Power Allocation for Space-Terrestrial Heterogeneous Networks with a Multibeam Satellite and Ground Relays

Jihwan P. Choi*

Daegu Gyeongbuk Institute of Science & Technology (DGIST), Korea

Changhee Joo†

Ulsan National Institute of Science and Technology (UNIST), Korea

Ji-Woong Choi‡

Daegu Gyeongbuk Institute of Science & Technology (DGIST), Korea

Narrow spotbeam scheduling and interbeam interference management can make the advanced multibeam satellite equipped with phased array antenna more effective to serve a large number of users over its coverage area. Further, if the satellite has an option of making use of ground stations to forward packets to end user terminals, path diversity can improve the overall system throughput. However, onboard resource management has been only analyzed without terrestrial resources into account. In this paper, we develop a power allocation scheme for a multibeam satellite and ground relay stations into joint consideration. We attempt to minimize the power consumption of a heterogeneous network by differently weighting power usage of the satellite and ground stations. The analytical solution is given in terms of channel conditions, quality of service (QoS) requirements, potential interbeam interference that is primarily determined by geographical distribution of users and relay locations, and the penalty cost for relaying. The optimum power allocation indicates that path selection between via relay station and directly to users, beamforming of multiple spotbeams, and user scheduling over a small number of beams should be decided in a cross-layer approach. Simulation results show that the proposed method seeks for a trade-off between the use of satellite and terrestrial resources. The formulation and analysis developed in this paper may hint at the direction of the cross-layer solutions in heterogeneous networks in general.

Nomenclature

M	Number of users
P	Power allocated to the direct user signal
Q	Power allocated to the ground relay signal
ψ	Power used in the terrestrial link
S	QoS funtion
$S^{\min}$	Minimum QoS required
ω	Weighting constant for terrestrial power consumption
λ	Weighting constant to scale the under-satisfied QoS
C	Shannon capacity for user

*Assistant Professor, Department of Information and Communication Engineering, Daegu Gyeongbuk Institute of Science & Technology (DGIST), Korea.

†Associate Professor, School of Electrical and Computer Engineering, Ulsan National Institute of Science and Technology (UNIST), Korea.

‡Associate Professor, Department of Information and Communication Engineering, Daegu Gyeongbuk Institute of Science & Technology (DGIST), Korea.

α^2	Signal power attenuation for user
W	Bandwidth used
N_0	Additive white Gaussian noise power density
H	Power loss incurred for user by deploying zero-forcing transmission strategy
C_g	Shannon capacity for ground relay
α_g^2	Signal power attenuation for ground relay
H_g	Power loss incurred for ground relay by deploying zero-forcing transmission strategy
G	Ground relay capacity allocated to user
κ	Scaling factor between ψ and Q
J	Lagrangian function
μ	Lagrangian variable
ν	Lagrangian variable
P^*	Optimum power allocated to the direct user signal
Q^*	Optimum power allocated to the ground relay signal
K	Number of active beams

Subscript

i	User index
j	User index
k	User index
m	User index

I. Introduction

Relaying has been an important research topic for 4G/5G cellular systems as it can extend the coverage to the cell-edge users and take traffic load off from macro base stations.[1–3] The seamless relaying through heterogeneous radio access technologies has yet to be realized in practice due to difficulty of controlling different protocols though different channels. Benefits and difficulties of realizing relay networks are equally applied to the space-terrestrial heterogeneous network. Effectiveness of multibeam satellites, which can serve a large number of users over its coverage area by beamforming narrow spotbeams and managing interbeam interference, can be enhanced when ground stations such as gateways, gap fillers or feeder antennas are also utilized. If the satellite has an option of making use of ground stations to relay packets to end user terminals, path diversity can improve the overall system throughput. If packets have to be retransmitted, the ground stations can fulfill the request, reducing retransmission delay from satellites. Thus, a scheduler of the space-terrestrial network should consider a possibility of packet retransmission when it makes a decision of packet routing and resource allocation. The hybrid automatic retransmission request (HARQ) can be used in the terrestrial link, and terminals may combine the retransmitted signals with the signals that have been previously received from the satellite but unsuccessful to decode.[4] Frequency bands in each channel of satellite to user, satellite to ground station, and ground station to user can overlap entirely or partially, in which case the allocation of the shared spectrum should be coordinated not to degrade the entire system performance. Recently, the standard community is interested in sharing spectrum (for example, around 2.1 GHz) between satellites and terrestrial cellular networks as the cellular communication technology advances rapidly and the demand for spectrum increases unexpectedly.

So far, onboard resource management has been only analyzed without terrestrial resources into account.[5,6] Resource allocation of the space and ground components should be jointly coordinated and user scheduling should be considered together with the decisions at other layers such as power/rate allocation and data routing path selection. By incorporating ground stations to receive satellite signals, the multibeam satellite has a choice of routing signals to gateways/feeder antennas and/or sending directly to end user terminals. Resource allocation and user scheduling solutions can be based on channel conditions, traffic demands, quality of service (QoS) requirements, traffic characteristics (voice, data, short messages, multimedia, unicast, multicast, broadcast, etc.), and potential interbeam interference that is primarily determined by geographical distribution of users and gateway locations. By comparing direct transmission to users and relaying via gateways, the scheduler chooses the users to be served at each transmit opportunity and the optimum transmission strategy in terms of transmit power allocation and signal path selection. The cross-layer

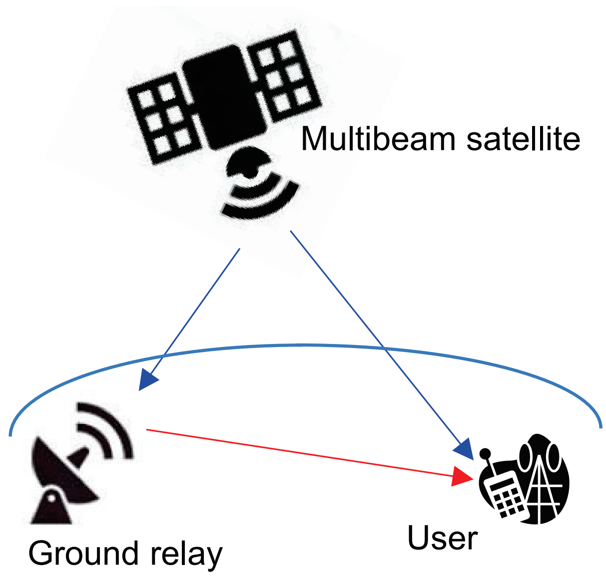

Figure 1. A simple space-terrestrial heterogeneous network with a multibeam satellite and a ground relay on the Earth.

optimized scheme takes into joint account the PHY layer power allocation, the MAC layer user scheduling, the Network layer routing path selection and Transport layer congestion control for system stability.

In this paper, we extend the aforementioned result[6] by considering the satellite and terrestrial power allocation altogether, and investigate a scheme of resource allocation and user scheduling when a choice of relaying signals to the ground relay station incurs the additional cost of terrestrial resource for relaying and potential retransmission. We formulate a minimum power consumption problem and give a resource allocation and user scheduling solution based on the terrestrial power consumption incurred by relaying. The simulation results show that the resource utilization and system performance heavily depend on the penalty cost that reflects the relay link power consumption.

II. Formulation

We consider a simple space-terrestrial heterogeneous network that has a multibeam satellite and a ground relay station to serve M users on the Earth (Fig. 1). Each user is assumed to receive its downlink signal either directly from the satellite or via a ground relay. The satellite scheduler faces a problem of choosing which signal path between the two per user. Let P_i denote the power allocated to the direct signal to user i and Q_i the power allocated to the ground relay that will forward the received data to user i using terrestrial power ψ_i. Given the QoS constraint $S_i^{\min} \leq S_i$ of each user with a QoS function S_i (representing, e.g., data rate, delay, throughput/goodput, etc.) and the required minimum QoS $S_i^{\min}$, we minimize the total weighted power consumption in the space-terrestrial heterogeneous network as follows:

$$\text{minimize} \quad \sum_{i=1}^{M}(P_i + Q_i) + \omega \sum_{i=1}^{M} \psi_i \tag{1}$$

$$\text{subject to} \quad S_i^{\min} \leq S_i \tag{2}$$

$$\text{and} \quad P_i \geq 0, \quad Q_i \geq 0 \quad \text{for every } i, \tag{3}$$

where ω is the weighting constant for terrestrial power consumption, compared to the satellite power. The total power constraint of the satellite is not explicitly specified since the problem is to minimize the power consumption.

One way to solve this problem is selecting the smaller amount between P_i and $Q_i + \omega\psi_i$ to satisfy $S_i^{\min} \leq S_i$ for every i, and deciding the amount of power allocation. However, in the practical multibeam satellite system, the problem may be infeasible since the number of available beams is less than that of users to be served and all the QoS constraints cannot be met all the time. We now add the penalty cost to the objective function to prevent the infeasible case. The QoS function S_i is modeled as a jointly concave function of P_i and Q_i, such that the concave property of S_i holds not only for each P_i and Q_i, but also for a pair of (P_i, Q_i) jointly. The terrestrial power usage ψ_i is assumed to be a function of Q_i only. A modified problem is as follows:

$$\text{minimize} \quad \sum_{i=1}^{M}(P_i + Q_i) + \omega \sum_{i=1}^{M} \psi_i + \sum_{i=1}^{M} \max\left\{\lambda_i \cdot \left(S_i^{\min} - S_i\right), 0\right\} \tag{4}$$

$$\text{subject to} \quad P_i \geq 0, \quad Q_i \geq 0 \quad \text{for every } i, \tag{5}$$

where λ_i is the weighting constant that scales the under-satisfied QoS.

We adopt the concave QoS function of the capacity, which is again a concave function of allocated power P_i or Q_i. We assume the data rate achieved by P_i as Shannon capacity C_i, i.e.,

$$C_i = W \log\left(1 + \frac{\alpha_i^2 H_i P_i}{W N_0}\right), \tag{6}$$

where α_i^2 (≤ 1) is signal power attenuation due to the weather effects that only change quasi-statically over the link to the i^{th} user, W is the bandwidth used, and N_0 is the additive white Gaussian noise power density. The value of H_i represents the power loss incurred by deploying zero-forcing transmission strategy and is given by a function of the beamwidth and the distance between active users. We assume that interbeam interference occurs when active users are located within one spotbeam width, which is defined as the mainlobe width of the narrowest sportbeam. The capacity of the multi-antnna ground relay is represented by

$$C_g = \sum_{i}^{M} G_i = \sum_{i}^{M} W \log\left(1 + \frac{\alpha_g^2 H_g Q_i}{W N_0}\right), \tag{7}$$

where $G_i = W \log\left(1 + \frac{\alpha_g^2 H_g Q_i}{W N_0}\right)$ is the capacity allocated to user i at the source link from the satellite to the ground relay, α_g^2 (≤ 1) is signal power attenuation over the link to the ground relay, and H_g represents the power loss incurred for the ground relay by deploying zero-forcing transmission strategy.

The provisioned capacity and bandwidth at the relay link from the ground station to the user are assumed to be the same as those at the source link from the satellite to the ground station, in order to guarantee the same QoS and to make the analysis tractable. The terrestrial power ψ_i is then given as a linearly scaled amount of Q_i:

$$\psi_i = \kappa_i \cdot Q_i, \tag{8}$$

where the scaling factor κ_i is given as the ratio of signal attenuation factors at each link, such as channel condition, free space loss and interferences. If the packet is retransmitted at the relay link, the scaling factor increases as more terrestrial power is consumed.

III. Analysis

As applying the Kuhn-Tucker condition,[7] we solve the problem for which P_i and/or Q_i have a non-zero value that indicates the scheduled user and the selected signal path. The corresponding Lagrangian function with $S_i^{\min} > S_i$ is given as follows:

$$J(P_i, Q_i) = \sum_i (P_i + Q_i) + \omega \sum_i \psi_i + \sum_i \lambda_i \cdot \left(S_i^{\min} - S_i\right) + \sum_i \mu_i \cdot (-P_i) + \sum_i \nu_i \cdot (-Q_i), \tag{9}$$

where the Lagrangian variables of μ_i (≥ 0) and ν_i (≥ 0) are for non-negative constraints of P_i and Q_i respectively, and added to see which users should be served with non-zero P_i and/or Q_i. Note that in case of $S_i^{\min} \leq S_i$, power does not have to be allocated. The Kuhn-Tucker condition yields $\mu_i \geq 0$ if $P_i = 0$ and $\mu_j = 0$ if $P_j > 0$, which applies to ν_i and Q_i in the same manner. Differentiating $J(P_i, Q_i)$ with respect to P_i and Q_i respectively gives

$$\frac{\partial J}{\partial P_i} = 1 - \lambda_i \frac{\partial S_i}{\partial P_i} - \mu_i, \tag{10}$$

and

$$\frac{\partial J}{\partial Q_i} = 1 + \omega \frac{\partial \psi_i}{\partial Q_i} - \lambda_i \frac{\partial S_i}{\partial Q_i} - \nu_i. \tag{11}$$

Here, as modeled in the formulation section, the terrestrial power ψ_i is a function of Q_i only, but not P_i, since relaying incurs a use of terrestrial power but direct transmission does not.

Following the method used in the previous works,[5,6] we first consider the case where only a single path is allowed for each user, in which there are three possibilities of optimum power allocation: user i has $P_i^* > 0$ and $Q_i^* = 0$, user j has $P_j^* = 0$ and $Q_j^* > 0$ with $\psi_j^* > 0$, and user k has $P_k^* = 0$ and $Q_k^* = 0$. In other words, user i receives a packet directly and user j's packet is sent to the relay station while user k is not served at this opportunity. At the optimum power $P_i^* > 0$ and $Q_i^* = 0$, user i has $\mu_i = 0$ and $\nu_i \geq 0$. This gives

$$\lambda_i \left.\frac{\partial S_i}{\partial P_i}\right|_{(P_i=0,Q_i=0)} > \lambda_i \left.\frac{\partial S_i}{\partial P_i}\right|_{(P_i=P_i^*,Q_i=0)} = 1, \tag{12}$$

where the inequality holds due to the concavity of S_i in terms of P_i and the equality comes from setting Eq. (10) equal to zero. Similarly at the optimum power $P_j^* = 0$ and $Q_j^* > 0$, user j has $\mu_j \geq 0$ and $\nu_j = 0$. Again from Eq. (11), we obtain

$$\left[\lambda_j \frac{\partial S_j}{\partial Q_j} - \omega \frac{\partial \psi_j}{\partial Q_j}\right]_{(P_j=0,Q_j=0)} > \left[\lambda_j \frac{\partial S_j}{\partial Q_j} - \omega \frac{\partial \psi_j}{\partial Q_j}\right]_{(P_j=0,Q_j=Q_j^*)} = 1. \tag{13}$$

If both P_k^* and Q_k^* have zero values, user k has $\mu_k \geq 0$ and $\nu_k \geq 0$, and thus,

$$1 \geq 1 - \mu_k = \lambda_k \left.\frac{\partial S_k}{\partial P_k}\right|_{(P_k=0,Q_k=0)}, \tag{14}$$

and

$$1 \geq 1 - \nu_k = \left[\lambda_k \frac{\partial S_k}{\partial Q_k} - \omega \frac{\partial \psi_k}{\partial Q_k}\right]_{(P_k=0,Q_k=0)}. \tag{15}$$

If there are only $K(< M)$ active beams available in the practical multibeam satellite, we should select K options out of $2M$ candidate pairs of user and path. The above results imply that the optimal user and path selection can be obtained based on the marginal returns of QoS at zero power of each path. Through Eq. (12) and (15), the threshold level of the partial derivative in this problem is 1, which is derived from satellite power consumption itself ($dP_i/dP_i = dQ_i/dQ_i = 1$). Hence, if only a single path is allowed for each user, we choose K highest values of $\lambda_i \left.\frac{\partial S_i}{\partial P_i}\right|_{(P_i=0,Q_i=0)}$ and $\left[\lambda_j \frac{\partial S_j}{\partial Q_j} - \omega \frac{\partial \psi_j}{\partial Q_j}\right]_{(P_j=0,Q_j=0)}$ such that $i \neq j$.

We now relax the single path assumption for each user, and consider the case that a user m has non-zero values for both P_m and Q_m. With both $P_m^* > 0$ and $Q_m^* > 0$, user m has $\mu_m = \nu_m = 0$, and thus,

$$\lambda_m \left.\frac{\partial S_m}{\partial P_m}\right|_{(P_m=P_m^*,Q_m=Q_m^*)} = \left[\lambda_m \frac{\partial S_m}{\partial Q_m} - \omega \frac{\partial \psi_m}{\partial Q_m}\right]_{(P_m=P_m^*,Q_m=Q_m^*)} = 1, \tag{16}$$

From the assumption of a jointly concave function S_m with respect to P_m and Q_m, user m should have

$$\lambda_m \left.\frac{\partial S_m}{\partial P_m}\right|_{(P_m=0,Q_m=0)} > \lambda_m \left.\frac{\partial S_m}{\partial P_m}\right|_{(P_m=P_m^*,Q_m=Q_m^*)} = 1, \tag{17}$$

and

$$\left[\lambda_m \frac{\partial S_m}{\partial Q_m} - \omega \frac{\partial \psi_m}{\partial Q_m}\right]_{(P_m=0,Q_m=0)} > \left[\lambda_j \frac{\partial S_j}{\partial Q_m} - \omega \frac{\partial \psi_m}{\partial Q_m}\right]_{(P_m=P_m^*,Q_m=Q_m^*)} = 1. \tag{18}$$

Again, we choose the K highest values of $\lambda_i \left.\frac{\partial S_i}{\partial P_i}\right|_{(P_i=0,Q_i=0)}$ and $\left[\lambda_j \frac{\partial S_j}{\partial Q_j} - \omega \frac{\partial \psi_j}{\partial Q_j}\right]_{(P_j=0,Q_j=0)}$, but without any restriction to i and j. To summarize, we see that the marginal return of S_i at zero power can be used for a metric to allocate power and decide the path for data to be sent over. It is reminded that the users with $S_i^{\min} > S_i$ are only considered for power allocation.

With the modeling of C_i with respect to P_i, and C_g and ψ_i with respect to Q_i as described in the previous section, the marginal returns at $P_i = 0$ and $Q_i = 0$ are as follow respectively:

$$\lambda_i \left.\frac{\partial S_i}{\partial P_i}\right|_{(P_i=0,Q_i=0)} = \lambda_i \cdot \frac{\alpha_i^2 H_i}{N_0} \cdot \left.\frac{\partial S_i}{\partial C_i}\right|_{(C_i=0,G_i=0)} \tag{19}$$

and

$$\left[\lambda_i \frac{\partial S_i}{\partial Q_i} - \omega \frac{\partial \psi_i}{\partial Q_i}\right]_{(P_i=0,Q_i=0)} = \lambda_i \cdot \frac{\alpha_g^2 H_g}{N_0} \cdot \left.\frac{\partial S_i}{\partial G_i}\right|_{(C_i=0,G_i=0)} - \omega\kappa_i, \tag{20}$$

where the chain rule is applied with $\left.\frac{\partial C_i}{\partial P_i}\right|_{P_i=0} = \frac{\alpha_i^2 H_i}{N_0}$ and $\left.\frac{\partial G_i}{\partial Q_i}\right|_{Q_i=0} = \frac{\alpha_g^2 H_g}{N_0}$. When we select users and signal paths, we compare the channel condition, potential interbeam interference, and the marginal return of the QoS function at zero capacity for both direct paths and relaying links. In addition, for the relaying links, the penalty for additional power consumption at the terrestrial links must be taken into account.

We note that the general utility maximization problem subject to the total power constraint can be formulated and solved in the same manner by adding the relay cost of terrestrial power consumption.

IV. Simulation Results

In this section, simulation results are provided to show system performance by changing the relay cost. The total of $M = 25$ users/gateway are located on the 5 by 5 planar grid with the maximum number of active beams $K = 20$. The ground relay is placed at one point of the grid. Each of the 25 users/relay is assumed in the moderately rainy channel condition at the Ka band, and modeled as an independent and identically distributed (i.i.d.) lognormally Gaussian random process with mean of 7.1205 dB and standard deviation of 1.8629 dB. The maximum number of the signals that can be transmitted simultaneously is assumed to be 8. The distance between adjacent users/relay is fixed at the 2/5 of the beamwidth, so that the effect of moderate interbeam interference is taken into account. The QoS function S_i to be used is a per-user date rate of $C_i + G_i$. Since we simulate the case when the minimum QoS of $S_i^{\min}$ is much greater than the capacity that can be provided from the total satellite power, our problem of minimizing power consumption with the under-satisfied QoS penalty is equivalent to that of maximizing the total QoS, which is the per-user data rate, for every user with the total satellite power consumed. We change the relay cost from 0 to the half of the maximum value of the weighted marginal return of the QoS function, which is equal to λ_i/N_0 in this case.

In Fig. 2, we plot the sum capacity gain of the satellite by using the ground relay, compared to the case of not using. At the zero penalty cost, the satellite is free to send packets to the ground relay, and the capacity gain is maximized as the ground relay link replaces the direct user links with the worst link conditions among the selected. As the penalty increases, a use of the link to the ground relay decreases and the sum capacity gain also degrades. Note that the capacity gain is below 1 for high relay penalty as the link to the ground relay is entirely disallowed, resulting in the loss of the one satellite downlink out of 25. Fig. 3 shows the ratio of the sum power used for a ground relay to that for direct transmissions. As the penalty increases, power allocation to the ground relay decreases, which leads to the sum capacity degradation. The ground relay consumes more than half of the total satellite power at the zero penalty cost, but no power is allocated as the normalized penalty cost approaches 0.5. The high penalty cost implies many retransmissions or other difficulties of utilizing terrestrial links, which prohibit the use of ground relay. Thus, the appropriate design of the terrestrial relay link is essential for maximizing the space-terrestrial heterogeneous network performance and resource utilization.

V. Conclusion

As deployed for emergency networks and utilized as a backbone for cellular systems, the space-terrestrial heterogeneous network can combine the advantages of satellite and wireless communications. The method of resource allocation should be jointly coordinated with power/rate allocation and data routing path selection in the cross-layer optimal approach. In this paper, we considered the satellite and terrestrial power allocation altogether, and investigated a scheme of resource allocation and user scheduling when a use of the ground relay results in the additional cost of terrestrial resource consumption and potential retransmission. We

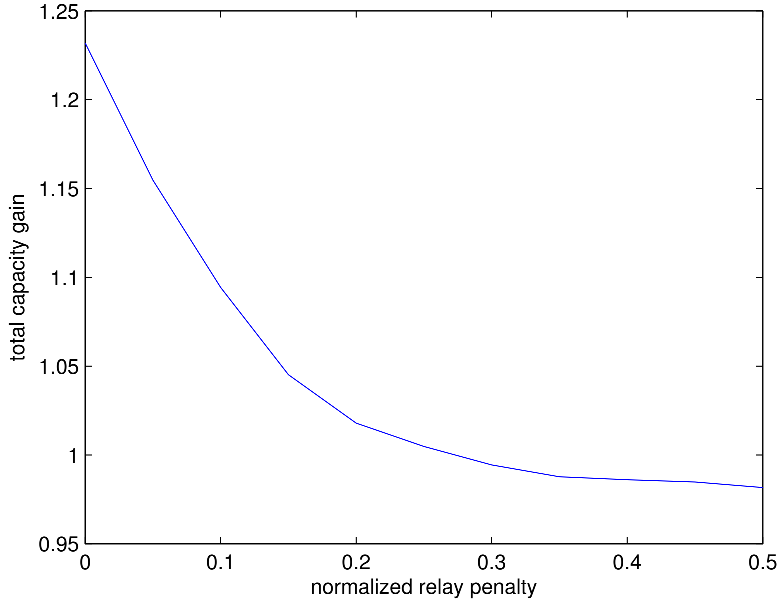

Figure 2. Sum capacity gain of using a ground relay, compared to not using, as a function of normalized relay penalty.

formulated a minimum power consumption problem and gave a solution based on the simple modeling of terrestrial power in terms of the satellite power allocated to the ground relay. Simulation results showed that resource utilization and system performance heavily depends on the penalty cost. Future work includes more rigorous and detailed modeling of the relay penalty cost, which can lead to more accurate performance evaluation of the space-terrestrial network. It is remarked that the formulation and analysis developed in this work is applicable to other generic heterogeneous networks.

Acknowledgments

This work was supported by the Daegu Gyeongbuk Institute of Science and Technology through the MIREBraiN program.

References

[1]X. Gong, S.A. Vorobyov, and C. Tellambura, "Joint bandwidth and power allocation with admission control in wireless multi-user networks with and without relaying," *IEEE Trans. Signal Processing*, vol. 59, no. 4, pp. 1801-1813, Apr. 2011.

[2]J. Tang, B. Mumey, K. Zhubayev and R. Wolff, "Leveraging cooperative, channel and multiuser diversities for efficient resource allocation in wireless relay networks," *IEEE Journal on Selected Areas in Communications (JSAC), Special Issue on Cooperative Networking Challenges and Applications*, Vol. 30, No. 9, pp. 1789-1797, Sept. 2012.

[3]Y. Ma, A. Liu and Y. Hua, "A Dual-Phase Power Allocation Scheme for Multicarrier Relay System with Direct Link," *IEEE Trans. Signal Processing*, vol. 62, no. 1, pp. 5-16, Jan. 2011.

[4]S. Kim, "Evaluation of Cooperative Techniques for Hybrid/Integrated Satellite Systems," *IEEE International Conference on Communications (ICC)*, June 2011.

[5]J. P. Choi and V. W. S. Chan, "Resource Management for Advanced Transmission Antenna Satellites," *IEEE Trans. on Wireless Communications*, vol. 8, no. 3, pp. 1308-1321, March 2009.

[6]J. P. Choi and C. Joo, "Resource Allocation and User Scheduling for a Multibeam Satellite with Signal Path Diversity," *32nd AIAA ICSSC 2013*, Oct. 2013.

[7]D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 1995.

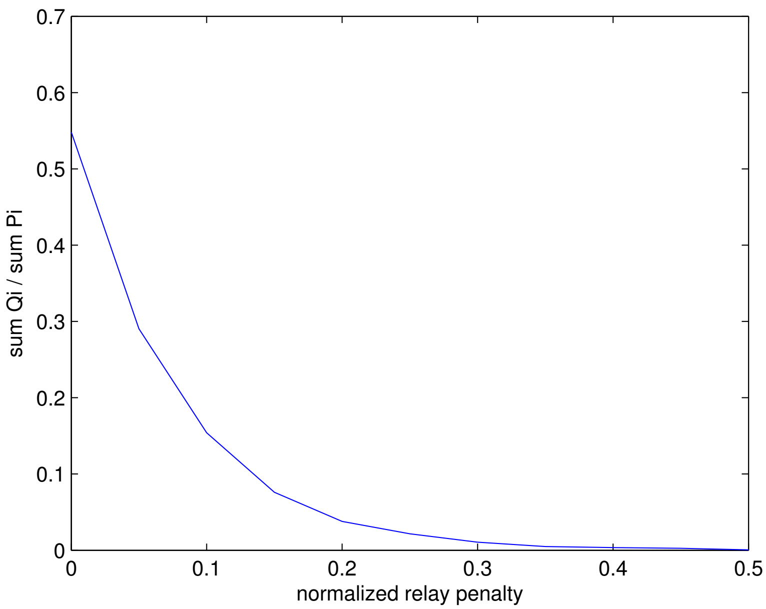

Figure 3. Ratio of power allocation to the ground relay and direct transmission as a function of normalized relay penalty.