}$ since $c_{2} \ll K$, the limiting behavior of $\mu'_{K}$ is given by

$$
\lim_{K \to \infty} \mu'_{K} = \lim_{K \to \infty} M_{s} \log_{2} \left( \frac{P_{a} \alpha c_{1}}{M_{s}} \log_{2} \left( \frac{M_{s}}{P_{a} \alpha} K^{\frac{1}{\alpha}} e^{\frac{M_{s}}{P_{a} \alpha}} \right) + O(1) \right) = \lim_{K \to \infty} M_{s} \log_{2} \log_{2} K + O(1)
$$

(25)

(26)

Next, the scale parameter $\sigma'_{K}$ can be expressed when $K \to \infty$ as

$$
\lim_{K \to \infty} \sigma'_{K} = \lim_{K \to \infty} M_{s} \log_{2} \left( \frac{P_{a} \alpha c_{1}}{M_{s}} (\log_{2} K + \log_{2} e) + O(1) \right) = 0.
$$

As a result, by substituting (26) and $\lim_{K \to \infty} \sigma'_{K} = 0$ into (17), we have (21).