Cross-system resource allocation based on random matrix theory

Samson Lasaulce
Lab. des Signaux et Systemes
CNRS - Supelec - Paris 11
91190, Gif-sur-Yvette, France
lasaulce@lss.supelec.fr

Alberto Suárez
Dpt. Communications Mobiles
Institut Eurecom
06904, Sophia Antipolis,
France
suarezr@eurecom.fr

Raul de Lacerda
Dpt. Communications Mobiles
Institut Eurecom
06904, Sophia Antipolis,
France
raul.de-lacerda@eurecom.fr

Merouane Debbah
Supelec
91190, Gif-sur-Yvette, France
merouane.debbah@supelec.fr

ABSTRACT
This paper investigates the situation where a (large) group of terminals can be connected simultaneously to several base stations using distinct wireless technologies. We introduce and solve the problem of optimally sharing the mobile transmit power between different systems. Key results from asymptotic random matrix theory (when the number of users and the dimensions of different systems increase) allow us to derive the best power allocation scheme in the sense of the sum-capacity of the overall system, for which the uplink is equivalent to a parallel fading multiple access channel. Moreover, we provide an iterative algorithm to solve the power allocation algorithm. Simulations for a finite number of users validate the asymptotic claims.

Categories and Subject Descriptors
H.4 [Information Systems Applications]: Miscellaneous;
D.2.8 [Software Engineering]: Metrics—complexity measures, performance measures

General Terms
Random matrix theory, Information theory

Keywords
Cross system optimization, multiple access channel, power allocation, iterative water-filling, heterogeneous networks.

1. INTRODUCTION
As the number of wireless systems has increased over the last decades, the idea of system convergence has been introduced (see e.g. [1][2]) in order to enable mobile terminals to operate with different standards. This convergence objective was one of the driving forces towards the design of reconfigurable terminals also known as Software Defined Radio or Flexible Radios [3]. Mobile phones currently available on the market are multi-mode, which means that they can work with different standards. In France for example, mobile phones typically implement both the GSM, GPRS and/or UMTS-FDD standards and even the 802.11a/b/g standards via the unlicensed mobile access (UMA) technology (see e.g. [4]). In fact, many other situations exist where a terminal can have access to several signals that are not in the same frequency band. For instance a GSM receiver is able to hear several GSM base stations, an UMTS terminal can hear WCDMA base stations but also possibly TD-CDMA base stations, a mobile terminal using the DVB-H standard can operate in a 3G standard or in the DVB standard, etc. For all these examples the terminal operates with only one standard at a time, depending on the user location and the type of service (Internet, TV, voice,...) asked by the user.

Although the present work is clearly based on an information-theoretic approach, it still aims at gaining more understanding on what could be done in the aforementioned situations to optimize the overall throughput by using all the systems simultaneously and not sequentially as it is the case in existing systems. More specifically, we consider several mobile users and different base stations. Each base station has a given frequency band with a given technology. The base stations are assumed to be connected through perfect communication links. In UMTS networks for example, the base stations are connected through a radio network controller and very reliable wirelines (e.g. optic fiber). In this context, we want to provide insights to the two issues mentioned be-
1. Assuming users having wireless links towards different base stations equipped with different technologies, we derive the optimal power and rate, splitting them over each link, with a fixed power constraint for each user. The corresponding power allocation scheme is optimized in order to maximize the sum-capacity (over the users and systems) of the overall network.

2. In a second step, we provide an iterative algorithm to solve the corresponding power allocation water-filling solution. In this respect, we prove the uniqueness and convergence of the algorithm.

There exist many works on how to optimally allocate transmit power between different sub-channels. To our knowledge, reference [5] is the closest work to the one presented here. The authors address the problem of jointly allocating power and subcarriers in the context of orthogonal frequency division multiple access (OFDMA) systems. Our work differs from it on several points: we consider a more general channel model (fading channels instead of Gaussian channels in particular), a very different context (cross-system optimization), a more general performance criterion (any point of the capacity region can be reached), all the sub-channels are used whereas [5] selects a subset of these. Finally, our approach exploits asymptotic random matrix theory in order to provide tractable expressions for the optimization problem. Hence, we will assume the dimensions of the systems as well as the number of users large enough in order to benefit from the self-averaging properties of the matrices under consideration. In particular, an interesting feature of these self-averaging properties shows that only the parameters of interest to the problem (system load, signal to noise ratio,...) are kept whereas all irrelevant parameters disappear [6]. This provides a neat analysis framework for multi-dimensional problems. Moreover, although the results are proved in the asymptotic regime, it turns out that the results (due to fast convergence properties) are true for very small systems.

This paper is structured as follows. The first section (Sec. 2) provides the signal model used to study the cross-system problem under investigation. In Section 3, we consider the simplest scenario (1 user, 2 one-dimensional base stations, systems with equal bandwidths, static links) in order to identify the nature of the problem of maximizing the overall sum-capacity. In Section 4, the sum-capacity achieving power allocation policy is provided in a more realistic framework: arbitrary numbers of users, base station dimensions, systems bandwidths and fading channels with e.g. transmit and receive correlations. Simulations are provided to assess the gain provided by the proposed approach in Section 5 and possible extensions of this work are given in Section 6.

2. SYSTEM MODEL

Notations: In this paper, the notations $s$, $\mathbf{u}$, $\mathbf{M}$ stand for scalar, vector and matrix respectively. Capital letters are used to denote index upper bounds. The superscripts $(\cdot)^T$ and $(\cdot)^H$ denote transpose and transpose conjugate, respectively. The trace of the matrix $\mathbf{M}$ is denoted by $\text{Tr}(\mathbf{M})$.

The global system under investigation is represented in Figure 1. It consists of $K$ mobile terminals and $S$ base stations using non-overlapping bands of frequency ($S = 3$ in Figure 1). Each mobile has only one antenna while the base station can possibly have multiple antennas depending on the system. The number of dimensions associated with the base station $s \in \{1, ..., S\}$ is denoted by $N_s$. For example, if a CDMA system is used, $N_s$ represents the spreading factor, on the other hand, if the base station is equipped with several antennas, $N_s$ would represent the number of receive antennas. The equivalent baseband signals received by the base stations can be written as

$$y_s(\tau) = \sum_{\ell=1}^{K} h_{s,\ell}(\tau)x_{\ell,1}(\tau) + Z_s(\tau)$$

The mathematical expectation operator is denoted by $E(.)$. The notation $\mathcal{N}(\mathbf{u}, \mathbf{M})$ stands for the complex multi-dimensional Gaussian random variable with mean $\mathbf{u}$ and covariance $\mathbf{M}$.

![Figure 1: Cross-System scenario](image)
reliable links to a central controller. Knowing the channels of all users, the central controller implements the algorithm and schedules the power of each user on different links.

As we will consider the overall system sum-capacity as a performance criterion, and assuming a large system in terms of numbers of users and dimensions at the base station (N_s), it is convenient to rewrite the received signal in the following matrix form:

\[
\begin{align*}
\mathbf{y}_1 &= \sqrt{\rho_1} \mathbf{H}_1 \mathbf{x}_1 + \mathbf{z}_1 \\
\mathbf{y}_2 &= \sqrt{\rho_2} \mathbf{H}_2 \mathbf{z}_2 + \mathbf{z}_2 \\
\vdots \\
\mathbf{y}_s &= \sqrt{\rho_s} \mathbf{H}_s \mathbf{x}_s + \mathbf{z}_s
\end{align*}
\]

where \( \forall s \in \{1, ..., s\} \), \( \rho_s \) is the signal-to-noise ratio, \( \mathbf{H}_s = [h_{1,s} \ldots h_{K,s}] \) and \( \mathbf{z}_s = (x_1, \ldots, x_{K,s})^T \) of system \( s \). In a general form, we will use the G-model where the channel matrix of a given system can be factorized, in the sense of the Hadamard product, as a product of two matrices

\[
\mathbf{H}_s = \mathbf{G}_s \circ \mathbf{W}_s,
\]

where \( \mathbf{W}_s \) and \( \mathbf{G}_s \) are respectively the instantaneous channel gains matrix and the pattern mask specific to a given technology. As one can see, this model is broad enough to incorporate several technologies. Here are three examples:

- OFDM systems: In this case, \( N_s \) represents the number of carriers and \( K \) the number of users. Assuming for simplicity an OFDMA system where each user uses one subcarrier, \( \mathbf{W}_s \) and \( \mathbf{G}_s \) are respectively an i.i.d. zero mean Gaussian matrix and the truncated identity matrix (as the matrix is not square). Note that if \( K < N_s \), some carriers are not used.
- MIMO systems: In this case, \( N_s \) represents the number of antennas at the base station and \( K \) the number of users (each equipped with a single antenna \( \mathbf{W}_s \) and \( \mathbf{G}_s \) are respectively an i.i.d. zero mean Gaussian matrix and an \( N_t \times K \) correlation matrix (i.e. the Kronecker model [7]) defined as \( g_{s}(i,j) = d_{s}^{-T} d_{j}^{T} \), where \( d_{s} \) are the eigenvalues of the correlation/path loss matrix at the receiver and \( d_{j}^{T} \) are the eigenvalues of the correlation/path loss matrix at the transmitter.
- Flat fading CDMA systems: In this case, \( N_s \) represents the spreading length and \( K \) the number of users. For a block fading channel, \( \mathbf{W}_s \) and \( \mathbf{G}_s \) are respectively the code matrix, where each column represents the code of a given user, and the channel gains matrix, where the columns are identical (due to the fact that we consider flat fading models).

In the rest of the paper, we restrict ourselves to matrices \( \mathbf{G}_s \) such as the entries are separable: \( g_{s}(i,j) = d_{s}(i) \cdot d_{j}^{T} \) (in other words, we restrict ourselves to CDMA and MIMO).

We will label a given scenario by the triplet \( (K, N_s, S) \).

### 3. ELEMENTARY SCENARIO

Here, we consider the extreme scenario (1,1,2) i.e. with 1 user, 1 dimension at the base stations, 2 base stations. The two frequency bands used by the two cells (possibly with spatial overlapping) are denoted by \( B_1 \) and \( B_2 \). For simplicity we assume that \( B_1 = B_2 = \frac{B}{2} \). The scalar channels are assumed to represent block fading channels. This case is very simple but captures some important features of the problem. The system of equations associated with the received signals is:

\[
\begin{align*}
&y_1 = \sqrt{\rho_1 b_1 x_1 + z_1} \\
&y_2 = \sqrt{\rho_2 b_2 x_2 + z_2}.
\end{align*}
\]

The power allocated to bands 1 and 2 are respectively denoted by \( P_1 = \alpha \) and \( P_2 = \beta P = (1 - \alpha) \). We also introduce \( \gamma_1 \) and \( \gamma_2 \) defined by: \( \gamma_1 = \frac{\rho_1 b_1^2}{\sigma^2} \) and \( \gamma_2 = \frac{\rho_2 b_2^2}{\sigma^2} \) and \( \sigma^2 = \frac{\rho_0 b_0^2}{\alpha} \). The system sum-capacity can then be written as:

\[
\begin{align*}
C &= \max_{\alpha} \left\{ \log_2 \left( 1 + \gamma_1 \alpha \right) + \log_2 \left( 1 + \gamma_2 (1 - \alpha) \right) \right\} \\
&= \max_{\alpha} \log_2 \left\{ 1 + \gamma_2 + \frac{\gamma_1 - \gamma_2 + \gamma_1 \gamma_2}{\alpha - \gamma_1 \gamma_2 \alpha^2} \right\} \left( \frac{1}{R(\alpha)} \right)
\end{align*}
\]

The function \( \alpha \mapsto R(\alpha) \) is strictly concave and its maximum is reached for:

\[
\alpha^* = \frac{1}{2} + \frac{\gamma_1 - \gamma_2}{2 \gamma_1 \gamma_2}.
\]

We see that the power fraction allocated to System 1 is merely linear in the difference between the two receive signal-to-noise ratios (SNR) associated with the two systems in presence. One can distinguish four operating regimes:

- when \( \gamma_1 \gg \gamma_2 : \alpha^* \rightarrow 1 \) that is to say that the dominant link receives all the power;
- when \( \gamma_1 \approx \gamma_2 : \alpha^* \approx \frac{1}{2} \), which means that there exists a better allocation power policy than uniformly allocating the transmit power between the two systems;
- when \( \frac{\gamma_1}{\gamma_2} \approx 1 \leq \gamma_2 : 0 \leq \alpha^* \leq 1 \), which means that there is a power allocation scheme outperforming the commonly used “hard handover” strategy;
- when \( \gamma_1 \leq \frac{\gamma_1}{\gamma_2} : \alpha^* = 0 \); as we did not impose the function \( R(\alpha) \) to meet the transmit power constraint, \( \alpha^* = 0 \) translates the water-filling solution when all the power has to be allocated to one single sub-channel.

### 4. LARGE SYSTEMS SCENARIO ANALYSIS

In this section, we consider a more realistic scenario for wireless communications. The different links between the transmitters and receivers are still block fading and the numbers of users, systems and the base station dimensions are arbitrary. Additionally the base stations can have different bandwidths \( B_1, \ldots, B_S \). The numbers of users and dimensions have to be large enough in order to make our asymptotic analysis sufficiently accurate. However, it is now well-known (see e.g. [8, 9]) that many asymptotic results from random matrix theory under the large system assumption apply for relatively small systems. More precisely, we consider a scenario where \( K \rightarrow +\infty \), \( N_s \rightarrow +\infty \) with \( \lim_{K \rightarrow +\infty} K \frac{N_s}{N} = c_s \) and \( 0 < c_s < +\infty \). Under these assumptions our objective is to derive the best power allocation scheme in the sense of the sum-capacity of the global system. This quantity coincides with the Shannon sum-capacity of the global system. Considering the sum-rate point of the system, instead of an arbitrary operating
point of the capacity region, has the advantage to simplify the technical problem. In particular, considering the sum-capacity as the performance criterion allows us to exploit some results obtained for fading multiple input multiple output (MIMO) single-user channels.

By considering the system of (orthogonal) equations (2), the total ergodic sum-capacity per user (over the users, systems and antennas) is expressed as:

\[
C = \max_{Q_1, \ldots, Q_K} \frac{1}{K} \left[ \sum_{k=1}^{S} B_k \log_2 \left( 1 + \rho_k \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^H \right) \right]
\]  

(6)

where \( \forall s \in \{1, \ldots, S\} \), \( \rho_k = \frac{P}{\gamma_k^{-1} B} \), \( N_0 \) is the power spectral density of the received noises at the base stations, \( \mathbf{Q}_k = \mathbb{E}(\mathbf{x}_k \mathbf{x}_k^H) \). As long as the signals transmitted by different users are independent, the matrices \( \mathbf{Q}_k \) are diagonal: \( \mathbf{Q}_k = \text{Diag}(\alpha_{k,1}, \ldots, \alpha_{K,s}) \). As the mobile terminals have identical transmit power, we have \( \forall k \in \{1, \ldots, K\}, \sum_{s=1}^{S} \alpha_{k,s} = 1 \).

So far we did not assume anything on the numbers of users and base station dimensions. From now on, in order to simplify the optimization problem associated with equation (6) we will assume the asymptotic regime as defined in the beginning of this section. In the asymptotic regime, Theorem 3.7 of [6], derived for fading single-user MIMO channels, can be applied to separable channels to show that the constrained optimization under consideration amounts to finding the approximate \( C \) of \( C \) defined by:

\[
C = \max_{\omega_1, \ldots, \omega_K} \frac{1}{K} \left[ \sum_{k=1}^{K} B_k \log_2 \left( 1 + \gamma_{k,s} \omega_{k,s} \right) \right]
\]  

(7)

\[
+ \frac{1}{K} \sum_{k=1}^{K} B_k \log_2 (1 + \beta_{k,s} q_k) - \sum_{k=1}^{K} B_k v_k q_k r_k \log_2 c - \sum_{k=1}^{K} \lambda_k \left( \sum_{s=1}^{S} \omega_{k,s} - 1 \right) \right].
\]

(8)

where \( \forall \ell \in \{1, \ldots, K\} \), \( \lambda_k \) is a Lagrange multiplier, \( \omega_k = (\omega_{k,1}, \ldots, \omega_{k,S}) \) and the pairs \( \{q_k, r_k\}_{k=1}^{K} \) are given by the systems of equations:

\[
\begin{aligned}
q_k &= \frac{1}{K v_k} \sum_{s=1}^{S} \frac{\gamma_{k,s} \omega_{k,s}}{1 + \beta_{k,s} q_s} \\
r_k &= \frac{1}{K v_k} \sum_{s=1}^{S} \frac{\gamma_{k,s} \omega_{k,s}}{1 + \gamma_{k,s} \omega_{k,s} r_s}
\end{aligned}
\]

(9)

where \( v_k = K \rho_k s \) denotes the index of a MIMO system and \( v_k = c_k \rho_k s \) if \( s \) denotes the index of a CDMA system.

Note that for a separable channel \( \mathbf{H}_k = \mathbf{R}_k^T \Theta \mathbf{T}_k^T \), \( \Theta \) is a matrix with i.i.d entries with unit-variance, \( \gamma_{k,s} = \mathbf{v}_s \mathbf{d}_{k,s}^H \), \( \mathbf{d}_{k,s} \) is the \( k \)-th eigenvalue of \( \mathbf{T}_s \); \( \beta_{k,s} = \mathbf{v}_s \mathbf{q}_s^H \), \( \mathbf{q}_s \) is the \( s \)-th eigenvalue of \( \mathbf{R}_s \). For CDMA systems with flat fading, \( \mathbf{R}_s = \mathbf{I}, \Theta \) represents the spreading matrix and \( \mathbf{T}_s \) the propagation channel matrix (fading, path loss, ...). For MIMO systems \( \mathbf{T}_s \) and \( \mathbf{R}_s \) represent the antenna correlations at the transmitter and receiver respectively.

In order to find the optimum power allocation scheme we need to derive the argument of the maximum in equation (7), which we refer as \( R(\alpha_1, \ldots, \alpha_K) \). Obviously for all \( s \in \{1, \ldots, S\} \), \( r_s \) and \( q_s \) are functions of the parameters to be optimized i.e. \( \alpha_1, \ldots, \alpha_K \). It turns out that the partial derivative with respect to \( \alpha_{k,s} \) is the same as it would be if \( r_s \) and \( q_s \) would be assumed to be independent of these parameters. This is proved in Appendix A. This result is useful because it allows us to cope with the convergence issue of the quantities \( r_s \), \( q_s \) towards strict constants as the numbers of users and dimensions grow. Based on this argument, the fact that \( (\alpha_1, \ldots, \alpha_K) \rightarrow R(\alpha_1, \ldots, \alpha_K) \) is a strictly concave function (its Hessian is strictly positive) and using the notation \( B_s = b_s \times B \) (where \( B = B_1 + \ldots + B_S \)) in order to use adimensional quantities, one can show that the optimum power fractions are given by

\[
\alpha_{k,s}^* = \left[ \sum_{s \in S_k^+} b_s \left( 1 + \sum_{i \in S_k^+} \frac{1}{\gamma_{k,s} r_i} \right) - \frac{1}{\gamma_{k,s} r_s} \right]^{-1}
\]

(10)

where for each user \( k \) the set \( S_k^+ \) represents the systems/sub-channels which receive a non-zero power; \( |S_k^+| \leq S \) by definition. Note that user \( k \) will allocate power to the system \( s \) only if the quantity \( \frac{1}{\gamma_{k,s} r_s} \) is strictly positive. We see that, thanks to the large system assumption, the analysis of the general system under consideration (with fading and arbitrary numbers of users and base station dimensions) leads to a solution similar to that obtained for the elementary system (equation (5)). Indeed we also have a water-filling equation for the optimum power allocation scheme. One can give two special cases of equation (9). The case where the base stations have the same bandwidth (e.g. UMTS-FDD + UMTS-TDD base stations):

\[
\alpha_{k,s}^* = \left[ \frac{1}{|S_k|} + \frac{1}{|S_k^+|} \sum_{i \in S_k^+} \frac{1}{\gamma_{k,s} r_i} - \frac{1}{\gamma_{k,s} r_s} \right]^{-1}
\]

(11)

and also the scenario (1, 1, 2) with different bandwidths:

\[
\alpha^* = \left[ \frac{a}{a + 1} + \frac{a \gamma_1 - \gamma_2}{(a + 1) \gamma_1 \gamma_2} \right]^{-1}
\]

(12)

with \( a = \frac{b_1}{b_2}, \gamma_i = \frac{P \gamma_i}{b_i^2} \).

Then we can state the following

Result: The capacity of the system under consideration is achieved if and only if all the waterfilling equations (eq. 9) are verified simultaneously.

Proof: For the only if part, it suffices to consider that at the sum rate optimum, some of the power fractions did not verify the waterfilling equation, then, by keeping fixed all the other users and setting this one to the waterfilling value, the total sum rate of the system would be increased, which contradicts the assumption that it was sum rate optimal, and at it, all users must satisfy the waterfilling equations for their power fractions.

The if part is obviously verified by the construction of the derivation of the waterfilling equations and the convexity of the problem under consideration.

The main issue we have to mention now is the way of implementing the proposed power allocation scheme. An iterative algorithm is proposed to achieve the optimal power allocation:
1. Initialization: assume a uniform power allocation scheme i.e. \( \forall (k, s) \in \{1, ..., K\} \times \{1, ..., S\}, \alpha_{k,s} = \frac{1}{S} \).

2. Compute the corresponding value for \( r_s \) by using the fixed-point method: the first equation of the system (8) can be written in the form: \( r_s = f_s(r_s) \).

3. Iterate the procedure while the desired accuracy on the power fractions is not reached.
   - For users 1, ..., \( K \)
     - Update the power fractions by using the water-filling equation (9).
     - Update the value of \( r_s \)

At each step of the iterative procedure, the total capacity of the system is increasing (or keeps constant if the power fraction is not modified), and since it is bounded above, it follows directly that it will converge to a limit. At the limit, all power fractions will verify the waterfilling equations. Then from the previous result it follows that the sum capacity of the system is achieved.

5. SIMULATION EXAMPLE
In this section simulation results are provided showing the performance of the proposed algorithm.

The main system under consideration will be composed by \( K = 2 \) users and \( S = 2 \) MIMO systems equipped with \( N = 2 \) and \( N = 8 \) antennas at each of the base stations. Its performance in terms of sum capacity (and not the sum capacity per user) will be a function of only 2 parameters: their respective power fractions. Note that in this particular case, the whole set of possibilities (in figure 2) can be plotted which allows us to verify the concavity of the function and therefore justifies the use of the iterative algorithm.

In figure 3, the average capacity per user is plotted versus the signal to noise ratio \( \rho \), in a system where all the path losses are equal. The optimal value obtained from the exhaustive search is also shown and identical to the one resulting from the proposed algorithm. It can be noticed that for rather small values of \( r \), the user closer to the better system (because of having more antennas) gets disconnected from the other subsystem, whereas the other user will continue to use both of them.

The influence of correlation at the receiver is analyzed in figure 6, where it is assumed that it only affects the system
with the 8 antennas and follows a typical and realistic antenna correlation profile: \( \forall (i, j) \in \{1, \ldots, n_r\}^2, \ R_{k}(i, j) = a^{(i-j)} \) where \( a \) is the correlation coefficient characterizing it. As the correlation increases, the system performance degrades, becoming closer for both systems, so the gain obtained by the power allocation compared to the uniform one reduces.

### 6. CONCLUSIONS

In this contribution, a cross-system power allocation algorithm has been provided in the context of MIMO and CDMA technologies. Interestingly, in the asymptotic regime, the solution turns out to be depending only on a limited number of parameters: dimensions of the system, number of users; channel gains, path loss, noise variance, correlation at the transmitter and the receiver. As a consequence, a simple cross-layer algorithm, analogous to the water-filling algorithm, can be implemented at the central controller to schedule the powers of all the users in order to maximize the system capacity, and this can be done in a simple, iterative way, which is shown to converge to the optimum.

### 7. REFERENCES


[10] A. Suarez, R. L. de Lacerda, M. Debbah and Linh-Trung Nguyen, “Power allocation under quality of
APPENDIX

A. DERIVATING THE ERGODIC SUM MUTUAL INFORMATION

We want to derivate the argument of the maximum in equation (7) with respect to \( \alpha_k \). First note from the system of equations (8) that \( r_t \) and \( q_t \) do not depend on \( \alpha_{k,s} \) for all \( t \neq s \). Based on this observation one just needs to consider the following auxiliary function:

\[
\phi(\alpha_k) = \log_2 \left\{ \prod_{\ell=1}^{K} \left[ 1 + \gamma_\ell \alpha_\ell r(\alpha_k) \right] \times \prod_{j=1}^{N} \left( 1 + \rho_d^2 j q \right) \times e^{-K \rho(r(\alpha_k) + q(\alpha_k))} \right\}
\]

where we dropped the system index \( s \) and receiver subscript \( (R) \) for sake of clarity.

Define \( u \triangleq \prod_{\ell=1}^{K} \left[ 1 + \gamma_\ell \alpha_\ell r(\alpha_k) \right] \) and

\[
v \triangleq \prod_{j=1}^{N} \left( 1 + \rho_d^2 j q \right) \times e^{-K \rho(r(\alpha_k) + q(\alpha_k))}.
\]

With these notations:

\[
\frac{\partial \phi(\alpha_k)}{\partial \alpha_k} = \frac{1}{\ln 2} \frac{1}{uv} \frac{\partial uv}{\partial \alpha_k}.
\]

It turns out that

\[
\frac{\partial (uv)}{\partial \alpha_k} = uv \times \frac{\gamma_k r}{1 + \gamma_k \alpha_k r}.
\]

This is what we want to show.

First step:
We want to derivate the function \( u \). As \( u \) is a product of functions \( u_\ell \), i.e. \( u = \prod_{\ell=1}^{K} u_\ell \), its derivative \( u' \) can be written as:

\[
u' = u \times \sum_{\ell=1}^{K} \frac{u'_\ell}{u_\ell}
\]

where

\[
u'_\ell = \begin{cases} 
\gamma_\ell \alpha_\ell r' & \text{if } \ell \neq k \\
\gamma_k (r + \alpha_k r') & \text{if } \ell = k.
\end{cases}
\]

Second step:
Using a similar reasoning for \( v \) one can check that

\[
v' = v \times \left[ \sum_{j=1}^{N} \frac{\rho d^2 j q'_j}{1 + \rho d^2 j q} - K \rho (q' r + q r') \right].
\]

Third step:
Now using the relations proved in the previous steps we have that

\[
\frac{\partial (uv)}{\partial \alpha_k} = uv \times \left[ \sum_{\ell=1}^{K} \frac{u'_\ell}{1 + \gamma_\ell \alpha_\ell r} + \sum_{j=1}^{N} \frac{\rho d^2 j q'_j}{1 + \rho d^2 j q} - K \rho (q' r + q r') \right].
\]
with $\psi$ expanding as

$$
\psi = \sum_{\ell \neq k} \frac{\gamma_{\ell} \alpha_{\ell} r'}{1 + \gamma_{\ell} \alpha_{\ell} r} + \frac{\gamma_k (r + \alpha_k r')}{1 + \gamma_k \alpha_k r} + \sum_{j=1}^N \frac{\rho d_j^2 q'}{1 + \rho d_j^2 q} - K \rho (q' r + qr').
$$

(12)

Now by observing that

$$
\begin{cases}
\sum_{\ell \neq k} \frac{\gamma_{\ell} \alpha_{\ell} r'}{1 + \gamma_{\ell} \alpha_{\ell} r} = \left( K \rho q - \frac{\gamma_k \alpha_k}{1 + \gamma_k \alpha_k r} \right) r' \\
\sum_{j=1}^N \frac{\rho d_j^2 q'}{1 + \rho d_j^2 q} = K \rho q' r
\end{cases}
$$

we find that

$$
\psi = \frac{\gamma_k r}{1 + \gamma_k \alpha_k r}
$$

which concludes the proof.