National Technical University of Athens
School of Electrical and Computer Engineering
Division of Information Transmission Systems and Material Technology

Optimization Algorithm Design and Performance Analysis for Next-Generation Wireless Networks

PhD Thesis

Christos N. Efrem
Diploma in Electrical and Computer Engineering,
National Technical University of Athens (2015)

Athens, Greece – April 2021

Εθνικο Μετσοβιο Πολυτεχνειο
Σχολη Ηλεκτρολογων Μηχανικων και
Μηχανικων Υπολογιστων
Τομεας Συστηματων Μεταδοσης Πληροφοριας
και Τεχνολογιας Υλικων

Σχεδίαση Αλγορίθμων Βελτιστοποίησης και Ανάλυση Επίδοσης για Ασύρματα Δίκτυα Επόμενης Γενιάς

Διδακτορική Διατριβή

Χρίστος Ν. Εφραίμ
Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών,
Εθνικό Μετσόβιο Πολυτεχνείο (2015)

Αθήνα, Ελλάδα – Απρίλιος 2021

National Technical University of Athens
School of Electrical and Computer Engineering
Division of Information Transmission Systems and Material Technology

Optimization Algorithm Design and Performance Analysis for Next-Generation Wireless Networks

PhD Thesis

Christos N. Efrem

Diploma in Electrical and Computer Engineering,
National Technical University of Athens (2015)

Advisory Committee: Athanasios D. Panagopoulos (*Supervisor*)
Panagiotis Cottis
George Fikioris

Approved by the seven-member examination committee on 12/04/2021.

....	....	....
Athanasios D. Panagopoulos Associate Professor, NTUA	Panagiotis Cottis Professor, NTUA	George Fikioris Professor, NTUA
....	....	....
Christos Capsalis Professor, NTUA	Symeon Papavassiliou Professor, NTUA	Ioanna Roussaki Assistant Professor, NTUA

....

George Efthymoglou
Professor, University of Piraeus

Athens, Greece – April 2021

. . . .

Χρίστος Ν. Εφραίμ

Διδάκτωρ Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών ΕΜΠ

"The road to wisdom? Well, it's plain and simple to express:
Err and err and err again, but less and less and less."

— Piet Hein

Abstract

The proliferation of connected devices has led to very strict requirements for next-generation wireless networks, taking into consideration *environmental* as well as *economic* concerns. In particular, one of the primary goals in the design of fifth-generation (5G) wireless networks is to satisfy the extremely high data rate (traffic demand) of users with the minimum energy consumption. For this purpose, a new performance indicator, namely, *energy efficiency (EE)*, has been proposed in the literature which is measured in bits/Joule and expresses the amount of information that can be reliably transmitted per unit of consumed energy.

This Dissertation deals with the design of *efficient optimization algorithms* for next-generation wireless networks, including terrestrial as well as satellite communication systems. More specifically, the theory of *sequential convex optimization (SCO)* is applied to solve challenging optimization problems, such as the maximization of several EE-metrics, so as to develop energy-efficient power allocation strategies. SCO is a powerful mathematical tool that can be used to solve nonconvex optimization problems by solving a sequence of convex optimization problems. This method is theoretically guaranteed to converge *for any initial feasible point* and, under suitable constraint qualifications, achieves a *stationary point* (i.e., a point that satisfies the Karush-Kuhn-Tucker (KKT) conditions) of the original problem.

Furthermore, we study some combinatorial optimization problems in satellite networks (SatNets), which are proven to be *NP-hard*. In particular, we focus on the optimum selection of ground stations (GSs) in SatNets with *site diversity (SD)*, satisfying given availability requirements. SD technique is used to improve the availability of satellite systems by mitigating the atmospheric impairments, such as rain (for radio frequencies) and cloud coverage (for optical frequencies). Moreover, we present *global optimization algorithms*, based on the branch-and-bound (B&B) method and dynamic programming (DP), as well as a *polynomial-time approximation algorithm* with provable performance guarantee.

Finally, we examine a *load-sharing smart gateway diversity (LS-SGD)* architecture in SatNets, which has been recently proposed in the literature. For this diversity scheme, we define the *system outage probability (SOP)* based on the Poisson binomial distribution (PBD) and taking into account the traffic demand as well as the gateway (GW) capacity. In addition, we present several methods for the exact and approximate calculation of SOP.

Keywords: wireless networks, satellite communications, energy efficiency, resource allocation, site diversity, smart gateway diversity, outage probability, ground station selection, sequential convex optimization, combinatorial optimization, computational complexity, NP-hardness, branch-and-bound method, dynamic programming.

Extended Abstract in Greek – Εκτεταμένη Περίληψη

Η ταχεία αύξηση των συνδεδεμένων συσκευών έχει οδηγήσει σε πολύ αυστηρές απαιτήσεις για τα ασύρματα δίκτυα επόμενης γενιάς, λαμβάνοντας υπόψη τόσο *περιβαλλοντικά* όσο και *οικονομικά* ζητήματα. Συγκεκριμένα, ένας από τους πρωταρχικούς στόχους στο σχεδιασμό των ασύρματων δικτύων πέμπτης γενιάς (5G) είναι η ικανοποίηση του εξαιρετικά υψηλού ρυθμού δεδομένων των χρηστών με την ελάχιστη κατανάλωση ενέργειας. Για το σκοπό αυτό, έχει προταθεί ένας νέος δείκτης επίδοσης στη βιβλιογραφία, που ονομάζεται *ενεργειακή απόδοση (ΕΑ)*, ο οποίος μετριέται σε bits/Joule και εκφράζει την ποσότητα πληροφορίας που μπορεί να μεταδοθεί αξιόπιστα ανά μονάδα καταναλισκόμενης ενέργειας.

Η παρούσα Διδακτορική Διατριβή ασχολείται με τη σχεδίαση *αποδοτικών αλγορίθμων βελτιστοποίησης* για ασύρματα δίκτυα επόμενης γενιάς, συμπεριλαμβανομένων των επίγειων καθώς και των δορυφορικών συστημάτων επικοινωνίας. Πιο συγκεκριμένα, εφαρμόζεται η θεωρία της *διαδοχικής κυρτής βελτιστοποίησης (ΔΚΒ)* για την επίλυση δύσκολων προβλημάτων βελτιστοποίησης, όπως η μεγιστοποίηση διαφόρων μετρικών ΕΑ, ώστε να αναπτυχθούν ενεργειακά-αποδοτικές στρατηγικές κατανομής ισχύος. Η ΔΚΒ είναι ένα ισχυρό μαθηματικό εργαλείο που μπορεί να χρησιμοποιηθεί για την επίλυση μη-κυρτών προβλημάτων βελτιστοποίησης επιλύοντας μια ακολουθία κυρτών προβλημάτων βελτιστοποίησης. Αυτή η μέθοδος είναι θεωρητικά εγγυημένη να συγκλίνει *για οποιοδήποτε αρχικό εφικτό σημείο* και, υπό κατάλληλες προϋποθέσεις, επιτυγχάνει ένα *στάσιμο σημείο* (δηλαδή, ένα σημείο που ικανοποιεί τις συνθήκες Karush-Kuhn-Tucker (KKT)) του αρχικού προβλήματος.

Επιπλέον, μελετάμε ορισμένα συνδυαστικά προβλήματα βελτιστοποίησης σε δορυφορικά δίκτυα, τα οποία αποδεικνύεται ότι είναι *NP-δύσκολα*. Συγκεκριμένα, εστιάζουμε στη βέλτιστη επιλογή επίγειων σταθμών (ΕΣ) σε δορυφορικά δίκτυα με *διαφορισμό θέσης (ΔΘ)*, ικανοποιώντας δεδομένες απαιτήσεις διαθεσιμότητας. Η τεχνική ΔΘ χρησιμοποιείται για τη βελτίωση της διαθεσιμότητας των δορυφορικών συστημάτων αμβλύνοντας τα ατμοσφαιρικά φαινόμενα, όπως η βροχή (για τις ραδιοσυχνότητες) και η κάλυψη από νέφη (για τις οπτικές συχνότητες). Ακόμη, παρουσιάζουμε *αλγόριθμους ολικής βελτιστοποίησης*, με βάση τη μέθοδο διακλάδωσης-και-φράγματος (Δ&Φ) και τον δυναμικό προγραμματισμό (ΔΠ), καθώς και έναν *προσεγγιστικό αλγόριθμο πολυωνυμικού χρόνου* με αποδεδειγμένη εγγύηση επίδοσης.

Τέλος, εξετάζουμε μια αρχιτεκτονική *διαφορισμού έξυπνων πυλών (ΔΕΠ) βασισμένη στο διαμοιρασμό φορτίου* σε δορυφορικά δίκτυα, η οποία έχει προταθεί πρόσφατα στη βιβλιογραφία. Για αυτό το σχήμα διαφορισμού, ορίζουμε την *πιθανότητα διακοπής του συστήματος (ΠΔΣ)* βάσει της διωνυμικής κατανομής Poisson και λαμβάνοντας υπόψη τη ζήτηση δεδομένων καθώς και τη χωρητικότητα κάθε πύλης. Επίσης, παρουσιάζουμε διάφορες μεθόδους για τον ακριβή και προσεγγιστικό υπολογισμό της ΠΔΣ.

Στη συνέχεια, παρουσιάζεται μια συνοπτική περιγραφή της Διδακτορικής Διατριβής και των σημαντικότερων συνεισφορών της.

Ενοποιημένη μεθοδολογία για τη μεγιστοποίηση της ενεργειακής απόδοσης σταθμισμένου αθροίσματος σε ασύρματα δίκτυα

Η *ενεργειακή απόδοση σταθμισμένου αθροίσματος (ΕΑΣΑ)* είναι μια βασική μετρική επίδοσης στα ετερογενή δίκτυα, όπου οι κόμβοι ενδέχεται να έχουν διαφορετικές απαιτήσεις ενεργειακής απόδοσης. Παρ' όλα αυτά, η μεγιστοποίηση της ΕΑΣΑ είναι ένα δύσκολο πρόβλημα λόγω της μη-κυρτής μορφής του. Σε αντίθεση με την υφιστάμενη έρευνα, παρουσιάζουμε μια συστηματική προσέγγιση για τη μεγιστοποίηση της ΕΑΣΑ θεωρώντας όχι μόνο περιορισμούς ισχύος, αλλά και περιορισμούς ρυθμού δεδομένων, χρησιμοποιώντας μια γενική έκφραση για το λόγο σήματος-προς-παρεμβολή-και-θόρυβο (SINR). Συγκεκριμένα, το αρχικό πρόβλημα μετασχηματίζεται σε ένα ισοδύναμο πρόβλημα και στη συνέχεια προτείνεται ένας αλγόριθμος *διαδοχικής κυρτής βελτιστοποίησης (ΔΚΒ)*. Αυτός ο αλγόριθμος είναι εγγυημένος να συγκλίνει για οποιοδήποτε αρχικό εφικτό σημείο και, υπό κατάλληλες προϋποθέσεις, επιτυγχάνει μια λύση που ικανοποιεί τις συνθήκες Karush-Kuhn-Tucker (KKT).

Επιπλέον, παρέχουμε αξιοσημείωτες επεκτάσεις της προτεινόμενης μεθοδολογίας, συμπεριλαμβανομένων των συστημάτων με πολλαπλά μπλοκ πόρων, καθώς και ένα γενικότερο μοντέλο κατανάλωσης ισχύος που δεν είναι απαραίτητα κυρτή συνάρτηση των ισχύων εκπομπής. Τέλος, σύμφωνα με την αριθμητική ανάλυση, ο αλγόριθμος παρουσιάζει γρήγορη σύγκλιση, χαμηλή πολυπλοκότητα και ευρωστία στα αρχικά σημεία.

Νέος συμβιβασμός μεταξύ δικαιοσύνης και συνολικής επίδοσης του συστήματος από την άποψη της ενεργειακής απόδοσης

Η *συνολική ενεργειακή απόδοση (ΣΕΑ)*, που ορίζεται ως ο λόγος του συνολικού ρυθμού δεδομένων προς τη συνολική κατανάλωση ενέργειας, θεωρείται η πιο σημαντική μετρική επίδοσης από την άποψη της ενεργειακής απόδοσης (ΕΑ). Ωστόσο, δεν εξαρτάται άμεσα από την ΕΑ κάθε ζεύξης και η μεγιστοποίηση της οδηγεί σε μη-δίκαιη κατανομή ισχύος. Από την άλλη πλευρά, η μεγιστοποίηση της *ελάχιστης ενεργειακής απόδοσης (ΕΕΑ)*, δηλαδή της ελάχιστης ΕΑ όλων των ζεύξεων, εγγυάται την πιο δίκαιη κατανομή ισχύος, όμως δεν περιέχει σαφή πληροφορία σχετικά με τη συνολική επίδοση του συστήματος.

Η κύρια τάση στην τρέχουσα έρευνα είναι η μεγιστοποίηση της ΣΕΑ ή της ΕΕΑ ξεχωριστά. Σε αντίθεση με τις προηγούμενες συνεισφορές, παρουσιάζουμε μια γενική πολυκριτηριακή προσέγγιση για τη βελτιστοποίηση της ΕΑ που λαμβάνει ταυτόχρονα υπόψη τόσο την ΣΕΑ όσο και την ΕΕΑ. Λόγω της μη-κυρτής μορφής του εξεταζόμενου προβλήματος, προτείνουμε έναν αλγόριθμο χαμηλής πολυπλοκότητας που βασίζεται στη θεωρία της *διαδοχικής κυρτής βελτιστοποίησης (ΔΚΒ)*. Τέλος, παρέχουμε ένα νέο θεωρητικό αποτέλεσμα για την πολυπλοκότητα των αλγορίθμων ΔΚΒ.

Ενεργειακά-αποδοτική κατανομή ισχύος σε δορυφορικά συστήματα με πολλαπλές δέσμες

Η κατανάλωση ενέργειας αποτελεί κύριο περιοριστικό παράγοντα για την κατερχόμενη ζεύξη (downlink) στα δορυφορικά συστήματα με πολλαπλές δέσμες, καθώς έχει σημαντικό αντίκτυπο στη μάζα και τη διάρκεια ζωής του δορυφόρου. Σε αυτό το πλαίσιο, μελετάμε ένα νέο πρόβλημα κατανομής ισχύος που στοχεύει στην από κοινού ελαχιστοποίηση της *μη-ικανοποιημένης χωρητικότητας συστήματος* και της *συνολικής ακτινοβολούμενης ισχύος* μέσω της πολυκριτηριακής βελτιστοποίησης.

Κατ' αρχάς, μετασχηματίζουμε το αρχικό μη-κυρτό μη-διαφορίσιμο πρόβλημα σε μια ισοδύναμη μη-κυρτή διαφορίσιμη μορφή εισάγοντας βοηθητικές μεταβλητές. Στη συνέχεια, σχεδιάζουμε έναν αλγόριθμο *διαδοχικής κυρτής προσέγγισης (ΔΚΠ)* προκειμένου να επιτύχουμε ένα *στάσιμο σημείο* με εύλογη πολυπλοκότητα. Λόγω της γρήγορης σύγκλισής του, αυτός ο αλγόριθμος είναι κατάλληλος για δυναμική κατανομή πόρων σε μελλοντικά συστήματα όπου ο δορυφόρος θα μπορεί να προσαρμόζει την ισχύ εκπομπής του. Επιπλέον, αποδεικνύουμε ένα νέο αποτέλεσμα σχετικά με την πολυπλοκότητα της μεθόδου ΔΚΠ, στη γενική περίπτωση, το οποίο συμπληρώνει την υπάρχουσα βιβλιογραφία όπου η πολυπλοκότητα αυτής της μεθόδου αναλύεται μόνο αριθμητικά.

Ολικά βέλτιστη επιλογή επίγειων σταθμών σε δορυφορικά συστήματα με διαφορισμό θέσης

Η διαθεσιμότητα των δορυφορικών συστημάτων επικοινωνίας περιορίζεται σε σημαντικό βαθμό από ατμοσφαιρικά φαινόμενα, όπως η βροχή (για τις ραδιοσυχνότητες) και η κάλυψη από νέφη (για τις οπτικές συχνότητες). Μια λύση σε αυτό το πρόβλημα είναι η τεχνική *διαφορισμού θέσης (ΔΘ)*, όπου ένα δίκτυο από γεωγραφικά κατανεμημένους επίγειους σταθμούς (ΕΣ) μπορεί να διασφαλίσει, με μεγάλη πιθανότητα, ότι τουλάχιστον ένας ΕΣ είναι διαθέσιμος για σύνδεση με τον δορυφόρο σε κάθε χρονική περίοδο. Ωστόσο, η εγκατάσταση περιττών ΕΣ επιφέρει μη-αναγκαίο πρόσθετο κόστος για τον διαχειριστή του δικτύου. Σε αυτό το πλαίσιο, μελετάμε ένα πρόβλημα βελτιστοποίησης που ελαχιστοποιεί τον αριθμό των απαιτούμενων ΕΣ, ικανοποιώντας συγκεκριμένους περιορισμούς διαθεσιμότητας.

Αρχικά, το πρόβλημα μετασχηματίζεται σε πρόβλημα δυαδικού (ακέραιου) γραμμικού προγραμματισμού, το οποίο αποδεικνύεται ότι είναι *NP-δύσκολο.* Στη συνέχεια, σχεδιάζουμε έναν αλγόριθμο *διακλάδωσης-και-φράγματος (Δ&Φ)*, με εγγύηση ολικής βελτιστοποίησης, ο οποίος βασίζεται στη χαλάρωση γραμμικού προγραμματισμού καθώς και σε μια άπληστη μέθοδο. Τέλος, τα αριθμητικά αποτελέσματα δείχνουν ότι ο προτεινόμενος αλγόριθμος υπερτερεί σημαντικά των υφιστάμενων μεθόδων και έχει χαμηλή πολυπλοκότητα μέσης-περίπτωσης.

Ελαχιστοποίηση του κόστους εγκατάστασης των επίγειων σταθμών σε δορυφορικά δίκτυα

Εδώ, μελετάμε τη βέλτιστη επιλογή επίγειων σταθμών (ΕΣ) σε RF/οπτικά δορυφορικά δίκτυα προκειμένου να ελαχιστοποιηθεί το συνολικό κόστος εγκατάστασης υπό δεδομένη απαίτηση πιθανότητας διακοπής, υποθέτοντας ανεξάρτητες καιρικές συνθήκες μεταξύ των ΕΣ. Πρώτα, δείχνουμε ότι το πρόβλημα βελτιστοποίησης μπορεί να διατυπωθεί ως πρόβλημα δυαδικού γραμμικού προγραμματισμού και μετά δίνουμε μια θεωρητική απόδειξη της *NP-σκληρότητας* του. Επιπλέον, σχεδιάζουμε έναν *αλγόριθμο δυναμικού προγραμματισμού* ψευδο-πολυωνυμικής πολυπλοκότητας με εγγύηση ολικής βελτιστοποίησης, καθώς και έναν *προσεγγιστικό αλγόριθμο* πολυωνυμικού χρόνου με αποδεδειγμένη εγγύηση επίδοσης. Τέλος, η επίδοση των προτεινόμενων αλγορίθμων επαληθεύεται μέσω αριθμητικών προσομοιώσεων.

Ακριβής και προσεγγιστικός υπολογισμός της πιθανότητας διακοπής σε δορυφορικά δίκτυα με διαφορισμό έξυπνων πυλών

Η χρησιμοποίηση εξαιρετικά υψηλών συχνοτήτων (ΕΥΣ) μπορεί να επιτύχει πολύ υψηλή ρυθμαπόδοση στα δορυφορικά δίκτυα. Ωστόσο, η σοβαρή εξασθένηση λόγω βροχής στις ΕΥΣ επιβάλλει αυστηρούς περιορισμούς στη διαθεσιμότητα του συστήματος. Ο *διαφορισμός έξυπνων πυλών (ΔΕΠ)* θεωρείται απαραίτητος προκειμένου να διασφαλιστεί η απαιτούμενη διαθεσιμότητα με εύλογο κόστος. Σε αυτό το πλαίσιο, εξετάζουμε μια αρχιτεκτονική *ΔΕΠ βασισμένη στο διαμοιρασμό φορτίου*, η οποία έχει προταθεί πρόσφατα στη βιβλιογραφία. Για αυτό το σχήμα διαφορισμού, ορίζουμε την *πιθανότητα διακοπής του συστήματος (ΠΔΣ)* χρησιμοποιώντας μια αυστηρή πιθανοτική ανάλυση βάσει της διωνυμικής κατανομής Poisson και λαμβάνοντας υπόψη τη ζήτηση δεδομένων καθώς και τη χωρητικότητα κάθε πύλης.

Επιπλέον, παρέχουμε διάφορες μεθόδους για τον ακριβή και προσεγγιστικό υπολογισμό της ΠΔΣ. Όσον αφορά τον ακριβή υπολογισμό της ΠΔΣ, δίνεται μια έκφραση κλειστής μορφής και ένας αλγόριθμος βασισμένος σε έναν αναδρομικό τύπο, και οι δύο με τετραγωνική πολυπλοκότητα ως προς τον αριθμό των πυλών. Τέλος, οι προσεγγιστικές μέθοδοι περιλαμβάνουν γνωστές κατανομές πιθανότητας (διωνυμική, Poisson, κανονική) και ένα φράγμα Chernoff. Σύμφωνα με τα αριθμητικά αποτελέσματα, η διωνυμική και η Poisson κατανομή είναι μακράν οι πιο ακριβείς προσεγγιστικές μέθοδοι.

Το κύριο μέρος αυτής της Διατριβής μπορεί να χωριστεί σε τρία μέρη. Το 1^o μέρος ασχολείται με *στρατηγικές κατανομής ισχύος* και περιλαμβάνει τα Κεφάλαια 2, 3 (μεγιστοποίηση ενεργειακής απόδοσης σε ασύρματα δίκτυα) καθώς και το Κεφάλαιο 4 (βελτιστοποίηση δορυφορικών συστημάτων λαμβάνοντας υπόψη την κατανάλωση ενέργειας). Το 2^o μέρος μελετά τη *βέλτιστη επιλογή επίγειων σταθμών* σε RF/οπτικά δορυφορικά δίκτυα με διαφορισμό θέσης και αποτελείται από τα Κεφάλαια 5 και 6. Το 3^o μέρος είναι το Κεφάλαιο 7, το οποίο καλύπτει την *τεχνική διαφορισμού έξυπνων πυλών (ΔΕΠ) βασισμένη στο διαμοιρασμό φορτίου* σε δορυφορικά συστήματα. Πιο συγκεκριμένα, η Διδακτορική Διατριβή είναι οργανωμένη ως εξής.

Αρχικά, το Κεφάλαιο 1 αποτελεί την εισαγωγή και περιλαμβάνει: το κίνητρο και το σκοπό, μια σύνοψη των κύριων συνεισφορών και τη δομή της Διδακτορικής Διατριβής.

Στο Κεφάλαιο 2, παρουσιάζουμε ένα πλαίσιο για τη μεγιστοποίηση της ενεργειακής απόδοσης σταθμισμένου αθροίσματος (ΕΑΣΑ) σε συστήματα ασύρματης επικοινωνίας, θεωρώντας μια γενική έκφραση του λόγου σήματος-προς-παρεμβολή-και-θόρυβο (SINR) που περιλαμβάνει παρεμβολή τόσο από τους υπόλοιπους χρήστες όσο και από τον ίδιο τον χρήστη (αυτο-παρεμβολή). Ειδικότερα, προτείνεται ένας αλγόριθμος διαδοχικής κυρτής βελτιστοποίησης (ΔΚΒ) και παρέχονται επίσης αξιοσημείωτες επεκτάσεις που απορρέουν από αυτήν τη μεθοδολογία.

Στο Κεφάλαιο 3, εισάγεται ένας νέος συμβιβασμός μεταξύ της δικαιοσύνης και της συνολικής επίδοσης του συστήματος σε όρους ενεργειακής απόδοσης. Συγκεκριμένα, παρουσιάζουμε μια γενική πολυκριτηριακή προσέγγιση για τη βελτιστοποίηση της ενεργειακής απόδοσης που λαμβάνει υπόψη τη συνολική ενεργειακή απόδοση (ΣΕΑ) καθώς και την ελάχιστη ενεργειακή απόδοση (ΕΕΑ). Επιπλέον, σχεδιάζουμε έναν αλγόριθμο χαμηλής πολυπλοκότητας χρησιμοποιώντας τη θεωρία της διαδοχικής κυρτής βελτιστοποίησης (ΔΚΒ) προκειμένου να αντιμετωπίσουμε το μη-κυρτό πρόβλημα. Τέλος, δίνεται ένα θεωρητικό αποτέλεσμα για την πολυπλοκότητα των αλγορίθμων ΔΚΒ.

Στο Κεφάλαιο 4, μελετάμε ένα πρόβλημα ενεργειακά-αποδοτικής κατανομής ισχύος σε δορυφορικά συστήματα με πολλαπλές δέσμες, το οποίο στοχεύει στην από κοινού ελαχιστοποίηση της μη-ικανοποιημένης χωρητικότητας συστήματος και της συνολικής ακτινοβολούμενης ισχύος μέσω της βελτιστοποίησης πολλαπλών κριτηρίων. Συγκεκριμένα, σχεδιάζουμε έναν αλγόριθμο διαδοχικής κυρτής προσέγγισης (ΔΚΠ) προκειμένου να επιτύχουμε ένα στάσιμο σημείο με χαμηλή πολυπλοκότητα. Επιπλέον, αποδεικνύουμε ένα νέο αποτέλεσμα σχετικά με την πολυπλοκότητα της μεθόδου ΔΚΠ (σημειώνεται ότι *η ΔΚΠ είναι μια ειδική περίπτωση της ΔΚΒ, όπου όλες οι συναρτήσεις του αρχικού προβλήματος μπορούν να εκφραστούν ως διαφορά δύο κυρτών συναρτήσεων*).

Στο Κεφάλαιο 5, εστιάζουμε στην ελαχιστοποίηση του αριθμού των επίγειων σταθμών ικανοποιώντας δεδομένες απαιτήσεις διαθεσιμότητας σε δορυφορικά συστήματα με διαφορισμό θέσης. Αρχικά, αποδεικνύουμε ότι το πρόβλημα βελτιστοποίησης είναι NP-δύσκολο και, στη συνέχεια, σχεδιάζουμε έναν αλγόριθμο διακλάδωσης-και-φράγματος (Δ&Φ) με εγγύηση ολικής βελτιστοποίησης και χαμηλή πολυπλοκότητα μέσης-περίπτωσης.

Το Κεφάλαιο 6 έχει να κάνει με τη βέλτιστη επιλογή επίγειων σταθμών σε δορυφορικά συστήματα ώστε να ελαχιστοποιηθεί το συνολικό κόστος εγκατάστασης, υπό δεδομένη απαίτηση πιθανότητας διακοπής. Αυτό το πρόβλημα βελτιστοποίησης αποδεικνύεται θεωρητικά ότι είναι NP-δύσκολο. Επιπλέον, παρουσιάζονται ένας αλγόριθμος ολικής βελτιστοποίησης που βασίζεται στο δυναμικό προγραμματισμό (ΔΠ) και ένας προσεγγιστικός αλγόριθμος πολυωνυμικού χρόνου.

Το Κεφάλαιο 7 αφιερώνεται στην ανάλυση του διαφορισμού έξυπνων πυλών (ΔΕΠ) βασισμένου στο διαμοιρασμό φορτίου σε δορυφορικά δίκτυα. Συγκεκριμένα, ορίζουμε την πιθανότητα διακοπής του συστήματος (ΠΔΣ) βάσει της διωνυμικής κατανομής Poisson και λαμβάνοντας υπόψη τόσο τη ζήτηση δεδομένων όσο και τη χωρητικότητα κάθε πύλης. Επιπλέον, παρέχουμε διάφορες μεθόδους για τον ακριβή και προσεγγιστικό υπολογισμό της ΠΔΣ.

Τέλος, το Κεφάλαιο 8 ολοκληρώνει τη Διατριβή με μια γενική περίληψη των συνεισφορών της και μια παρουσίαση ανοιχτών προβλημάτων, ανοίγοντας το δρόμο για μελλοντική έρευνα.

Εν κατακλείδι, παρουσιάζουμε τα γενικά συμπεράσματα της Διδακτορικής Διατριβής.

Σε αυτή τη Διατριβή, θέτοντας τον *άνθρωπο* και τη *φύση* ως βασικούς πυλώνες, έχουμε ασχοληθεί με τη σχεδίαση των ασύρματων δικτύων έχοντας επίγνωση του φυσικού περιβάλλοντος, πράγμα το οποίο αδιαμφισβήτητα αποτελεί μια νέα κατεύθυνση έρευνας. Συγκεκριμένα, οι προτεινόμενοι αλγόριθμοι βελτιστοποίησης της ενεργειακής απόδοσης στοχεύουν στην ικανοποίηση του ρυθμού δεδομένων των χρηστών με την ελάχιστη κατανάλωση ενέργειας. Με άλλα λόγια, ο πρωταρχικός στόχος είναι η δημιουργία *πράσινων δικτύων επικοινωνίας* τα οποία μπορούν να παρέχουν υψηλής-ποιότητας υπηρεσίες, διατηρώντας παράλληλα την ηλεκτρομαγνητική ακτινοβολία σε ασφαλή επίπεδα και μειώνοντας τις εκπομπές διοξειδίου του άνθρακα (*χαμηλό αποτύπωμα άνθρακα*). Επιπλέον, οι λειτουργικές δαπάνες των παρόχων τηλεπικοινωνιακών υπηρεσιών καθώς και η μάζα των δορυφόρων μπορούν να μειωθούν σημαντικά. Επίσης, οι σχεδιαζόμενοι αλγόριθμοι έχουν τη δυνατότητα να παρατείνουν τη διάρκεια ζωής της μπαταρίας των συσκευών των χρηστών και μπορούν να χρησιμοποιηθούν σε εφαρμογές με αυστηρές απαιτήσεις υπολογιστικού χρόνου (λόγω της *χαμηλής πολυπλοκότητας* και της *γρήγορης σύγκλισης* τους). Σε κάθε περίπτωση, η συνεισφορά της Διατριβής είναι μόνο *ένα μικρό κομμάτι του παζλ* και θα πρέπει να συνδυαστεί με περαιτέρω έρευνα ώστε αυτή η επιστημονική πρόκληση να γίνει πραγματικότητα.

Επιπρόσθετα, έχουμε αποδείξει ότι η βέλτιστη επιλογή επίγειων σταθμών σε RF/οπτικά δορυφορικά δίκτυα με διαφορισμό θέσης (υπό περιορισμούς διαθεσιμότητας) είναι ένα *NP-δύσκολο* πρόβλημα. Ακόμη, έχουμε αναπτύξει *αλγόριθμους ολικής βελτιστοποίησης (μέθοδος Δ&Φ και ΔΠ)* καθώς και έναν *προσεγγιστικό αλγόριθμο πολυωνυμικού χρόνου* με αποδεδειγμένη εγγύηση επίδοσης. Αυτοί οι αλγόριθμοι θα μπορούσαν να είναι χρήσιμοι κατά τον αρχικό σχεδιασμό του δικτύου, αφού είναι σε θέση να παρέχουν σημαντική εξοικονόμηση κόστους όσον αφορά την εγκατάσταση των επίγειων σταθμών. Τέλος, έχουμε μελετήσει λεπτομερώς την επίδοση του *διαφορισμού έξυπνων πυλών βασισμένου στο διαμοιρασμό φορτίου* σε δορυφορικά δίκτυα, παρουσιάζοντας διάφορες μεθόδους για τον υπολογισμό της *πιθανότητας διακοπής του συστήματος*.

Λέξεις Κλειδιά: ασύρματα δίκτυα, δορυφορικές επικοινωνίες, ενεργειακή απόδοση, κατανομή πόρων, διαφορισμός θέσης, διαφορισμός έξυπνων πυλών, πιθανότητα διακοπής, επιλογή επίγειου σταθμού, διαδοχική κυρτή βελτιστοποίηση, συνδυαστική βελτιστοποίηση, υπολογιστική πολυπλοκότητα, NP-σκληρότητα, μέθοδος διακλάδωσης-και-φράγματος, δυναμικός προγραμματισμός.

Glossary of Technical Terms – Γλωσσάρι Τεχνικών Όρων

5G Wireless Networks: Ασύρματα Δίκτυα 5ης Γενιάς

Adaptive Coding and Modulation: Προσαρμοστική Κωδικοποίηση και Διαμόρφωση

Approximation Algorithm: Προσεγγιστικός Αλγόριθμος

Availability: Διαθεσιμότητα

Base Station: Σταθμός Βάσης

Branch-and-Bound Method: Μέθοδος Διακλάδωσης-και-Φράγματος

Computational Complexity: Υπολογιστική Πολυπλοκότητα

Cumulative Distribution Function: Αθροιστική Συνάρτηση Κατανομής

Data Rate: Ρυθμός Δεδομένων

Device-to-Device: Συσκευή-προς-Συσκευή

Downlink: Κατερχόμενη Ζεύξη

Dynamic Programming: Δυναμικός Προγραμματισμός

Energy Efficiency: Ενεργειακή Απόδοση

Fade Mitigation Techniques: Τεχνικές Άμβλυνσης Διαλείψεων

Generating Function: Γεννήτρια Συνάρτηση

Geostationary (GEO) Satellite: Γεωστατικός Δορυφόρος

Ground Station: Επίγειος Σταθμός

Information and Communications Technology: Τεχνολογία της Πληροφορικής και των Επικοινωνιών

Integer Programming: Ακέραιος Προγραμματισμός

Linear Programming: Γραμμικός Προγραμματισμός

Low-Earth-Orbit (LEO) Satellite: Δορυφόρος Χαμηλής Τροχιάς

Medium-Earth-Orbit (MEO) Satellite: Δορυφόρος Μέσης Τροχιάς

Multibeam Satellite Systems: Δορυφορικά Συστήματα με Πολλαπλές Δέσμες

Multi-Objective Optimization: Πολυκριτηριακή Βελτιστοποίηση

Orthogonal Frequency Division Multiple Access: Ορθογώνια Διαίρεση Συχνότητας Πολλαπλής Πρόσβασης

Outage Probability: Πιθανότητα Διακοπής

Probability Density Function: Συνάρτηση Πυκνότητας Πιθανότητας

Probability Mass Function: Συνάρτηση Μάζας Πιθανότητας

Quality of Service: Ποιότητα Υπηρεσίας

Radio Frequency: Ραδιοσυχνότητα

Random Variable: Τυχαία Μεταβλητή

Recursive Formula: Αναδρομικός Τύπος

Satellite Communications: Δορυφορικές Επικοινωνίες

Sequential Convex Optimization: Διαδοχική Κυρτή Βελτιστοποίηση

Signal-to-Interference-plus-Noise Ratio: Λόγος Σήματος-προς-Παρεμβολή-και-Θόρυβο

Site Diversity: Διαφορισμός Θέσης

Smart Gateway Diversity: Διαφορισμός Έξυπνων Πυλών

Successive Convex Approximation: Διαδοχική Κυρτή Προσέγγιση

Unmet System Capacity: Μη-ικανοποιημένη Χωρητικότητα Συστήματος

Uplink: Ανερχόμενη Ζεύξη

User Equipment: Συσκευή Χρήστη

Acknowledgements

During my PhD journey, I have learned two things. Firstly, it is important to succeed, but even more important to struggle and keep trying, because we gain new insights through this process. Secondly, failure is (almost always) the inevitable road to success; in other words, we usually face some (or, sometimes, many) obstacles and difficulties until we reach the desired goal. I think that scientific research, and so this Dissertation, is no exception. At this point, I want to express my gratitude to the people who have contributed in making the last four and a half years a unique experience and helped me, in their own way, to complete my PhD.

First of all, I am deeply grateful to my supervisor, Prof. Athanasios D. Panagopoulos (or, simply, Thanasis), for supporting me during this journey and giving me the necessary freedom and guidance to follow my own research interests. His enthusiasm, encouragement, advice, understanding and his belief in me were essential to overcome several obstacles and accomplish this work. But most importantly, I thank him for being there as a friend, who was always willing to help me in every possible way.

Moreover, I am indebted to Profs. Panagiotis Cottis, George Fikioris, Christos Capsalis, Symeon Papavassiliou, Ioanna Roussaki and George Efthymoglou for being members of my examination committee, their useful advice and insightful comments during the PhD Thesis defense. I would also like to thank the A. G. Leventis Foundation and the Special Account for Research Funding (ELKE) of NTUA for their financial support during my research (2016–2017 and 2017–2021, respectively) as well as the Computer Center of NTUA for providing High Performance Computing (HPC) infrastructure required for some numerical simulations.

In addition, I would like to thank my colleagues, collaborators and friends at NTUA, Charilaos Kourogiorgas, Stavros Sagkriotis, Nikolaos Lyras, Apostolos Papafragkakis, Argyris Roumeliotis, Theodore Kapsis, Antonis Savva, Neoklis Hadjigeorgiou, Michalis Panayi and Theofanis Petroudis, as well as all my close friends in Cyprus for interesting discussions, their help and support, and the precious moments they shared with me throughout this journey. Moreover, I feel very grateful to my high-school teacher, Mr. Alexandros Baros, for inspiring my love for mathematics.

Furthermore, I am deeply thankful to my (girl)friend, Agkita, for her substantial support in both good and bad moments as well as her continuous encouragement and constant belief in me. Thank you for your love, patience, and the beautiful moments we spent together. I would also like to express my sincere thanks to my close relatives (grandparents, uncles, aunts and cousins) who kindly supported me all these years.

Finally, I am deeply grateful to my family, since this Dissertation would not have been possible without them. Special thanks to my parents, Nikos and Maria, and to my sister, Charoula, for their unconditional love, understanding, support and continuous encouragement since the day I was born. I believe that my parents have taught me (through their actions and lifestyle, not just words) the importance of simplicity, perseverance, strength and patience that I really appreciate, because they particularly help me to achieve my goals. Thank you all for the wonderful memories, the time we spent thinking together, and your invaluable presence in my life!

Christos N. Efrem

March 2021

To my parents, Nikos and Maria,
and to my sister, Charoula.

Contents

List of Figures

List of Tables

List of Algorithms

Abbreviations

5G: 5^{th} Generation

ACM: Adaptive Coding and Modulation

B&B: Branch-and-Bound

BA: Binomial Approximation

BER: Bit Error Rate

BILP: Binary Integer Linear Programming

BS: Base Station

CB: Chernoff Bound

CC: Concave-Convex

CCDF: Complementary Cumulative Distribution Function

CCI: Co-Channel Interference

CDF: Cumulative Distribution Function

CF: Cost Function

CFE: Closed-Form Expression

CLT: Central Limit Theorem

D2D: Device-to-Device

DFT: Discrete Fourier Transform

DP: Dynamic Programming

DPAA: DP-based Approximation Algorithm

DRA: Data Rate Adaptation

EE: Energy Efficiency

EHF: Extremely High Frequency

FFT: Fast Fourier Transform

FIFO: First-In First-Out

FMT: Fade Mitigation Technique

FPTAS: Fully Polynomial-Time Approximation Scheme

FSPL: Free-Space Path Loss

GEO: Geostationary Orbit

GF: Generating Function

GS: Ground Station

GW: Gateway

HTS: High Throughput Satellite

i.i.d.: independent and identically distributed

ICT: Information and Communications Technology

IEEE: Institute of Electrical and Electronics Engineers

iff: if and only if

JFI: Jain's Fairness Index

KKT: Karush-Kuhn-Tucker

LEO: Low Earth Orbit

LP: Linear Programming

MAE: Mean Absolute Error

maxAE: maximum Absolute Error

MEE: Minimum Energy Efficiency

MEO: Medium Earth Orbit

MIMO: Multiple-Input Multiple-Output

MISO: Multiple-Input Single-Output

MNCP: Minimum Node Cover Problem

MOO: Multi-Objective Optimization

MRC: Maximum-Ratio Combining

MSS: Multibeam Satellite System

NA: Normal Approximation

NCC: Network Control Center

NTUA: National Technical University of Athens

OFDMA: Orthogonal Frequency Division Multiple Access

OPEX: Operational Expenditure

PA: Poisson Approximation

PAPR: Peak-to-Average Power Ratio

PBD: Poisson Binomial Distribution

PDF: Probability Density Function

PEE: Product Energy Efficiency

PMF: Probability Mass Function

QoS: Quality of Service

RB: Resource Block

RF: Radio Frequency / Recursive Formula

RMSE: Root-Mean-Square Error

RNA: Refined Normal Approximation

RV: Random Variable

s.t.: subject to

SA: System Availability

SatNet: Satellite Network

SC: Subcarrier

SCA/CCP: Successive Convex Approximation / Concave-Convex Procedure

SCO: Sequential Convex Optimization

SD: Site Diversity

SEE: Sum Energy Efficiency

SGD: Smart Gateway Diversity

SINR: Signal-to-Interference-plus-Noise Ratio

SOP: System Outage Probability

TDD: Time Division Duplex

TEE/GEE: Total/Global Energy Efficiency

UE: User Equipment

ULPC: Uplink Power Control

UPA: Uniform Power Allocation

USC: Unmet System Capacity

WM: Weighted Minimum

WP: Weighted Product

WSEE: Weighted-Sum Energy Efficiency

WSR: Weighted-Sum Rate

Chapter 1

Introduction

1.1 Motivation and Scope

As the number of connected devices is expected to increase significantly in the next few years, energy consumption has become a fundamental issue in the design of fifth-generation (5G) wireless networks. Specifically, one of the primary requirements is to achieve extremely higher data rates compared to the existing cellular systems [1]. Obviously, increasing accordingly the transmit power would give rise to prohibitively high energy demand. As a result, the network energy efficiency has to be considerably improved in order to achieve this goal.

Furthermore, *environmental concerns* impose power control strategies that take into account the energy consumption of wireless communication systems. In particular, information and communications technology (ICT) causes a significant amount of the global carbon-dioxide (CO_2) emissions nowadays [2, 3]. The situation may deteriorate, since the number of connected devices grows exponentially. Moreover, given the high capacity requirements of 5G networks, electromagnetic radiation will exceed safety limits if the appropriate measures are not taken.

Apart from the ecological concerns, *economic reasons* related to energy cost are crucial for both telecommunication service providers and users. In this context, energy efficiency optimization plays an important role, because it can reduce the operational expenditure (OPEX) and prolong the battery lifetime of users' devices as well. In addition, energy consumption is a major limitation in the downlink of satellite systems, since it has a great impact on the mass and lifetime of satellites.

Energy efficiency (EE) is a key performance indicator for 5G networks which is measured in bits/Joule and expresses the amount of information that can be reliably transmitted per unit of consumed energy [1, 4]. This performance indicator is widely used in the literature for several types of wireless networks. In [5], for instance, the EE is maximized in order to determine the transmit powers in a multi-carrier system. Further studies that consider the concept of EE are [6–12] for orthogonal-frequency-division-multiple-access (OFDMA) networks, [13–18] for multiple-input multiple-output (MIMO) systems, [19] and [20] for device-to-device (D2D) communications, [21] for relay-assisted systems, [22] and [23] for cognitive networks, and [24] for distributed antenna systems.

Moreover, a unified framework for the design of both centralized and decentralized (distributed) energy-efficient power allocation strategies is proposed in [25]. A distributed approach for EE optimization is also presented in [26] and [27]. As regards spectrum-sharing networks with one common frequency channel, the authors in [28] investigate power control mechanisms for maximizing proportional, max-min and harmonic fair EE. By applying appropriate transformation, each of the three problems, which is initially nonconvex, can be converted into an equivalent convex problem and then globally solved by standard convex optimization methods.

The *total/global EE* is defined as the network benefit-cost ratio (the total data rate divided by the total power consumption) and is considered the most meaningful EE metric. However, the total EE does not depend directly on each link's EE and its maximization results in low fairness between the links from the perspective of EE, because it tends to favor the links with better propagation conditions [4,29]. An alternative way to study EE is through multi-objective optimization (MOO) by defining a goal function that explicitly depends on the links' energy efficiencies. MOO is a mathematical tool to solve optimization problems with multiple conflicting objectives [30]. Following this approach, we can define the *weighted-sum EE*, the *weighted-product EE* and the *weighted-minimum EE*. If all weights are equal, then the fairest optimal solution in terms of EE is achieved by maximizing the weighted-minimum EE [29]. Nevertheless, none of these three goal functions contains explicit information about the total system performance, i.e., the total EE.

In addition to terrestrial networks, *satellite communication systems* offer worldwide coverage and connectivity by providing telecommunication services to users in rural and remote areas, where the terrestrial networks are not able to do so; maritime and aerial users benefit from this large coverage as well. In mobile-user scenarios, satellite broadcasting can be used to offload the terrestrial network, thus reducing the backhaul requirements. Furthermore, satellites are a key technology for broadband services (e.g., distance learning/education, especially for developing countries with limited terrestrial internet access) and e-health that ensures high-quality care for patients (i.e., accurate diagnosis by experts in a short time). Other applications include earth observation (e.g., remote sensing, landscape imaging, weather forecasting), global navigation/positioning/tracking (for cars, ships and aircrafts) as well as emergency management services for better responses to natural and man-made disasters (e.g., timely alerts, pre-event preparation, and post-event recovery) [31,32].

Recently, the traffic demand in high throughput satellite (HTS) systems has approached the Tbps, so the challenge is to utilize more wisely the available resources (e.g., bandwidth, transmit power). Shifting to higher frequency bands in order to achieve more spectrum, the signal experiences higher attenuation. The standard fade mitigation techniques (FMTs), such as uplink power control (ULPC) and adaptive coding and modulation (ACM), are not sufficient to cope with the severe signal degradation. So far, there are two solutions to this problem: the (classical) *site diversity (SD)* and the *smart gateway diversity (SGD)* techniques, which can achieve very high network availability at the expense of installing additional ground stations (GSs). An extremely high availability is of paramount importance, especially for safety/security applications and critical/emergency communications in order to support timely rescue efforts during natural disasters (e.g., floods, earthquakes and hurricanes), where human life is in danger. Since terrestrial networks are frequently affected and disrupted in case of natural disasters, satellites are the preferred medium for communication due to their resilience to ground events.

In this PhD Thesis, we leverage the theory of *sequential convex optimization (SCO)* in order to tackle nonconvex optimization problems in next-generation wireless networks, including terrestrial and satellite systems; this technique has been used in [33, 34] for data rate maximization and in [12, 15, 18, 25, 35] for EE maximization as well. In particular, SCO is a powerful mathematical tool that can be used to solve difficult (nonconvex) optimization problems by solving a sequence of easier (convex) optimization problems [36]. Although this method does not guarantee global optimality, it converges to a point that satisfies the Karush-Kuhn-Tucker (KKT) conditions (i.e., a stationary point) of the original problem with affordable computational complexity. Furthermore, some advanced algorithm design techniques, the *branch-and-bound (B&B)* method and *dynamic programming (DP)*, are used in order to find globally optimal solutions to NP-hard combinatorial problems in satellite networks (SatNets) with SD. Finally, we analyze the performance of load-sharing SGD-based SatNets using *probability theory*.

1.2 Overview of Main Contributions

The main contributions of this Dissertation are summarized as follows:

1. ***A framework for weighted-sum energy efficiency maximization in wireless networks:*** *Weighted-sum energy efficiency (WSEE)* is a key performance metric in heterogeneous networks, where the nodes may have different energy-efficiency requirements. Nevertheless, WSEE maximization is a challenging problem due to its nonconvex sum-of-ratios form. Unlike previous work, we present a systematic approach to WSEE maximization under not only power constraints, but also data rate constraints, using a general signal-to-interference-plus-noise-ratio (SINR) expression. In particular, a *sequential convex optimization (SCO)* algorithm is proposed, which is theoretically guaranteed to converge for any initial feasible point, and, under suitable constraint qualifications, achieves a Karush-Kuhn-Tucker (KKT) solution with low complexity.

Furthermore, we provide remarkable extensions of the proposed methodology, including systems with multiple resource blocks as well as a general power consumption model, which is not necessarily a convex function of the transmit powers. Finally, numerical analysis reveals that the proposed algorithm exhibits fast convergence, low complexity, and robustness (insensitivity to initial points).

2. ***A new trade-off between fairness and total system performance in terms of energy efficiency:*** The *total energy efficiency (TEE)*, defined as the ratio between the total data rate and the total power consumption, is considered the most meaningful performance metric. Nevertheless, it does not depend directly on the EE of each link and its maximization leads to unfairness between the links. On the other hand, the maximization of the *minimum energy efficiency (MEE)*, i.e., the minimum of the EEs of all links, guarantees the fairest power allocation, but it does not contain any explicit information about the total system performance. The main trend in current research is to maximize TEE and MEE separately. Unlike previous contributions, we present a general multi-objective approach for EE optimization that takes into account both TEE and MEE at the same time, and thus achieves various trade-off points in the MEE-TEE plane. In this way, network designers are able to make a compromise between fairness and total system performance according to their needs and preferences. Due to the nonconvex form of the resulting problem, we propose a low-complexity algorithm using the theory of *sequential convex optimization (SCO)*. Last but not least, we provide a novel theoretical result for the complexity of SCO algorithms.

3. ***Dynamic energy-efficient power allocation in multibeam satellite systems:*** Power consumption is a major limitation in the downlink of multibeam satellite systems, since it has a significant impact on the mass and lifetime of the satellite. In this context, we study a new energy-aware power allocation problem that aims to jointly minimize the *unmet system capacity (USC)* and

the *total radiated power* by means of multi-objective optimization. First, we transform the original nonconvex-nondifferentiable problem into an equivalent nonconvex-differentiable form by introducing auxiliary variables. Subsequently, we design a *successive convex approximation (SCA)* algorithm in order to attain a stationary point with reasonable complexity. Due to its fast convergence, this algorithm is suitable for dynamic resource allocation in emerging on-board processing technologies. In addition, we formally prove a new result about the complexity of the SCA method, in the general case, that complements the existing literature where the complexity of this method is only numerically analyzed.

4. ***Globally optimal selection of ground stations in satellite systems with site diversity:*** The availability of satellite communication systems is extremely limited by atmospheric impairments, such as rain (for radio frequencies) and cloud coverage (for optical frequencies). A solution to this problem is the *site diversity* technique, where a network of geographically distributed ground stations (GSs) can ensure, with high probability, that at least one GS is available for connection to the satellite at each time period. However, the installation of redundant GSs induces unnecessary additional costs for the network operator. In this context, we study an optimization problem that minimizes the number of required GSs, subject to availability constraints. First, the problem is transformed into a binary-integer-linear-programming (BILP) problem, which is proven to be *NP-hard*. Subsequently, we design a *branch-and-bound (B&B)* algorithm, with global-optimization guarantee, based on the linear-programming (LP) relaxation and a greedy method as well. Finally, numerical results show that the proposed algorithm significantly outperforms state-of-the-art methods and has low complexity in the average case.

5. ***Minimization of the installation cost of ground stations in satellite networks:*** Here, we study the optimum selection of ground stations (GSs) in RF/optical satellite networks (SatNets) in order to minimize the overall installation cost under an outage probability requirement, assuming independent

weather conditions between sites. First, we show that the optimization problem can be formulated as a binary-linear-programming problem, and then we give a formal proof of its *NP-hardness*. Furthermore, we design a *dynamic-programming algorithm* of pseudo-polynomial complexity with global optimization guarantee as well as an efficient (polynomial-time) *approximation algorithm* with provable performance guarantee on the distance of the achieved objective value from the global optimum. Finally, the performance of the proposed algorithms is verified through numerical simulations.

6. ***Computation and approximation of outage probability in satellite networks with smart gateway diversity:*** The utilization of extremely high frequency (EHF) bands can achieve very high throughput in satellite networks (SatNets). Nevertheless, the severe rain attenuation at EHF bands imposes strict limitations on the system availability. *Smart gateway diversity (SGD)* is considered indispensable in order to guarantee the required availability with reasonable cost. In this context, we examine a *load-sharing SGD (LS-SGD)* architecture, which has been recently proposed in the literature. For this diversity scheme, we define the *system outage probability (SOP)* using a rigorous probabilistic analysis based on the Poisson binomial distribution (PBD), and taking into consideration the traffic demand as well as the gateway (GW) capacity. Furthermore, we provide several methods for the exact and approximate calculation of SOP. As concerns the exact computation of SOP, a closed-form expression and an efficient algorithm based on a recursive formula are given, both with quadratic worst-case complexity in the number of GWs. Finally, the proposed approximation methods include well-known probability distributions (binomial, Poisson, normal) and a Chernoff bound. According to the numerical results, binomial and Poisson distributions are by far the most accurate approximation methods.

1.3 Thesis Outline

The core of this Dissertation can be divided into three parts. The 1st part deals with *power allocation strategies* and includes Chapters 2, 3 (energy efficiency maximization in wireless networks) as well as Chapter 4 (energy-aware optimization in satellite systems). The 2nd part studies the *optimum selection of ground stations* in RF/optical satellite networks with site diversity and consists of Chapters 5 and 6. The 3rd part is Chapter 7, which covers the *load-sharing smart gateway diversity (LS-SGD)* technique in satellite systems. More specifically, the rest of this PhD Thesis is organized as follows.

In Chapter 2, we present a framework for weighted-sum energy efficiency (WSEE) maximization in wireless communication systems, considering a general signal-to-interference-plus-noise-ratio (SINR) expression which includes inter-user interference as well as self-interference terms. In particular, a sequential convex optimization (SCO) algorithm is proposed and remarkable extensions stemming from this methodology are also provided.

In Chapter 3, a new trade-off between fairness and total system performance, in terms of EE, is introduced. In particular, we present a general multi-objective approach for EE optimization that takes into consideration both the total energy efficiency (TEE) and the minimum energy efficiency (MEE) at the same time. Moreover, we develop a low-complexity algorithm using the theory of sequential convex optimization (SCO) in order to address the resulting (nonconvex) problem. Finally, a theoretical result for the complexity of SCO algorithms is given.

In Chapter 4, we study an energy-efficient power allocation problem in multibeam satellite systems, which aims to jointly minimize the unmet system capacity (USC) and the total radiated power by means of multi-objective optimization. Specifically, we design a successive convex approximation (SCA) algorithm in order to achieve a stationary point with reasonable complexity. In addition, we prove a new result about the complexity of the SCA method (note that *SCA is a special case of SCO, where the objective and constraint functions of the original problem can be written as* *the difference of two convex functions*).

In Chapter 5, we focus on the minimization of the number of ground stations (GSs) satisfying given availability requirements in satellite systems with site diversity. First, we show that the optimization problem is NP-hard, and then we design a branch-and-bound (B&B) algorithm with global-optimization guarantee and low average-case complexity.

Chapter 6 has to do with the optimal selection of GSs in satellite systems so as to minimize the total installation cost, under a given outage probability requirement. This optimization problem is theoretically proven to be NP-hard. Furthermore, a global optimization algorithm based on dynamic programming (DP) and a polynomial-time approximation algorithm are presented.

Chapter 7 is devoted to the analysis of a load-sharing smart gateway diversity (LS-SGD) scheme in satellite networks. In particular, we define the system outage probability (SOP) based on the Poisson binomial distribution (PBD) and taking into consideration the traffic demand as well as the gateway (GW) capacity. Moreover, we provide several methods for the exact and approximate calculation of SOP.

Finally, Chapter 8 concludes the Dissertation with a general summary of its contributions and a presentation of open problems, paving the way for future work.

Bibliography

[1] J. G. Andrews, S. Buzzi,W. Choi, S. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065-1082, June 2014.

[2] A. Fehske, J. Malmodin, G. Biczok, and G. Fettweis, "The global footprint of mobile communications: The ecological and economic perspective," *IEEE Commun. Mag.*, pp. 55-62, August 2011.

[3] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, O. Blume, and A. Fehske, "How much energy is needed to run a wireless network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40-49, 2011.

[4] S. Buzzi, I. Chih-Lin, T. E. Klein, V. Poor, C. Yang, and A. Zappone, "A survey of energy-efficient techniques for 5G networks and challenges ahead," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 697-709, 2016.

[5] C. Isheden, Z. Chong, E. A. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2946-2957, August 2012.

[6] K. T. K. Cheung, S. Yang and L. Hanzo, "Achieving maximum energy-efficiency in multi-relay OFDMA cellular networks: A fractional programming approach," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2746-2757, July 2013.

[7] Q. Xu, X. Li, H. Ji and X. Du, "Energy-efficient resource allocation for heterogeneous services in OFDMA downlink networks: Systematic perspective," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2071-2082, Jun 2014.

[8] D. W. K. Ng, E. S. Lo and R. Schober, "Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3292-3304, September 2012.

[9] Y. Li, M. Sheng, C. W. Tan, Y. Zhang, Y. Sun, X. Wang, Y. Shi and J. Li, "Energy-efficient subcarrier assignment and power allocation in OFDMA systems with max-min fairness guarantees," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3183-3195, Sept. 2015.

[10] D. W. K. Ng, E. S. Lo and R. Schober, "Energy-efficient resource allocation in multi-cell OFDMA systems with limited backhaul capacity," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3618-3631, October 2012.

[11] S. Kim, B. G. Lee and D. Park, "Energy-per-bit minimized radio resource allocation in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1862-1873, April 2014.

[12] L. Venturino, A. Zappone, C. Risi and S. Buzzi, "Energy-efficient scheduling and power allocation in downlink OFDMA networks with base station coordination," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 1-14, Jan. 2015.

[13] S. He, Y. Huang, S. Jin and L. Yang, "Coordinated beamforming for energy efficient transmission in multicell multiuser systems," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 4961-4971, December 2013.

[14] S. He, Y. Huang, L. Yang and B. Ottersten, "Coordinated multicell multiuser precoding for maximizing weighted sum energy efficiency," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 741-751, Feb.1, 2014.

[15] A. Zappone, E. A. Jorswieck and S. Buzzi, "Energy efficiency and interference neutralization in two-hop MIMO interference channels," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6481-6495, Dec.15, 2014.

[16] J. Xu and L. Qiu, “Energy efficiency optimization for MIMO broadcast channels,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 690-701, February 2013.

[17] Q. Shi, C. Peng, W. Xu, M. Hong and Y. Cai, “Energy efficiency optimization for MISO SWIPT systems with zero-forcing beamforming,” *IEEE Trans. Signal Process.*, vol. 64, no. 4, pp. 842-854, Feb.15, 2016.

[18] D. Nguyen, L. N. Tran, P. Pirinen and M. Latva-aho, “Precoding for full duplex multiuser MIMO systems: Spectral and energy efficiency maximization,” *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4038-4050, Aug.15, 2013.

[19] K. Yang, S. Martin, C. Xing, J. Wu and R. Fan, “Energy-efficient power control for device-to-device communications,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3208-3220, Dec. 2016.

[20] A. Zappone, F. D. Stasio, S. Buzzi, and E. Jorswieck, “Energy efficient resource allocation in 5G with application to D2D,” in *Signal Processing for 5G: Algorithms and Implementations*, Wiley, 2016, ch. 19.

[21] A. Zappone, P. Cao and E. A. Jorswieck, “Energy efficiency optimization in relay-assisted MIMO systems with perfect and statistical CSI,” *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 443-457, Jan.15, 2014.

[22] A. Zappone, B. Matthiesen and E. A. Jorswieck, “Energy efficiency in MIMO underlay and overlay device-to-device communications and cognitive radio systems,” *IEEE Trans. Signal Process.*, vol. 65, no. 4, pp. 1026-1041, Feb.15, 2017.

[23] Y. Wang, W. Xu, K. Yang and J. Lin, “Optimal energy-efficient power allocation for OFDM-based cognitive radio networks,” *IEEE Commun. Lett.*, vol. 16, no. 9, pp. 1420-1423, September 2012.

[24] W. Feng, Y. Chen, N. Ge and J. Lu, “Optimal energy-efficient power allocation for distributed antenna systems with imperfect CSI,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 7759-7763, Sept. 2016.

[25] A. Zappone, L. Sanguinetti, G. Bacci, E. A. Jorswieck, and M. Debbah, “Energy-efficient power control: A look at 5G wireless technologies,” *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1668-1683, April 2016.

[26] G. Miao, N. Himayat, G. Y. Li, and S. Talwar, “Distributed interference-aware energy-efficient power optimization,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1323-1333, April 2011.

[27] B. Du, C. Pan, W. Zhang, and M. Chen, “Distributed energy-efficient power optimization for CoMP systems with max-min fairness,” *IEEE Commun. Lett.*, vol. 18, no. 6, pp. 999-1002, 2014.

[28] C. Guo, B. Liao, L. Huang, X. Lin and J. Zhang, "On convexity of fairness-aware energy-efficient power allocation in spectrum-sharing networks," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 534-537, March 2016.

[29] A. Zappone and E. Jorswieck, "Energy efficiency in wireless networks via fractional programming theory," *Foundations and Trends in Communications and Information Theory*, vol. 11, no. 3-4, pp. 185-396, 2015.

[30] E. Bjornson, E. A. Jorswieck, M. Debbah and B. Ottersten, "Multiobjective signal processing optimization: The way to balance conflicting metrics in 5G systems," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 14-23, Nov. 2014.

[31] O. Acker, F. Pötscher and T. Lefort, "Why satellites matter: The relevance of commercial satellites in the 21st century – a perspective 2012-2020," Booz & Company, [Online]. Available: `https://www.esoa.net/cms-data/positions/Why-Satellites-Matter-Full-Report.pdf`.

[32] O. Kodheli *et al.*, "Satellite communications in the new space era: A survey and future challenges," *IEEE Communications Surveys & Tutorials*, doi: 10.1109/COMST.2020.3028247.

[33] M. Chiang, C. Wei, D. P. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640-2651, Jul. 2007.

[34] L. Venturino, N. Prasad, and X. Wang, "Coordinated scheduling and power allocation in downlink multicell OFDMA networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 6, pp. 2835-2848, Jul. 2009.

[35] A. Zappone, E. Bjornson, L. Sanguinetti and E. Jorswieck, "Globally optimal energy-efficient power control and receiver design in wireless networks," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2844-2859, June 1, 2017.

[36] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for non-convex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681-683, 1978.

Chapter 2

A Framework for Weighted-Sum Energy Efficiency Maximization in Wireless Networks[1]

Weighted-sum energy efficiency (WSEE) is a key performance metric in heterogeneous networks, where the nodes may have different energy efficiency (EE) requirements. Nevertheless, WSEE maximization is a challenging problem due to its nonconvex sum-of-ratios form. Unlike previous work, this chapter presents a systematic approach to WSEE maximization under not only power constraints, but also data rate constraints, using a general SINR expression. In particular, the original problem is transformed into an equivalent form, and then a sequential convex optimization (SCO) algorithm is proposed. This algorithm is theoretically guaranteed to converge for any initial feasible point, and, under suitable constraint qualifications, achieves a Karush-Kuhn-Tucker (KKT) solution. Furthermore, we provide remarkable extensions to the proposed methodology, including systems with multiple resource blocks as well as a more general power consumption model which is not necessarily a convex function of the transmit powers. Finally, numerical analysis reveals that the proposed algorithm exhibits fast convergence, low complexity, and robustness (insensitivity to initial points).

2.1 Introduction

Recently, energy efficiency (EE) maximization has become a primary issue in the design of next generation wireless networks due to economic, operational and environmental concerns. Although the network global energy efficiency (GEE), namely, the ratio between the total achievable data rate and the total power consumption, has the most meaningful interpretation as a benefit-cost ratio of the whole network, it does not contain any explicit information about the individual energy efficiencies of the links. An alternative approach in order to overcome this limitation, while maintaining high global performance, is to maximize the WSEE defined as the weighted sum of the links' energy efficiencies [1].

WSEE maximization belongs to the family of sum-of-ratios optimization problems, which are often difficult to solve. In the special case where all the ratios are in concave-convex (CC) form (assuming the case of maximization problems) and the feasible set is convex, the optimization method presented in [2] can be used to globally solve the problem. On the other hand, if at least one ratio of the sum is not in CC form and/or the feasible set is nonconvex, the optimization problem becomes more challenging. In this case, the use of standard global optimization algorithms is quite limited in practice, since they exhibit high computational complexity (generally exponential in the worst case).

An energy efficient multicell multiuser precoding technique is presented in [3], where the WSEE maximization problem is transformed into a parametrized subtractive form, and then a two-layer optimization is used to solve the problem. Later, the authors in [4] investigate the design of centralized and distributed energy-efficient coordinated beamforming in multiple-input single-output (MISO) systems with a general rate-dependent power consumption model. Furthermore, a pricing-based distributed algorithm for WSEE maximization in Ad hoc networks is given in [5].

Moreover, the authors in [6] consider the downlink of a cellular OFDMA (orthogonal frequency-division multiple-access) network with base station coordination, and propose a joint scheduling and power allocation algorithm to maximize the WSEE under maximum power constraints. Finally, the joint downlink and uplink resource allocation in time division duplex (TDD) systems with carrier aggregation is studied in [7].

The remainder of this chapter is organized as follows. In Section 2.2 we introduce the system model and formulate the WSEE maximization problem. An optimization algorithm is developed in Section 2.3, and then interesting extensions are reported in Section 2.4. Finally, simulation results are provided in Section 2.5, while Section 2.6 concludes this chapter.

2.2 System Model and Problem Formulation

We consider a wireless network with N transmitters (users), Λ receivers, and communication bandwidth B. Without loss of generality, we assume that each transmitter is associated to exactly one receiver, and thus $N \geq \Lambda$. Based on [1], the signal-to-interference-plus-noise-ratio (SINR) experienced by user i ($1 \leq i \leq N$) at its intended receiver is given by the following general expression:

$$\gamma_i(\mathbf{p}) = \omega_{i,i} p_i \Big/ \left(\sum\nolimits_{j \neq i} \omega_{j,i} p_j + \phi_i p_i + \mathcal{N}_i \right) \tag{2.1}$$

where $\mathbf{p} = [p_1, p_2, \ldots, p_N]^T$ is the vector of users' transmit powers, $\mathcal{N}_i$ is the equivalent noise power, while $\omega_{j,i}$ and ϕ_i are *non-negative parameters* that do not depend on $\mathbf{p}$ (note that the self-interference term $\phi_i p_i$ may be zero). Next, the achievable data rate and power consumption (assuming the power amplifier operates in the linear region) of the i^{th} user are given respectively by: $R_i(\mathbf{p}) = B \log_2 (1 + \gamma_i(\mathbf{p}))$ and $P_{c,i}(p_i) = \mu_i p_i + P_{st,i}$, where $\mu_i = 1/\eta_i$, with $0 < \eta_i \leq 1$ the power amplifier efficiency, and $P_{st,i} > 0$ is the static dissipated power in all other circuit blocks of the i^{th} transmitter and its intended receiver. Moreover, the EE of user i (measured in bit/Joule) is defined as follows: $EE_i(\mathbf{p}) = R_i(\mathbf{p})/P_{c,i}(p_i)$. Now, we can formulate the WSEE maximization problem:

$$\max_{\mathbf{p}\in S} \quad \text{WSEE}(\mathbf{p}) = \sum_{i=1}^{N} w_i EE_i(\mathbf{p}) \tag{2.2}$$

with feasible set $S = \{\mathbf{p} \in \mathbb{R}^N : \ 0 \leq p_i \leq P_i^{\max} \text{ and } R_i(\mathbf{p}) \geq R_i^{\min},\ 1 \leq i \leq N\}$, where w_i, $P_i^{\max}$ and $R_i^{\min}$ are the priority weight, the maximum transmit power and minimum required data rate of user i, respectively (note that $w_i \geq 0$ and $\sum_{i=1}^{N} w_i = 1$). It can be observed that the objective function is not in sum-of-CC-ratios form ($R_i(\mathbf{p})$ is not concave), and therefore the optimization method in [2] cannot be used. Nevertheless, by applying the variable transformation $\mathbf{p} = 2^{\mathbf{q}}$ ($p_i = 2^{q_i}$, $1 \leq i \leq N$ with $\mathbf{q} = [q_1, q_2, \ldots, q_N]^T$), and due to the fact that the objective is an increasing function of each user's EE, we can equivalently reformulate problem (2.2) as follows:

$$\max_{(\mathbf{q},\mathbf{v})\in Z} \quad f(\mathbf{v}) = \sum_{i=1}^{N} w_i 2^{v_i} \tag{2.3}$$

with feasible set $Z = \{(\mathbf{q},\mathbf{v}) \in \mathbb{R}^{2N} : \ 2^{q_i} \leq P_i^{\max},\ R_i(2^{\mathbf{q}}) \geq R_i^{\min}$ and $EE_i(2^{\mathbf{q}}) \geq 2^{v_i},\ 1 \leq i \leq N\}$, where $\mathbf{v} = [v_1, v_2, \ldots, v_N]^T$ is the vector of auxiliary variables. In addition, after some mathematical operations we get $Z = \{(\mathbf{q},\mathbf{v}) \in \mathbb{R}^{2N} : \ 2^{q_i} \leq P_i^{\max},\ \vartheta_i(\mathbf{q}) \geq 0 \text{ and } \varphi_i(\mathbf{q}, v_i) \geq 0,\ 1 \leq i \leq N\}$, where $\vartheta_i(\mathbf{q}) = \log_2(\omega_{i,i}/\gamma_i^{\min}) + q_i - \log_2\left(\sum_{j\neq i} \omega_{j,i} 2^{q_j} + \phi_i 2^{q_i} + \mathcal{N}_i\right)$, with $\gamma_i^{\min} = 2^{\left(R_i^{\min}/B\right)} - 1$ ($\gamma_i^{\min} \geq 0$, since $R_i^{\min} \geq 0$), and $\varphi_i(\mathbf{q}, v_i) = R_i'(\mathbf{q}) - \mu_i 2^{q_i+v_i} - P_{st,i} 2^{v_i}$, with $R_i'(\mathbf{q}) = R_i(2^{\mathbf{q}})$. The first and the second constraints in Z are convex (the log-sum-exp function is convex [8]), whereas the third constraint is nonconvex, and $f(\mathbf{v})$ is a *strictly convex* function.

2.3 WSEE Maximization Algorithm

In the sequel, we leverage the theory of SCO, [9, 10], in order to solve problem (2.3). In particular, if we have a nonconvex maximization problem $\mathcal{G}$ with objective $g_0(\mathbf{x})$ and compact feasible set $\{\mathbf{x} \in \mathbb{R}^n : \ g_i(\mathbf{x}) \geq 0,\ 1 \leq i \leq I\}$, then we can achieve a KKT solution of $\mathcal{G}$ by solving a sequence of convex maximization problems $\{\widetilde{\mathcal{G}}_j\}_{j\geq 1}$ with objective $\widetilde{g}_{0,j}(\mathbf{x})$, compact feasible set $\{\mathbf{x} \in \mathbb{R}^n : \ \widetilde{g}_{i,j}(\mathbf{x}) \geq 0, \ 1 \leq i \leq I\}$, and global maximum $\mathbf{x}_j^*$ ($\mathbf{x}_0^*$ is any feasible point of $\mathcal{G}$). Moreover, we would like to emphasize that $g_i(\mathbf{x})$, $\widetilde{g}_{i,j}(\mathbf{x})$ ($0 \leq i \leq I$ and $j \geq 1$) are differentiable functions that satisfy *three basic properties*: 1) $g_i(\mathbf{x}) \geq \widetilde{g}_{i,j}(\mathbf{x})$, $\forall \mathbf{x} \in \mathbb{R}^n$, 2) $g_i(\mathbf{x}_{j-1}^*) = \widetilde{g}_{i,j}(\mathbf{x}_{j-1}^*)$, and 3) $\nabla g_i(\mathbf{x}_{j-1}^*) = \nabla \widetilde{g}_{i,j}(\mathbf{x}_{j-1}^*)$.

In order to lower-bound the function $\varphi_i(\mathbf{q}, v_i)$ we use the following logarithmic inequality [11]:

$$\log_2(1+\gamma) \geq \alpha \log_2 \gamma + \beta, \quad \forall \gamma, \gamma' \geq 0 \tag{2.4}$$

where $\alpha = \gamma'/(1+\gamma')$ and $\beta = \log_2(1+\gamma') - \alpha \log_2 \gamma'$. Notice that $\alpha \geq 0$, while the left-hand side and the right-hand side of inequality have equal values and first-derivatives (with respect to γ) at $\gamma = \gamma'$. Therefore, it holds that $R_i'(\mathbf{q}) \geq \widetilde{R_i'}(\mathbf{q})$, with $\widetilde{R_i'}(\mathbf{q}) = B\left[\beta_i + \alpha_i \log_2(\omega_{i,i})\right] + B\alpha_i \left[q_i - \log_2\left(\sum_{j\neq i} \omega_{j,i} 2^{q_j} + \ \phi_i 2^{q_i} + \mathcal{N}_i\right)\right]$, which implies that $\varphi_i(\mathbf{q}, v_i) \geq \widetilde{\varphi_i}(\mathbf{q}, v_i)$, where $\widetilde{\varphi_i}(\mathbf{q}, v_i) = \widetilde{R_i'}(\mathbf{q}) - \mu_i 2^{q_i + v_i} - P_{st,i} 2^{v_i}$. Due to the convexity of the log-sum-exp function and $2^{h(\mathbf{x})}$ (assuming $h(\mathbf{x})$ is convex) [8], both $\widetilde{R_i'}(\mathbf{q})$ and $\widetilde{\varphi_i}(\mathbf{q}, v_i)$ are concave functions. Furthermore, it is known that any convex and differentiable function is lower-bounded by its first-order Taylor expansion at any point [8], and therefore we have $f(\mathbf{v}) \geq f(\mathbf{v}') + \nabla f(\mathbf{v}')^T(\mathbf{v} - \mathbf{v}') = \widetilde{f}(\mathbf{v})$, $\forall \mathbf{v}, \mathbf{v}' \in \mathbb{R}^N$ (observe that $f(\mathbf{v}') = \widetilde{f}(\mathbf{v}')$ and $\nabla f(\mathbf{v}') = \nabla \widetilde{f}(\mathbf{v}')$). More precisely, the affine (and thus concave) function $\widetilde{f}(\mathbf{v})$ is expressed as follows:

$$\widetilde{f}(\mathbf{v}) = \sum\nolimits_{i=1}^N w_i 2^{v_i'} + \ln(2) \sum\nolimits_{i=1}^N w_i 2^{v_i'}(v_i - v_i') \tag{2.5}$$

Consequently, we can formulate the following convex maximization problem which depends on the parameters $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_N]^T$, $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots, \beta_N]^T$, and the point $\mathbf{v}' = [v_1', v_2', \ldots, v_N']^T$:

$$\max_{(\mathbf{q},\mathbf{v})\in\Omega} \widetilde{f}(\mathbf{v}) \quad \Leftrightarrow \quad \max_{(\mathbf{q},\mathbf{v})\in\Omega} \pi(\mathbf{v}) = \sum\nolimits_{i=1}^N w_i 2^{v_i'} v_i \tag{2.6}$$

with feasible set $\Omega = \{(\mathbf{q}, \mathbf{v}) \in \mathbb{R}^{2N} : \ 2^{q_i} \leq P_i^{\max}, \ \vartheta_i(\mathbf{q}) \geq 0$ and $\widetilde{\varphi_i}(\mathbf{q}, v_i) \geq 0$,

Algorithm 2.1 WSEE Maximization

1: Choose a sufficiently small tolerance $\varepsilon > 0$, and a feasible point $\mathbf{p}$
2: Set $\ell = 0$, $v_i = \log_2(EE_i(\mathbf{p}))$ for $1 \leq i \leq N$, and $f^{(0)} = f(\mathbf{v})$
3: **repeat**
4: Compute the parameter vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ with $\boldsymbol{\gamma}' = \boldsymbol{\gamma}(\mathbf{p})$
5: Solve the convex maximization problem (2.6) with parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\mathbf{v}' = \mathbf{v}$ in order to obtain a global maximum $(\mathbf{q}^*, \mathbf{v}^*)$
6: Set $\ell = \ell + 1$, $\mathbf{q} = \mathbf{q}^*$, $\mathbf{v} = \mathbf{v}^*$, $\mathbf{p} = 2^{\mathbf{q}}$, and $f^{(\ell)} = f(\mathbf{v})$
7: **until** $\left|f^{(\ell)} - f^{(\ell-1)}\right| \Big/ \left|f^{(\ell-1)}\right| < \varepsilon$

$1 \leq i \leq N\}$. It is noted that the two problems in (2.6) are equivalent, since in the second problem we omit the constant terms of the objective $\widetilde{f}(\mathbf{v})$. In Algorithm 2.1, we provide an iterative procedure to solve problem (2.3), which is equivalent to the initial WSEE problem (2.2), using the notation $\boldsymbol{\gamma}' = [\gamma_1', \gamma_2', \ldots, \gamma_N']^T$ and $\boldsymbol{\gamma}(\mathbf{p}) = [\gamma_1(\mathbf{p}), \gamma_2(\mathbf{p}), \ldots, \gamma_N(\mathbf{p})]^T$.

According to [9] and [10], *Algorithm 2.1 monotonically increases the value of the objective function $f(\mathbf{v})$ in each iteration (i.e., $f^{(\ell)} \geq f^{(\ell-1)}$) and converges. In addition, assuming suitable constraint qualifications (e.g., Slater's condition for convex problems), the final solution $(\mathbf{q}, \mathbf{v})$ satisfies the KKT optimality conditions of problem* (2.3). It is noted that Algorithm 2.1 does not necessarily achieve the global optimum, since KKT are only necessary (provided that some regularity conditions are satisfied), but not sufficient conditions for optimality in the case of nonconvex problems.

2.4 Extensions to the Proposed Approach

2.4.1 Systems with Multiple Resource Blocks

Firstly, the previous analysis can be straightforwardly extended to wireless networks with multiple ($K > 1$) resource blocks of bandwidth B_{RB} (e.g., OFDMA systems). Based on [1], the only difference is that the QoS (quality-of-service) constraints $B_{RB} \sum_{k=1}^{K} \log_2\left(1 + \gamma_i^{[k]}\right) \geq R_i^{\min}$, with $\gamma_i^{[k]} = \omega_{i,i}^{[k]} 2^{q_i^{[k]}} \Big/ \left(\sum_{j \neq i} \omega_{j,i}^{[k]} 2^{q_j^{[k]}} + \phi_i^{[k]} 2^{q_i^{[k]}} + \mathcal{N}_i^{[k]}\right)$, are not convex now, and they should be approximated by the convex constraints $B_{RB} \sum_{k=1}^{K} \left(\alpha_i^{[k]} \log_2 \gamma_i^{[k]} + \beta_i^{[k]}\right) \geq R_i^{\min}$.

2.4.2 General Power Consumption Model

Secondly, we consider a more general rate-dependent power consumption model with non-linear power terms:

$$P_{c,i}(\mathbf{p}) = \sum_{m=1}^{M} \mu_{i,m} p_i^m + \xi_i (R_i(\mathbf{p}))^{\delta_i} + P_{st,i} \tag{2.7}$$

where M is the order of non-linear power terms, $\mu_{i,m} \geq 0$ measured in W^{1-m} ($\mu_{i,1} = \mu_i = 1/\eta_i$), $0 < \delta_i \leq 1$, and $\xi_i \geq 0$ measured in $\mathrm{W}/(\mathrm{bit/s})^{\delta_i}$. In conventional systems, we have $M = 1$ (absence of non-linear power terms) and $\xi_i = 0$, i.e., $P_{c,i}(p_i) = \mu_i p_i + P_{st,i}$. The term $\sum_{m=2}^{M} \mu_{i,m} p_i^m$ is useful in the case of transmit signals with high peak-to-average power ratio (PAPR), and/or power amplifiers with very narrow linear region. Now, the WSEE maximization problem is formulated as follows:

$$\max_{\mathbf{p} \in S} \quad \mathrm{WSEE}'(\mathbf{p}) = \sum_{i=1}^{N} w_i \, \psi_i \left(p_i, R_i(\mathbf{p})\right) \tag{2.8}$$

where $\psi_i(p_i, \rho_i) = \rho_i \Big/ \left(\sum_{m=1}^{M} \mu_{i,m} p_i^m + \xi_i \rho_i^{\delta_i} + P_{st,i}\right)$. Notice that $\psi_i(p_i, \rho_i)$ is a strictly increasing function of ρ_i for $p_i, \rho_i \geq 0$, since:

$$\frac{\partial \psi_i(p_i, \rho_i)}{\partial \rho_i} = \frac{\sum_{m=1}^{M} \mu_{i,m} p_i^m + \xi_i (1-\delta_i) \rho_i^{\delta_i} + P_{st,i}}{\left(\sum_{m=1}^{M} \mu_{i,m} p_i^m + \xi_i \rho_i^{\delta_i} + P_{st,i}\right)^2} > 0 \tag{2.9}$$

(recall that $1 - \delta_i \geq 0$ and $P_{st,i} > 0$). Hence, we can rewrite problem (2.8) in the following form:

$$\max_{(\mathbf{p}, \boldsymbol{\rho}) \in \Gamma} \quad \sum_{i=1}^{N} w_i \, \psi_i(p_i, \rho_i) \tag{2.10}$$

with feasible set $\Gamma = \{(\mathbf{p}, \boldsymbol{\rho}) \in \mathbb{R}^{2N} : \ \mathbf{p} \in S \ \text{ and } \ R_i(\mathbf{p}) \geq \rho_i \geq 0, \ 1 \leq i \leq N\}$, where $\boldsymbol{\rho} = [\rho_1, \, \rho_2, \ldots, \rho_N]^T$ is the vector of additional variables. Using the variable transformation $\mathbf{p} = 2^{\mathbf{q}}$, $\boldsymbol{\rho} = 2^{\mathbf{y}}$ ($\rho_i = 2^{y_i}$, $1 \leq i \leq N$ with $\mathbf{y} = [y_1, y_2, \ldots, y_N]^T$), and because the objective is an increasing function of each $\psi_i(p_i, \rho_i)$, problem (2.10) is equivalent to:

$$\max_{(\mathbf{q}, \mathbf{y}, \mathbf{v}) \in \mathrm{T}} \quad f(\mathbf{v}) = \sum_{i=1}^{N} w_i \, 2^{v_i} \tag{2.11}$$

with feasible set $\mathrm{T} = \{(\mathbf{q}, \mathbf{y}, \mathbf{v}) \in \mathbb{R}^{3N} : \ 2^{q_i} \le P_i^{\max}, \ \vartheta_i(\mathbf{q}) \ge 0, \ R_i'(\mathbf{q}) \ge 2^{y_i}$ and $\varepsilon_i(q_i, y_i, v_i) \le 0, \ 1 \le i \le N\}$, where $\varepsilon_i(q_i, y_i, v_i) = \sum_{m=1}^{M} \mu_{i,m} 2^{m\, q_i + v_i - y_i} + \xi_i 2^{v_i - (1-\delta_i) y_i} + P_{st,i} 2^{v_i - y_i} - 1$ (the fourth constraint is derived from $\psi_i(2^{q_i}, 2^{y_i}) \ge 2^{v_i}$). Note that only the third constraint in T is nonconvex. Therefore, we can obtain a KKT solution for problem (2.11), which is equivalent to (2.8), by solving a sequence of convex problems of the following form:

$$\max_{(\mathbf{q},\mathbf{y},\mathbf{v}) \in \Psi} \widetilde{f}(\mathbf{v}) \ \Leftrightarrow \ \max_{(\mathbf{q},\mathbf{y},\mathbf{v}) \in \Psi} \pi(\mathbf{v}) = \sum_{i=1}^{N} w_i 2^{v_i'} v_i \tag{2.12}$$

with feasible set $\Psi = \{(\mathbf{q}, \mathbf{y}, \mathbf{v}) \in \mathbb{R}^{3N} : \ 2^{q_i} \le P_i^{\max}, \ \vartheta_i(\mathbf{q}) \ge 0, \ \widetilde{R_i'}(\mathbf{q}) \ge 2^{y_i}$ and $\varepsilon_i(q_i, y_i, v_i) \le 0, \ 1 \le i \le N\}$.

2.5 Numerical Results

Consider a relay-assisted multiple-input multiple-output (MIMO) network, where N transmitters communicate with N receivers through a single-antenna amplify-and-forward relay (receiver i is the intended receiver of transmitter i). We denote by L_T, L_R the number of antennas at each transmitter and receiver, respectively. Moreover, $\mathbf{b}_i$ (with $\|\mathbf{b}_i\| = 1$) is the $L_T \times 1$ beamforming vector of transmitter i (assume that p_i is equally divided between the transmit antennas, i.e., $\mathbf{b}_i = \left(1 \big/ \sqrt{L_T}\right) \mathbf{1}_{L_T \times 1}$, with $\mathbf{1}_{L_T \times 1}$ the $L_T \times 1$ vector of ones), $\mathbf{h}_i$ is the $1 \times L_T$ channel vector from transmitter i to the relay, $\mathbf{g}_i$ is the $L_R \times 1$ channel vector from the relay to receiver i, and $\mathbf{c}_i$ is the $L_R \times 1$ combining vector of receiver i. Also, suppose the receivers perform maximum-ratio combining (MRC), i.e., $\mathbf{c}_i = \mathbf{g}_i \mathbf{h}_i \mathbf{b}_i$. The received signal at the relay is given by $x_r = \sum_{j=1}^{N} \sqrt{p_j}\, \mathbf{h}_j \mathbf{b}_j s_j + n_r$, where s_j is the information symbol of transmitter j ($E\{s_j\} = 0$, $E\{|s_j|^2\} = 1$), and $n_r \sim \mathcal{CN}\left(0, \sigma_r^2\right)$ is the relay thermal noise. Thus, the total input power at the relay is $P_{r,in} = \sum_{j=1}^{N} p_j |\mathbf{h}_j \mathbf{b}_j|^2 + \sigma_r^2$. Then, the received signal at the relay is normalized by $\sqrt{P_{r,in}}$, before being amplified by a factor $\sqrt{P_r}$ (P_r is the relay transmit power) and forwarded to the receivers, in order to ensure that the relay power amplifier operates within the linear region (the signal transmitted by the relay is $y_r = \sqrt{P_r} x_r \big/ \sqrt{P_{r,in}})$. The signals at receiver i before and after the diversity combining unit are $\mathbf{x}_i' = \mathbf{g}_i y_r + \mathbf{n}_i$ and $x_i = \mathbf{c}_i^H \mathbf{x}_i'$, respectively, where $\mathbf{n}_i \sim \mathcal{CN}\left(\mathbf{0}_{L_R\times 1}, \sigma_i^2 \mathbf{I}_{L_R}\right)$ is the receiver thermal noise ($\mathbf{0}_{L_R\times 1}$ is the $L_R \times 1$ zero vector, and $\mathbf{I}_{L_R}$ is the $L_R \times L_R$ identity matrix). Finally, the SINR takes the form in (2.1) with $\omega_{i,i} = \left|\mathbf{c}_i^H \mathbf{g}_i \mathbf{h}_i \mathbf{b}_i\right|^2$, $\omega_{j,i} = \left|\mathbf{c}_i^H \mathbf{g}_i \mathbf{h}_j \mathbf{b}_j\right|^2 + \sigma_i^2 \|\mathbf{c}_i\|^2 |\mathbf{h}_j \mathbf{b}_j|^2 \big/ P_r$ ($j \neq i$), $\phi_i = \sigma_i^2 \|\mathbf{c}_i\|^2 |\mathbf{h}_i \mathbf{b}_i|^2 \big/ P_r$, and $\mathcal{N}_i = \left(\left|\mathbf{c}_i^H \mathbf{g}_i\right|^2 + \sigma_i^2 \|\mathbf{c}_i\|^2 \big/ P_r\right) \sigma_r^2$.

As concerns the simulation parameters, we set $N = 5$, $L_T = L_R = 2$, $P_r = 30$ dBm, $\varepsilon = 10^{-4}$, carrier frequency 2 GHz, $B = 2$ MHz, $\sigma_i^2 = \sigma_r^2 = F\mathcal{N}_0 B$ (with noise figure $F = 3$ dB and power spectral density $\mathcal{N}_0 = -174$ dBm/Hz), $\mu_i = \mu = 5$, $P_i^{\max} = P_{\max}$, $P_{st,i} = P_{st} = 375$ mW, and $w_i = 1/N$ for $1 \leq i \leq N$. The distance of each transmitter/receiver from the relay is uniformly distributed in the interval [200,300] m. A path loss model with reference distance 100 m, path-loss-exponent 3.5, and standard deviation of log-normal shadowing 8 dB has been used, assuming Rayleigh fading. In addition, the QoS requirements are set as follows: $R_i^{\min} = r_i \bar{R}_i$, where $r_i \geq 0$ (for simplicity, $r_i = r$ for $1 \leq i \leq N$), and $\bar{R}_i = B \log_2\left(1 + \bar{\gamma}_i\right)$ with $\bar{\gamma}_i = \gamma_i(p\mathbf{1}_{N\times 1})|_{\mathcal{N}_i = 0} = \omega_{i,i} \big/ \left(\sum_{j\neq i} \omega_{j,i} + \phi_i\right)$ the SINR of user i when all the transmit powers are equal and the equivalent noise power is zero. Unless otherwise stated, the initial point is selected as $\mathbf{p} = P_{\max} \mathbf{1}_{N\times 1}$ (we assume $0 \leq r < 1$, since this point is infeasible when $r \geq 1$). All the results are derived from the statistical average of 10^4 independent problem instances.

First of all, we examine the convergence speed of Algorithm 2.1 through numerical analysis, since it is difficult to be studied analytically. Fig. 2-1 shows that Algorithm 2.1 always generates an increasing sequence and converges very fast within only a few iterations. Thus, Algorithm 2.1 exhibits low complexity because the number of iterations until convergence is quite small and the convex problem in each iteration can be globally solved in polynomial time using standard convex optimization techniques, such as interior-point methods [8]. Furthermore, Algorithm 2.1 is robust since different initialization points achieve slightly different final objective values, and also the convergence speed remains almost the same.

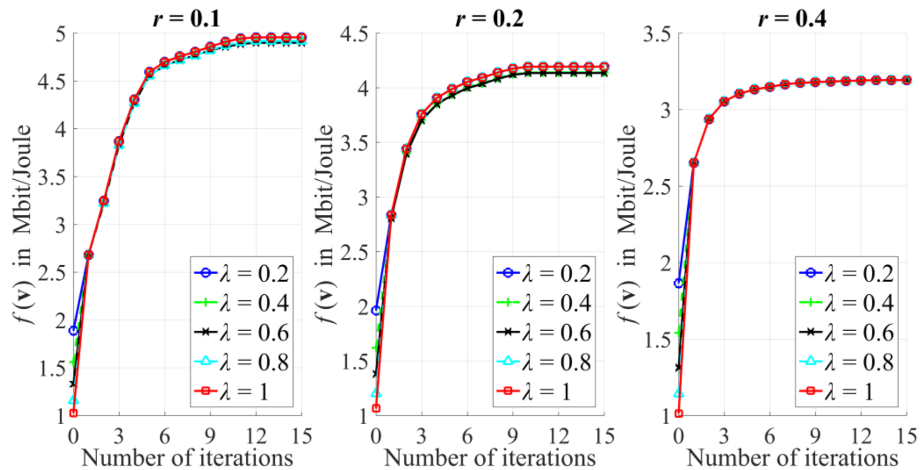

Figure 2-1: Convergence of Algorithm 2.1 (WSEE maximization), with $P_{\max} =$ 20 dBm, for different QoS requirements and initial point $\mathbf{p} = \lambda P_{\max}\mathbf{1}_{N\times 1}$.

Subsequently, for the sake of comparison, we introduce a baseline scheme, namely, weighted-sum rate (WSR) maximization defined as follows:

$$\max_{\mathbf{p}\in S} \quad \mathrm{WSR}(\mathbf{p}) = \sum\nolimits_{i=1}^{N} w_i R_i(\mathbf{p}) \tag{2.13}$$

This problem is solved by SCO, using again the transformation $\mathbf{p} = 2^{\mathbf{q}}$, where the convex problems take the form:

$$\max_{\mathbf{q}\in\Theta} \quad \sum\nolimits_{i=1}^{N} w_i \widetilde{R}_i'(\mathbf{q}) \tag{2.14}$$

with feasible set $\Theta = \{\mathbf{q} \in \mathbb{R}^N : \ 2^{q_i} \le P_i^{\max} \text{ and } \vartheta_i(\mathbf{q}) \ge 0, \ 1 \le i \le N\}$. Figs. 2-2 and 2-3 illustrate respectively the achieved WSEE and WSR versus $P_{\max}$ for different QoS requirements. In Fig. 2-2, we can observe that: 1) for each scheme, the increase of QoS requirements leads to the decrease of WSEE because the feasible set becomes smaller, and 2) for low $P_{\max}$, WSEE and WSR maximization are almost equivalent, since $\mathrm{WSEE}(\mathbf{p}) \approx (1/P_{st})\mathrm{WSR}(\mathbf{p})$ $(\mu p_i \le \mu P_{\max} \ll P_{st} \Rightarrow P_{c,i}(p_i) \approx P_{st})$, while WSEE increases with $P_{\max}$. Similar observations can be made in Fig. 2-3. Nevertheless, for larger values of $P_{\max}$, it can be seen that: 1) in Fig. 2-2, WSEE remains constant when maximizing the WSEE, whereas decreases with $P_{\max}$ when maximizing the WSR because of the higher required transmit power, and 2) in

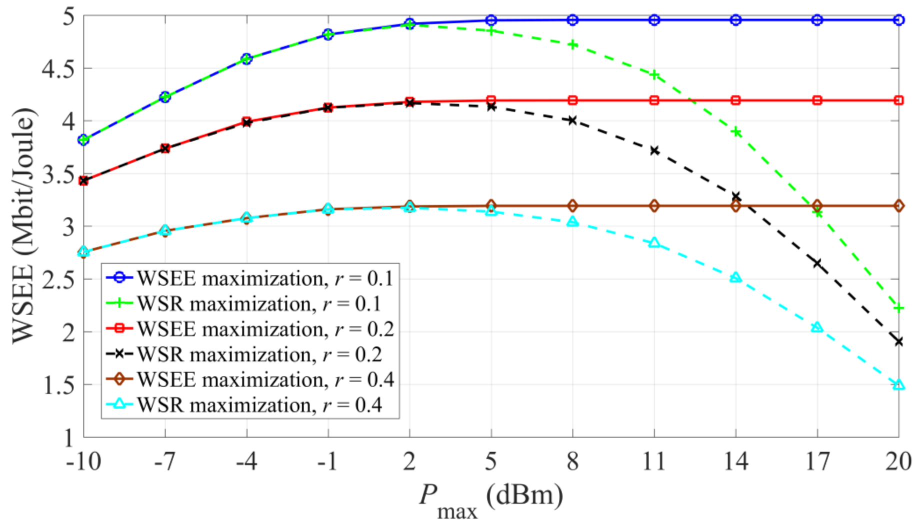

Figure 2-2: Achieved WSEE versus $P_{\max}$ by maximizing: a) the WSEE (Algorithm 2.1), and b) the WSR (baseline scheme) for different QoS requirements.

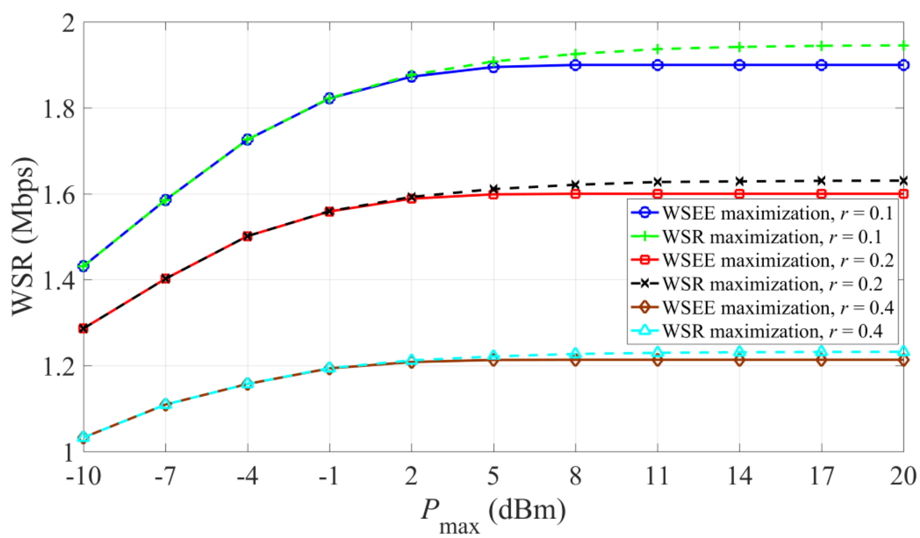

Figure 2-3: Achieved WSR versus $P_{\max}$ by maximizing: a) the WSEE (Algorithm 2.1), and b) the WSR (baseline scheme) for different QoS requirements.

Fig. 2-3, WSR maximization achieves slightly higher WSR than WSEE maximization, while both schemes reach a peak value (note that WSR is upper-bounded when $\phi_i \neq 0$: $\mathrm{WSR}(\mathbf{p}) \leq \sum_{i=1}^{N} w_i B \log_2 (1 + \gamma_i^{\max})$ with $\gamma_i^{\max} = \lim_{p_i \to \infty} \gamma_i (\mathbf{p}) = \omega_{i,i}/\phi_i$).

2.6 Conclusion

In this chapter, we have presented a general methodology for WSEE maximization in wireless networks. More specifically, we have developed a low-complexity and robust algorithm that is theoretically guaranteed to converge and is able to achieve a KKT solution. Finally, we have studied notable extensions of the proposed approach to systems with multiple resource blocks and general power consumption model as well.

Bibliography

[1] A. Zappone and E. Jorswieck, "Energy efficiency in wireless networks via fractional programming theory," *Foundations and Trends in Communications and Information Theory*, vol. 11, no. 3-4, pp. 185-396, 2015.

[2] Y. Jong, "An efficient global optimization algorithm for nonlinear sum-of-ratios problem," May 2012. [Online]. Available: `http://www.optimization-online.org/DB_FILE/2012/08/3586.pdf`

[3] S. He, Y. Huang, L. Yang and B. Ottersten, "Coordinated multicell multiuser precoding for maximizing weighted sum energy efficiency," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 741-751, Feb. 2014.

[4] O. Tervo, A. Tölli, M. Juntti and L. N. Tran, "Energy-efficient beam coordination strategies with rate-dependent processing power," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 6097-6112, Nov. 2017.

[5] C. Pan, W. Xu, W. Zhang, J. Wang, H. Ren and M. Chen, "Weighted sum energy efficiency maximization in Ad hoc networks," *IEEE Wireless Commun. Lett.*, vol. 4, no. 3, pp. 233-236, June 2015.

[6] L. Venturino, A. Zappone, C. Risi and S. Buzzi, "Energy-efficient scheduling and power allocation in downlink OFDMA networks with base station coordination," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 1-14, Jan. 2015.

[7] G. Yu, Q. Chen, R. Yin, H. Zhang and G. Y. Li, "Joint downlink and uplink resource allocation for energy-efficient carrier aggregation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3207-3218, June 2015.

[8] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge Univ. Press, 2004.

[9] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for non-convex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681-683, 1978.

[10] B. K. Sriperumbudur and G. R. Lanckriet, "On the convergence of the concave-convex procedure," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 1759–1767.

[11] J. Papandriopoulos and J. S. Evans, "SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3711-3724, Aug. 2009.

Chapter 3

Energy Efficiency Optimization: A New Trade-Off Between Fairness and Total System Performance[2]

The total energy efficiency (TEE), defined as the ratio between the total data rate and the total power consumption, is considered the most meaningful performance metric in terms of energy efficiency (EE). Nevertheless, it does not depend directly on the EE of each link and its maximization leads to unfairness between the links. On the other hand, the maximization of the minimum EE (MEE), i.e., the minimum of the EEs of all links, guarantees the fairest power allocation, but it does not contain any explicit information about the total system performance. The main trend in current research is to maximize TEE and MEE separately. Unlike previous contributions, this chapter presents a general multi-objective approach for EE optimization that takes into account both TEE and MEE at the same time, and thus achieves various trade-off points in the MEE-TEE plane. Due to the nonconvex form of the resulting problem, we propose a low-complexity algorithm leveraging the theory of sequential convex optimization (SCO). Last but not least, we provide a novel theoretical result for the complexity of SCO algorithms.

3.1 Introduction

Energy efficiency expresses the amount of information that can be reliably transmitted per Joule of consumed energy (measured in bit/Joule), and is recently characterized as a key performance indicator for 5G networks. Zappone *et al.* [1] propose a unified framework for the design of both centralized and distributed energy-efficient power control algorithms. Furthermore, power allocation strategies for maximizing the proportional, max-min, and harmonic fair EE in spectrum-sharing networks are given in [2]. The optimization of various EE performance metrics is also investigated in [3] and [4] for MIMO (multiple-input multiple-output) and OFDMA (orthogonal frequency division multiple access) systems, respectively. Finally, the recent study [5] presents a systematic approach to weighted-sum EE maximization in wireless networks.

In summary, the existing approaches maximize the total/global, sum, product and minimum EE individually. The TEE, albeit the most important EE metric, does not depend directly on the links' EEs and its maximization results in low fairness. On the other hand, the last three EE metrics explicitly depend on the links' EEs, but none of them contains specific information about the total system performance (i.e., TEE). Moreover, the fairest resource allocation is achieved by maximizing the MEE. Consequently, in this chapter, we introduce a new multi-objective approach that takes into consideration the two extremes (TEE and MEE) at the same time, and thus providing a set of MEE-TEE operating points which are not achievable with existing approaches.

The remainder of this chapter is organized as follows. Section 3.2 introduces the system model and formulates the general EE optimization problem. Subsequently, an EE optimization algorithm is developed and analyzed in Section 3.3. Finally, numerical results are provided in Section 3.4, while concluding remarks are given in Section 3.5.

3.2 System Model and Problem Formulation

We consider a wireless network with N transmitters/users, M receivers and K mutually orthogonal resource blocks of bandwidth B_{RB}. In addition, we assume that each transmitter is associated to exactly one receiver (its intended receiver), and therefore it holds that $N \geq M$.[3] The Signal-to-Interference-plus-Noise-Ratio (SINR) experienced by user i $(1 \leq i \leq N)$ at its intended receiver on resource block k $(1 \leq k \leq K)$ is given by the following formula:[4]

$$\gamma_i^{(k)} = \omega_{i,i}^{(k)} p_i^{(k)} \Big/ \left(\sum\nolimits_{j \neq i} \omega_{j,i}^{(k)} p_j^{(k)} + \mathcal{N}_i^{(k)} \right) \tag{3.1}$$

where $p_j^{(k)}$ is the transmit power of user j, $\mathcal{N}_i^{(k)}$ is the noise power at the i^{th} user's intended receiver, and $\omega_{j,i}^{(k)}$ is the channel gain between j^{th} transmitter and i^{th} user's intended receiver, all on resource block k. For convenience, we denote the vector of transmit powers by $\mathbf{p} = \left[\mathbf{p}_1^T, \mathbf{p}_2^T, \ldots, \mathbf{p}_N^T\right]^T$, where $\mathbf{p}_i = \left[p_i^{(1)}, p_i^{(2)}, \ldots, p_i^{(K)}\right]^T$ with $1 \leq i \leq N$.

The i^{th} user's and total achievable data rate (in bit/s) are given respectively by: $R_i(\mathbf{p}) = B_{RB} \sum_{k=1}^{K} \log_2 \left(1 + \gamma_i^{(k)}\right)$ and $R_{tot}(\mathbf{p}) = \sum_{i=1}^{N} R_i(\mathbf{p})$. Next, assuming that the transmit power amplifiers operate in the linear region and the hardware dissipated power is fixed, the i^{th} user's and total power consumption can be modeled respectively as follows: $P_{c,i}(\mathbf{p}_i) = \mu_i \sum_{k=1}^{K} p_i^{(k)} + P_{st,i}$ and $P_{c,tot}(\mathbf{p}) = \sum_{i=1}^{N} P_{c,i}(\mathbf{p}_i)$, where $\mu_i = 1/\xi_i$, with $\xi_i \in (0,1]$ the efficiency of the power amplifier of transmitter i, and $P_{st,i}$ is the static dissipated power in all other circuit blocks of the i^{th} transmitter and its intended receiver (e.g., cooling, filtering, signal up and down conversion, digital-to-analog and analog-to-digital conversion). Furthermore, the i^{th} user's and total EE (in bit/Joule) are defined respectively as the following ratios: $EE_i(\mathbf{p}) = R_i(\mathbf{p})/P_{c,i}(\mathbf{p}_i)$ and $EE_{tot}(\mathbf{p}) = R_{tot}(\mathbf{p})/P_{c,tot}(\mathbf{p})$.

[3]Without loss of generality, we make this assumption to reduce the amount of notation needed to express the SINR in (3.1). Similar formula can be obtained when each transmitter is associated to more than one receiver.

[4]The proposed methodology can be straightforwardly modified to include a self-interference term in the denominator of (3.1), as in [1] and [5].

Now, we introduce the following nonconvex maximization problem, based on the multi-objective optimization theory:

$$\max_{\mathbf{p}\in S_{\mathbf{p}}} \quad G(\mathbf{p}) = F\left(EE_{tot}(\mathbf{p}), \min_{1\leq i\leq N} EE_i(\mathbf{p})\right) \tag{3.2}$$

with feasible set $S_{\mathbf{p}} = \{\mathbf{p} \in \mathbb{R}_{+}^{NK} : \sum_{k=1}^{K} p_i^{(k)} \leq P_i^{\max}$, and $R_i(\mathbf{p}) \geq R_i^{th}$ for $1 \leq i \leq N\}$, where $P_i^{\max}$ and R_i^{th} are the i^{th} user's maximum transmit power and minimum required data rate, respectively. Moreover, we assume that: *1) the objective $F(x, y)$ is an increasing function of x and y, 2) $F(2^u, 2^v) > 0$, $\forall(u, v) \in \mathbb{R}^2$*, and *3) $f(u, v) = \log_2 F(2^u, 2^v)$ is a differentiable concave function.*

In the sequel, we transform the original nonconvex problem (3.2) into an equivalent problem in a more tractable form. Due to the fact that $F(x, y)$ is an increasing function and $EE_{tot}(\mathbf{p}), \min_{1\leq i\leq N} EE_i(\mathbf{p}) \geq 0$, $\forall \mathbf{p} \in \mathbb{R}_{+}^{NK}$, problem (3.2) can be equivalently written as follows:

$$\max_{(\mathbf{p},\eta_{tot}^{th},\eta_{\min}^{th})\in \mathrm{T}} \quad F\left(\eta_{tot}^{th}, \eta_{\min}^{th}\right) \tag{3.3}$$

with feasible set $\mathrm{T} = \{(\mathbf{p}, \eta_{tot}^{th}, \eta_{\min}^{th}) \in \mathbb{R}_{+}^{NK+2} : \mathbf{p} \in S_{\mathbf{p}}$, $EE_{tot}(\mathbf{p}) \geq \eta_{tot}^{th}$, and $EE_i(\mathbf{p}) \geq \eta_{\min}^{th}$ for $1 \leq i \leq N\}$, where η_{tot}^{th} and $\eta_{\min}^{th}$ are auxiliary variables. Notice that the set of constraints $EE_i(\mathbf{p}) \geq \eta_{\min}^{th}$ $(1 \leq i \leq N)$ is equivalent to $\min_{1\leq i\leq N} EE_i(\mathbf{p}) \geq \eta_{\min}^{th}$, and the maximum objective value is obtained when $EE_{tot}(\mathbf{p}) = \eta_{tot}^{th}$ and $\min_{1\leq i\leq N} EE_i(\mathbf{p}) = \eta_{\min}^{th}$.

Subsequently, by applying the variable transformation $\mathbf{p} = 2^{\mathbf{q}}$ $(p_i^{(k)} = 2^{q_i^{(k)}}$, $1 \leq i \leq N$ and $1 \leq k \leq K)$, $\eta_{tot}^{th} = 2^u$, $\eta_{\min}^{th} = 2^v$, and after a few mathematical operations, we get the following nonconvex problem (note that the maximization of F is equivalent to the maximization of $\log_2 F$):

$$\max_{(\mathbf{q},u,v)\in Z} \quad f(u, v) = \log_2 F(2^u, 2^v) \tag{3.4}$$

with feasible set $Z = \{(\mathbf{q}, u, v) \in \mathbb{R}^{NK+2} : \sum_{k=1}^{K} 2^{q_i^{(k)}} \leq P_i^{\max}$, $R_i'(\mathbf{q}) \geq R_i^{th}$,

$\psi_i(\mathbf{q}, v) \geq 0$ for $1 \leq i \leq N$, and $g(\mathbf{q}, u) \geq 0\}$, where $R'_i(\mathbf{q}) = R_i(2^{\mathbf{q}})$, $R'_{tot}(\mathbf{q}) = R_{tot}(2^{\mathbf{q}})$, $\psi_i(\mathbf{q}, v) = R'_i(\mathbf{q}) - \mu_i \sum_{k=1}^{K} 2^{q_i^{(k)}+v} - P_{st,i}2^v$, and $g(\mathbf{q}, u) = R'_{tot}(\mathbf{q}) - \sum_{i=1}^{N} \mu_i \sum_{k=1}^{K} 2^{q_i^{(k)}+u} - \left(\sum_{i=1}^{N} P_{st,i}\right) 2^u$.

3.3 Energy Efficiency Optimization Algorithm

In this section, we leverage the theory of SCO (see Appendix 3.6) so as to achieve a Karush-Kuhn-Tucker (KKT) solution for the equivalent problem (3.4).

3.3.1 Algorithm Design and Complexity

In order to satisfy the properties of Theorem 3.1 in the Appendix 3.6, we use the following inequality with logarithms [5] ($\log_2 0 = -\infty$ and $0 \cdot \log_2 0 = 0$): $A(\gamma) = \log_2(1+\gamma) \geq a \cdot \log_2\gamma + b = B(\gamma, \gamma')$, $\forall \gamma, \gamma' \geq 0$, where a, b are given by:

$$a = \gamma'/(1+\gamma'), \quad b = \log_2(1+\gamma') - a \cdot \log_2\gamma' \tag{3.5}$$

Observe that $a \geq 0$, $A(\gamma)|_{\gamma=\gamma'} = B(\gamma, \gamma')|_{\gamma=\gamma'}$, and $\frac{dA(\gamma)}{d\gamma}\Big|_{\gamma=\gamma'} = \frac{\partial B(\gamma,\gamma')}{\partial \gamma}\Big|_{\gamma=\gamma'}$. Consequently, we can construct the following lower bounds: $R'_i(\mathbf{q}) \geq B_{RB} \cdot \sum_{k=1}^{K} \left[b_i^{(k)} + a_i^{(k)} \log_2\left(\omega_{i,i}^{(k)}\right) + a_i^{(k)} q_i^{(k)} - a_i^{(k)} \log_2\left(\sum_{j\neq i} \omega_{j,i}^{(k)} 2^{q_j^{(k)}} + \mathcal{N}_i^{(k)}\right)\right] = \widetilde{R'_i}(\mathbf{q})$, $R'_{tot}(\mathbf{q}) \geq \sum_{i=1}^{N} \widetilde{R'_i}(\mathbf{q}) = \widetilde{R'_{tot}}(\mathbf{q})$, $\psi_i(\mathbf{q}, v) \geq \widetilde{R'_i}(\mathbf{q}) - \mu_i \sum_{k=1}^{K} 2^{q_i^{(k)}+v} - P_{st,i}2^v = \widetilde{\psi_i}(\mathbf{q}, v)$, and $g(\mathbf{q}, u) \geq \widetilde{R'_{tot}}(\mathbf{q}) - \sum_{i=1}^{N} \mu_i \sum_{k=1}^{K} 2^{q_i^{(k)}+u} - \left(\sum_{i=1}^{N} P_{st,i}\right) 2^u = \widetilde{g}(\mathbf{q}, u)$, where $a_i^{(k)}$ and $b_i^{(k)}$ are given by (3.5) with $\gamma' = {\gamma'}_i^{(k)}$. Notice that $\widetilde{R'_i}(\mathbf{q})$, $\widetilde{R'_{tot}}(\mathbf{q})$, $\widetilde{\psi_i}(\mathbf{q}, v)$, and $\widetilde{g}(\mathbf{q}, u)$ are all concave functions (the log-sum-exp, 2^{x+y}, and 2^x are convex functions [6]). Based on the previous analysis, we can formulate the following convex problem which depends on the parameters $a_i^{(k)}$ and $b_i^{(k)}$:

$$\max_{(\mathbf{q},u,v)\in\Omega} \quad f(u, v) = \log_2 F(2^u, 2^v) \tag{3.6}$$

with feasible set $\Omega = \{(\mathbf{q}, u, v) \in \mathbb{R}^{NK+2} : \sum_{k=1}^{K} 2^{q_i^{(k)}} \leq P_i^{\max}$, $\widetilde{R'_i}(\mathbf{q}) \geq R_i^{th}$, $\widetilde{\psi_i}(\mathbf{q}, v) \geq 0$ for $1 \leq i \leq N$, and $\widetilde{g}(\mathbf{q}, u) \geq 0\}$.

Algorithm 3.1 Energy Efficiency Optimization

1: Choose a tolerance $\varepsilon > 0$, and an initial point $\mathbf{p} \in S_{\mathbf{p}}$
2: Set $l = 0$, $u = \log_2 \left(EE_{tot}(\mathbf{p}) \right)$, $v = \log_2 \left(\min_{1 \le i \le N} EE_i(\mathbf{p}) \right)$,
and $f_0 = f(u, v)$
3: **repeat**
4: Compute the SINR vector $\boldsymbol{\gamma}$ according to (3.1), and then the parameter vectors $\boldsymbol{a}, \boldsymbol{b}$ according to (3.5) with $\boldsymbol{\gamma'} = \boldsymbol{\gamma}$
5: Solve the convex optimization problem (3.6) with parameters $\boldsymbol{a}, \boldsymbol{b}$ in order to obtain a globally optimal solution $(\mathbf{q}^*, u^*, v^*)$
6: Set $l = l + 1$, $\mathbf{q} = \mathbf{q}^*$, $u = u^*$, $v = v^*$, $\mathbf{p} = 2^{\mathbf{q}}$ and $f_l = f(u, v)$
7: **until** $|f_l - f_{l-1}| / |f_{l-1}| < \varepsilon$

Algorithm 3.1 provides an iterative SCO procedure using the following notation: $\boldsymbol{\sigma} = \left[\boldsymbol{\sigma}_1^T, \boldsymbol{\sigma}_2^T, \ldots, \boldsymbol{\sigma}_N^T \right]^T$ for $\boldsymbol{\sigma} \in \{\mathbf{p}, \mathbf{q}, \boldsymbol{\gamma}, \boldsymbol{\gamma'}, \boldsymbol{a}, \boldsymbol{b}\}$, where $\boldsymbol{\sigma}_i = \left[\sigma_i^{(1)}, \sigma_i^{(2)}, \ldots, \sigma_i^{(K)} \right]^T$ with $1 \le i \le N$. Based on Theorem 3.1 in the Appendix 3.6, Algorithm 3.1 monotonically increases the objective $f(u, v)$ in each iteration and, under suitable constraint qualifications, converges to a point that satisfies the KKT conditions of problem (3.4).

Finally, the complexity of Algorithm 3.1 depends on the number of iterations until convergence as well as on the complexity of each iteration (which is mainly restricted by the optimization of a convex problem). According to Theorem 3.2 in the Appendix 3.6, the overall complexity of Algorithm 3.1 is $O\left((\lambda/\varepsilon) \, \phi(N, K) \right)$, where $\lambda = f_*/f_0 \ge 1$, with f_* being the globally maximum value of problem (3.4), and $\phi(N, K)$ is the complexity of problem (3.6). If this convex problem is solved by an interior-point method, then $\phi(N, K)$ is polynomial in the number of variables and constraints (which are $NK + 2$ and $3N + 1$, respectively), and thus polynomial in N and K.

3.3.2 Applications

Afterwards, we examine two special applications of Algorithm 3.1, namely, the weighted product (WP) and the weighted minimum (WM) of TEE and MEE, which are respectively defined by: $F_{WP}(x, y) = x^w y^{1-w}$ and $F_{WM}(x, y) = \min\left(x/w, y/(1 - w) \right)$, with $x = EE_{tot}(\mathbf{p})$ and $y = \min_{1 \le i \le N} EE_i(\mathbf{p})$. Note that w and $1 - w$ are the priority weights of TEE and MEE, respectively ($0 \le w \le 1$).

Specifically, $w = 1$ corresponds to TEE maximization, while $w = 0$ corresponds to MEE maximization. Moreover, we have that: $f_{WP}(u,v) = w\,u + (1-w)\,v$, and $f_{WM}(u,v) = \min\left(u - \log_2 w, v - \log_2(1-w)\right)$ since $\min(2^r, 2^s) = 2^{\min(r,s)}$. Observe that $f_{WP}(u,v)$ and $f_{WM}(u,v)$ are both concave functions (the minimum of concave functions is also a concave function [6]).

Concerning the WM maximization, we cannot consider the KKT conditions of problem (3.4) directly, since the objective $f_{WM}(u,v)$ is not differentiable. However, Algorithm 3.1 converges to a point that satisfies the KKT conditions of the following problem (equivalent epigraph form of problem (3.4)): $\max_{(\mathbf{q},u,v,t)\in\Gamma} t$ with feasible set $\Gamma = \{(\mathbf{q},u,v,t) \in \mathbb{R}^{NK+3} :\ (\mathbf{q},u,v) \in Z,\ u - \log_2 w \geq t,\ \text{and}\ v - \log_2(1-w) \geq t\}$. This statement can be easily proved if we write problem (3.6) in its equivalent epigraph form: $\max_{(\mathbf{q},u,v,t)\in\Theta} t$ with feasible set $\Theta = \{(\mathbf{q},u,v,t) \in \mathbb{R}^{NK+3} :\ (\mathbf{q},u,v) \in \Omega,$ $u - \log_2 w \geq t,\ \text{and}\ v - \log_2(1-w) \geq t\}$, and observe that the properties of Theorem 3.1 in the Appendix 3.6 are satisfied.

3.4 Numerical Results

Consider the uplink of a cellular network with a single micro-cell, where $K = 5$ resource blocks allocated to one cellular UE (User Equipment) are reused by 4 D2D (Device-to-Device) transmitter/receiver-pairs ($N = 5$). The cellular UE is associated to the BS (Base Station) and each D2D transmitter is associated to its intended D2D receiver ($M = N$). In addition, the D2D link distance, namely, the distance between the transmitter and receiver of one D2D pair, is considered the same for all D2D pairs and is denoted by d_{D2D}. As concerns the simulation parameters, the cellular UE as well as the D2D pairs are uniformly distributed in [30,100] m from the BS. Moreover, we assume a carrier frequency of 5 GHz, $\varepsilon = 10^{-3}$, $B_{RB} = 500$ KHz, $\mathcal{N}_i^{(k)} = F\mathcal{N}_0 B_{RB}$ (with receiver noise figure $F = 3$ dB, and power spectral density of the thermal noise $\mathcal{N}_0 = -174$ dBm/Hz), $\mu_i = \mu = 1$, $P_{st,i} = P_{st} = 10$ dBm, $P_i^{\max} = P_{\max} = 23$ dBm, and $R_i^{th} = R_{th} = 0$ for $1 \leq i \leq N$ (in the sequel we study the fairness, so it is preferable not to consider the data rate constraints). Unless

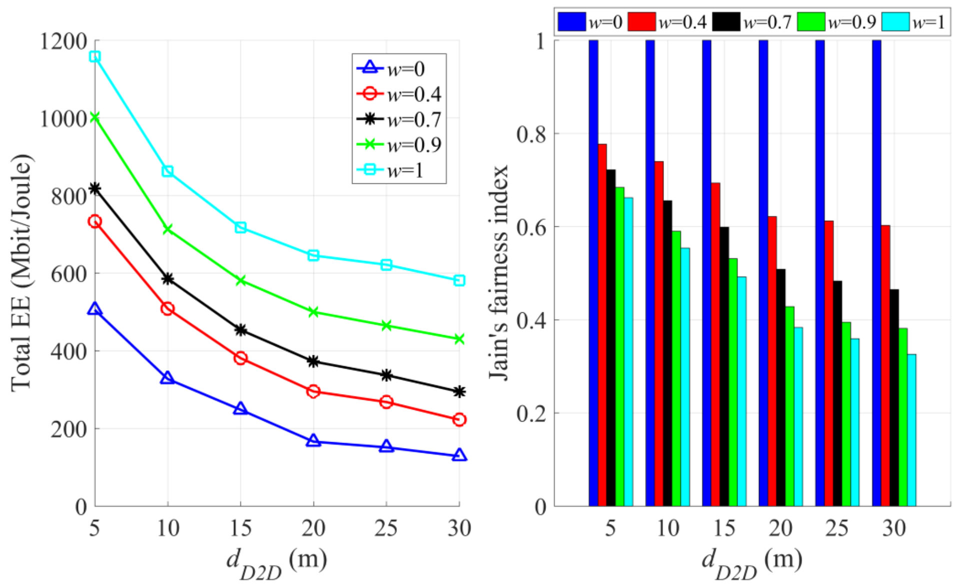

Figure 3-1: TEE and JFI versus D2D link distance for different priority weights.

otherwise stated, the initial feasible point is selected as $\mathbf{p} = (P_{\max}/K)\,\mathbf{1}_{NK\times 1}$, where $\mathbf{1}_{NK\times 1}$ is the $NK \times 1$ vector of ones. Furthermore, all the results (except for Fig. 3-2) are obtained by averaging over 10^3 independent simulations, and the following analysis refers to Algorithm 3.1 specialized to maximize the WP of TEE and MEE.

For the evaluation of fairness, we make use of Jain's fairness index (JFI) as a function of users' EEs: $\mathcal{J} = \frac{\left(\sum_{i=1}^{N} EE_i\right)^2}{N\sum_{i=1}^{N} EE_i^2}$ with $0 \leq \mathcal{J} \leq 1$. In general, the closer JFI is to 1, the fairer the power allocation is in terms of EE. In the special case where $w = 0$ (MEE maximization) all the EEs are equal at the maximum point [7], and therefore $\mathcal{J} = 1$ and TEE=MEE.

First of all, Fig. 3-1 shows the TEE and JFI versus the D2D link distance for different weights. For fixed d_{D2D}, it is clear that TEE increases while JFI decreases as the weight w increases, since higher priority is given to TEE and lower to MEE. According to the left figure, TEE decreases with the D2D link distance for all w. In addition, as shown in the right figure, JFI decreases with the D2D link distance for $w \neq 0$, whereas it remains equal to 1 for $w = 0$ as already mentioned.

Afterwards, Fig. 3-2 illustrates the Pareto operating points in the MEE-TEE plane achieved by: a) the proposed approach for 200 equally-spaced values of the weight w in $[0, 1]$, b) product-EE maximization [4] with $PEE(\mathbf{p}) = \prod_{i=1}^{N} EE_i(\mathbf{p})$, and c) sum-

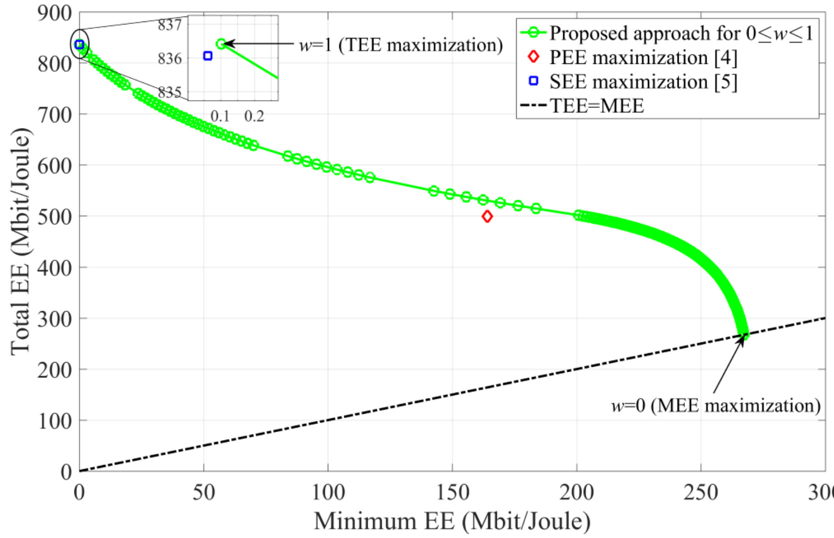

Figure 3-2: Pareto operating points in the MEE-TEE plane for a specific simulation scenario with $d_{D2D} = 10$ m.

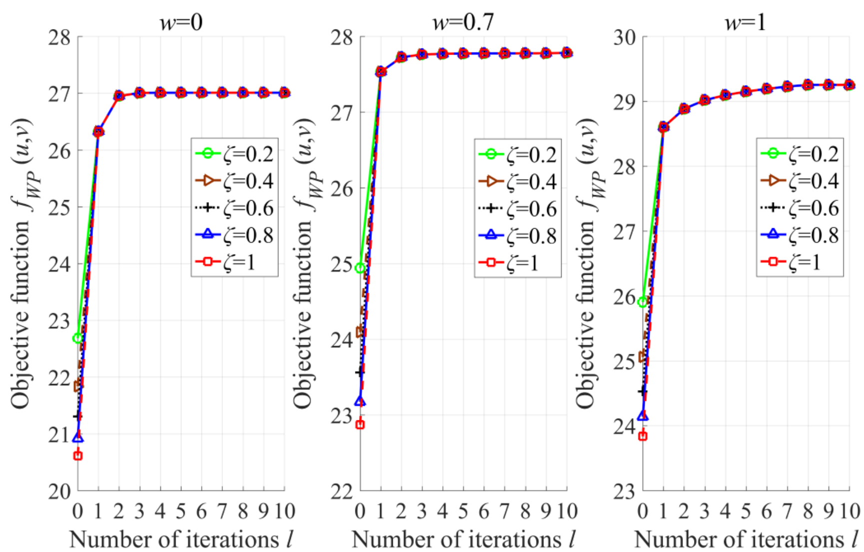

Figure 3-3: Convergence of Algorithm 3.1 (WP maximization) for different priority weights and initial point $\mathbf{p} = \zeta \left(P_{\max}/K\right) \mathbf{1}_{NK\times 1}$ with $d_{D2D} = 20$ m.

EE maximization [5] with $SEE(\mathbf{p}) = \sum_{i=1}^{N} EE_i(\mathbf{p})$. As can be seen, the proposed approach for $0 < w < 1$ achieves several trade-off points which are not attainable by maximizing TEE, SEE, PEE and MEE individually. Moreover, we can observe that all the Pareto points lie on or above the line TEE=MEE, since it can be easily proved that $EE_{tot}(\mathbf{p}) \geq \min\limits_{1\leq i\leq N} EE_i(\mathbf{p})$, $\forall \mathbf{p} \in \mathbb{R}_+^{NK}$.

Finally, we examine the convergence of Algorithm 3.1 for different priority weights and initial points. According to Fig. 3-3, Algorithm 3.1 exhibits fast convergence and insensitivity to initial points for all simulation scenarios, and requires a quite small number of iterations to converge. In particular, given the tolerance $\varepsilon = 10^{-3}$ ($\varepsilon = 10^{-4}$), it converges within approximately 4, 5 and 9 (5, 6 and 10) iterations for $w = 0,\ 0.7$ and 1, respectively.

3.5 Conclusion

In this chapter, we have developed a unified methodology for EE optimization that incorporates a new trade-off between fairness and total system performance. Furthermore, an efficient SCO algorithm has been proposed which can be applied to practical scenarios of wireless networks. Finally, we have presented a general complexity analysis for SCO algorithms.

3.6 Appendix: Sequential Convex Optimization

Let $\mathcal{F}$ be a nonconvex maximization problem with objective $f_0(\mathbf{x})$, and nonempty, compact feasible set $S = \{\mathbf{x} \in \mathbb{R}^n :\ f_i(\mathbf{x}) \geq 0,\ 1 \leq i \leq I\}$. Also, let $\{\mathcal{H}_j\}_{j\geq 1}$ be a sequence of convex maximization problems with objective $h_{0,j}(\mathbf{x}, \mathbf{x}^*_{j-1})$, compact feasible set $S_j = \{\mathbf{x} \in \mathbb{R}^n :\ h_{i,j}(\mathbf{x}, \mathbf{x}^*_{j-1}) \geq 0,\ 1 \leq i \leq I\}$, and global maximum $\mathbf{x}^*_j$. Let $\mathbf{x}^*_0$ be any feasible point of problem $\mathcal{F}$, that is, $\mathbf{x}^*_0 \in S$. Moreover, assume that $f_i(\mathbf{x})$ and $h_{i,j}(\mathbf{x}, \mathbf{x}^*_{j-1})$, $0 \leq i \leq I$ and $j \geq 1$, are differentiable functions. The next theorem follows directly from [8].

Theorem 3.1 (Convergence)**.** *Suppose that the functions $h_{i,j}(\mathbf{x}, \mathbf{x}^*_{j-1})$, $0 \leq i \leq I$ and $j \geq 1$, satisfy the following three properties (where $\nabla = [\partial/\partial x_1, \partial/\partial x_2, \ldots, \partial/\partial x_n]^T$):*

(a) $h_{i,j}(\mathbf{x}, \mathbf{x}^*_{j-1}) \leq f_i(\mathbf{x}),\ \forall \mathbf{x} \in S_j$

(b) $h_{i,j}(\mathbf{x}, \mathbf{x}^*_{j-1})\Big|_{\mathbf{x}=\mathbf{x}^*_{j-1}} = f_i(\mathbf{x}^*_{j-1})$

(c) $\nabla h_{i,j}(\mathbf{x}, \mathbf{x}^*_{j-1})\Big|_{\mathbf{x}=\mathbf{x}^*_{j-1}} = \nabla f_i(\mathbf{x}^*_{j-1})$

Then, the sequence $\left\{f_0(\mathbf{x}^*_j)\right\}_{j\geq 0}$ is monotonically increasing ($f_0(\mathbf{x}^*_j) \geq f_0(\mathbf{x}^*_{j-1})$, $j \geq 1$) and converges to a finite value L ($\lim\limits_{j\to\infty} f_0(\mathbf{x}^*_j) = L < \infty$). In addition, every *accumulation/limit point* $\bar{\mathbf{x}}$ *of the sequence* $\left\{\mathbf{x}_j^*\right\}_{j\geq 0}$ *achieves the objective value* L ($f_0(\bar{\mathbf{x}}) = L$) *and, assuming suitable constraint qualifications, satisfies the KKT conditions of the initial problem* $\mathcal{F}$.

A rigorous mathematical analysis for the complexity of SCO is very challenging since the convergence rate depends on the particular structure of the problem, and no theoretical results are available so far. Nevertheless, we provide the following general theorem exploiting the monotonicity of SCO.

Theorem 3.2 (Complexity)**.** *Assume that: 1) the properties of Theorem 3.1 are satisfied, 2) SCO terminates when* $\left|f_0(\mathbf{x}_j^*) - f_0(\mathbf{x}_{j-1}^*)\right| / \left|f_0(\mathbf{x}_{j-1}^*)\right| < \varepsilon$, *where* $\varepsilon > 0$ *is a predefined tolerance, and 3)* $f_0(\mathbf{x}_0^*) > 0$. *Then, the number of iterations until convergence is* $O\left(\lambda/\varepsilon\right)$, *where* $\lambda = f_0(\mathbf{x}^*)/f_0(\mathbf{x}_0^*) \geq 1$ *and* $\mathbf{x}^*$ *is a global maximum of problem* $\mathcal{F}$. *In addition, the overall complexity of SCO is* $O\left(\left(\lambda/\varepsilon\right)\varphi(n, I)\right)$, *where* $\varphi(n, I)$ *is the complexity of the method used to solve each convex problem with* n *variables and* I *constraints.*

Proof. By virtue of Theorem 3.1, we have that $f_0(\mathbf{x}_j^*) \geq f_0(\mathbf{x}_{j-1}^*) \geq f_0(\mathbf{x}_0^*) > 0$, $j \geq 1$. Next, let $k \geq 1$ be the number of iterations until convergence, that is, the smallest integer for which $\left(f_0(\mathbf{x}_k^*) - f_0(\mathbf{x}_{k-1}^*)\right) / f_0(\mathbf{x}_{k-1}^*) < \varepsilon$. Hence, before the termination of the algorithm, it holds that $\varepsilon \leq \left(f_0(\mathbf{x}_j^*) - f_0(\mathbf{x}_{j-1}^*)\right) / f_0(\mathbf{x}_{j-1}^*) \leq \left(f_0(\mathbf{x}_j^*) - f_0(\mathbf{x}_{j-1}^*)\right) / f_0(\mathbf{x}_0^*)$, and thus $\varepsilon f_0(\mathbf{x}_0^*) \leq f_0(\mathbf{x}_j^*) - f_0(\mathbf{x}_{j-1}^*)$ for $1 \leq j \leq k-1$ (if $k = 1$, there is no such j). Now, by taking the sum from $j = 1$ to $k-1$, we get $\sum_{j=1}^{k-1} \varepsilon f_0(\mathbf{x}_0^*) \leq \sum_{j=1}^{k-1} f_0(\mathbf{x}_j^*) - \sum_{j=1}^{k-1} f_0(\mathbf{x}_{j-1}^*) \Rightarrow (k-1)\varepsilon f_0(\mathbf{x}_0^*) \leq \sum_{j=1}^{k-1} f_0(\mathbf{x}_j^*) - \sum_{j=0}^{k-2} f_0(\mathbf{x}_j^*) = f_0(\mathbf{x}_{k-1}^*) - f_0(\mathbf{x}_0^*)$. Due to Property (a) of Theorem 3.1, every feasible point of problem $\mathcal{H}_j$ is also feasible for problem $\mathcal{F}$ ($S_j \subseteq S$, $j \geq 1$), and therefore $f_0(\mathbf{x}_j^*) \leq f_0(\mathbf{x}^*)$ for $j \geq 0$ ($\mathbf{x}^*$ is a global maximum of problem $\mathcal{F}$). This implies that $f_0(\mathbf{x}_{k-1}^*) \leq f_0(\mathbf{x}^*)$, and thus $(k-1)\varepsilon f_0(\mathbf{x}_0^*) \leq f_0(\mathbf{x}^*) - f_0(\mathbf{x}_0^*)$ $\Rightarrow k \leq 1 + (\lambda - 1)/\varepsilon = O\left(\lambda/\varepsilon\right)$, where $\lambda = f_0(\mathbf{x}^*)/f_0(\mathbf{x}_0^*) \geq 1$. Since the number of iterations until convergence is $O\left(\lambda/\varepsilon\right)$ and in each iteration a convex problem is solved with complexity $\varphi(n, I)$, Theorem 3.2 follows immediately. □

In general, a global optimum of a convex problem can be obtained in polynomial time, using standard convex optimization techniques such as interior-point methods [6] (i.e., $\varphi(n, I)$ is a polynomial function of n and I). Note that the best upper-complexity-bound for a generic convex problem, known so far, is $O(n^4)$ and is yielded by interior-point methods [9].

Bibliography

[1] A. Zappone, L. Sanguinetti, G. Bacci, E. A. Jorswieck, and M. Debbah, "Energy-efficient power control: A look at 5G wireless technologies," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1668-1683, April 2016.

[2] C. Guo, B. Liao, L. Huang, X. Lin and J. Zhang, "On convexity of fairness-aware energy-efficient power allocation in spectrum-sharing networks," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 534-537, March 2016.

[3] Y. Yang and M. Pesavento, "Energy efficiency in MIMO interference channels: Social optimality and max-min fairness," in *Proc. IEEE ICASSP*, Calgary, AB, Canada, 2018, pp. 3689-3693.

[4] L. Venturino, A. Zappone, C. Risi, and S. Buzzi, "Energy-efficient scheduling and power allocation in downlink OFDMA networks with base station coordination," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 1-14, Jan. 2015.

[5] C. N. Efrem and A. D. Panagopoulos, "A framework for weighted-sum energy efficiency maximization in wireless networks," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 153-156, Feb. 2019.

[6] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, UK: Cambridge Univ. Press, 2004.

[7] E. Bjornson, E. A. Jorswieck, M. Debbah, and B. Ottersten, "Multiobjective signal processing optimization: The way to balance conflicting metrics in 5G systems," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 14-23, Nov. 2014.

[8] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for non-convex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681-683, 1978.

[9] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications.* Philadelphia, PA, USA: SIAM, 2001.

Chapter 4

Dynamic Energy-Efficient Power Allocation in Multibeam Satellite Systems[5]

Power consumption is a major limitation in the downlink of multibeam satellite systems, since it has a significant impact on the mass and lifetime of the satellite. In this context, we study a new energy-aware power allocation problem that aims to jointly minimize the unmet system capacity (USC) and total radiated power by means of multi-objective optimization. First, we transform the original nonconvex-nondifferentiable problem into an equivalent nonconvex-differentiable form by introducing auxiliary variables. Subsequently, we design a successive convex approximation (SCA) algorithm in order to attain a stationary point with reasonable complexity. Due to its fast convergence, this algorithm is suitable for dynamic resource allocation in emerging on-board processing technologies. In addition, we formally prove a new result about the complexity of the SCA method, in the general case, that complements the existing literature where the complexity of this method is only numerically analyzed.

4.1 Introduction

Multibeam satellite systems (MSS) provide flexibility and efficient exploitation of the available resources in order to satisfy the (potentially asymmetric) traffic demand of users. Due to the fact that the satellite power is quite limited, resource allocation mechanisms should take into consideration not only the co-channel interference (CCI), but also the satellite power consumption in the downlink transmission.

The joint problem of routing and power allocation in MSS is examined in [1], using Lyapunov stability theory. Moreover, the studies [2] and [3] deal with several resource allocation problems in MSS with and without CCI, respectively. In [4], a dynamic power allocation algorithm is proposed exploiting a rain attenuation stochastic model. A comparison between non-orthogonal frequency reuse (NOFR) and beam-hopping (BH) systems is presented in [5], where various capacity optimization schemes are reported. Furthermore, linear and nonlinear precoding techniques are investigated in [6] and [7].

Unlike previous works, a multi-objective approach that minimizes the USC together with the satellite power consumption is presented in [8]. In particular, a two-stage optimization is proposed to attain a set of Pareto optimal solutions using metaheuristics. However, these algorithms do not provide any optimality guarantee, and their performance is heavily affected by the optimization parameters. Besides, although this method is suitable for offline power allocation, it is rather inappropriate for online/real-time power allocation since it requires a lot of computation time to find nearly-optimal solutions.

In this chapter, we introduce a new performance metric, which has not been systematically studied so far, including both the USC and total power consumption. This is in contrast to the majority of recent studies that solely minimize either the former or the latter objective. Moreover, we develop an optimization algorithm which always converges and, assuming appropriate constraint qualifications, achieves a stationary point (first-order optimality guarantee) with relatively low complexity. In addition, numerical results show that the algorithm performance is almost independent of the initialization point. Consequently, the proposed algorithm can be used in dynamic wireless environments where the resource allocation should be decided in a very short time. Finally, a formal proof about the complexity of the SCA method is also given.

The rest of this study is organized as follows. In Section 4.2, the optimization problem is formulated and then transformed into an equivalent differentiable form. Afterwards, based on the SCA method, we design an energy-efficient power allocation algorithm in Section 4.3. The performance of this algorithm is analyzed through simulations in Section 4.4, and some conclusions are provided in Section 4.5.

4.2 Problem Formulation and Transformation

Consider a multibeam satellite system with a geostationary satellite using N beams ($\mathcal{N} = \{1, 2, \ldots, N\}$) and K subcarriers (SCs) of bandwidth B_{SC} ($\mathcal{K} = \{1, 2, \ldots, K\}$). For notation simplicity and without loss of generality, it is assumed that: 1) the total bandwidth, $B_{tot} = KB_{SC}$, is reused by all beams, i.e., the frequency reuse factor is equal to 1 (worst-case scenario), and 2) during a specific time slot, each beam serves only one user within its coverage area (user i is served by the i^{th} satellite beam, $\forall i \in \mathcal{N}$). Moreover, we focus on the downlink (data transmission from the satellite to users) considering ideal, without noise and interference, feeder links between the gateways and the satellite.

The signal to interference-and-noise ratio (SINR) of the i^{th} user ($i \in \mathcal{N}$) on the k^{th} SC ($k \in \mathcal{K}$) is expressed by: $\gamma_i^{[k]} = g_{i,i}^{[k]} p_i^{[k]} \Big/ \left(\sum_{j \in \mathcal{N} \setminus i} g_{j,i}^{[k]} p_j^{[k]} + \sigma_{i,k}^2 \right)$, where $p_j^{[k]}$ is the transmit power of the j^{th} satellite beam, $\sigma_{i,k}^2$ is the thermal noise power of the i^{th} user, and $g_{j,i}^{[k]}$ is the channel power gain between the j^{th} satellite beam and the i^{th} user, all over the k^{th} SC. More precisely, $g_{j,i}^{[k]}$ includes free-space path loss (FSPL), rain attenuation, transmit antenna gain of satellite beam as well as receive antenna gain of user. For the sake of convenience, the transmit power vector is denoted by $\mathbf{p} = \left[\mathbf{p}^{[1]}, \mathbf{p}^{[2]}, \ldots, \mathbf{p}^{[K]}\right]$, where $\mathbf{p}^{[k]} = \left[p_1^{[k]}, p_2^{[k]}, \ldots, p_N^{[k]}\right]$, $\forall k \in \mathcal{K}$. In addition, the

USC [9] is defined by:

$$USC(\mathbf{p}) = \sum_{i\in\mathcal{N}} \max\left(C_i^{req} - C_i(\mathbf{p}), 0\right) \tag{4.1}$$

where C_i^{req} and $C_i(\mathbf{p}) = B_{SC} \sum_{k\in\mathcal{K}} \log_2\left(1+\gamma_i^{[k]}\right)$ are the i^{th} user's requested and offered capacity (in bps), respectively.[6] Moreover, the total radiated power is given by:

$$P_{tot}(\mathbf{p}) = \sum_{i\in\mathcal{N}} \sum_{k\in\mathcal{K}} p_i^{[k]} \tag{4.2}$$

Focusing on the multi-objective optimization, we study the following nonconvex minimization problem:

$$\min_{\mathbf{p}\in Z} \quad f(\mathbf{p}) = USC(\mathbf{p}) + w\,P_{tot}(\mathbf{p}) \tag{4.3}$$

with convex feasible set $Z = \{\mathbf{p} \in \mathbb{R}_+^{NK} : \sum_{k\in\mathcal{K}} p_i^{[k]} \leq P_i^{\max}, \ \forall i \in \mathcal{N}$ and $\sum_{i\in\mathcal{N}} \sum_{k\in\mathcal{K}} p_i^{[k]} \leq P_{tot}^{\max}\}$, where $P_i^{\max}$ is the maximum transmit power of the i^{th} satellite beam, and $P_{tot}^{\max}$ is the maximum total radiated power of the satellite.[7] The fixed/predefined weight $w \in [0, +\infty)$ is measured in bps/W, and expresses the priority of the total radiated power with respect to USC. Consequently, a trade-off between the USC and total power consumption (which is proportional to the total radiated power) can be achieved for a specific value of w. In particular, $w = 0$ corresponds to USC minimization. Moreover, it can be proved that problem (4.3) is NP-hard by following similar arguments as in [8]. Nevertheless, as will be seen later, we can obtain a stationary point of the equivalent differentiable problem with reasonable complexity.

Afterwards, by applying the transformation $\mathbf{p} = 2^{\mathbf{y}}$ ($p_i^{[k]} = 2^{y_i^{[k]}}$, $\forall i \in \mathcal{N}$, $k \in \mathcal{K}$), where $\mathbf{y} = \left[\mathbf{y}^{[1]}, \mathbf{y}^{[2]}, \ldots, \mathbf{y}^{[K]}\right]$ with $\mathbf{y}^{[k]} = \left[y_1^{[k]}, y_2^{[k]}, \ldots, y_N^{[k]}\right]$, $\forall k \in \mathcal{K}$, we obtain the

[6] In case of *adaptive coding and modulation (ACM)*, the offered capacity can be approximated by $C_i^{ACM}(\mathbf{p}) \approx B_{SC} \sum_{k\in\mathcal{K}} \log_2\left(1+\zeta\gamma_i^{[k]}\right)$ without altering the methodology, where $\zeta \in (0,1)$ is obtained through curve fitting (offered capacity versus SINR).

[7] It is possible to have additional *minimum-capacity constraints for each user* ($C_i(\mathbf{p}) \geq C_i^{min}$, $\forall i \in \mathcal{N}$) in order to increase the system availability (the methodology remains the same).

equivalent nonconvex problem:

$$\min_{\mathbf{y}\in S} \quad f(2^{\mathbf{y}}) = USC(2^{\mathbf{y}}) + w\,P_{tot}(2^{\mathbf{y}}) \tag{4.4}$$

with convex feasible set $S = \{\mathbf{y} \in \mathbb{R}^{NK} : \sum_{k\in\mathcal{K}} 2^{y_i^{[k]}} \leq P_i^{\max},\ \forall i \in \mathcal{N}$ and $\sum_{i\in\mathcal{N}}\sum_{k\in\mathcal{K}} 2^{y_i^{[k]}} \leq P_{tot}^{\max}\}$. Notice that the above transformation reduces the number of constraints by NK (lower complexity), since $\mathbf{p} \in \mathbb{R}_+^{NK}$ becomes $\mathbf{y} \in \mathbb{R}^{NK}$.

Finally, in order to remove the non-differentiability of the objective function, we rewrite problem (4.4) in its *equivalent epigraph-form* [10] using the auxiliary variable $\mathbf{t} = [t_1, t_2, \ldots, t_N]$:

$$\min_{(\mathbf{y},\mathbf{t})\in\Omega} \quad F(\mathbf{y},\mathbf{t}) = \sum_{i\in\mathcal{N}} t_i + w\sum_{i\in\mathcal{N}}\sum_{k\in\mathcal{K}} 2^{y_i^{[k]}} \tag{4.5}$$

with nonconvex feasible set $\Omega = \{(\mathbf{y},\mathbf{t}) \in \mathbb{R}^{NK+N} :\ t_i \geq 0,\ t_i \geq C_i^{req} - C_i(2^{\mathbf{y}}),$ $\forall i \in \mathcal{N}$ and $\mathbf{y} \in S\}$. Observe that the new objective $F(\mathbf{y},\mathbf{t})$ is convex now, and the first two constraints in Ω are equivalent to $t_i \geq \max\left(C_i^{req} - C_i(2^{\mathbf{y}}), 0\right)$, $\forall i \in \mathcal{N}$. Furthermore, problem (4.5) is equivalent to problem (4.4) in the following sense: $(\mathbf{y},\mathbf{t})$ is a global optimum of (4.5) if and only if $\mathbf{y}$ is a global optimum of (4.4) and $t_i = \max\left(C_i^{req} - C_i(2^{\mathbf{y}}), 0\right)$, $\forall i \in \mathcal{N}$.

4.3 Energy-Efficient Power Allocation

Subsequently, we utilize the mathematical tool of SCA (refer to the Appendix 4.6) in order to tackle problem (4.5) with relatively low complexity. Firstly, the offered capacity can be written as follows: $C_i(2^{\mathbf{y}}) = B_{SC}\sum_{k\in\mathcal{K}}\left[\varphi_i^{[k]}\left(\mathbf{y}^{[k]}\right) - \vartheta_i^{[k]}\left(\mathbf{y}^{[k]}\right)\right]$, where $\varphi_i^{[k]}\left(\mathbf{y}^{[k]}\right)$ and $\vartheta_i^{[k]}\left(\mathbf{y}^{[k]}\right)$ are convex functions given by (note that the log-sum-exp function is convex [10]):

$$\varphi_i^{[k]}\left(\mathbf{y}^{[k]}\right) = \log_2\left(\sum_{j\in\mathcal{N}} g_{j,i}^{[k]} 2^{y_j^{[k]}} + \sigma_{i,k}^2\right) \tag{4.6}$$

Algorithm 4.1 Energy-Efficient Power Allocation

1: Select a starting point $\mathbf{p} \in Z$, and a tolerance $\epsilon > 0$
2: Set $\ell = 0$, $\mathbf{y} = \log_2(\mathbf{p})$, $t_i = \max\left(C_i^{req} - C_i(\mathbf{p}), 0\right)$, $\forall i \in \mathcal{N}$ and $F_0 = F(\mathbf{y}, \mathbf{t})$
3: **repeat**
4: Solve the convex minimization problem (4.8) with approximation point $\bar{\mathbf{y}} = \mathbf{y}$ in order to achieve a global optimum $(\mathbf{y}^*, \mathbf{t}^*)$
5: Set $\ell = \ell + 1$, $\mathbf{y} = \mathbf{y}^*$, $\mathbf{t} = \mathbf{t}^*$, $\mathbf{p} = 2^{\mathbf{y}}$ and $F_\ell = F(\mathbf{y}, \mathbf{t})$
6: **until** $|F_\ell - F_{\ell-1}| \leq \epsilon\, |F_{\ell-1}|$

$$\vartheta_i^{[k]}\left(\mathbf{y}^{[k]}\right) = \log_2\left(\sum_{j \in \mathcal{N} \setminus i} g_{j,i}^{[k]} 2^{y_j^{[k]}} + \sigma_{i,k}^2\right) \tag{4.7}$$

Now, for a given approximation point $\bar{\mathbf{y}} \in \mathbb{R}^{NK}$, we can construct the next convex minimization problem:

$$\min_{(\mathbf{y},\mathbf{t}) \in \Theta(\bar{\mathbf{y}})} \quad F(\mathbf{y}, \mathbf{t}) = \sum_{i \in \mathcal{N}} t_i + w \sum_{i \in \mathcal{N}} \sum_{k \in \mathcal{K}} 2^{y_i^{[k]}} \tag{4.8}$$

with convex feasible set $\Theta(\bar{\mathbf{y}}) = \{(\mathbf{y}, \mathbf{t}) \in \mathbb{R}^{NK+N} : \; t_i \geq 0, \; t_i \geq C_i^{req} - \widetilde{C}_i(\mathbf{y}, \bar{\mathbf{y}}),$ $\forall i \in \mathcal{N}$ and $\mathbf{y} \in S\}$, where:

$$\widetilde{C}_i(\mathbf{y}, \bar{\mathbf{y}}) = B_{SC} \sum_{k \in \mathcal{K}} \left[\widetilde{\varphi_i^{[k]}}\left(\mathbf{y}^{[k]}, \bar{\mathbf{y}}^{[k]}\right) - \vartheta_i^{[k]}\left(\mathbf{y}^{[k]}\right)\right] \tag{4.9}$$

$$\widetilde{\varphi_i^{[k]}}\left(\mathbf{y}^{[k]}, \bar{\mathbf{y}}^{[k]}\right) = \varphi_i^{[k]}\left(\bar{\mathbf{y}}^{[k]}\right) + \nabla \varphi_i^{[k]}\left(\bar{\mathbf{y}}^{[k]}\right) \cdot \left(\mathbf{y}^{[k]} - \bar{\mathbf{y}}^{[k]}\right)^T \tag{4.10}$$

Observe that $\widetilde{C}_i(\mathbf{y}, \bar{\mathbf{y}})$ is a concave function of $\mathbf{y}$. In addition, the elements of $\nabla \varphi_i^{[k]}\left(\bar{\mathbf{y}}^{[k]}\right)$ are given by:

$$\frac{\partial \varphi_i^{[k]}\left(\bar{\mathbf{y}}^{[k]}\right)}{\partial y_l^{[k]}} = \frac{g_{l,i}^{[k]} 2^{\bar{y}_l^{[k]}}}{\sum_{j \in \mathcal{N}} g_{j,i}^{[k]} 2^{\bar{y}_j^{[k]}} + \sigma_{i,k}^2}, \quad \forall l \in \mathcal{N} \tag{4.11}$$

Algorithm 4.1 presents an iterative process based on the SCA method. In particular, we provide the next proposition which readily follows from Theorems 4.1 and 4.2 in the Appendix 4.6. Note that the number of variables and constraints of problem (4.8) is polynomial in N and K ($NK + N$ and $3N + 1$, respectively).

Proposition 4.1. *Algorithm 4.1 generates a monotonically decreasing sequence* $\{F_\ell\}_{\ell\geq 0}$ *(i.e.,* $F_{\ell+1} \leq F_\ell$*) and converges to a finite value* L *(*$\lim_{\ell\to\infty} F_\ell = L > -\infty$*). Moreover, assuming suitable constraint qualifications,* $L = \lim_{\ell\to\infty} F_\ell = F\left(\hat{\mathbf{y}}, \hat{\mathbf{t}}\right)$ *for some stationary point* $\left(\hat{\mathbf{y}}, \hat{\mathbf{t}}\right)$ *of problem (4.5). Finally, the complexity of Algorithm 4.1 is* $\mathcal{O}\left(\left(\xi/\epsilon\right) h(N,K)\right)$*, where* $\xi = F_0/F_* \geq 1$*, with* F_* *being the globally minimum objective value of problem (4.5), and* $h(N,K)$ *is the complexity of the convex problem (4.8) which is polynomial in* N *and* K*.*

4.4 Numerical Simulations and Discussion

In this section, we examine a MSS with the parameters given in Table 4.1. Unless otherwise specified, the tolerance and the starting point of Algorithm 4.1 are selected as $\epsilon = 10^{-3}$ and $\mathbf{p} = \left(P_{tot}^{\max}/(NK)\right) \mathbf{1}_{1\times NK}$, where $\mathbf{1}_{1\times NK}$ is the all-ones $1\times NK$ vector. As concerns the requested capacities of the users, we have assumed an asymmetric traffic distribution according to the linear model: $C_i^{req} = r\,i$, $\forall i \in \mathcal{N}$, where r is the traffic slope measured in bps. Furthermore, each satellite beam antenna has the following radiation pattern [6], [8]: $G(\theta) = G_{\max}\left(\frac{J_1(u)}{2u} + 36\frac{J_3(u)}{u^3}\right)^2$, where θ is the angle between the corresponding beam center and the user location with respect to the satellite, $G_{\max}$ is the maximum satellite beam antenna gain ($G(0) = G_{\max}$), $u = 2.07123\frac{\sin(\theta)}{\sin(\theta_{3\text{dB}})}$ with $\theta_{3\text{dB}}$ the 3-dB angle ($G(\theta_{3\text{dB}}) = G_{\max}/2$), and $J_1(u)$, $J_3(u)$ are respectively the first and third order Bessel functions of the first kind.

All graphs, except for Fig. 4-3, present statistical averages derived from 200 independent Monte Carlo simulations, where each user is uniformly distributed within its beam coverage area. For the sake of comparison, we have used a conventional scheme, namely, uniform power allocation (UPA), where $p_{i,UPA}^{[k]} = P_{tot}^{\max}/(NK)$, $\forall i \in \mathcal{N}$, $k \in \mathcal{K}$.

Firstly, we investigate the convergence speed of the proposed algorithm for $w = 0$, $10\,\text{Mbps/W}$ and different starting points. As shown in Fig. 4-1, Algorithm 4.1 achieves nearly the same convergence rate and final objective value regardless of the starting point. Given the tolerance $\epsilon = 10^{-3}$, the proposed algorithm requires about

Table 4.1: System Parameters

Parameter	**Value**
Beam radius	150 km
Frequency band	Ka (20 GHz)
Number of beams and SCs	$N = 7$, $K = 4$
Subcarrier bandwidth (B_{SC})	125 MHz
Thermal noise power ($\sigma_{i,k}^2 = \sigma^2$, $\forall i \in \mathcal{N}$, $k \in \mathcal{K}$)	-124 dBW
Maximum beam power ($P_i^{\max} = P_{\max}$, $\forall i \in \mathcal{N}$)	100 W
Maximum total power ($P_{tot}^{\max}$)	500 W
Free-space path loss	210 dB
Rain attenuation mean and standard deviation	2.6 dB, 1.63 dB
User antenna gain	41.7 dBi
Maximum satellite beam antenna gain ($G_{\max}$)	52 dBi
3-dB angle ($\theta_{3\text{dB}}$)	0.2°

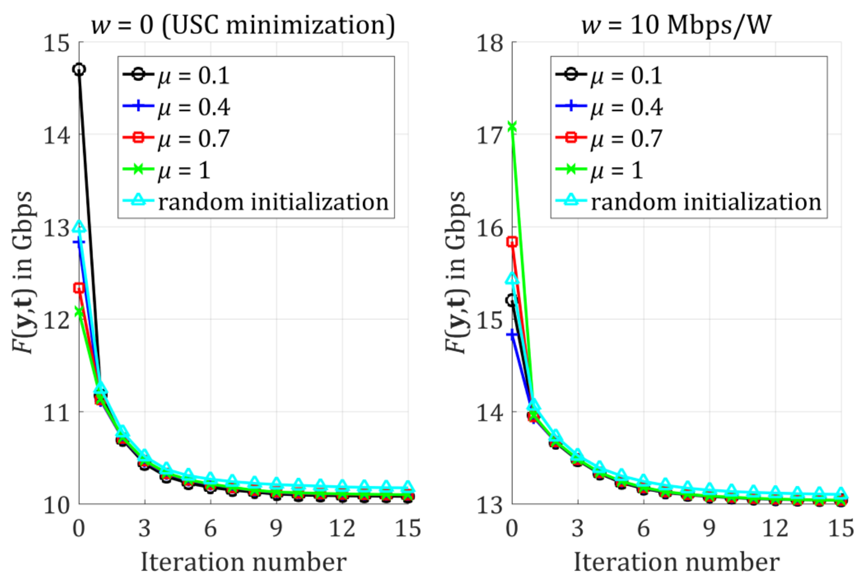

Figure 4-1: Convergence of Algorithm 4.1 for $r = 0.7$ Gbps, and starting point $\mathbf{p} = \mu\,(P_{tot}^{\max}/(NK))\,\mathbf{1}_{1\times NK}$ or random initialization.

10 iterations to converge for both values of w and for all the starting points under consideration.

Secondly, Fig. 4-2 illustrates the USC and total radiated power achieved by the conventional scheme and Algorithm 4.1 (for two different weights) versus the traffic slope. Although the UPA scheme makes full use of the available power, it has the highest USC. On the other hand, for $w = 0$ (USC minimization) we have the lowest USC using less power than UPA. In addition, the last scheme with $w = 10$ Mbps/W

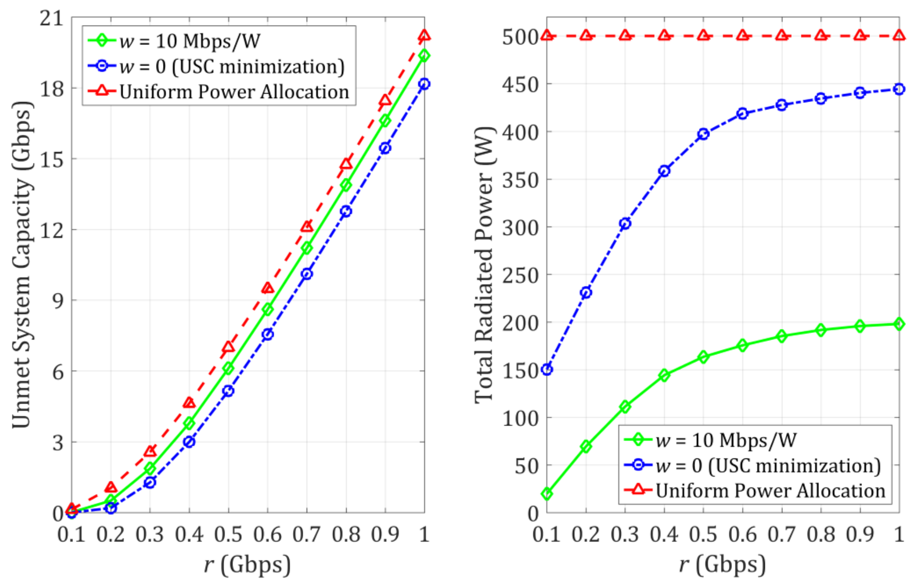

Figure 4-2: USC and total radiated power versus the traffic slope.

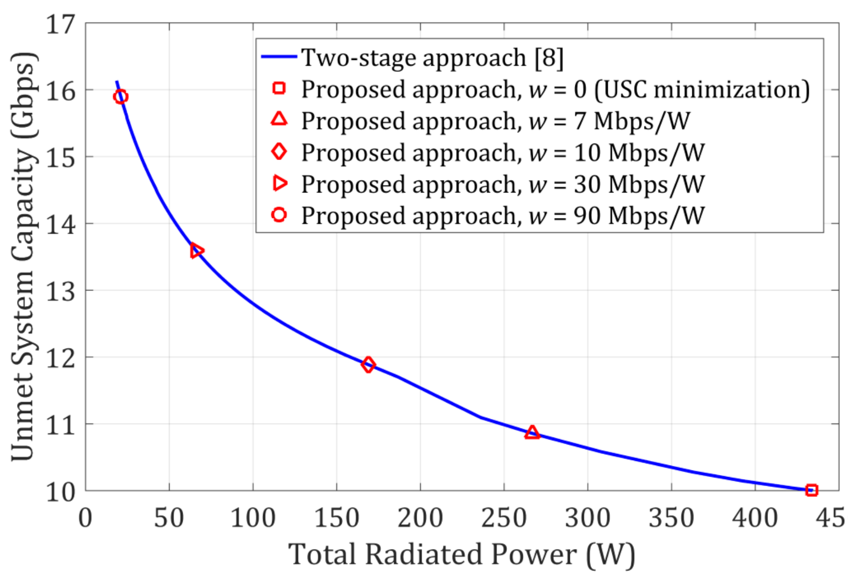

Figure 4-3: Performance comparison with the two-stage approach [8] for a particular system configuration with $r = 0.7\,$Gbps.

achieves an USC that lies between the other two schemes, but with much less power (high energy savings). This is expected because higher priority is given to the total radiated power as the weight w increases.

Last but not least, Fig. 4-3 compares the performance of the proposed method with the two-stage approach [8]. In particular, the 5 operating points attained by the proposed approach belong to the Pareto boundary obtained from [8]. It has been observed that many values of w achieve operating points on the Pareto boundary, but we only present 5 points for better illustration. Therefore, the proposed method shows similar performance with [8]. Note that in multi-objective optimization, there is no objectively optimal solution, but only Pareto/subjectively optimal solutions.

In summary, [8] presents *a posteriori method* where the network designer selects an operating point after the computation/visualization of the Pareto boundary, while this chapter introduces *a priori method* where the weight w is specified before any computation, and then a single solution is obtained. Finally, we would like to emphasize that the former approach is appropriate for offline power allocation (no strict limitations on processing time), whereas the latter approach is suitable for online/dynamic power allocation due to its rapid convergence.

4.5 Conclusion

In this chapter, we have designed a SCA-based optimization algorithm with high convergence speed, which is suitable for real-time power allocation in MSS with strict computation/processing-time requirements. The proposed multi-objective approach enables network designers to achieve a compromise between the USC and total power consumption. Numerical simulations have also verified the advantage of this approach. Moreover, the complexity of the SCA method, in its general form, has been studied theoretically.

4.6 Appendix: Successive Convex Approximation

SCA is an iterative method that attains a stationary point of a nonconvex optimization problem by solving a sequence of convex problems [11]. Despite the fact that the achieved solution may or may not be globally optimal, this technique has reasonable computational complexity. More specifically, the following theorem is provided, where *all the functions are assumed to be differentiable (and therefore continuous).*

Theorem 4.1 ([11]). *Let $\mathcal{P}$ be a nonconvex minimization problem with objective $\psi_0(\boldsymbol{x})$, and nonempty-compact feasible set $D = \{\boldsymbol{x} \in \mathbb{R}^n : \psi_i(\boldsymbol{x}) \leq 0,\ 1 \leq i \leq m\}$, with $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]$. Moreover, suppose that $\psi_i(\boldsymbol{x}) = u_i(\boldsymbol{x}) - v_i(\boldsymbol{x})$ for $0 \leq i \leq m$, where $u_i(\boldsymbol{x})$ and $v_i(\boldsymbol{x})$ are convex functions. Let $\left\{\widetilde{\mathcal{P}_j}\right\}_{j\geq 1}$ be a sequence of convex minimization problems with objective $\widetilde{\psi_{0,j}}(\boldsymbol{x}, \boldsymbol{x}^*_{j-1})$, compact feasible set $D_j = \{\boldsymbol{x} \in \mathbb{R}^n : \widetilde{\psi_{i,j}}(\boldsymbol{x}, \boldsymbol{x}^*_{j-1}) \leq 0,\ 1 \leq i \leq m\}$, and global minimum $\boldsymbol{x}^*_j$ (with $\boldsymbol{x}^*_0 \in D$). If $\widetilde{\psi_{i,j}}(\boldsymbol{x}, \boldsymbol{x}^*_{j-1}) = u_i(\boldsymbol{x}) - \widetilde{v_i}(\boldsymbol{x}, \boldsymbol{x}^*_{j-1})$ for $0 \leq i \leq m$ and $j \geq 1$, where $\widetilde{v_i}(\boldsymbol{x}, \boldsymbol{x}^*_{j-1}) = v_i(\boldsymbol{x}^*_{j-1}) + \nabla v_i(\boldsymbol{x}^*_{j-1}) \cdot \left(\boldsymbol{x} - \boldsymbol{x}^*_{j-1}\right)^T$, with $\nabla v_i(\boldsymbol{x}) = [\partial v_i(\boldsymbol{x})/\partial x_1, \ldots, \partial v_i(\boldsymbol{x})/\partial x_n]$, then: (a) $\boldsymbol{x}^*_{j-1} \in D_j \subseteq D$ and $\psi_0(\boldsymbol{x}^*_j) \leq \psi_0(\boldsymbol{x}^*_{j-1})$, $\forall j \geq 1$, (b) $\lim_{j\to\infty} \psi_0(\boldsymbol{x}^*_j) = \psi_0(\widehat{\boldsymbol{x}}) = L > -\infty$ for all the accumulation/limit points $\widehat{\boldsymbol{x}}$ of the sequence $\left\{\boldsymbol{x}^*_j\right\}_{j\geq 0}$, and (c) assuming suitable constraint qualifications, all the accumulation points $\widehat{\boldsymbol{x}}$ are stationary points of $\mathcal{P}$ (i.e., satisfy the corresponding Karush-Kuhn-Tucker conditions), and $L = \lim_{j\to\infty} \psi_0(\boldsymbol{x}^*_j) = \psi_0(\hat{\boldsymbol{x}})$, where $\hat{\boldsymbol{x}}$ is some stationary point of $\mathcal{P}$.*

Taking advantage of the fact that SCA generates a monotonically decreasing sequence of objective values, and using *the property of telescoping sums*: $\sum_{l=1}^{M} (a_{l-1} - a_l) = a_0 - a_M$ for any integer $M \geq 1$, we introduce and prove the following result concerning the complexity of the SCA method.

Theorem 4.2. *Suppose that the SCA method terminates when $\left|\psi_0(\boldsymbol{x}^*_j) - \psi_0(\boldsymbol{x}^*_{j-1})\right| \leq \epsilon \left|\psi_0(\boldsymbol{x}^*_{j-1})\right|$ for some predefined tolerance $\epsilon > 0$, and $\psi_0(\boldsymbol{x}^*) > 0$, where $\boldsymbol{x}^*$ is a global minimum of $\mathcal{P}$. Then, the complexity of the SCA method is $\mathcal{O}\left((\xi/\epsilon)\, h(n, m)\right)$, where $\xi = \psi_0(\boldsymbol{x}^*_0)/\psi_0(\boldsymbol{x}^*) \geq 1$ and $h(n, m)$ is the complexity of each convex optimization problem which is a polynomial function of the number of variables and constraints (n and m, respectively).*

Proof. According to Theorem 4.1, it holds that $\psi_0(\boldsymbol{x}^*_0) \geq \psi_0(\boldsymbol{x}^*_{j-1}) \geq \psi_0(\boldsymbol{x}^*_j) \geq \psi_0(\boldsymbol{x}^*) > 0$, $\forall j \geq 1$. As concerns the number of iterations until convergence, if we denote by ν the smallest integer such that $\left|\psi_0(\boldsymbol{x}^*_\nu) - \psi_0(\boldsymbol{x}^*_{\nu-1})\right| \leq \epsilon \left|\psi_0(\boldsymbol{x}^*_{\nu-1})\right| \Leftrightarrow \psi_0(\boldsymbol{x}^*_{\nu-1}) - \psi_0(\boldsymbol{x}^*_\nu) \leq \epsilon\, \psi_0(\boldsymbol{x}^*_{\nu-1})$, then for all integers less than ν the last inequality does not hold: $\psi_0(\boldsymbol{x}^*_{l-1}) - \psi_0(\boldsymbol{x}^*_l) > \epsilon\, \psi_0(\boldsymbol{x}^*_{l-1}) \geq \epsilon\, \psi_0(\boldsymbol{x}^*) \Rightarrow \psi_0(\boldsymbol{x}^*_{l-1}) - \psi_0(\boldsymbol{x}^*_l) > \epsilon\, \psi_0(\boldsymbol{x}^*)$, $\forall l \in \{1, 2, \ldots, \nu - 1\}$. By summing from 1 to $\nu - 1$, we obtain

$\sum_{l=1}^{\nu-1}\left(\psi_0(\boldsymbol{x}_{l-1}^*) - \psi_0(\boldsymbol{x}_l^*)\right) > \sum_{l=1}^{\nu-1}\epsilon\,\psi_0(\boldsymbol{x}^*) \Rightarrow \psi_0(\boldsymbol{x}_0^*) - \psi_0(\boldsymbol{x}_{\nu-1}^*) > (\nu-1)\,\epsilon\,\psi_0(\boldsymbol{x}^*)$. Since $\psi_0(\boldsymbol{x}_{\nu-1}^*) \geq \psi_0(\boldsymbol{x}^*)$, we get $(\nu-1)\,\epsilon\,\psi_0(\boldsymbol{x}^*) < \psi_0(\boldsymbol{x}_0^*) - \psi_0(\boldsymbol{x}^*)$, and therefore $\nu < 1 + (\xi-1)/\epsilon < 1 + \xi/\epsilon = \mathcal{O}(\xi/\epsilon)$. Hence, the SCA method requires $\mathcal{O}(\xi/\epsilon)$ iterations to converge. Moreover, each convex optimization problem can be globally solved with polynomial complexity in the number of variables and constraints [10], and thus Theorem 4.2 follows directly. □

Bibliography

[1] M. J. Neely, E. Modiano and C. E. Rohrs, "Power allocation and routing in multibeam satellites with time-varying channels," *IEEE/ACM Transactions on Networking*, vol. 11, no. 1, pp. 138-152, Feb. 2003.

[2] J. P. Choi and V. W. S. Chan, "Resource management for advanced transmission antenna satellites," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1308-1321, March 2009.

[3] J. P. Choi and V. W. S. Chan, "Optimum power and beam allocation based on traffic demands and channel conditions over satellite downlinks," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2983-2993, Nov. 2005.

[4] A. Destounis and A. D. Panagopoulos, "Dynamic power allocation for broadband multi-beam satellite communication networks," *IEEE Communications Letters*, vol. 15, no. 4, pp. 380-382, April 2011.

[5] J. Lei and M. Á. Vázquez-Castro, "Multibeam satellite frequency/time duality study and capacity optimization," *Journal of Communications and Networks*, vol. 13, no. 5, pp. 472-480, Oct. 2011.

[6] D. Christopoulos *et al.*, "Linear and nonlinear techniques for multibeam joint processing in satellite communications," *EURASIP Journal on Wireless Communications and Networking*, 2012.

[7] G. Zheng, S. Chatzinotas and B. Ottersten, "Generic optimization of linear precoding in multibeam satellite systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2308-2320, June 2012.

[8] A. I. Aravanis *et al.*, "Power allocation in multibeam satellite systems: A two-stage multi-objective optimization," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3171-3182, June 2015.

[9] J. Anzalchi *et al.*, "Beam hopping in multi-beam broadband satellite systems: System simulation and performance comparison with non-hopped systems," *5th Advanced Satellite Multimedia Systems Conference and the 11th Signal Processing for Space Communications Workshop*, 2010.

[10] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge Univ. Press, 2004.

[11] B. K. Sriperumbudur and G. R. Lanckriet, "On the convergence of the concave-convex procedure," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 1759–1767.

Chapter 5

Globally Optimal Selection of Ground Stations in Satellite Systems with Site Diversity[8]

The availability of satellite communication systems is extremely limited by atmospheric impairments, such as rain (for radio frequencies) and cloud coverage (for optical frequencies). A solution to this problem is the site diversity technique, where a network of geographically distributed ground stations (GSs) can ensure, with high probability, that at least one GS is available for connection to the satellite at each time period. However, the installation of redundant GSs induces unnecessary additional costs for the network operator. In this context, we study an optimization problem that minimizes the number of required GSs, subject to availability constraints. First, the problem is transformed into a binary-integer-linear-programming (BILP) problem, which is proven to be NP-hard. Subsequently, we design a branch-and-bound (B&B) algorithm, with global-optimization guarantee, based on the linear-programming (LP) relaxation and a greedy method as well. Finally, numerical results show that the proposed algorithm significantly outperforms state-of-the-art methods and has low

[8]Copyright © 2020 IEEE. Reprinted, with permission, from: C. N. Efrem and A. D. Panagopoulos, "Globally optimal selection of ground stations in satellite systems with site diversity," *IEEE Wireless Communications Letters*, vol. 9, no. 7, pp. 1101-1104, July 2020. Personal use of this material is permitted, but republication/redistribution requires IEEE permission.

complexity in the average case.

5.1 Introduction

Site diversity technique is used to improve the availability of satellite communication systems by mitigating the atmospheric effects [1]. In particular, multiple GSs separated over long distances receive the same signal from the satellite, and in this way the probability of all GSs experiencing severe weather conditions simultaneously is reduced. A joint optimization method for the design of optical satellite networks is proposed in [2], which consists of two parts. The first part is the optical-GS positioning optimization performed by an iterative greedy procedure, while the second part is the backbone network optimization taking into consideration the optical fiber cost. In [3], a network optimization method with reduced complexity is presented, exploiting the single-site availabilities as well as the correlation between sites.

Furthermore, the optimal location of optical GSs for low-earth-orbit (LEO) satellite missions is examined in [4] through multi-objective optimization, using genetic algorithms (GAs) and considering three performance metrics: system availability, latency, and network cost. GAs are also used in [5] to minimize two different objective functions in extremely-high-frequency (EHF) satellite networks with smart-gateway (SG) diversity. In addition, the selection of the minimum number of GSs in optical satellite networks with a medium-earth-orbit (MEO) or a geostationary (GEO) satellite is investigated in [6] and [7], respectively. Both studies present heuristic algorithms of low complexity, taking into account the spatial correlation as well as the monthly variability of cloud coverage.

The *main contributions* of this chapter compared to existing approaches are the following: 1) rigorous mathematical formulation of the optimization problem with a formal proof of its NP-hardness, 2) system availability guarantee for several time periods (e.g., months), not only for a year, and 3) unlike existing methods that provide suboptimal solutions without any performance guarantee, the designed B&B algorithm achieves global optimality with low average-case complexity (i.e., good trade-off between performance and complexity).

The remainder of this chapter is structured as follows. Section 5.2 presents the system model and formulates the optimization problem, which is transformed into an NP-hard BILP problem in Section 5.3. Afterwards, a global optimization algorithm is given in Section 5.4, while its performance is numerically analyzed in Section 5.5. Finally, Section 5.6 concludes this chapter.

Mathematical notation: The absolute value of a real number x is denoted by $|x|$, while $|\mathcal{D}| = D$ represents the cardinality of a set $\mathcal{D}$. Also, $\mathbf{0}_N/\mathbf{1}_N$ stands for the N-dimensional zero/all-ones vector respectively, and $\lceil\cdot\rceil$ is the ceiling function.

5.2 System Model and Problem Formulation

We consider a satellite system with a geostationary satellite and a ground station network employing site diversity. Specifically, $\mathcal{K} = \{1, 2, \dots, K\}$ is the set of available locations for installing a GS (or, equivalently, the set of candidate GSs), and $\mathcal{T} = \{1, 2, \dots, T\}$ denotes the set of time periods (e.g., months). In addition, $p_{k,t}^{\text{out}}$ is the outage probability of GS k in time period t,[9] and $P_t^{\text{out,req}}$ is the maximum required system outage probability in time period t.

Moreover, we make the following *assumptions*: a) $\{p_{k,t}^{\text{out}}\}_{k\in\mathcal{K}}$ are probabilities of *mutually independent* events, $\forall t \in \mathcal{T}$,[10] b) the system availability is defined as *the probability of having at least one GS available*, c) $\{p_{k,t}^{\text{out}}\}_{k\in\mathcal{K},\, t\in\mathcal{T}}$ are supposed to be accurate (i.e., *without uncertainty*); the uncertainty in the calculation of outage probabilities is beyond the scope of this chapter, and d) without loss of generality we assume that $p_{k,t}^{\text{out}}, P_t^{\text{out,req}} > 0$, $\forall k \in \mathcal{K}$ and $\forall t \in \mathcal{T}$.

In order to reduce the cost of installing and operating the GSs, we study the

[9] In *radio-frequency (RF) satellite systems*, a GS is in outage when the rain attenuation exceeds a specific threshold [5], which is determined by the required bit-error-rate (BER). In *optical satellite networks*, a GS is in outage when experiencing cloud blockage [8, 9]. Otherwise, the GS is available.

[10] This can be achieved if the distance between any two distinct GSs is sufficiently large, and therefore the spatial correlation of weather conditions is negligible. Furthermore, this case is quite common and preferable in practice, so as to take full advantage of site diversity by attaining the highest availability.

following *cardinality minimization problem*:

$$\begin{aligned} \min_{\mathcal{S}\subseteq\mathcal{K}} \quad & |\mathcal{S}| = S \\ \text{s.t.} \quad & P_t^{\text{avl}}(\mathcal{S}) \geq P_t^{\text{avl,req}}, \ \ \forall t \in \mathcal{T} \end{aligned} \tag{5.1}$$

where $\mathcal{S}$ denotes the set of selected GSs, $P_t^{\text{avl}}(\mathcal{S}) = 1 - \prod_{s\in\mathcal{S}} p_{s,t}^{\text{out}}$ is the system availability in time period t achieved by the set $\mathcal{S}$ of GSs (or, equivalently, the probability of having at least one GS in $\mathcal{S}$ available in time period t), and $P_t^{\text{avl,req}} = 1 - P_t^{\text{out,req}}$ is the minimum required system availability in time period t. Notice that $P_t^{\text{avl}}(\mathcal{S}) \geq P_t^{\text{avl,req}}$ $\Leftrightarrow \prod_{s\in\mathcal{S}} p_{s,t}^{\text{out}} \leq P_t^{\text{out,req}}$, $\forall t \in \mathcal{T}$.

5.3 Equivalent BILP Problem and NP-hardness

Subsequently, we introduce the vector $\mathbf{z} = [z_1, z_2, \ldots, z_K]$ of binary (0-1) variables. In particular, $z_k = 1$ if $k \in \mathcal{S}$, i.e., the k^{th} GS is selected (or, equivalently, a GS is installed at the k^{th} location), otherwise $z_k = 0$. Based on this definition, we have that $|\mathcal{S}| = \sum_{k\in\mathcal{K}} z_k$ and $\prod_{s\in\mathcal{S}} p_{s,t}^{\text{out}} = \prod_{k\in\mathcal{K}} (p_{k,t}^{\text{out}})^{z_k}$. As a result, problem (5.1) can be written as follows:

$$\begin{aligned} \min_{\mathbf{z}} \quad & \sum_{k\in\mathcal{K}} z_k \\ \text{s.t.} \quad & \prod_{k\in\mathcal{K}} (p_{k,t}^{\text{out}})^{z_k} \leq P_t^{\text{out,req}}, \ \ \forall t \in \mathcal{T} \\ & z_k \in \{0,1\}, \ \ \forall k \in \mathcal{K} \end{aligned} \tag{5.2}$$

By taking the logarithms on both sides of the inequality-constraints and then multiplying by -1, we obtain an equivalent BILP problem:

$$\begin{aligned} \min_{\mathbf{z}} \quad & g(\mathbf{z}) = \sum_{k\in\mathcal{K}} z_k \\ \text{s.t.} \quad & \sum_{k\in\mathcal{K}} \alpha_{t,k} z_k \geq \beta_t, \ \ \forall t \in \mathcal{T} \\ & z_k \in \{0,1\}, \ \ \forall k \in \mathcal{K} \end{aligned} \tag{5.3}$$

with $\alpha_{t,k} = \log(1/p_{k,t}^{\text{out}})$ and $\beta_t = \log(1/P_t^{\text{out,req}})$, $\forall t \in \mathcal{T}$ and $\forall k \in \mathcal{K}$. Note that

$\alpha_{t,k}, \beta_t \geq 0$, since $0 < p_{k,t}^{\text{out}}, P_t^{\text{out,req}} \leq 1$.

Theorem 5.1. *The equivalent BILP problem (5.3) is NP-hard.*

Proof. In order to prove the NP-hardness of problem (5.3), the following property is exploited: if a special case of a problem is NP-hard, so is the general problem. Now, we consider the *minimum node cover problem* (MNCP): Given a graph $G(\mathcal{N}, \mathcal{E})$, with $\mathcal{N}$ and $\mathcal{E}$ being the sets of nodes and edges respectively, find a minimum-cardinality set of nodes $\mathcal{N}' \subseteq \mathcal{N}$ such that $\{n, m\} \in \mathcal{E} \Rightarrow n \in \mathcal{N}'$ or $m \in \mathcal{N}'$. Furthermore, the MNCP is known to be NP-hard [10] and can be formulated as the following BILP problem:

$$\begin{aligned} \min_{\mathbf{z}} \quad & \sum_{n \in \mathcal{N}} z_n \\ \text{s.t.} \quad & z_n + z_m \geq 1, \;\; \forall \{n, m\} \in \mathcal{E} \\ & z_n \in \{0, 1\}, \;\; \forall n \in \mathcal{N} \end{aligned} \tag{5.4}$$

Obviously, the NP-hard problem (5.4) constitutes a special case of the general problem (5.3), and so we have Theorem 5.1. □

5.4 Global Optimization Algorithm

Since problem (5.3) is proven to be NP-hard, it cannot be solved in polynomial time unless P=NP. In other words, it is rather unlikely that there is an algorithm which finds an optimal solution and has polynomial complexity in the worst case. Nevertheless, we will design a global optimization B&B algorithm of low average-case complexity. B&B is an intelligent technique which recursively splits the search space into smaller spaces (*branching*), and uses appropriate bounds on the optimum value (*bounding*) to avoid, as much as possible, the exhaustive enumeration of candidate solutions [10].

Next, consider problem (5.3) with some variables being fixed:

$$\begin{aligned}\min_{\mathbf{z}_{\mathcal{V}}} \quad & g(\mathbf{z}_{\mathcal{V}};\bar{\mathbf{z}}_{\mathcal{C}}) = \sum_{v\in\mathcal{V}} z_v + \sum_{c\in\mathcal{C}} \bar{z}_c \\ \text{s.t.} \quad & \sum_{v\in\mathcal{V}} \alpha_{t,v} z_v \geq \beta'_t, \ \ \forall t \in \mathcal{T} \\ & z_v \in \{0,1\}, \ \ \forall v \in \mathcal{V}\end{aligned} \tag{5.5}$$

where the sets $\mathcal{V}$ and $\mathcal{C}$ contain the indices of free and constant variables respectively ($\mathcal{V} \cup \mathcal{C} = \mathcal{K}$, $\mathcal{V} \cap \mathcal{C} = \emptyset$), $\mathbf{z}_{\mathcal{V}} = [z_v]_{v\in\mathcal{V}}$, $\bar{\mathbf{z}}_{\mathcal{C}} = [\bar{z}_c]_{c\in\mathcal{C}}$ (with $\bar{z}_c \in \{0,1\}$, $\forall c \in \mathcal{C}$), and $\beta'_t = \beta_t - \sum_{c\in\mathcal{C}} \alpha_{t,c}\bar{z}_c$, $\forall t \in \mathcal{T}$. Notice that when $\mathcal{V} = \mathcal{K}$ and $\mathcal{C} = \emptyset$, problem (5.5) is identical to the original problem (5.3). Also, $\mathbf{z}_{\mathcal{V}}^{\text{opt}}$ denotes an optimal solution of problem (5.5), and $|\mathcal{V}| = V \leq K$.

Moreover, the following statements can be easily proven: a) $g^* \leq g(\mathbf{z}_{\mathcal{V}}^{\text{opt}};\bar{\mathbf{z}}_{\mathcal{C}})$, where g^* is the optimum value of problem (5.3), b) if $\mathbf{z}_{\mathcal{V}}$ is a feasible solution of problem (5.5), then $[\mathbf{z}_{\mathcal{V}};\bar{\mathbf{z}}_{\mathcal{C}}]$ is a feasible solution of problem (5.3), and c) *necessary-and-sufficient feasibility condition*: problem (5.5) is feasible $\Leftrightarrow \sum_{v\in\mathcal{V}} \alpha_{t,v} \geq \beta'_t$, $\forall t \in \mathcal{T}$ (i.e., $\mathbf{1}_V$ is a feasible solution).

Now, in order to construct a *lower bound* on the optimum value of problem (5.5), the *LP relaxation* is exploited, where the binary constraints ($z_v \in \{0,1\}$, $\forall v \in \mathcal{V}$) are relaxed:

$$\begin{aligned}\min_{\mathbf{z}_{\mathcal{V}}} \quad & g(\mathbf{z}_{\mathcal{V}};\bar{\mathbf{z}}_{\mathcal{C}}) = \sum_{v\in\mathcal{V}} z_v + \sum_{c\in\mathcal{C}} \bar{z}_c \\ \text{s.t.} \quad & \sum_{v\in\mathcal{V}} \alpha_{t,v} z_v \geq \beta'_t, \ \ \forall t \in \mathcal{T} \\ & 0 \leq z_v \leq 1, \ \ \forall v \in \mathcal{V}\end{aligned} \tag{5.6}$$

An optimal solution $\mathbf{z}_{\mathcal{V}}^{\text{LP}}$ of the LP relaxation (problem (5.6)) can be obtained in polynomial time, using interior-point methods [11]. In addition, note that: a) the feasibility of problem (5.5) implies the feasibility of the LP relaxation, b) if $\mathbf{z}_{\mathcal{V}}^{\text{LP}} \in \{0,1\}^V$, then $g(\mathbf{z}_{\mathcal{V}}^{\text{opt}};\bar{\mathbf{z}}_{\mathcal{C}}) = g(\mathbf{z}_{\mathcal{V}}^{\text{LP}};\bar{\mathbf{z}}_{\mathcal{C}})$, and c) $g(\mathbf{z}_{\mathcal{V}}^{\text{LP}};\bar{\mathbf{z}}_{\mathcal{C}}) \leq g(\mathbf{z}_{\mathcal{V}}^{\text{opt}};\bar{\mathbf{z}}_{\mathcal{C}})$, and because $g(\mathbf{z}_{\mathcal{V}}^{\text{opt}};\bar{\mathbf{z}}_{\mathcal{C}})$ is an integer, we have that $\left\lceil g(\mathbf{z}_{\mathcal{V}}^{\text{LP}};\bar{\mathbf{z}}_{\mathcal{C}}) \right\rceil \leq g(\mathbf{z}_{\mathcal{V}}^{\text{opt}};\bar{\mathbf{z}}_{\mathcal{C}})$.

In the sequel, we develop a greedy method to provide an *upper bound* on the optimum value of problem (5.5). This method is based on the following *cost function*

Algorithm 5.1 CF-based Greedy Method

Input: The BILP problem (5.5) with $\sum_{v\in\mathcal{V}} \alpha_{t,v} \geq \beta'_t$, $\forall t \in \mathcal{T}$

Output: A feasible solution $\mathbf{z}_{\mathcal{V}}$ of problem (5.5)

$\mathbf{z}_{\mathcal{V}} := \mathbf{0}_V$, $\mathcal{R} := \mathcal{V}$, $d_t = \beta'_t \ \ \forall t \in \mathcal{T}$, $f := \sum_{t\in\mathcal{T}} \max(d_t, 0)$

while $f > 0$ **do**

  $n := \arg\min_{r\in\mathcal{R}} \sum_{t\in\mathcal{T}} \max(d_t - \alpha_{t,r}, 0)$, $z_n := 1$, $\mathcal{R} := \mathcal{R}\backslash\{n\}$

  $d_t := d_t - \alpha_{t,n} \ \ \forall t \in \mathcal{T}$, $f := \sum_{t\in\mathcal{T}} \max(d_t, 0)$

end while

(CF): $f(\mathbf{z}_{\mathcal{V}}) = \sum_{t\in\mathcal{T}} \max(d_t, 0)$, with $d_t = \beta'_t - \sum_{v\in\mathcal{V}} \alpha_{t,v} z_v$, $\forall t \in \mathcal{T}$, which quantifies the total violation of inequality-constraints induced by the vector $\mathbf{z}_{\mathcal{V}}$. Observe that: a) $f(\mathbf{z}_{\mathcal{V}}) \geq 0$, and b) $f(\mathbf{z}_{\mathcal{V}}) = 0 \Leftrightarrow \sum_{v\in\mathcal{V}} \alpha_{t,v} z_v \geq \beta'_t$, $\forall t \in \mathcal{T}$.

Algorithm 5.1 presents the CF-based greedy method, where $\mathcal{R} = \{v \in \mathcal{V} : \ z_v = 0\}$. In particular, $\mathbf{z}_{\mathcal{V}}$ is initialized to the zero vector, and in each iteration we find the index in $\mathcal{R}$ which minimizes the CF when the corresponding 0-variable changes to 1. Then, this variable is set equal to 1 and its index is removed from the set $\mathcal{R}$. The algorithm terminates when the CF equals 0, i.e., all the inequality-constraints are satisfied. In addition, $g(\mathbf{z}_{\mathcal{V}}^{\text{opt}}; \bar{\mathbf{z}}_{\mathcal{C}}) \leq g(\mathbf{z}_{\mathcal{V}}^{\text{CF}}; \bar{\mathbf{z}}_{\mathcal{C}})$, where $\mathbf{z}_{\mathcal{V}}^{\text{CF}}$ is a feasible solution of problem (5.5) obtained from Algorithm 5.1.

Complexity of Algorithm 5.1: The complexity of the i^{th} iteration is $\Theta(T(V + 1 - i))$, since $|\mathcal{R}| = V + 1 - i$. From the input assumption of Algorithm 5.1 (feasibility condition), we have that $f(\mathbf{1}_V) = 0$, and therefore Algorithm 5.1 requires a maximum of V iterations to terminate. Consequently, the worst-case complexity of Algorithm 5.1 is $\sum_{i=1}^{V} T(V + 1 - i) = T \sum_{j=1}^{V} j = TV(V + 1)/2 = \Theta(TV^2)$, i.e., polynomial in the size of the input.

As concerns the branching procedure in the B&B method, we choose a *branching variable* z_b ($b \in \mathcal{V}$) such that z_b^{LP} is the most "uncertain" fractional variable, i.e., closer to 0.5 than any other variable in $\mathbf{z}_{\mathcal{V}}^{\text{LP}}$. Afterwards, problem (5.5) is decomposed into two *subproblems* by setting either $z_b = 0$ or $z_b = 1$:

$$\begin{aligned} \min_{\mathbf{z}_{\mathcal{V}\setminus\{b\}}} \quad & g(\mathbf{z}_{\mathcal{V}\setminus\{b\}}; \bar{\mathbf{z}}_{\mathcal{C}\cup\{b\}}) = \sum_{v\in\mathcal{V}\setminus\{b\}} z_v + \sum_{c\in\mathcal{C}} \bar{z}_c \\ \text{s.t.} \quad & \sum_{v\in\mathcal{V}\setminus\{b\}} \alpha_{t,v} z_v \geq \beta'_t, \ \ \forall t \in \mathcal{T} \\ & z_v \in \{0,1\}, \ \ \forall v \in \mathcal{V}\setminus\{b\} \end{aligned} \tag{5.7}$$

$$\begin{aligned} \min_{\mathbf{z}_{\mathcal{V}\setminus\{b\}}} \quad & g(\mathbf{z}_{\mathcal{V}\setminus\{b\}}; \bar{\mathbf{z}}_{\mathcal{C}\cup\{b\}}) = \sum_{v\in\mathcal{V}\setminus\{b\}} z_v + \sum_{c\in\mathcal{C}} \bar{z}_c + 1 \\ \text{s.t.} \quad & \sum_{v\in\mathcal{V}\setminus\{b\}} \alpha_{t,v} z_v \geq \beta'_t - \alpha_{t,b}, \ \ \forall t \in \mathcal{T} \\ & z_v \in \{0,1\}, \ \ \forall v \in \mathcal{V}\setminus\{b\} \end{aligned} \tag{5.8}$$

These subproblems have the same form as problem (5.5), with $\bar{z}_b = 0/1$ for subproblem (5.7)/(5.8), respectively. Moreover, if g_0^{opt} and g_1^{opt} are respectively the optimum values of subproblems (5.7) and (5.8) (assuming that the optimum value of an infeasible problem equals $+\infty$), then $g(\mathbf{z}_{\mathcal{V}}^{\text{opt}}; \bar{\mathbf{z}}_{\mathcal{C}}) = \min(g_0^{\text{opt}}, g_1^{\text{opt}})$.

The proposed B&B method is given in Algorithm 5.2, where U is the best *global upper bound* found so far by the algorithm ($g^* \leq U$), and $\mathcal{L}$ is the *list of active subproblems* that controls the order in which the subproblems are examined (a generated subproblem is called active if it has not been examined yet). Note that $\mathcal{L}$ is a first-in-first-out (FIFO) list; this is preferable when "good" upper bounds are available in order to "prune" the search space as early as possible.

Furthermore, the B&B method performs three *fundamental operations*, where no further investigation is needed for the examined subproblem: 1) *Infeasibility*: the examined subproblem is infeasible, 2) *Pruning*: the examined subproblem cannot produce a better solution ($U \leq \left\lceil g(\mathbf{z}_{\mathcal{V}}^{\text{LP}}; \bar{\mathbf{z}}_{\mathcal{C}}) \right\rceil$), and 3) *Fathoming*: an optimal solution of the examined subproblem is found; this occurs when the solution of the LP relaxation is integer ($\mathbf{z}_{\mathcal{V}}^{\text{LP}} \in \{0,1\}^V$), or when $\left\lceil g(\mathbf{z}_{\mathcal{V}}^{\text{LP}}; \bar{\mathbf{z}}_{\mathcal{C}}) \right\rceil = g(\mathbf{z}_{\mathcal{V}}^{\text{CF}}; \bar{\mathbf{z}}_{\mathcal{C}})$ which implies $g(\mathbf{z}_{\mathcal{V}}^{\text{opt}}; \bar{\mathbf{z}}_{\mathcal{C}}) = g(\mathbf{z}_{\mathcal{V}}^{\text{CF}}; \bar{\mathbf{z}}_{\mathcal{C}})$. Finally, Algorithm 5.2 produces a *nonincreasing sequence of global upper bounds* U, and after its termination $U = g^*$ since all the generated subproblems have been examined ($\mathcal{L} = \emptyset$).

Algorithm 5.2 LP&CF-based B&B Method

Input: The original BILP problem (5.3) with $\sum_{k\in\mathcal{K}} \alpha_{t,k} \geq \beta_t$, $\forall t \in \mathcal{T}$

Output: A (globally) optimal solution $\mathbf{z}^*$ of problem (5.3)

```
z* := 1_K, U := K, L := {problem (5.3)}
while L ≠ ∅ do
    Remove the front subproblem from the list L,
    which has the form of problem (5.5)
    if ∃t ∈ T : Σ_{v∈V} α_{t,v} < β'_t then {continue} end if          ▷ Infeasibility
    Compute an optimal solution z_V^LP of the LP relaxation (in the form
    of problem (5.6)), using a LP-solver of polynomial complexity
    if U ≤ ⌈g(z_V^LP; z̄_C)⌉ then {continue} end if                     ▷ Pruning
    if z_V^LP ∈ {0,1}^V then        ▷ Fathoming (integer solution), given that
                         U > ⌈g(z_V^LP; z̄_C)⌉ = g(z_V^LP; z̄_C) = g(z_V^opt; z̄_C)
        U := g(z_V^LP; z̄_C), z* := [z_V^LP; z̄_C], continue
    end if
    Compute a feasible solution z_V^CF of the examined subproblem,
    using the CF-based greedy method (Algorithm 5.1)
    if g(z_V^CF; z̄_C) < U then {U := g(z_V^CF; z̄_C), z* := [z_V^CF; z̄_C]} end if
    if ⌈g(z_V^LP; z̄_C)⌉ = g(z_V^CF; z̄_C) then {continue} end if       ▷ Fathoming
    Select a branching variable z_b (b := argmin_{v∈V} |z_v^LP − 0.5|), and then
    generate two new subproblems in the form of problems (5.7) and (5.8)
    Insert the generated subproblems at the end of the list L
end while
```

Complexity of Algorithm 5.2: The complexity of each iteration is mainly restricted by the LP-solver (polynomial complexity $O((T+V)^{1.5}V^2)$ [11]) as well as Algorithm 5.1, so it is $O((T+V)^{1.5}V^2 + TV^2) = O((T+V)^{1.5}V^2) = O((T+K)^{1.5}K^2)$. Furthermore, in each iteration we examine one subproblem, while we generate at most two new subproblems by fixing one of the free variables. Therefore, the number of iterations/subproblems is $\leq \sum_{j=0}^{K} 2^j = 2^{K+1} - 1 = O(2^K)$. Overall, the worst-case complexity of Algorithm 5.2 is $O(2^K(T+K)^{1.5}K^2)$, i.e., exponential in the size of the input. Although the original BILP problem (5.3) is probably intractable in the worst case (due to its NP-hardness), the most difficult problem instances may rarely occur in practice (because of their special structure), so the *average-case complexity* may be a more appropriate measure of an algorithm's efficiency. Specifically, assuming a probability distribution over problem instances, the average-case complexity of Algorithm 5.2 is $O(M(T+K)^{1.5}K^2)$, where M is the mean/average number of iterations. Observe that if $M = \text{poly}(T,K)$, where $\text{poly}(T,K)$ is some polynomial in T and K, then Algorithm 5.2 will have polynomial-time complexity in the average case.[11] Nevertheless, the average-case complexity of the B&B method is very challenging to study theoretically, so we resort to a numerical analysis in Section 5.5.

5.5 Numerical Results and Discussion

In this section, the performance of the designed B&B algorithm is evaluated through a series of problem instances. More specifically, the following simulation parameters have been considered: $K \in \{10, 15, 20, 25, 30\}$, $T = 12$, $P_t^{\text{avl,req}} = 99.9\%$, $\forall t \in \mathcal{T}$, and 200 independent scenarios (for each value of K) with the outage probabilities $\{p_{k,t}^{\text{out}}\}_{k\in\mathcal{K},\, t\in\mathcal{T}}$ being uniformly distributed in the interval $[0.1, 1]$.

Firstly, we compare Algorithm 5.2 with state-of-the-art methods. As shown in Table 5.1, GHA exhibits the lowest performance, while Algorithm 5.2 achieves exactly the same performance with ESA and significantly outperforms GHA and CHA. Moreover, for $K \in \{15, 20, 25, 30\}$, GHA and CHA attain a globally optimal solution in less than 70% of cases.[12] On the other hand, Algorithm 5.2 finds the global optimum in all cases, since it is theoretically guaranteed to do so.

Furthermore, we examine the complexity of Algorithm 5.2 in terms of the required iterations (recall that each iteration has polynomial-time complexity). According to

[11] Note that an *exhaustive-enumeration algorithm*, despite its global optimality, requires $T \sum_{j=1}^{K} \binom{K}{j} j = TK2^{K-1} = \Theta(2^K KT)$ comparisons in all cases, thus having exponential complexity in both the worst and the average case.

[12] Although the worst-case complexity of both GHA and CHA is $\Theta(TK^2)$, these heuristic methods do not provide any performance guarantee.

Table 5.1: Performance Comparison with Existing Methods: Average # of Selected GSs & Percentage of Problems Optimally Solved[a]

K \ METHOD	ESA [7]	GHA[b] [7]	CHA[b] [7]	Algorithm 5.2
10	9.84	9.89 (96%)	9.87 (97%)	9.84 (100%)
15	11.36	11.89 (55%)	11.70 (68%)	11.36 (100%)
20	10.33	11.15 (33%)	10.74 (60%)	10.33 (100%)
25	9.62	10.62 (17%)	10.00 (63%)	9.62 (100%)
30	9.23	10.14 (24%)	9.65 (58%)	9.23 (100%)

[a] This percentage is calculated using the global minimum obtained from the exhaustive-search algorithm (ESA) given in [7].
[b] GHA and CHA select up to 3 and up to 2 redundant GSs, respectively.

Table 5.2: Iterations Required by Algorithm 5.2

K	Total # of iterations [mean (standard deviation)]	# of iterations until a global minimum is found for the 1st time [mean (standard deviation)]	Upper bound on the total # of iterations [$= 2^{K+1} - 1$]
10	1.93 (2.37)	0.22 (0.71)	$> 2 \times 10^3$
15	14.23 (15.61)	6.04 (10.30)	$> 6 \times 10^4$
20	41.37 (57.01)	15.86 (38.91)	$> 2 \times 10^6$
25	87.74 (117.14)	27.52 (75.55)	$> 6 \times 10^7$
30	117.90 (204.85)	28.42 (121.32)	$> 2 \times 10^9$

Table 5.2, the B&B method requires extremely few iterations on average (with small standard deviation) compared to the upper bound $2^{K+1} - 1$. Thus, Algorithm 5.2 has low average-case complexity.

Finally, Fig. 5-1 illustrates the progress of the B&B method for a specific problem. In particular, we can observe: 1) the nonincreasing sequence of global upper bounds U, and 2) that the number of active subproblems $|\mathcal{L}| = L$ is equal to 1 at the beginning of the algorithm, and becomes 0 in the last iteration.

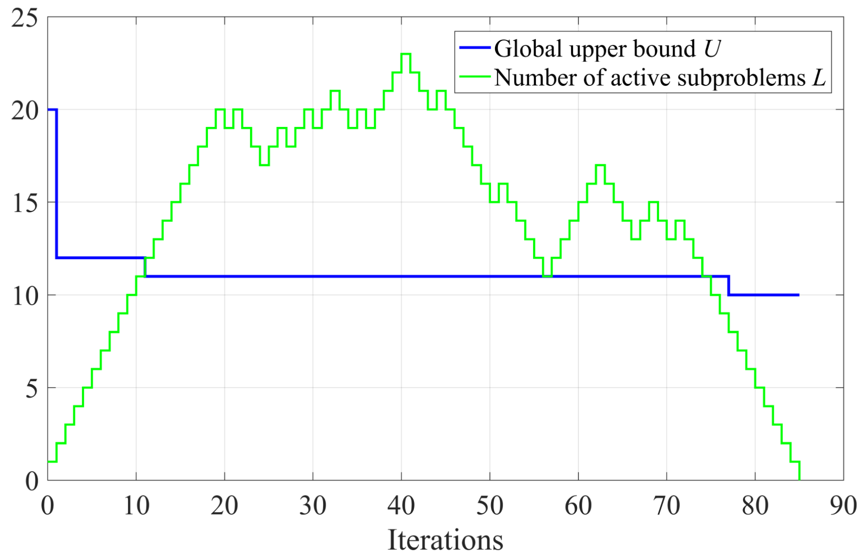

Figure 5-1: Progress of Algorithm 5.2 for a simulation scenario with $K = 20$. Global minimum = 10, total number of iterations = 85, and number of iterations until a global minimum is found for the 1st time = 77.

5.6 Conclusion

In this chapter, we have studied the optimal selection of GSs in satellite systems with site diversity. Furthermore, we have developed a global optimization algorithm, which can provide significant cost savings for the network operator. Finally, according to the numerical results, the proposed B&B method exhibits low average-case complexity, while achieving much higher performance than existing algorithms.

Bibliography

[1] A. D. Panagopoulos *et al.*, "Satellite communications at Ku, Ka, and V bands: Propagation impairments and mitigation techniques," *IEEE Commun. Surv. Tutor.*, vol. 6, no. 3, pp. 2-14, Third Quarter 2004.

[2] S. Poulenard, M. Crosnier and A. Rissons, "Ground segment design for broadband geostationary satellite with optical feeder link," *IEEE/OSA J. Optical Commun. and Network.*, vol. 7, no. 4, pp. 325-336, April 2015.

[3] C. Fuchs and F. Moll, "Ground station network optimization for space-to-ground optical communication links," *IEEE/OSA J. Optical Commun. and Network.*, vol. 7, no. 12, pp. 1148-1159, Dec. 2015.

[4] I. del Portillo *et al.*, “Optimal location of optical ground stations to serve LEO spacecraft,” *IEEE Aerospace Conf.*, Big Sky, MT, 2017, pp. 1-16.

[5] T. Rossi *et al.*, “Smart gateway diversity optimization for EHF satellite networks,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 1, pp. 130-141, Feb. 2020.

[6] N. K. Lyras *et al.*, “Optimizing the ground network of optical MEO satellite communication systems,” *IEEE Systems Journal*, doi: 10.1109/JSYST.2019.2945838.

[7] N. K. Lyras *et al.*, “Optimum monthly-based selection of ground stations for optical satellite networks,” *IEEE Commun. Lett.*, vol. 22, no. 6, pp. 1192-1195, June 2018.

[8] M. Sanchez Net *et al.*, “Approximation methods for estimating the availability of optical ground networks,” *IEEE/OSA J. Optical Commun. and Network.*, vol. 8, no. 10, pp. 800-812, Oct. 2016.

[9] N. K. Lyras, C. I. Kourogiorgas and A. D. Panagopoulos, “Cloud free line of sight prediction modeling for optical satellite communication networks,” *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1537-1540, July 2017.

[10] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Englewood Cliffs, NJ: Prentice-Hall Inc., 1982.

[11] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Philadelphia, PA, USA: SIAM, 2001.

Chapter 6

Minimizing the Installation Cost of Ground Stations in Satellite Networks: Complexity, Dynamic Programming and Approximation Algorithm[13]

In this chapter, we study the optimum selection of ground stations (GSs) in RF/optical satellite networks (SatNets) in order to minimize the overall installation cost under an outage probability requirement, assuming independent weather conditions between sites. First, we show that the optimization problem can be formulated as a binary-linear-programming problem, and then we give a formal proof of its NP-hardness. Furthermore, we design a dynamic-programming algorithm of pseudo-polynomial complexity with global optimization guarantee as well as an efficient (polynomial-time) approximation algorithm with provable performance guarantee on the distance of the achieved objective value from the global optimum. Finally, the performance of the proposed algorithms is verified through numerical simulations.

6.1 Introduction

The availability of satellite networks (SatNets) is heavily affected by atmospheric impairments, especially rain in radio-frequency (RF) and clouds in optical SatNets. Site diversity techniques are able to improve the network availability by mitigating the extremely high attenuation induced by rain and clouds [1]. An optimization method for selecting optical GSs is proposed in [2], taking into consideration the single-site availabilities and the spatial-correlation between sites as well. In [3], a joint optimization algorithm for the design of optical SatNets is presented, which is divided into two parts: the GS positioning and the backbone network optimization considering the optical fiber cost.

Moreover, [4] and [5] present low-complexity heuristic algorithms, which exploit the spatial correlation and the monthly variability of cloud coverage, in order to select the minimum number of GSs in optical SatNets with a geostationary (GEO) or a medium-earth-orbit (MEO) satellite, respectively. A multi-objective optimization approach that achieves various tradeoffs between availability, latency and cost is examined in [6], so as to determine the optimal location of optical GSs for low-earth-orbit (LEO) SatNets. In addition, as concerns the smart gateway diversity optimization in extremely-high-frequency (EHF) SatNets, [7] presents another multi-objective approach using genetic algorithms.

Recently, [8] provides an efficient gradient-projection method to select a given number of GSs maximizing the availability of free-space optical (FSO) SatNets. Finally, a branch-and-bound (B&B) algorithm with global optimization guarantee and low average-case complexity is developed in [9] to select the minimum number of GSs under availability requirements for each time period.

In this chapter, we develop useful optimization algorithms for selecting GSs with the minimum installation cost satisfying an outage probability constraint. More specifically, the main contributions of this chapter are summarized as follows:

- Mathematical formulation of the optimization problem in binary-linear-programming form with a rigorous proof of its computational complexity (*NP-hardness*).
- Design of a *dynamic-programming algorithm with pseudo-polynomial complexity*, which is theoretically guaranteed to find the global optimum.
- Design of a *polynomial-time approximation algorithm with provable performance guarantee* on the distance between the objective value of the achieved solution and the global optimum (thus achieving a reasonable performance-complexity tradeoff).
- Unlike existing approaches that minimize just the number of GSs (cardinality minimization problem, assuming implicitly the same cost for each GS), the proposed algorithms minimize the overall installation cost allowing *possibly different costs of GSs*.

The remainder of this chapter is organized as follows. Section 6.2 presents the formulation of the optimization problem with a theoretical proof of its NP-hardness. Subsequently, a global optimization algorithm using dynamic programming is given in Section 6.3, while a polynomial-time approximation algorithm is presented in Section 6.4. Finally, Section 6.5 provides some numerical results and Section 6.6 concludes this chapter.

Mathematical notation: The set of positive integers is denoted by $\mathbb{Z}_+ = \{1, 2, 3, \ldots\}$, while $\mathbf{0}_K$ and $\mathbf{1}_K$ are respectively the K-dimensional all-zeros and all-ones vectors. Moreover, $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ stand for the floor and ceiling functions, respectively.

6.2 Problem Formulation & NP-Hardness

Consider an RF/optical SatNet with site diversity, consisting of a GEO satellite and a network of geographically distributed GSs. In particular, $\mathcal{K} = \{1, 2, \ldots, K\}$

denotes the set of candidate locations/sites for installing a GS ($K \in \mathbb{Z}_+$). In addition, we assume that: 1) the *network outage probability* is defined as the probability of having *all GSs in outage* and 2) *the distance between any two distinct locations is large enough* so that the spatial correlation between sites can be ignored, without significant error on the calculation of network outage probability; this implies (approximately) *independent weather conditions* between the candidate locations.

In this context, we study the minimization of the total installation cost of GSs satisfying a given outage probability requirement:

$$\underset{\mathcal{S}}{\text{minimize}} \quad \sum_{s \in \mathcal{S}} c_s \tag{6.1a}$$

$$\text{subject to} \quad P_{\text{out}}(\mathcal{S}) \leq P_{\text{out}}^{\text{th}} \tag{6.1b}$$

$$\mathcal{S} \subseteq \mathcal{K} \tag{6.1c}$$

where $\mathcal{S}$ is the set of selected locations, $c_k \in \mathbb{Z}_+$ denotes the cost of installing a GS at the k^{th} location , $\forall k \in \mathcal{K}$ (without loss of generality, we assume that $c_1 \leq c_2 \leq \cdots \leq c_K$; this requires an extra complexity of $O(K \log K)$ for sorting the sites in ascending-cost order),[14] $P_{\text{out}}(\mathcal{S}) = \prod_{s \in \mathcal{S}} p_s$ is the network outage probability achieved by the set $\mathcal{S}$, with $p_k \in (0, 1]$ being the outage probability of a GS installed at the k^{th} location, $\forall k \in \mathcal{K}$,[15] and $P_{\text{out}}^{\text{th}} \in (0, 1]$ is the (network) outage probability threshold. Herein, $[p_k]_{k \in \mathcal{K}}$ and $P_{\text{out}}^{\text{th}}$ are defined on an annual basis, and therefore the proposed approach does not take into account the monthly/seasonal variability of weather conditions. Note that in the special case where $c_k = 1$, $\forall k \in \mathcal{K}$, we have a cardinality minimization problem.

Afterwards, by introducing the vector $\mathbf{z} = [z_1, z_2, \ldots, z_K]$ of binary (0/1) variables ($z_k = 1$ if and only if $k \in \mathcal{S}$), we can equivalently formulate problem (6.1) as follows

[14] Note that the coefficient c_k may include the cost of fiber-optic cables needed to connect the k^{th} GW to the existing access points (points of presence) of the terrestrial backbone network.

[15] The outage probability of each GS can be obtained from experimental data (when available) or using time-series synthesizers. Moreover, in RF SatNets a GS is in outage when the rain attenuation is higher than a specific threshold [7], whereas in optical SatNets when experiencing cloud blockage [2].

(note that $\sum_{s\in\mathcal{S}} c_s = \sum_{k\in\mathcal{K}} c_k z_k$ and $P_{\text{out}}(\mathcal{S}) = \prod_{k\in\mathcal{K}} (p_k)^{z_k}$):

$$\begin{aligned} \underset{\mathbf{z}}{\text{minimize}} \quad & f(\mathbf{z}) = \sum_{k\in\mathcal{K}} c_k z_k && \text{(6.2a)} \\ \text{subject to} \quad & \prod_{k\in\mathcal{K}} (p_k)^{z_k} \le P_{\text{out}}^{\text{th}} && \text{(6.2b)} \\ & z_k \in \{0,1\},\ \forall k \in \mathcal{K} && \text{(6.2c)} \end{aligned}$$

Exploiting the fact that $x \le y \Leftrightarrow \log(x) \le \log(y)$, $\forall x, y > 0$, the constraint $\prod_{k\in\mathcal{K}} (p_k)^{z_k} \le P_{\text{out}}^{\text{th}}$ is equivalent to $\sum_{k\in\mathcal{K}} z_k \log(p_k) \le \log(P_{\text{out}}^{\text{th}})$. Consequently, problem (6.2) can be written as a *binary-linear-programming problem*:

$$\begin{aligned} \underset{\mathbf{z}}{\text{minimize}} \quad & f(\mathbf{z}) = \sum_{k\in\mathcal{K}} c_k z_k && \text{(6.3a)} \\ \text{subject to} \quad & \sum_{k\in\mathcal{K}} a_k z_k \ge b && \text{(6.3b)} \\ & z_k \in \{0,1\},\ \forall k \in \mathcal{K} && \text{(6.3c)} \end{aligned}$$

where $a_k = -\log(p_k) \ge 0$, $\forall k \in \mathcal{K}$, and $b = -\log(P_{\text{out}}^{\text{th}}) \ge 0$. Let $\mathcal{F} = \left\{\mathbf{z} \in \{0,1\}^K : \prod_{k\in\mathcal{K}} (p_k)^{z_k} \le P_{\text{out}}^{\text{th}}\right\}$, or equivalently $\mathcal{F} = \left\{\mathbf{z} \in \{0,1\}^K : \sum_{k\in\mathcal{K}} a_k z_k \ge b\right\}$, be the feasible set and $\mathbf{z}^* \in \arg\min_{\mathbf{z}} \{f(\mathbf{z}) : \mathbf{z} \in \mathcal{F}\}$ be a (globally) optimal solution of problem (6.2)/(6.3). Since $a_k \ge 0$, $\forall k \in \mathcal{K}$, the following *necessary and sufficient feasibility condition* applies: problem (6.2)/(6.3) is feasible (i.e., $\mathcal{F} \neq \emptyset$) *if and only if* $\prod_{k\in\mathcal{K}} p_k \le P_{\text{out}}^{\text{th}}$ or, equivalently, $\sum_{k\in\mathcal{K}} a_k \ge b$ (i.e., $\mathbf{1}_K \in \mathcal{F}$).

Theorem 6.1 (NP-hardness)**.** *The binary-linear-programming problem* (6.3) *is NP-hard.*

Proof. In order to prove the NP-hardness of problem (6.3), it is sufficient to show that a special case of this problem is NP-hard. Firstly, let consider the *0-1 knapsack problem* which is a well-known NP-hard problem [10]:

$$\begin{aligned} \underset{\mathbf{x}}{\text{maximize}} \quad & \sum_{k\in\mathcal{K}} v_k x_k && \text{(6.4a)} \\ \text{subject to} \quad & \sum_{k\in\mathcal{K}} w_k x_k \leq W && \text{(6.4b)} \\ & x_k \in \{0,1\},\ \forall k \in \mathcal{K} && \text{(6.4c)} \end{aligned}$$

where $W \in \mathbb{Z}_+$ is the knapsack capacity, and $v_k, w_k \in \mathbb{Z}_+$ are the value and weight of the k^{th} item, respectively, $\forall k \in \mathcal{K}$. Moreover, applying the *polynomial-time,* $\Theta(K)$, *variable transformation* $x_k = 1 - z_k$, $\forall k \in \mathcal{K}$, we get the following equivalent problem:[16]

$$\begin{aligned} \underset{\mathbf{z}}{\text{minimize}} \quad & \sum_{k\in\mathcal{K}} v_k z_k && \text{(6.5a)} \\ \text{subject to} \quad & \sum_{k\in\mathcal{K}} w_k z_k \geq W' && \text{(6.5b)} \\ & z_k \in \{0,1\},\ \forall k \in \mathcal{K} && \text{(6.5c)} \end{aligned}$$

where $W' = \sum_{k\in\mathcal{K}} w_k - W$. Without loss of generality, we can assume that the integer $W' \geq 0$; otherwise the optimal solution of problem (6.5) is trivially equal to $\mathbf{0}_K$. Obviously, the NP-hard problem (6.5) is a subcase of problem (6.3), and this completes the proof. □

6.3 Global Optimization Using Dynamic Programming

Due to the fact that problem (6.3) is NP-hard, it cannot be (globally) solved in polynomial time unless P=NP. Nevertheless, we can use a powerful optimization technique, namely, *dynamic programming (DP)*, in order to achieve the global minimum with pseudo-polynomial complexity.

[16]Note that the optimum objective values of problems (6.4) and (6.5) differ only by a constant, i.e., $\sum_{k\in\mathcal{K}} v_k x_k^* = \sum_{k\in\mathcal{K}} v_k - \sum_{k\in\mathcal{K}} v_k z_k^*$.

DP performs an intelligent enumeration of all the feasible solutions, thus providing a *global optimization guarantee*. In particular, DP follows a *bottom-up approach* by decomposing the problem into "smaller" subproblems and combining their optimal solutions (using a *recursive formula*) in order to find an optimal solution to the original problem; this is known as *the principle of optimality* and such problems are said to have *optimal substructure* [10]. Furthermore, DP is a *tabular method* where each subproblem is solved only once and then its solution is stored in a table, so that it can be readily used (without re-computation) by "larger" problems when needed.

Let C be an *integer upper bound* on the optimum value of problem (6.3), i.e., $f(\mathbf{z}^*) \leq C$, where $C \in \{0, 1, \dots, \overline{C}\}$ with $\overline{C} = \sum_{k \in \mathcal{K}} c_k$ (this is *the "worst" upper bound* that can be used). In addition, we define the following bivariate function $\forall i \in \mathcal{K}_0 = \{0, 1, \dots, K\}$ and $\forall j \in \mathcal{C}_0 = \{0, 1, \dots, C\}$:

$$R(i, j) = \max_{\mathbf{z}_{\mathcal{I}}} \left\{ \sum_{k \in \mathcal{I}} a_k z_k : \sum_{k \in \mathcal{I}} c_k z_k = j, \ \mathbf{z}_{\mathcal{I}} \in \{0, 1\}^i \right\} \tag{6.6}$$

where $\mathcal{I} = \{1, 2, \dots, i\}$ and $\mathbf{z}_{\mathcal{I}} = [z_1, z_2, \dots, z_i]$, with $i = 0 \Rightarrow \mathcal{I} = \emptyset$ and $\sum_{k \in \emptyset} a_k z_k = \sum_{k \in \emptyset} c_k z_k = 0$. If this maximization problem is infeasible, then $R(i, j) = -\infty$.

Theorem 6.2 (Computation of the global optimum)**.** *Assuming that problem* (6.2)/(6.3) *is feasible, its global minimum can be found as follows:* $f(\mathbf{z}^*) = \min \{j \in \mathcal{C}_0 : R(K, j) \geq b\}$.

Proof. Firstly, observe that when $i = K$, we have $\mathcal{I} = \mathcal{K}$ and $\mathbf{z}_{\mathcal{I}} = \mathbf{z}$. Secondly, we know that $f(\mathbf{z}^*) \in \mathcal{C}_0$ and $R(K, f(\mathbf{z}^*)) \geq \sum_{k \in \mathcal{K}} a_k z_k^* \geq b$. Now, suppose that $f(\mathbf{z}^*) \neq j^*$, where $j^* = \min \{j \in \mathcal{C}_0 : R(K, j) \geq b\}$. Let examine two cases: 1) $f(\mathbf{z}^*) < j^*$ and 2) $f(\mathbf{z}^*) > j^*$. In the former case, we would have that $R(K, f(\mathbf{z}^*)) < b$, which leads to a contradiction. Moreover, the latter case contradicts the global optimality of $f(\mathbf{z}^*)$. Hence, $f(\mathbf{z}^*) = j^*$ and Theorem 6.2 has been proven. $\square$

Subsequently, we partition the feasible set of problem (6.6), by setting $z_i = 0$ and $z_i = 1$, respectively (note that $\mathcal{I} \backslash \{i\} = \{1, 2, \dots, i - 1\}$):

$$\begin{aligned}
&\max_{\mathbf{z}_{\mathcal{I}}}\left\{\sum_{k\in\mathcal{I}} a_k z_k : \sum_{k\in\mathcal{I}} c_k z_k = j,\ \mathbf{z}_{\mathcal{I}} \in \{0,1\}^i,\ z_i = 0\right\} = \\
&= \max_{\mathbf{z}_{\mathcal{I}\setminus\{i\}}}\left\{\sum_{k\in\mathcal{I}\setminus\{i\}} a_k z_k : \sum_{k\in\mathcal{I}\setminus\{i\}} c_k z_k = j,\ \mathbf{z}_{\mathcal{I}\setminus\{i\}} \in \{0,1\}^{i-1}\right\} = \\
&= R(i-1,j)
\end{aligned} \tag{6.7}$$

$$\begin{aligned}
&\max_{\mathbf{z}_{\mathcal{I}}}\left\{\sum_{k\in\mathcal{I}} a_k z_k : \sum_{k\in\mathcal{I}} c_k z_k = j,\ \mathbf{z}_{\mathcal{I}} \in \{0,1\}^i,\ z_i = 1\right\} = \\
&= a_i + \max_{\mathbf{z}_{\mathcal{I}\setminus\{i\}}}\left\{\sum_{k\in\mathcal{I}\setminus\{i\}} a_k z_k : \sum_{k\in\mathcal{I}\setminus\{i\}} c_k z_k = j - c_i,\ \mathbf{z}_{\mathcal{I}\setminus\{i\}} \in \{0,1\}^{i-1}\right\} = \\
&= a_i + R(i-1,j-c_i)
\end{aligned} \tag{6.8}$$

Therefore, we have the following *recursive formula* $\forall i \in \mathcal{K} = \{1,2,\ldots,K\}$ and $\forall j \in \mathcal{C}_0 = \{0,1,\ldots,C\}$:

$$R(i,j) = \begin{cases} \max\left\{R(i-1,j), a_i + R(i-1,j-c_i)\right\}, & \text{if } j \geq c_i \\ R(i-1,j), & \text{otherwise} \end{cases} \tag{6.9}$$

with *initial conditions*: a) $R(0,0) = 0$ and b) $R(0,j) = -\infty$, $\forall j \in \mathcal{C} = \{1,2,\ldots,C\}$. Observe that if $j < c_i$, then problem (6.8) is definitely infeasible, so $R(i-1,j-c_i) = -\infty$; this explains the 2$^{\text{nd}}$ branch in (6.9).

Algorithm 6.1 presents a DP procedure based on the previous analysis. First, we compute the coefficients $[a_k]_{k\in\mathcal{K}}$ and b (line 1), and then a *greedy method* is used in order to calculate the upper bound C (lines 2-5). In essence, this method sequentially selects GSs in ascending-cost order and finds a feasible solution to problem (6.3), which is certainly an upper bound on the optimum value. Afterwards, the algorithm stores the $R(i,j)$ values in a $(K+1)\times(C+1)$ table, whose entries are computed in row order from left to right (lines 6-15). Moreover, the global optimum can be found by checking the last row, since $f(\mathbf{z}^*) = j^* = \min\{j \in \mathcal{C}_0 : R(K,j) \geq b\}$ according to Theorem 6.2. Finally, an optimal solution can be deduced from the generated table by starting at $R(K,j^*)$ and tracing where the optimal values come from (lines 16-23). In particular, if $R(i,j) = R(i-1,j)$, then $z_i^* = 0$, and we continue tracing with

Algorithm 6.1 Dynamic Programming (DP)

Input: $K \in \mathbb{Z}_+$, $\mathbf{c} = [c_1, c_2, \dots, c_K] \in \mathbb{Z}_+^K$ where $c_1 \leq c_2 \leq \cdots \leq c_K$,
$\mathbf{p} = [p_1, p_2, \dots, p_K] \in (0,1]^K$, $P_{\text{out}}^{\text{th}} \in (0,1]$ with $\prod_{k \in \mathcal{K}} p_k \leq P_{\text{out}}^{\text{th}}$

Output: $\mathbf{z}^* \in \arg\min_{\mathbf{z}} \left\{ \sum_{k \in \mathcal{K}} c_k z_k : \mathbf{z} \in \mathcal{F} \right\}$

$a_k \coloneqq -\log(p_k)$, $\forall k \in \mathcal{K}$, $b \coloneqq -\log(P_{\text{out}}^{\text{th}})$
$A \coloneqq 0$, $C \coloneqq 0$, $k \coloneqq 1$
while $A < b$ **do** ▷ *Calculation of the upper bound C*
  $A \coloneqq A + a_k$, $C \coloneqq C + c_k$, $k \coloneqq k + 1$
end while
$R(0,0) \coloneqq 0$, $R(0,j) \coloneqq -\infty$, $\forall j \in \mathcal{C}$
for $i \coloneqq 1$ **to** K **step** $+1$ **do** ▷ *Computation of the table R*
  for $j \coloneqq 0$ **to** C **step** $+1$ **do**
    if $j \geq c_i$ **then**
      $R(i,j) \coloneqq \max\{R(i-1,j), a_i + R(i-1, j-c_i)\}$
    else
      $R(i,j) \coloneqq R(i-1,j)$
    end if
  end for
end for
$j^* \coloneqq \min\{j \in \mathcal{C}_0 : R(K,j) \geq b\}$, $q \coloneqq R(K, j^*)$, $j \coloneqq j^*$
for $i \coloneqq K$ **to** 1 **step** -1 **do** ▷ *Reconstruction of an optimal solution*
  if $q = R(i-1,j)$ **then**
    $z_i^* \coloneqq 0$, $q \coloneqq R(i-1,j)$
  else
    $z_i^* \coloneqq 1$, $q \coloneqq R(i-1, j-c_i)$, $j \coloneqq j - c_i$
  end if
end for

$R(i-1,j)$. Otherwise $z_i^* = 1$, and we continue tracing with $R(i-1, j-c_i)$. This process is repeated for each i from K down to 1 (with step -1). Therefore, *Algorithm 6.1 is theoretically guaranteed to find a (globally) optimal solution.*

Complexity of Algorithm 6.1: The complexity of computing the coefficients $[a_k]_{k \in \mathcal{K}}$

and b is $\Theta(K)$. Moreover, the greedy method used to find an upper bound on the optimum value requires at most K iterations (since $\mathbf{1}_K \in \mathcal{F}$), thus having $O(K)$ complexity. In addition, the computation of the table R requires $\Theta(KC)$ arithmetic operations in total. Finally, the computation of j^* requires $O(C)$ comparisons, while the complexity of reconstructing/tracing the solution is $\Theta(K)$ since it starts in row K of the table and moves up one row at each step. Ultimately, *the overall complexity of Algorithm 6.1 is* $\Theta(KC) = O(K\overline{C}) = O(K^2 c_{\max})$, because $C \leq \overline{C} \leq K c_{\max}$ where $c_{\max} = \max_{k\in\mathcal{K}}\{c_k\} = c_K$. As a result, the proposed DP algorithm has *pseudo-polynomial time complexity* [10].

Remark 6.1. Strictly speaking, Algorithm 6.1 is an *exponential-time algorithm*, since the size of the input is upper bounded by $O(K \log c_{\max}) = O(K \log \overline{C})$, because $c_{\max} \leq \overline{C}$. Nevertheless, under certain conditions, this algorithm is practical despite its exponential worst-case complexity. For example, if $\overline{C} = O(K^d)$ for some constant $d \geq 0$ (which is usually the case in practice), then the running time of Algorithm 6.1 will be *polynomial in* K. In any case, the optimization problem under consideration does not need to be solved in real time, but during the initial network design.

Remark 6.2. In Algorithm 6.1, due to the fact that $C = \sum_{u\in\mathcal{U}} c_u$ for some $\mathcal{U} \subseteq \mathcal{K}$ depending on $[a_k]_{k\in\mathcal{K}}$ and b, we can divide all coefficients $[c_k]_{k\in\mathcal{K}}$ with their *greatest common divisor* (i.e., $c'_k = c_k/\zeta \in \mathbb{Z}_+$, $\forall k \in \mathcal{K}$, where $\zeta = \gcd(c_1, c_2, \ldots, c_K) \in \mathbb{Z}_+$) without altering the set of optimal solutions. In this way, *the complexity of Algorithm 6.1 can be reduced*, since $C' = \sum_{u\in\mathcal{U}} c'_u = C/\zeta \leq C$.

6.4 Polynomial-Time Approximation Algorithm

Subsequently, a practical and efficient (polynomial-time) approximation algorithm with *provable performance guarantee* is given. The design of the approximation algorithm is based on the idea of trading accuracy for running time, thus achieving a *reasonable tradeoff between performance and complexity.*

The approximation algorithm utilizes Algorithm 6.1 and is shown in Algorithm 6.2. Specifically, Algorithm 6.2 is similar to the fully polynomial-time approximation

Algorithm 6.2 DP-based Approximation Algorithm (DPAA)

Input: $K \in \mathbb{Z}_+$, $\mathbf{c} = [c_1, c_2, \ldots, c_K] \in \mathbb{Z}_+^K$ where $c_1 \leq c_2 \leq \cdots \leq c_K$,
$\mathbf{p} = [p_1, p_2, \ldots, p_K] \in (0,1]^K$, $P_{\text{out}}^{\text{th}} \in (0,1]$ with $\prod_{k\in\mathcal{K}} p_k \leq P_{\text{out}}^{\text{th}}$, $\epsilon > 0$

Output: $\tilde{\mathbf{z}}^* \in \mathcal{F}$ such that $f(\mathbf{z}^*) \leq f(\tilde{\mathbf{z}}^*) \leq f(\mathbf{z}^*) + \min(\lfloor \epsilon c_{\max} \rfloor, \overline{C})$

1: $\vartheta \coloneqq \epsilon c_{\max}/K$, where $c_{\max} = \max_{k\in\mathcal{K}}\{c_k\} = c_K$

2: $\tilde{c}_k \coloneqq \lceil c_k/\vartheta \rceil$, $\forall k \in \mathcal{K}$

3: Run Algorithm 6.1 with input $[K, \tilde{\mathbf{c}}, \mathbf{p}, P_{\text{out}}^{\text{th}}]$ and return the optimal solution $\tilde{\mathbf{z}}^*$, where $\tilde{\mathbf{c}} = [\tilde{c}_1, \tilde{c}_2, \ldots, \tilde{c}_K] \in \mathbb{Z}_+^K$ $(\tilde{c}_1 \leq \tilde{c}_2 \leq \cdots \leq \tilde{c}_K)$

scheme (FPTAS) for the knapsack problem provided in [11], which is inspired by the work of Ibarra and Kim [12]. Moreover, note that $\vartheta > 0$, and therefore $\tilde{c}_k \in \mathbb{Z}_+$, $\forall k \in \mathcal{K}$.

Theorem 6.3 (Performance guarantee)**.** *Assuming that problem (6.2)/(6.3) is feasible, Algorithm 6.2 takes a parameter $\epsilon > 0$ as input and produces an approximate solution $\tilde{\mathbf{z}}^* \in \mathcal{F}$ such that $f(\mathbf{z}^*) \leq f(\tilde{\mathbf{z}}^*) \leq f(\mathbf{z}^*) + \min(\lfloor \epsilon c_{\max} \rfloor, \overline{C})$, where the term $\min(\lfloor \epsilon c_{\max} \rfloor, \overline{C})$ is the absolute-error bound. In addition, for any $0 < \epsilon < 1/c_{\max}$, Algorithm 6.2 always finds an optimal solution, i.e., it becomes an exact optimization algorithm.*

Proof. Obviously, $\tilde{\mathbf{z}}^* \in \mathcal{F}$ and thus $f(\mathbf{z}^*) \leq f(\tilde{\mathbf{z}}^*)$. Now, it is sufficient to prove that $f(\tilde{\mathbf{z}}^*) \leq f(\mathbf{z}^*) + \min(\lfloor \epsilon c_{\max} \rfloor, \overline{C})$. First, we will show that $f(\tilde{\mathbf{z}}^*) \leq f(\mathbf{z}^*) + \epsilon c_{\max}$. Due to the fact that $x \leq \lceil x \rceil < x + 1$, we have $c_k/\vartheta \leq \tilde{c}_k < c_k/\vartheta + 1 \Rightarrow c_k \leq \vartheta\tilde{c}_k < c_k + \vartheta$. Also, let us define the function $g(\mathbf{z}) = \sum_{k\in\mathcal{K}} \tilde{c}_k z_k$. From $\vartheta\tilde{c}_k < c_k + \vartheta$, we deduce that $\vartheta\tilde{c}_k z_k^* \leq c_k z_k^* + \vartheta z_k^*$, $\forall k \in \mathcal{K}$ (because $z_k^* \geq 0$). By taking the sum for all $k \in \mathcal{K}$, we obtain $\vartheta g(\mathbf{z}^*) \leq f(\mathbf{z}^*) + \vartheta \sum_{k\in\mathcal{K}} z_k^* \leq f(\mathbf{z}^*) + \vartheta K = f(\mathbf{z}^*) + \epsilon c_{\max}$. Since $\mathbf{z}^* \in \mathcal{F}$, we conclude that $g(\tilde{\mathbf{z}}^*) \leq g(\mathbf{z}^*) \Rightarrow \vartheta g(\tilde{\mathbf{z}}^*) \leq \vartheta g(\mathbf{z}^*)$ because $\vartheta > 0$, and therefore $\vartheta g(\tilde{\mathbf{z}}^*) \leq f(\mathbf{z}^*) + \epsilon c_{\max}$. In addition, from $c_k \leq \vartheta\tilde{c}_k \Rightarrow c_k \tilde{z}_k^* \leq \vartheta\tilde{c}_k \tilde{z}_k^*$, $\forall k \in \mathcal{K}$ (because $\tilde{z}_k^* \geq 0$). By taking the sum for all $k \in \mathcal{K}$ once more, we get $f(\tilde{\mathbf{z}}^*) \leq \vartheta g(\tilde{\mathbf{z}}^*)$. Consequently, $f(\tilde{\mathbf{z}}^*) \leq f(\mathbf{z}^*) + \epsilon c_{\max}$. Afterwards, due to the fact that $f(\tilde{\mathbf{z}}^*)$ and $f(\mathbf{z}^*)$ are *integers*, we have $f(\tilde{\mathbf{z}}^*) - f(\mathbf{z}^*) \leq \lfloor \epsilon c_{\max} \rfloor$. Moreover, since $f(\tilde{\mathbf{z}}^*) \leq \overline{C}$ and $f(\mathbf{z}^*) \geq 0$, we obtain $f(\tilde{\mathbf{z}}^*) - f(\mathbf{z}^*) \leq \overline{C}$. Hence, $f(\tilde{\mathbf{z}}^*) - f(\mathbf{z}^*) \leq \min(\lfloor \epsilon c_{\max} \rfloor, \overline{C})$,

Table 6.1: Performance & Complexity of Optimization Algorithms

Optimization Algorithm	**Performance Guarantee**	**Computational Complexity**
Exhaustive Search	Global Optimization	$\Theta(2^K K)$
Dynamic Programming (DP)	Global Optimization	$\Theta(KC) = O(K\overline{C}) =$ $= O(K^2 c_{\max})$
DP-based Approximation Algorithm (DPAA)	$f(\mathbf{z}^*) \le f(\tilde{\mathbf{z}}^*) \le$ $\le f(\mathbf{z}^*) + \min(\lfloor \epsilon c_{\max} \rfloor, \overline{C})$	$O(K^2 \lceil K/\epsilon \rceil) =$ $= O(K^3/\epsilon)$

because it holds that: $x \le u$ and $x \le v \Leftrightarrow x \le \min(u, v)$.

Furthermore, if $0 < \epsilon < 1/c_{\max} \Rightarrow 0 < \epsilon c_{\max} < 1 \Rightarrow \lfloor \epsilon c_{\max} \rfloor = 0$, and thus $f(\mathbf{z}^*) \le f(\tilde{\mathbf{z}}^*) \le f(\mathbf{z}^*) \Rightarrow f(\tilde{\mathbf{z}}^*) = f(\mathbf{z}^*)$. In other words, for any $0 < \epsilon < 1/c_{\max}$, the approximation algorithm will be forced to produce an optimal solution. □

Complexity of Algorithm 6.2: The complexity of DPAA is mainly due to Algorithm 6.1, so it is $O(K^2 \tilde{c}_{\max}) = O(K^2 \lceil K/\epsilon \rceil) = O(K^3/\epsilon)$, where $\tilde{c}_{\max} = \max_{k \in \mathcal{K}}\{\tilde{c}_k\} = \lceil c_{\max}/\vartheta \rceil = \lceil K/\epsilon \rceil$. As a result, *Algorithm 6.2 has polynomial complexity in K and* $1/\epsilon$. Observe that, for any fixed $\epsilon > 0$, DPAA has cubic complexity $O(K^3)$.

Finally, the performance and complexity of all optimization algorithms are summarized in Table 6.1. The exhaustive search algorithm simply checks all subsets of $\mathcal{K}$ and selects that with the minimum objective value satisfying the outage probability constraint. Therefore, it requires $\sum_{i=0}^{K} \binom{K}{i} i = K 2^{K-1} = \Theta(2^K K)$ arithmetic operations to find the global minimum.

6.5 Numerical Simulations and Discussion

In this section, we examine the performance of the proposed optimization algorithms through numerical simulations. In particular, the following system parameters have been used: $K = 25$ and $c_k = \lceil k/5 \rceil$, $\forall k \in \mathcal{K}$ ($\overline{C} = \sum_{k \in \mathcal{K}} c_k = 75$ and $c_{\max} = 5$). Moreover, we generate 100 independent (feasible) optimization problems where the outage probabilities of GSs, $[p_k]_{k \in \mathcal{K}}$, are uniformly distributed in the

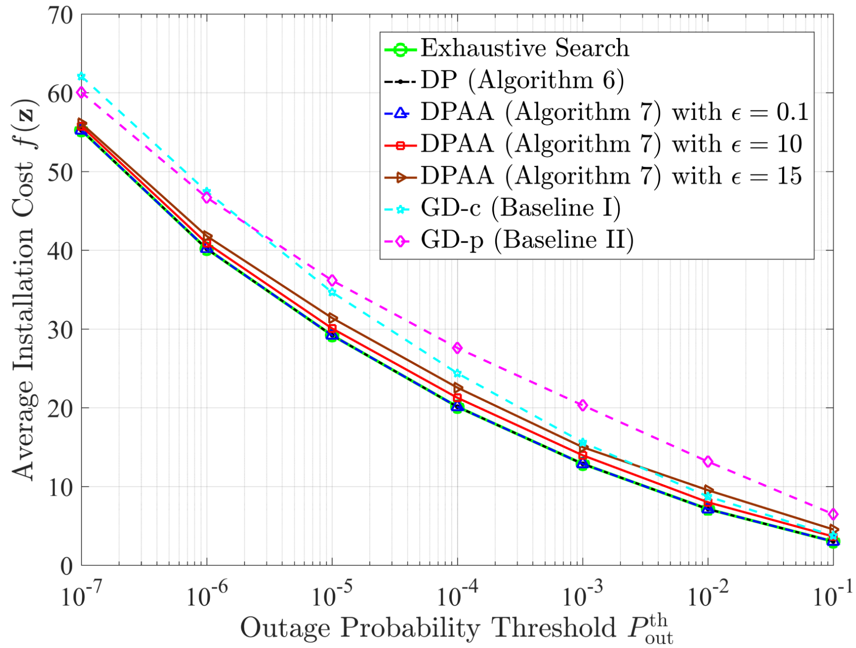

Figure 6-1: Performance comparison between optimization algorithms.

interval $(0.25, 0.75)$. For the sake of comparison, we consider two *baseline greedy algorithms*, namely, *GD-c* and *GD-p*: first, sort the candidate locations in ascending order of installation cost (respectively, outage probability), and then select the locations $\{1, 2, \ldots, n\}$ so that n is the smallest integer for which the outage probability threshold is met.

Fig. 6-1 illustrates the average installation cost, versus the outage probability threshold, achieved by a) the exhaustive search, b) DP (Algorithm 6.1), c) DPAA (Algorithm 6.2) for different values of the parameter ϵ, and d) the baseline algorithms. More specifically, DP and DPAA with $\epsilon = 0.1$ have identical performance with the exhaustive search; this is in agreement with the theory presented in the previous sections, since DP is a global optimization algorithm and DPAA is forced to produce an optimal solution when $0 < \epsilon < 1/c_{\max} = 0.2$ (see Theorem 6.3). Furthermore, as expected, DPAA leads to higher installation cost (with lower complexity) by increasing the parameter ϵ. It is interesting to note that, for $\epsilon \in \{10, 15\}$, the actual distance of the objective value achieved by DPAA from the global minimum is much less than the absolute-error bound, i.e., $f(\tilde{\mathbf{z}}^*) - f(\mathbf{z}^*) \ll \min(\lfloor \epsilon c_{\max} \rfloor, \overline{C})$. Finally, for relatively small outage probability thresholds, the baseline algorithms have lower performance than the proposed algorithms, even for large values of ϵ (e.g., $\epsilon = \overline{C}/c_{\max} = 15$).

Indicatively, for $P_{\text{out}}^{\text{th}} = 10^{-4}$ and using a computer with Intel Core i7-4790 CPU (3.6 GHz) and 16 GB RAM, the average runtime of the exhaustive search is 2.85 minutes, whereas that of all the other algorithms shown in Fig. 6-1 is less than 0.025 seconds.

6.6 Conclusion

In this chapter, we have dealt with the minimization of the installation cost of GSs in RF/optical SatNets satisfying an outage probability constraint. In particular, the examined problem has been theoretically proven to be NP-hard. Moreover, we have presented a global optimization algorithm with pseudo-polynomial complexity as well as a polynomial-time approximation algorithm with provable performance guarantee.

Bibliography

[1] A. D. Panagopoulos, "Propagation phenomena and modeling for fixed satellite systems: Evaluation of fade mitigation techniques," in *Radio wave propagation and channel modeling for earth-space systems*, Taylor & Francis Group, CRC Press, 2016.

[2] C. Fuchs and F. Moll, "Ground station network optimization for space-to-ground optical communication links," *IEEE/OSA J. Optical Commun. Network.*, vol. 7, no. 12, pp. 1148-1159, Dec. 2015.

[3] S. Poulenard, M. Crosnier and A. Rissons, "Ground segment design for broadband geostationary satellite with optical feeder link," *IEEE/OSA J. Optical Commun. Network.*, vol. 7, no. 4, pp. 325-336, April 2015.

[4] N. K. Lyras *et al.*, "Optimum monthly-based selection of ground stations for optical satellite networks," *IEEE Commun. Lett.*, vol. 22, no. 6, pp. 1192-1195, June 2018.

[5] N. K. Lyras *et al.*, "Optimizing the ground network of optical MEO satellite communication systems," *IEEE Systems Journal*, doi: 10.1109/JSYST.2019.2945838.

[6] I. del Portillo *et al.*, "Optimal location of optical ground stations to serve LEO spacecraft," *IEEE Aerosp. Conf.*, Big Sky, MT, USA, 2017, pp. 1-16.

[7] T. Rossi *et al.*, "Smart gateway diversity optimization for EHF satellite networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 130-141, Feb. 2020.

[8] S. Gong *et al.*, "Network availability maximization for free-space optical satellite communications," *IEEE Wireless Commun. Lett.*, vol. 9, no. 3, pp. 411-415, March 2020.

[9] C. N. Efrem and A. D. Panagopoulos, "Globally optimal selection of ground stations in satellite systems with site diversity," *IEEE Wireless Commun. Lett.*, to be published, doi: 10.1109/LWC.2020.2982139.

[10] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Englewood Cliffs, NJ: Prentice-Hall Inc., 1982.

[11] V. V. Vazirani, *Approximation Algorithms*. Springer-Verlag Berlin Heidelberg, 2003.

[12] O. H. Ibarra and C. E. Kim, "Fast approximation algorithms for the knapsack and sum of subset problems," *Journal of the ACM*, vol. 22, no. 4, pp. 463-468, Oct. 1975.

Chapter 7

On the Computation and Approximation of Outage Probability in Satellite Networks with Smart Gateway Diversity[17]

The utilization of extremely high frequency (EHF) bands can achieve very high throughput in satellite networks (SatNets). Nevertheless, the severe rain attenuation at EHF bands imposes strict limitations on the system availability. Smart gateway diversity (SGD) is considered indispensable in order to guarantee the required availability with reasonable cost. In this context, we examine a load-sharing SGD (LS-SGD) architecture, which has been recently proposed in the literature. For this diversity scheme, we define the system outage probability (SOP) using a rigorous probabilistic analysis based on the Poisson binomial distribution (PBD), and taking into consideration the traffic demand as well as the gateway (GW) capacity. Furthermore, we provide several methods for the exact and approximate calculation of SOP.

As concerns the exact computation of SOP, a closed-form expression and an efficient algorithm based on a recursive formula are given, both with quadratic worst-case complexity in the number of GWs. Finally, the proposed approximation methods include well-known probability distributions (binomial, Poisson, normal) and a Chernoff bound. According to the numerical results, binomial and Poisson distributions are by far the most accurate approximation methods.

7.1 Introduction

Next-generation broadband SatNets require very high data-rates (up to 1 Tbps) that can be accomplished by utilizing EHF bands (above 30 GHz) in the feeder links. Although the frequency shift from Ka (20/30 GHz) to Q/V (40/50 GHz) or W (75-110 GHz) bands provides more spectrum, the high levels of rain attenuation (tens of dB) cannot be tackled by the standard fade mitigation techniques (FMTs), such as uplink power control (ULPC), adaptive coding and modulation (ACM) and data rate adaptation (DRA). As a result, gateway diversity (GD) is necessary to achieve high system availability, since it is a more effective and powerful FMT (at the expense of installing additional GWs) [1–5]. Nevertheless, the conventional GD (where the same signal is transmitted by a group of GWs) is economically prohibitive for reaching the Tbps due to the large number of required GWs [6]. An alternative solution to achieve high availability with reasonable cost is the smart gateway diversity (SGD), where a user beam can be served by different GWs depending on the propagation conditions and the traffic load. In particular, if a GW experiences deep fades then its traffic can be rerouted to other GWs with better propagation conditions.

7.1.1 Related Work

In [6], two SGD techniques are examined, namely, the frequency multiplexing diversity and the $N + P$ diversity. The performance analysis of these schemes is based on a simple probabilistic model, assuming the same outage probability for each GW (although unusual in practice) as well as independent propagation conditions over the GW locations. Moreover, the authors in [7] study the N-active diversity (with time or frequency multiplexing, taking into account the spatial correlation between the GWs) and the $N+P$ diversity (where there are N active plus P redundant or idle GWs). In the former scheme, all the N GWs are active and each user beam is served by a group of GWs, whereas in the latter scheme each user beam is served by only one GW and switches to a redundant GW in case of outage.

A novel GW switching scheme for the $N+P$ scenario is proposed in [8], using a dynamic rain attenuation model and considering two key performance indicators: the average outage probability and the average switching rate. Furthermore, a different SGD scheme, where there are no redundant GWs but each GW should have some spare capacity, is analyzed in [9]. Specifically, in nominal clear-sky conditions all GWs are active and operate using a maximum fraction of their full capacity, while if some GWs experience heavy rain attenuation then their traffic is served by the remaining GWs using their extra capacity. Finally, an extension of the well-known N-active and $N+P$ diversity schemes to multiple-input-multiple-output (MIMO) architectures is presented in [10].

7.1.2 Contribution

The main contributions of this chapter, in comparison with existing approaches, are as follows:

- In this chapter, we analyze in detail a SGD architecture operating in *load-sharing mode*, where the GWs do not necessarily have equal outage probabilities. To the best of our knowledge, the concept of LS-SGD has been firstly introduced in [9], assuming that all GWs utilize the same fraction of their full capacity in clear-sky conditions; our analysis, however, does not make such an assumption.

- Unlike previous research, we present a *system-level approach* taking into account the *traffic demand* as well as the *GW capacity*. In particular, we are interested in the *system outage probability (SOP)*, defined as the probability of not satisfying the overall traffic demand, which is a *stricter performance metric* than the *user outage probability (UOP)*, i.e., the probability of not satisfying the traffic demand of a specific user.

- Furthermore, we study the performance improvement (in terms of SOP) that can be achieved by increasing the number of GWs in the LS-SGD scheme. For this purpose, we define two comparative metrics, namely, the *SOP-improvement factor* and the *generalized SOP-improvement factor*.

- In addition, exact methods for the computation of SOP are given, including a *closed-form expression* and an *efficient algorithm based on a recursive formula*. The worst-case complexity of both methods is quadratic in the number of GWs.

- Finally, we provide some approximation methods for the estimation of SOP. More specifically, the SOP can be approximated by various probability distributions (binomial, Poisson, normal) as well as a Chernoff bound. Ultimately, we conclude that *binomial* and *Poisson* distributions are the most appropriate approximation methods for SGD systems operating in EHF bands.

7.1.3 Chapter Organization

The remainder of this chapter is organized as follows. Section 7.2 describes and analyzes in detail the LS-SGD architecture. Moreover, Sections 7.3 and 7.4 present exact and approximation methods for calculating the SOP, respectively. In addition, the performance of LS-SGD as well as the accuracy of approximation methods are examined in Section 7.5. Finally, concluding remarks are given in Section 7.6.

7.1.4 Mathematical Notation & Conventions

Mathematical notation: $\mathbb{Z}^+ = \{1, 2, 3, \ldots\}$, $\mathbb{Z}_0^+ = \{0, 1, 2, \ldots\}$, $\mathcal{N} = \{1, 2, \ldots, N\}$ and $\mathcal{N}_0 = \{0, 1, \ldots, N\}$, where $N \in \mathbb{Z}^+$. Moreover, $\mathbb{P}(\cdot)$ and $\mathbb{E}(\cdot)$ denote probability and expectation, respectively. $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are respectively the floor and ceiling functions. In addition, $|x|$ represents the absolute value of a real number x, while

$|\mathcal{S}|$ stands for the cardinality of a set $\mathcal{S}$. $\mathbf{0}_N$ and $\mathbf{1}_N$ denote the N-dimensional all-zeros and all-ones vectors, respectively. Furthermore, $\varphi(x) = \left(\sqrt{2\pi}\right)^{-1} e^{-0.5x^2}$ is the probability density function (PDF), $\Phi(x) = \int_{-\infty}^{x} \varphi(u)du$ is the cumulative distribution function (CDF), and $Q(x) = 1 - \Phi(x)$ is the complementary CDF (CCDF) of the standard normal distribution. Finally, the *total variation distance* between two (discrete) random variables (RVs) X and Y on $\mathbb{Z}_0^+$ is defined as follows:

$$\begin{aligned} d_{\mathrm{TV}}(X,Y) = \sup_{\mathcal{A}\subseteq\mathbb{Z}_0^+} |\mathbb{P}(X \in \mathcal{A}) - \mathbb{P}(Y \in \mathcal{A})| = \\ = \tfrac{1}{2} \sum_{m\in\mathbb{Z}_0^+} |\mathbb{P}(X = m) - \mathbb{P}(Y = m)| \end{aligned} \tag{7.1}$$

Mathematical conventions: $\sum_{i\in\emptyset} a_i = 0$ and $\prod_{i\in\emptyset} a_i = 1$.

7.1.5 Preliminaries on Discrete Probability Distributions

7.1.5.1 Bernoulli Distribution

A binary (0/1) RV follows a Bernoulli distribution with parameter $p \in [0,1]$, $X \sim \mathrm{Bern}(p)$, if and only if (iff) its probability mass function (PMF) is given by: $\mathbb{P}(X = 1) = 1 - \mathbb{P}(X = 0) = p$.

7.1.5.2 Binomial Distribution

A discrete (integer-valued) RV $X \sim \mathrm{Bin}(N,p)$, where $N \in \mathbb{Z}^+$ and $p \in [0,1]$, iff its PMF is:

$$\mathbb{P}(X = m) = \binom{N}{m} p^m (1-p)^{N-m}, \ \ \forall m \in \mathcal{N}_0 \tag{7.2}$$

The binomial distribution is a generalization of the Bernoulli distribution, because $\mathrm{Bin}(1,p) \equiv \mathrm{Bern}(p)$. Furthermore, if $\{X_n\}_{n\in\mathcal{N}}$ is a set of *independent and identically distributed (i.i.d.)* Bernoulli RVs ($X_n \sim \mathrm{Bern}(p)$, $\forall n \in \mathcal{N}$), then $S = \sum_{n\in\mathcal{N}} X_n \sim \mathrm{Bin}(N,p)$.

7.1.5.3 Poisson Binomial Distribution

A discrete RV $X \sim \text{PoisBin}(\mathbf{p})$, where $\mathbf{p} = [p_1, p_2, \ldots, p_N] \in [0,1]^N$ with $N \in \mathbb{Z}^+$, iff its PMF is given by:

$$\mathbb{P}(X = m) = \sum_{\mathcal{A} \in \mathcal{C}_m} \prod_{i \in \mathcal{A}} p_i \prod_{j \in \mathcal{N} \setminus \mathcal{A}} (1 - p_j), \quad \forall m \in \mathcal{N}_0 \tag{7.3}$$

where $\mathcal{C}_m = \{\mathcal{A} \subseteq \mathcal{N} : |\mathcal{A}| = m\}$ (i.e., the set of all subsets of $\mathcal{N}$ having m elements) with $|\mathcal{C}_m| = \binom{N}{m} = \frac{N!}{m!(N-m)!}$. The binomial distribution is a special case of the PBD, since $\text{PoisBin}(p\mathbf{1}_N) \equiv \text{Bin}(N, p)$. Moreover, if $\{X_n\}_{n \in \mathcal{N}}$ is a set of *independent, but not necessarily identically distributed,* Bernoulli RVs ($X_n \sim \text{Bern}(p_n)$, $\forall n \in \mathcal{N}$), then $S = \sum_{n \in \mathcal{N}} X_n \sim \text{PoisBin}(\mathbf{p})$.

7.1.5.4 Poisson Distribution

A discrete RV $X \sim \text{Pois}(\lambda)$, where $\lambda \geq 0$, iff its PMF is expressed by: $\mathbb{P}(X = m) = e^{-\lambda} \lambda^m (m!)^{-1}$, $\forall m \in \mathbb{Z}_0^+$.

7.2 Smart Gateway Diversity Architecture

In this section, we describe and analyze a *load-sharing SGD (LS-SGD)* scheme, where the unused capacity of available (not in outage) GWs can be exploited to serve the users of the remaining GWs (which are in outage). To the best of our knowledge, this SGD architecture has been firstly proposed and analyzed in [9]. Nevertheless, our approach is somewhat different, since it explicitly takes into consideration the *traffic demand* as well as the *GW capacity.*

7.2.1 System Model

Consider a SatNet consisting of a geostationary satellite and a ground network of $N \in \mathbb{Z}^+$ (geographically distributed) GWs, which are denoted by the set $\mathcal{N} = \{1, 2, \ldots, N\}$. All the GWs are connected to a *network control center (NCC)* through dedicated terrestrial links. The NCC performs, when necessary (in case of deep fading), the traffic switching/rerouting between the GWs.[18] Furthermore, the following analysis focuses on the *feeder links* (data transmission from the GWs to the satellite), considering ideal (without noise and interference) satellite-user links.[19]

In addition, *the distance between any two different GWs is large enough* (some hundreds of km), and thus the spatial correlation of the propagation impairments at the GW locations is extremely small [6, 13]. As a result, the rain attenuations/fades experienced by the GWs can be considered *(mutually) independent.* It is also assumed that there is no ACM, so each feeder link is either available at full capacity or completely unavailable.[20] Therefore, the feeder links can be mathematically modeled as a set $\{X_n\}_{n\in\mathcal{N}}$ of *independent, but not necessarily identically distributed,* Bernoulli RVs ($X_n \sim \text{Bern}(p_n)$, $\forall n \in \mathcal{N}$), where $p_n \in [0,1]$ is the outage/exceedance probability of the n^{th} link/GW (i.e., the probability that the rain attenuation exceeds a specific threshold); some methods for calculating p_n are discussed in [9]. Moreover, we define the RV $S_{\mathcal{N}} = \sum\limits_{n\in\mathcal{N}} X_n \sim \text{PoisBin}(\mathbf{p}_{\mathcal{N}})$, with $\mathbf{p}_{\mathcal{N}} = [p_1, p_2, \ldots, p_N]$, which is *the total number of GWs that are in outage in the set* $\mathcal{N}$.[21] The expectation, the standard deviation, and the 3rd central moment of $S_{\mathcal{N}}$ are given respectively by:

$$\mu_{\mathcal{N}} = \mathbb{E}(S_{\mathcal{N}}) = \sum_{n\in\mathcal{N}} p_n \tag{7.4}$$

$$\sigma_{\mathcal{N}} = \sqrt{\mathbb{E}\left((S_{\mathcal{N}} - \mu_{\mathcal{N}})^2\right)} = \sqrt{\sum_{n\in\mathcal{N}} p_n(1-p_n)} \tag{7.5}$$

[18]The details on the switching/handover procedure are beyond the scope of this chapter; see [6, 8, 11] for more information on this important topic.

[19]As concerns the downlink of multibeam satellite systems, an energy-efficient power allocation in order to jointly minimize the unmet system capacity and the total radiated power is proposed in [12].

[20]Classical FMTs, such as ULPC, ACM and DRA, can tackle impairments of a few dB (e.g., gaseous absorption and cloud attenuation). However, in EHF bands these techniques alone are no longer effective, because the rain attenuation can reach tens of dB. Hence, SGD has to be used in order to keep SOP at the required levels. In essence, due to the intense rain attenuation in EHF bands, *SGD is the primary FMT*, whereas ULPC, ACM and DRA are secondary/supplementary FMTs. As a result, the absence of ACM in the analysis of SGD is quite reasonable. In any case, our approach provides a *lower bound on the performance* of a more realistic system that utilizes SGD together with standard FMTs.

[21]According to Section 7.1.5.2, if $p_n = p$, $\forall n \in \mathcal{N}$ (i.i.d. Bernoulli RVs), then $S_{\mathcal{N}} \sim \text{Bin}(N, p)$. Note that this is rarely the case in practice.

$$\nu_{\mathcal{N}} = \mathbb{E}\left((S_{\mathcal{N}} - \mu_{\mathcal{N}})^3\right) = \sum_{n\in\mathcal{N}} p_n(1-p_n)(1-2p_n) \tag{7.6}$$

Note that $\mu_{\mathcal{N}} \geq \sigma^2_{\mathcal{N}}$, $\mu_{\mathcal{N}} \in [0, N]$, $\sigma^2_{\mathcal{N}} \in [0, N/4]$, and $\nu_{\mathcal{N}} \in [-N/(6\sqrt{3}), N/(6\sqrt{3})]$.

7.2.2 System Outage Probability

In the sequel, suppose that the n^{th} GW can offer a *maximum data-rate (capacity)* $R_n^{\max} > 0$, and the *total requested data-rate (traffic demand)* is $R_{\text{tot}}^{\text{req}} = \sum_{u\in\mathcal{U}} R_u^{\text{req}} > 0$, where $\mathcal{U} = \{1, 2, \ldots, U\}$ is the set of users and $R_u^{\text{req}} \geq 0$ is the requested data-rate of user u. Moreover, the operation of NCC ensures the following *load-sharing property*: all users receive their requested data-rate if and only if (iff) the overall capacity of the available (not in outage) GWs is greater than or equal to the traffic demand. Equivalently, there is at least one user that receives inadequate data-rate iff the overall capacity of the available GWs is less than the traffic demand.

Definition 7.1 (General SOP expression)**.** *The SOP is defined as follows:*

$$P_{\text{out}}^{\text{sys}} = \sum_{\mathcal{A}\in\mathcal{F}} \prod_{i\in\mathcal{A}} p_i \prod_{j\in\mathcal{N}\setminus\mathcal{A}} (1-p_j) \tag{7.7}$$

where $\mathcal{F} = \left\{\mathcal{A} \subseteq \mathcal{N} : \sum_{j\in\mathcal{N}\setminus\mathcal{A}} R_j^{\max} < R_{\text{tot}}^{\text{req}}\right\}$*. In other words,* $\mathcal{F}$ *contains all the subsets* $\mathcal{A}$ *of the* N *GWs such that: if the GWs in* $\mathcal{A}$ *are all in outage and the remaining GWs in* $\mathcal{N}\setminus\mathcal{A}$ *are all available (not in outage), then the traffic demand cannot be satisfied by the latter group of GWs. In essence, the SOP expresses the probability of not satisfying the traffic demand of all users (or, equivalently, the probability that there is at least one user that receives inadequate data-rate). Similarly, we can define the* system availability (SA) *as the probability of the complementary event:* $P_{\text{avail}}^{\text{sys}} = 1 - P_{\text{out}}^{\text{sys}}$.

For simplicity, we assume that *all GWs have the same capacity,* $R_{\text{GW}}^{\max} > 0$*, in the rest of the chapter*; this is not such a strong assumption in practice, since the same frequency band is fully reused in each feeder link and the clear-sky link budget is almost identical for all GWs.

Theorem 7.1 (Special SOP expression)**.** *Suppose that all GWs have the same capacity, i.e.,* $R_n^{\max} = R_{\text{GW}}^{\max} > 0$, $\forall n \in \mathcal{N}$*. Then,* (7.7) *reduces to the following expression:*

$$P_{\text{out}}^{\text{sys}} = P_{\text{out}}^{\text{sys}}(L, N) = \sum_{m=L}^{N} \sum_{\mathcal{A} \in \mathcal{C}_m} \prod_{i \in \mathcal{A}} p_i \prod_{j \in \mathcal{N} \setminus \mathcal{A}} (1 - p_j) \tag{7.8}$$

where $\mathcal{C}_m = \{\mathcal{A} \subseteq \mathcal{N} : |\mathcal{A}| = m\}$ *and* L *is given by:*

$$L = N - \lceil r \rceil + 1 \tag{7.9}$$

where $r > 0$ *is* the ratio of the traffic demand to the GW capacity*, that is:*

$$r = R_{\text{tot}}^{\text{req}} / R_{\text{GW}}^{\max} \tag{7.10}$$

Proof. Under the condition of equal GW capacities, we have that $\mathcal{F} = \{\mathcal{A} \subseteq \mathcal{N} : (N - |\mathcal{A}|) R_{\text{GW}}^{\max} < R_{\text{tot}}^{\text{req}}\}$. In addition, $(N - |\mathcal{A}|) R_{\text{GW}}^{\max} < R_{\text{tot}}^{\text{req}} \Leftrightarrow N - |\mathcal{A}| < r \Leftrightarrow N - |\mathcal{A}| < \lceil r \rceil \Leftrightarrow N - |\mathcal{A}| \leq \lceil r \rceil - 1 \Leftrightarrow |\mathcal{A}| \geq N - \lceil r \rceil + 1$. Consequently, $\mathcal{F} = \{\mathcal{A} \subseteq \mathcal{N} : |\mathcal{A}| \geq L\} = \bigcup_{m=L}^{N} \mathcal{C}_m$ and then (7.8) follows immediately from (7.7). □

Remark 7.1*.* According to Section 7.1.5.3, $P_{\text{out}}^{\text{sys}}(L, N) = \sum_{m=L}^{N} \mathbb{P}(S_{\mathcal{N}} = m) = \mathbb{P}(S_{\mathcal{N}} \geq L)$, i.e., the SOP is *the probability of having at least* L *out of* N *GWs in outage.*[22]

Although in general $L \in \mathcal{N}_0$, for the diversity system under consideration $L \in \mathcal{N}$ due to the fact that $\lceil r \rceil \in \mathcal{N}$, since a) $r > 0 \Leftrightarrow \lceil r \rceil \geq 1$, and b) $N R_{\text{GW}}^{\max} \geq R_{\text{tot}}^{\text{req}} \Leftrightarrow N \geq r \Leftrightarrow N \geq \lceil r \rceil$ (note that $N_{\min} = \lceil r \rceil$ is *the minimum required number of GWs*). Finally, we provide a result about the monotonicity of SOP.

[22] Similar formula is also given in [9] and [14], however, without explicit dependence on the traffic demand and the GW capacity. Herein, this dependence is clearly expressed by (7.9) and (7.10). Note that this SOP definition is a generalization of *the classical SOP* (i.e., the probability of having all GWs in outage), which is obtained when $\lceil r \rceil = 1 \Rightarrow L = N \Rightarrow P_{\text{out}}^{\text{sys}} = \prod_{n \in \mathcal{N}} p_n$; the classical SOP is used in [15] to select the (globally) minimum number of GWs satisfying SOP-requirements.

Proposition 7.1 (SOP monotonicity)**.** *For a given set $\mathcal{N}$ of GWs, the SOP is an increasing function of r.*

Proof. Let $r_1 \geq r_2 \Rightarrow \lceil r_1 \rceil \geq \lceil r_2 \rceil \Rightarrow L_1 \leq L_2 \Rightarrow P_{\text{out}}^{\text{sys}}(L_1, N) \geq P_{\text{out}}^{\text{sys}}(L_2, N)$. □

7.2.3 SOP-Improvement Factor

Subsequently, we study the performance improvement (in terms of SOP) achieved by an N-GW diversity system in comparison with a single-GW system.

Definition 7.2 (SOP-improvement factor)**.** *Assuming the same $\lceil r \rceil = 1$ and that $P_{\text{out}}^{\text{sys}}(N, N) > 0$, the SOP-improvement factor is defined as follows:*

$$I = \frac{P_{\text{out}}^{\text{sys}}(1,1)}{P_{\text{out}}^{\text{sys}}(N,N)} = \frac{p_1}{\prod\limits_{n \in \mathcal{N}} p_n} = \left(\prod_{n=2}^{N} p_n \right)^{-1} \tag{7.11}$$

Obviously, it holds that $I \geq 1$.

Next, consider a diversity system with $N + K$ GWs ($K \in \mathbb{Z}_0^+$) all of which have the same capacity $R_{\text{GW}}^{\max} > 0$, and $\lceil r \rceil \in \mathcal{N}$ (since $1 \leq \lceil r \rceil \leq \min(N, N+K) = N$). Furthermore, let $\mathcal{K} = \{N+1, N+2, \ldots, N+K\}$ be the set of additional GWs, and $\mathbf{p}_{\mathcal{N}\cup\mathcal{K}} = [\mathbf{p}_{\mathcal{N}}, \mathbf{p}_{\mathcal{K}}] = [p_1, p_2, \ldots, p_{N+K}]$ be the vector of GW outage probabilities, where $\mathbf{p}_{\mathcal{K}} = [p_{N+1}, p_{N+2}, \ldots, p_{N+K}]$. Suppose also that $\{X_i\}_{i \in \mathcal{N}\cup\mathcal{K}}$ is a set of *independent, but not necessarily identically distributed,* Bernoulli RVs ($X_i \sim \text{Bern}(p_i)$, $\forall i \in \mathcal{N}\cup\mathcal{K}$). Besides $S_{\mathcal{N}}$, we define the RVs $S_{\mathcal{K}} = \sum\limits_{k \in \mathcal{K}} X_k \sim \text{PoisBin}(\mathbf{p}_{\mathcal{K}})$ and $S_{\mathcal{N}\cup\mathcal{K}} = \sum\limits_{i \in \mathcal{N}\cup\mathcal{K}} X_i = S_{\mathcal{N}} + S_{\mathcal{K}} \sim \text{PoisBin}(\mathbf{p}_{\mathcal{N}\cup\mathcal{K}})$ denoting the total number of GWs which are in outage in the sets $\mathcal{K}$ and $\mathcal{N}\cup\mathcal{K}$, respectively. For this diversity system $L' = N + K - \lceil r \rceil + 1 = L + K$, with $L' \in \{K+1, K+2, \ldots, K+N\}$.

Proposition 7.2 (SOP reduction)**.** *Let $P_{\text{out}}^{\mathcal{N}} = P_{\text{out}}^{\text{sys}}(L, N) = \mathbb{P}(S_{\mathcal{N}} \geq L)$ and $P_{\text{out}}^{\mathcal{N}\cup\mathcal{K}} = P_{\text{out}}^{\text{sys}}(L', N+K) = \mathbb{P}(S_{\mathcal{N}\cup\mathcal{K}} \geq L')$ stand for the SOP of the N-GW and $(N+K)$-GW diversity systems, respectively. Then, it holds that $P_{\text{out}}^{\mathcal{N}\cup\mathcal{K}} \leq P_{\text{out}}^{\mathcal{N}}$.*

Proof. See Appendix 7.7. □

In view of this fact, we can generalize the definition of SOP-improvement factor.

Definition 7.3 (Generalized SOP-improvement factor). *Assuming the same* $\lceil r \rceil \in \mathcal{N}$ *and that* $P_{\text{out}}^{\mathcal{N}\cup\mathcal{K}} > 0$, *we define the generalized SOP-improvement factor of the* $(N+K)$*-GW over the* N*-GW diversity system as follows:*[23]

$$I_{\text{g}} = \frac{P_{\text{out}}^{\mathcal{N}}}{P_{\text{out}}^{\mathcal{N}\cup\mathcal{K}}} = \left.\frac{P_{\text{out}}^{\text{sys}}(L,N)}{P_{\text{out}}^{\text{sys}}(L+K,N+K)}\right|_{L=N-\lceil r \rceil+1} \tag{7.12}$$

According to Proposition 7.2, we have that $I_{\text{g}} \geq 1$.

Notice that by setting $N = 1$ and $K = N' - 1$ (thus $\lceil r \rceil = 1$ and $L = 1$), we obtain $I_{\text{g}} = \frac{P_{\text{out}}^{\text{sys}}(1,1)}{P_{\text{out}}^{\text{sys}}(N',N')} = I$. Finally, we would like to emphasize that by increasing the number of GWs the SOP decreases, but higher GW connectivity is required; such connectivity issues are very important in the design and optimization of SatNets [16]. In other words, there is a trade-off between performance improvement and connectivity complexity.

7.3 Exact Methods for Computing SOP

In the sequel, several techniques for the exact computation of SOP are presented. The time complexity of these methods is summarized in Table 7.1.

7.3.1 Direct Computation

The direct computation of SOP is based on the analytic formula (7.8), which requires $\sum_{m=L}^{N} |\mathcal{C}_m|\, N = N \sum_{m=L}^{N} \binom{N}{m} \leq N \sum_{m=0}^{N} \binom{N}{m} = 2^N N = O(2^N N)$ arithmetic operations. Because of its exponential worst-case complexity, this method is practicable only for very small N.

[23] The generalized SOP-improvement factor I_{g} can be estimated using the approximation methods provided in Section 7.4.

Table 7.1: Complexity Comparison Between Exact Methods

Exact Method	Direct Computation	CFE	RF (Algorithm 7.1)	FFT-based Algorithm [20]
Time Complexity	$O(2^N N)$	$\Theta(N^2)$	$\Theta(L(N-L+1)) = O(N^2)$	$O(N(\log N)^2)$

7.3.2 Closed-Form Expression

According to [17], the SOP can be calculated, using polynomial interpolation and discrete Fourier transform (DFT), by the following *closed-form expression (CFE)*:

$$P_{\text{out}}^{\text{sys}}(L,N) = 1 - \tfrac{1}{N+1}\left(L + \sum_{n\in\mathcal{N}} \tfrac{1-c^{-nL}}{1-c^{-n}} \prod_{m\in\mathcal{N}} (1+(c^n-1)p_m)\right) \tag{7.13}$$

where $c = e^{j2\pi/(N+1)}$, with $j = \sqrt{-1}$ being the imaginary unit. It is interesting to note that the CFE comprises a sum of complex numbers, but the overall outcome is a real number in $[0,1]$. The same formula is also derived in [18], using the characteristic function of the PBD as well as the DFT. Furthermore, the computational complexity of (7.13) is $\Theta(N^2)$.

7.3.3 Recursive Formula

In this part, we explore the power and beauty of recursion.

Theorem 7.2 (SOP recursive formula)**.** *The SOP is given by the following recursive formula (RF):*

$$P_{\text{out}}^{\text{sys}}(L,N) = (1-p_N)P_{\text{out}}^{\text{sys}}(L,N-1) + p_N P_{\text{out}}^{\text{sys}}(L-1,N-1) \tag{7.14}$$

with initial/boundary conditions: a) $P_{\text{out}}^{\text{sys}}(0,N) = 1$ *and b)* $P_{\text{out}}^{\text{sys}}(N+1,N) = 0$, $\forall N \in \mathbb{Z}^+$.

Proof. See Appendix 7.8. □

It can be verified, using mathematical induction, that (7.8) is the solution of (7.14). To the best of our knowledge, this RF is derived for the first time in [19],

Algorithm 7.1 Exact Computation of SOP

Input: $N \in \mathbb{Z}^+$, $L \in \mathcal{N}_0$, and $\mathbf{p} = [p_1, p_2, \ldots, p_N] \in [0,1]^N$

Output: $P_{\text{out}}^{\text{sys}} = P_{\text{out}}^{\text{sys}}(L, N)$

$D \coloneqq N - L$, $M \coloneqq L + 1$, $\boldsymbol{\alpha} \coloneqq \mathbf{0}_M$, $\alpha_1 \coloneqq 1$, $\ell \coloneqq 1$

for $i \coloneqq 1$ **to** N **step** $+1$ **do**

  $h \coloneqq i$

  if $i > D + 1$ **then** $\ell \coloneqq i - D$ **end if**

  if $i > L$ **then** $h \coloneqq L$ **end if**

  for $j \coloneqq h$ **to** ℓ **step** -1 **do** ▷ *h,ℓ: high/low index*

    $\alpha_{j+1} \coloneqq (1 - p_i) \cdot \alpha_{j+1} + p_i \cdot \alpha_j$

  end for

end for

$P_{\text{out}}^{\text{sys}} \coloneqq \alpha_M$

making use of symmetric switching functions. Our proof, however, is much simpler.

Algorithm 7.1 presents an efficient method to compute the SOP using the RF, which follows directly from the algorithm given in [19]. The *time complexity* of Algorithm 7.1 is $\Theta(L(N-L+1)) = O(N^2)$, with best-case complexity $\Theta(1)$ for $L = 0$, and worst-case complexity $\Theta(N^2)$ for $L = \lfloor N/2 \rfloor$ and $L = \lceil N/2 \rceil$. Moreover, notice that the complexity is $\Theta(N)$ for $L = 1$ and $L = N$. As a result, Algorithm 7.1 has lower complexity in some cases than the CFE which requires $\Theta(N^2)$ operations regardless of L. Finally, the *space complexity* of Algorithm 7.1 is $\Theta(N + L) = \Theta(N)$.

7.3.4 FFT-based Algorithm

An even more efficient and advanced algorithm for computing the SOP is provided in [20]. This method recursively applies the fast Fourier transform (FFT) to compute *generating function (GF)* products, thus achieving an overall complexity of $O(N(\log N)^2)$.

In particular, the PMF of $S_{\mathcal{N}} \sim \text{PoisBin}(\mathbf{p}_{\mathcal{N}})$ can be written in the following form:

$$\begin{aligned}&[\mathbb{P}(S_{\mathcal{N}}=0)\ \ \mathbb{P}(S_{\mathcal{N}}=1)\ \ \cdots\ \ \mathbb{P}(S_{\mathcal{N}}=N)] = \\ &= [q_1\ \ p_1] * [q_2\ \ p_2] * \cdots * [q_N\ \ p_N]\end{aligned} \tag{7.15}$$

where $*$ stands for the *convolution* operation and $q_n = 1 - p_n$, $\forall n \in \mathcal{N}$. In addition, the GF of the Poisson-binomial PMF is given by:

$$\begin{aligned}g(z) &= \sum_{n\in\mathcal{N}_0} \mathbb{P}(S_{\mathcal{N}}=n)\, z^n = \prod_{n\in\mathcal{N}} (q_n + p_n z) = \\ &= g_\pi \prod_{n\in\mathcal{N}} (1 + a_n z) = g_\pi\,(1 + A(z))\end{aligned} \tag{7.16}$$

where $g_\pi = \prod_{n\in\mathcal{N}} q_n$ and $a_n = p_n/q_n$, $\forall n \in \mathcal{N}$. Obviously, the SOP is simply the sum of the coefficients of z^m from $m = L$ to N (see Remark 7.1). Since the product of two GF is equivalent to the convolution of two sequences formed from the GF coefficients, the FFT can be used to compute GF products more efficiently compared to the term-by-term calculation. The basic idea of the algorithm proposed in [20] is to apply the FFT to compute the GF $A(z)$ using a *divide-and-conquer* approach. More details on the implementation of the algorithm can be found therein.

Remark 7.2. Despite the fact that the FFT-based algorithm is more sophisticated and has lower asymptotic complexity, CFE and Algorithm 7.1 are sufficient in the context of SGD, where the number of GWs N is relatively small.

7.4 Approximation Methods for Estimating SOP

Afterwards, we introduce some useful methods to approximate the SOP, exploiting the fact that $P_{\text{out}}^{\text{sys}}(L, N) = \mathbb{P}(S_{\mathcal{N}} \geq L) = 1 - \mathbb{P}(S_{\mathcal{N}} \leq L-1)$, $\forall L \in \mathcal{N}_0$. These techniques consist of *probability distributions (binomial, Poisson, normal)* as well as a *Chernoff bound*. For convenience, a summary of approximation methods is given in Table 7.2.

Table 7.2: Summary of Approximation Methods

Approximation Method	SOP Approximation Formula $\widetilde{P}_{\text{out}}^{\text{sys}}(L,N)$	Parameters/Range of L	Condition for Higher Accuracy
Binomial Approximation (BA)[a]	$1-\sum_{m=0}^{L-1}\binom{N}{m}\bar{p}^m\bar{q}^{N-m}$	$\bar{p}=\frac{1}{N}\sum_{n\in\mathcal{N}}p_n,\ \ \bar{q}=1-\bar{p}$	$(N\bar{p}\bar{q})^{-1}\sigma_{\mathcal{N}}^2\to 1$
Poisson Approximation (PA)[b]	$1-e^{-\mu_{\mathcal{N}}}\sum_{m=0}^{L-1}\mu_{\mathcal{N}}^m(m!)^{-1}$	—	$\sum_{n\in\mathcal{N}}p_n^2\to 0$
Normal Approximation (NA)	$1-\Phi\left(\zeta\right)=Q\left(\zeta\right)$	$\zeta=(L-\mu_{\mathcal{N}}-0.5)\sigma_{\mathcal{N}}^{-1}$	$\sigma_{\mathcal{N}}^2\to\infty$
Refined Normal Approximation (RNA)	$\min\left(\max\left(1-G(\zeta),0\right),1\right)$		
Chernoff Bound (CB)	$(\mu_{\mathcal{N}}/L)^Le^{L-\mu_{\mathcal{N}}}$	$\forall L\in\{\lfloor\mu_{\mathcal{N}}\rfloor+1,\lfloor\mu_{\mathcal{N}}\rfloor+2,\ldots,N\}$	—

[a,b] According to the numerical results (Section 7.5), BA and PA are the most appropriate approximation methods for SGD systems operating in EHF bands.

7.4.1 Binomial Approximation (BA)

The PBD can be approximated by the binomial distribution [24] in the following sense, defining $\bar{p}=\frac{1}{N}\sum_{n\in\mathcal{N}}p_n$, $\bar{q}=1-\bar{p}$, and assuming $\bar{p}\in(0,1)$: a) $d_{\text{TV}}(S_{\mathcal{N}},Y)\leq(N/(N+1))(1-\bar{p}^{N+1}-\bar{q}^{N+1})\delta_{\mathcal{N}}$, where $Y\sim\text{Bin}(N,\bar{p})$ and $\delta_{\mathcal{N}}=1-(N\bar{p}\bar{q})^{-1}\sigma_{\mathcal{N}}^2$, and b) $d_{\text{TV}}(S_{\mathcal{N}},Y)\to 0$ *if and only if (iff)* $\delta_{\mathcal{N}}\to 0$ (or, equivalently, $(N\bar{p}\bar{q})^{-1}\sigma_{\mathcal{N}}^2\to 1$). It is interesting to note that when $p_n=p$, $\forall n\in\mathcal{N}$, it holds that: $\bar{p}=p$, $\bar{q}=1-p$ and $\sigma_{\mathcal{N}}^2=N\bar{p}\bar{q}\Rightarrow\delta_{\mathcal{N}}=0\Rightarrow d_{\text{TV}}(S_{\mathcal{N}},Y)=0\Rightarrow S_{\mathcal{N}}\sim\text{Bin}(N,p)$, which is in agreement with Section 7.1.5.2. Hence, the BA is given by:

$$P_{\text{out}}^{\text{sys}}(L,N)\approx 1-\mathbb{P}(Y\leq L-1)=1-\sum_{m=0}^{L-1}\binom{N}{m}\bar{p}^m(1-\bar{p})^{N-m} \tag{7.17}$$

7.4.2 Poisson Approximation (PA)

In 1960, Le Cam [25] established a remarkable inequality: $d_{\text{TV}}(S_{\mathcal{N}},Z)\leq\sum_{n\in\mathcal{N}}p_n^2$, where $Z\sim\text{Pois}(\mu_{\mathcal{N}})$. It is obvious that if $\sum_{n\in\mathcal{N}}p_n^2\to 0$, then $d_{\text{TV}}(S_{\mathcal{N}},Z)\to 0$. As reported in [26], *Le Cam's theorem/inequality* admits various proofs using different techniques. Consequently, we have that:

$$P_{\text{out}}^{\text{sys}}(L,N)\approx 1-\mathbb{P}(Z\leq L-1)=1-e^{-\mu_{\mathcal{N}}}\sum_{m=0}^{L-1}\mu_{\mathcal{N}}^m(m!)^{-1} \tag{7.18}$$

7.4.3 Normal Approximation (NA)

According to [21], *the central limit theorem (CLT)* for the PBD states that: $\lim_{N\to\infty} \Delta_{\mathcal{N}} = 0$ (asymptotic normality of $(S_{\mathcal{N}} - \mu_{\mathcal{N}})\sigma_{\mathcal{N}}^{-1}$) *iff* $\lim_{N\to\infty} \sigma_{\mathcal{N}}^2 = \infty$, where $\Delta_{\mathcal{N}} = \sup_{s\in\mathbb{R}} \left|\mathbb{P}(S_{\mathcal{N}} \leq s) - \Phi\left((s - \mu_{\mathcal{N}})\sigma_{\mathcal{N}}^{-1}\right)\right|$. Therefore, by applying a *continuity correction*,[24] the SOP can be approximated by:

$$P_{\text{out}}^{\text{sys}}(L, N) \approx 1 - \Phi(\zeta) = Q(\zeta) \tag{7.19}$$

where $\zeta = (L - \mu_{\mathcal{N}} - 0.5)\sigma_{\mathcal{N}}^{-1}$.

7.4.4 Refined Normal Approximation (RNA)

Consider the following function:

$$G(x) = \Phi(x) + \nu_{\mathcal{N}}(6\sigma_{\mathcal{N}}^3)^{-1}(1 - x^2)\varphi(x) \tag{7.20}$$

According to [21–23], there exists a constant $C < \infty$ such that $\Delta'_{\mathcal{N}} = \sup_{s\in\mathbb{R}} \left|\mathbb{P}(S_{\mathcal{N}} \leq s) - G\left((s - \mu_{\mathcal{N}})\sigma_{\mathcal{N}}^{-1}\right)\right| \leq C\sigma_{\mathcal{N}}^{-2} = O(\sigma_{\mathcal{N}}^{-2})$. Observe that $\lim_{N\to\infty} \Delta'_{\mathcal{N}} = 0$, when $\lim_{N\to\infty} \sigma_{\mathcal{N}}^2 = \infty$. As a result, by applying the *continuity correction* once more, we obtain the following approximation:

$$P_{\text{out}}^{\text{sys}}(L, N) \approx \min\left(\max\left(\widehat{P}_{\text{out}}^{\text{sys}}(L, N), 0\right), 1\right) \tag{7.21}$$

where $\widehat{P}_{\text{out}}^{\text{sys}}(L, N) = 1 - G(\zeta)$ and $\zeta = (L - \mu_{\mathcal{N}} - 0.5)\sigma_{\mathcal{N}}^{-1}$. Note that we make use of the above min-max formula in order to ensure that $P_{\text{out}}^{\text{sys}}(L, N) \in [0, 1]$, because $\widehat{P}_{\text{out}}^{\text{sys}}(L, N)$ may be outside the interval $[0, 1]$ in some cases.

[24] In probability theory, a *continuity correction* is an adjustment that is made when a discrete (probability) distribution is approximated by a continuous distribution. In particular, suppose that the continuous RV Y approximates the discrete RV X. Then, $\mathbb{P}(X \leq m) = \mathbb{P}(X \leq m + 0.5) \approx \mathbb{P}(Y \leq m + 0.5)$, $\forall m \in \mathbb{Z}$.

7.4.5 Chernoff Bound (CB)

A Chernoff (upper) bound can be constructed using a result given in [27] which states that: $\mathbb{P}\left(S_{\mathcal{N}} \geq (1+\delta)\mu_{\mathcal{N}}\right) \leq \left(e^{\delta}/(1+\delta)^{1+\delta}\right)^{\mu_{\mathcal{N}}}$, $\forall \delta > 0$. Specifically, by setting $(1+\delta)\mu_{\mathcal{N}} = L$ and assuming $\mu_{\mathcal{N}} > 0$, we obtain:

$$P_{\text{out}}^{\text{sys}}(L, N) \leq (\mu_{\mathcal{N}}/L)^{L} e^{L-\mu_{\mathcal{N}}} \tag{7.22}$$

which holds $\forall L \in \{\lfloor \mu_{\mathcal{N}} \rfloor + 1, \lfloor \mu_{\mathcal{N}} \rfloor + 2, \ldots, N\}$, since $\delta > 0 \Leftrightarrow L > \mu_{\mathcal{N}} \Leftrightarrow L > \lfloor \mu_{\mathcal{N}} \rfloor$ $\Leftrightarrow L \geq \lfloor \mu_{\mathcal{N}} \rfloor + 1$.

7.5 Numerical Results and Discussion

In this section, all results present statistical averages derived from 10^3 independent system configurations, where the GW outage probabilities $\{p_i\}_{i \in \mathcal{N} \cup \mathcal{K}}$ are uniformly distributed in $(0, 0.02)$, i.e., 98% to 100% link availability.

7.5.1 SOP Analysis

Firstly, we study the SOP as a function of the number of GWs, N, and the ratio of the traffic demand to the GW capacity, r. As shown in Fig. 7-1, the SOP increases with $\lceil r \rceil$ for all values of N, which is in accordance with Proposition 7.1. Moreover, for any fixed $\lceil r \rceil$, we can observe that the SOP decreases with the increase of N (see Proposition 7.2). Nevertheless, as mentioned at the end of Section 7.2.3, this SOP improvement is achieved in exchange for higher connectivity complexity.

Secondly, we examine the performance enhancement achieved by a $(5+K)$-GW compared to a 5-GW diversity system by means of the generalized SOP-improvement factor (where $K \in \{1, 2, 3, 4\}$ is the number of additional GWs). Specifically, as illustrated in Fig. 7-2, I_{g} decreases with the increase of $\lceil r \rceil$ for every value of K. Furthermore, for a given $\lceil r \rceil$, larger number of additional GWs results in higher performance improvement.

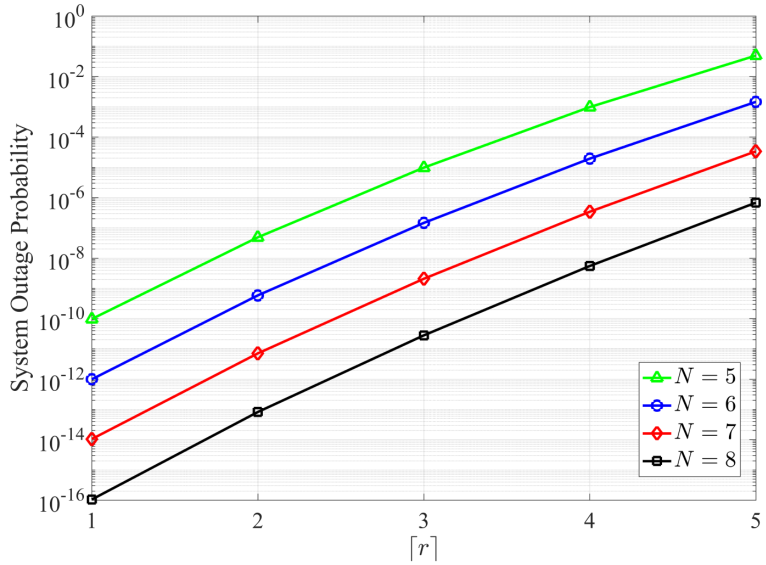

Figure 7-1: System outage probability, $P_{\text{out}}^{\text{sys}}$, (calculated using Algorithm 7.1) versus the ceiling of r (the ratio of the traffic demand to the GW capacity).

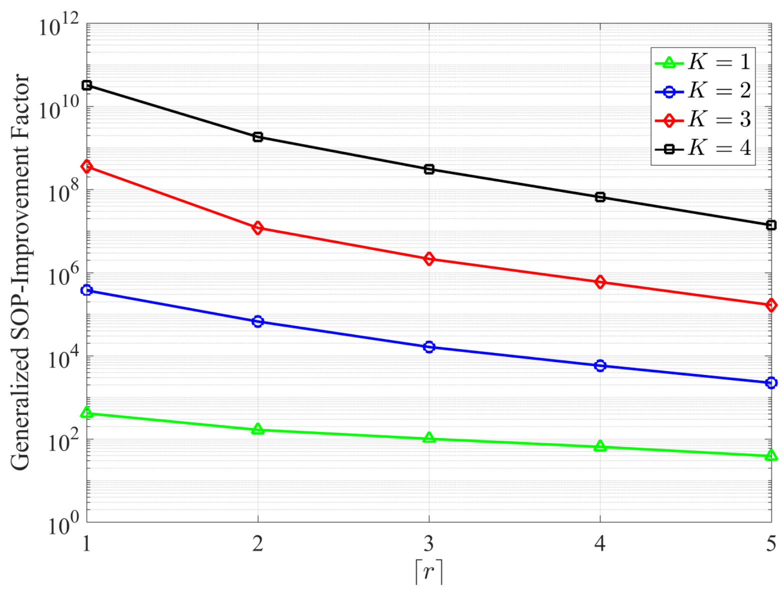

Figure 7-2: Generalized SOP-improvement factor, I_{g}, (computed using Algorithm 7.1), in comparison with a diversity system consisting of $N = 5$ GWs, versus the ceiling of r (the ratio of the traffic demand to the GW capacity).

7.5.2 Performance of Approximation Methods

In order to evaluate the accuracy of a probability distribution and the tightness/sharpness of the Chernoff bound, we define the maximum absolute error (maxAE), the root-mean-square error (RMSE), and the mean absolute error (MAE) as follows:

$$\epsilon_{\max}(N) = \max_{L\in\mathcal{S}} \left| P_{\rm out}^{\rm sys}(L,N) - \widetilde{P}_{\rm out}^{\rm sys}(L,N) \right| \tag{7.23}$$

$$\epsilon_{\rm rms}(N) = \sqrt{\tfrac{1}{|\mathcal{S}|}\sum_{L\in\mathcal{S}} \left(P_{\rm out}^{\rm sys}(L,N) - \widetilde{P}_{\rm out}^{\rm sys}(L,N) \right)^2} \tag{7.24}$$

$$\epsilon_{\rm mean}(N) = \tfrac{1}{|\mathcal{S}|}\sum_{L\in\mathcal{S}} \left| P_{\rm out}^{\rm sys}(L,N) - \widetilde{P}_{\rm out}^{\rm sys}(L,N) \right| \tag{7.25}$$

where $\widetilde{P}_{\rm out}^{\rm sys}(L,N)$ is the approximate SOP. Moreover, for probability distributions $\mathcal{S} = \mathcal{N}_0$ (with $|\mathcal{S}| = N+1$), while for CB $\mathcal{S} = \{\lfloor \mu_{\mathcal{N}} \rfloor + 1, \lfloor \mu_{\mathcal{N}} \rfloor + 2, \ldots, N\}$ (with $|\mathcal{S}| = N - \lfloor \mu_{\mathcal{N}} \rfloor \geq 1$). In general, it holds that $\epsilon_{\max}(N) \geq \epsilon_{\rm rms}(N) \geq \epsilon_{\rm mean}(N)$.

Fig. 7-3 presents the accuracy of approximation methods, in terms of maxAE, RMSE and MAE, versus the number of GWs. It can be observed that the approximation methods in descending-performance (or, equivalently, ascending-error) order are as follows: {BA, PA, NA, RNA, CB}. More specifically, BA and PA significantly outperform the other methods (the achieved errors are of the order of 10^{-4} or 10^{-5}), while CB exhibits the lowest accuracy. At this point, we would like to give an explanation of the performance of BA, PA, NA and RNA. In practice, *the number of GWs is relatively small* ($N \approx 4$ to 7) and *all the GW outage probabilities are very close to zero* (i.e., $p_n \approx 0, \forall n \in \mathcal{N} \Rightarrow p_1 \approx p_2 \approx \cdots \approx p_N$). As a result, the variance $\sigma_{\mathcal{N}}^2 = \sum_{n\in\mathcal{N}} p_n(1-p_n)$ and the quantity $\sum_{n\in\mathcal{N}} p_n^2$ are quite small, while $\sigma_{\mathcal{N}}^2 \approx N\bar{p}\bar{q}$ (see Section 7.4.1). Finally, according to Table 7.2, it is clear that the condition for higher accuracy of BA/PA is well satisfied, whereas that of NA/RNA is not. In summary, BA and PA are the most suitable approximation methods for SGD systems.

7.6 Conclusion

In this chapter, we have studied in depth the LS-SGD scheme, which has been recently introduced in SatNets. Furthermore, a number of useful mathematical tools have been presented in order to compute and approximate the SOP. Finally, based on the numerical results, we conclude that the SOP can be well approximated by BA and

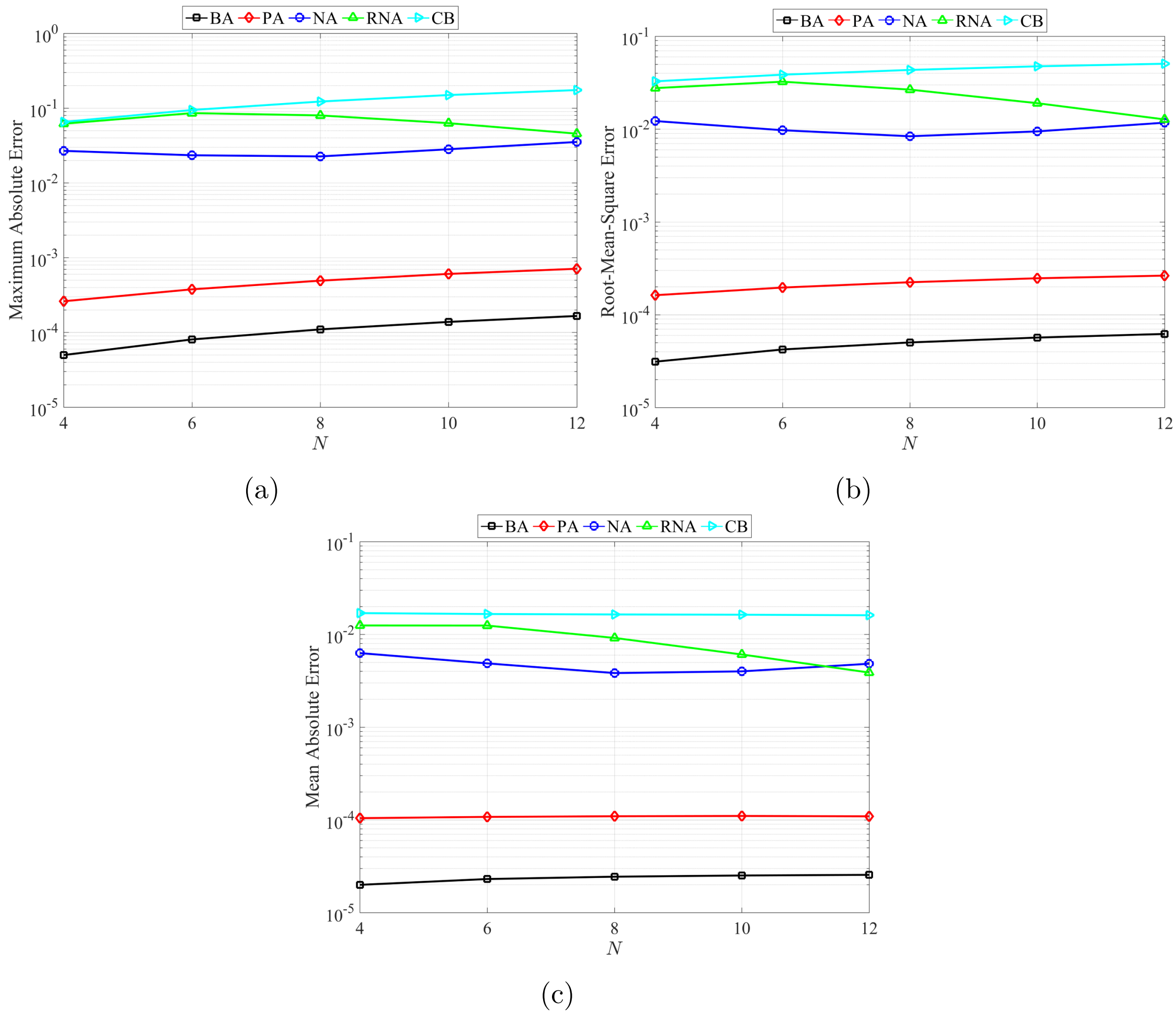

Figure 7-3: Accuracy comparison of approximation methods: (a) maximum absolute error, (b) root-mean-square error, and (c) mean absolute error versus the number of GWs.

PA, since these methods achieve remarkable accuracy. Such approximations may be useful for simplifying and solving hard optimization problems with SOP-constraints in SGD-based SatNets.

7.7 Appendix-A: Proof of Proposition 7.2

By virtue of *the law/theorem of total probability*, we obtain:

$$\begin{aligned}
&P_{\text{out}}^{\mathcal{N}\cup\mathcal{K}} = \mathbb{P}(S_{\mathcal{N}} + S_{\mathcal{K}} \geq L + K) = \\
&= \sum_{j=0}^{K} \mathbb{P}(S_{\mathcal{K}} = j)\mathbb{P}(S_{\mathcal{N}} + S_{\mathcal{K}} \geq L + K | S_{\mathcal{K}} = j) = \\
&= \sum_{j=0}^{K} \mathbb{P}(S_{\mathcal{K}} = j)\mathbb{P}(S_{\mathcal{N}} \geq L + K - j) = \\
&= \sum_{j=0}^{K} \mathbb{P}(S_{\mathcal{K}} = j)\left[\mathbb{P}(S_{\mathcal{N}} \geq L) - \mathbb{P}(L \leq S_{\mathcal{N}} \leq L + K - j - 1)\right] \leq \\
&\leq \mathbb{P}(S_{\mathcal{N}} \geq L) \underbrace{\sum_{j=0}^{K} \mathbb{P}(S_{\mathcal{K}} = j)}_{=1} = \mathbb{P}(S_{\mathcal{N}} \geq L) = P_{\text{out}}^{\mathcal{N}}
\end{aligned} \tag{7.26}$$

and the proposition follows.

7.8 Appendix-B: Proof of Theorem 7.2

Firstly, the initial conditions of the RF are trivially true. Secondly, from *the law/theorem of total probability*, the SOP $P_{\text{out}}^{\text{sys}}(L, N) = \mathbb{P}(S_{\mathcal{N}} \geq L)$ can be written as follows:

$$\begin{aligned}
P_{\text{out}}^{\text{sys}}(L, N) &= \sum_{j=0}^{1} \mathbb{P}(X_N = j)\mathbb{P}(S_{\mathcal{N}} \geq L | X_N = j) = \\
&= \sum_{j=0}^{1} \mathbb{P}(X_N = j)\mathbb{P}(S_{\mathcal{N}\setminus\{N\}} \geq L - j) = \\
&= \mathbb{P}(X_N = 0)\mathbb{P}(S_{\mathcal{N}\setminus\{N\}} \geq L) + \mathbb{P}(X_N = 1)\mathbb{P}(S_{\mathcal{N}\setminus\{N\}} \geq L - 1)
\end{aligned} \tag{7.27}$$

where $S_{\mathcal{N}\setminus\{N\}} = \sum_{n\in\mathcal{N}\setminus\{N\}} X_n = S_{\mathcal{N}} - X_N$. Due to the fact that $\mathbb{P}(X_N = 0) = 1 - p_N$ and $\mathbb{P}(X_N = 1) = p_N$, we get (7.14) and this completes the proof.

Bibliography

[1] A. D. Panagopoulos *et al.*, "Satellite communications at Ku, Ka, and V bands: Propagation impairments and mitigation techniques," *IEEE Commun. Surv. & Tutor.*, vol. 6, no. 3, pp. 2-14, Third Quarter 2004.

[2] A. D. Panagopoulos *et al.*, "Long-term rain attenuation probability and site diversity gain prediction formulas," *IEEE Trans. Antennas Propag.*, vol. 53, no. 7, pp. 2307-2313, July 2005.

[3] C. I. Kourogiorgas *et al.*, "On the earth-space site diversity modeling: A novel physical-mathematical outage prediction model," *IEEE Trans. Antennas Propag.*, vol. 60, no. 9, pp. 4391-4397, Sept. 2012.

[4] G. A. Karagiannis *et al.*, "Multidimensional rain attenuation stochastic dynamic modeling: Application to earth-space diversity systems," *IEEE Trans. Antennas Propag.*, vol. 60, no. 11, pp. 5400-5411, Nov. 2012.

[5] A. D. Panagopoulos, "Propagation phenomena and modeling for fixed satellite systems: Evaluation of fade mitigation techniques," in *Radio wave propagation and channel modeling for earth-space systems*, Taylor & Francis Group, CRC Press, 2016.

[6] N. Jeannin *et al.*, "Smart gateways for terabit/s satellite," *Int. J. Satell. Commun. Netw.*, vol. 32, no. 2, pp. 93-106, March 2014.

[7] A. Kyrgiazos, B. G. Evans and P. Thompson, "On the gateway diversity for high throughput broadband satellite systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5411-5426, Oct. 2014.

[8] A. Gharanjik *et al.*, "Multiple gateway transmit diversity in Q/V band feeder links," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 916-926, March 2015.

[9] T. Rossi, M. De Sanctis and F. Maggio, "Evaluation of outage probability for satellite systems exploiting smart gateway configurations," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1541-1544, July 2017.

[10] T. Delamotte and A. Knopp, "Smart diversity through MIMO satellite Q/V-band feeder links," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 285-300, Feb. 2020.

[11] M. Muhammad, G. Giambene and T. de Cola, "QoS support in SGD-based high throughput satellite networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8477-8491, Dec. 2016.

[12] C. N. Efrem and A. D. Panagopoulos, "Dynamic energy-efficient power allocation in multibeam satellite systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 228-231, Feb. 2020.

[13] A. Z. Papafragkakis *et al.*, "ALPHASAT site diversity experiments in Greece and the UK at Ka band: Comparison of 2-years' results," *ITU Journal: ICT Discoveries*, vol. 2, no. 1, Nov. 2019.

[14] T. Rossi *et al.*, "Smart gateway diversity optimization for EHF satellite networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 130-141, Feb. 2020.

[15] C. N. Efrem and A. D. Panagopoulos, "Globally optimal selection of ground stations in satellite systems with site diversity," *IEEE Wireless Commun. Lett.*, vol. 9, no. 7, pp. 1101-1104, July 2020.

[16] J. Pérez-Trufero *et al.*, "High throughput satellite system with Q/V-band gateways and its integration with terrestrial broadband communication networks," in *Proc. 32nd AIAA International Communications Satellite Systems Conference*, Aug. 2014, pp. 1-10.

[17] M. Fernandez and S. Williams, "Closed-form expression for the Poisson-binomial probability density function," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 2, pp. 803-817, April 2010.

[18] Y. Hong, "On computing the distribution function for the Poisson binomial distribution," *Computational Statistics & Data Analysis*, vol. 59, pp. 41-51, March 2013.

[19] A. M. Rushdi, "Utilization of symmetric switching functions in the computation of k-out-of-n system reliability." *Microelectron. Reliab.*, vol. 26, no. 5, pp. 973-987, 1986.

[20] L. A. Belfore, "An $O(n(\log_2 n)^2)$ algorithm for computing the reliability of k-out-of-n:G and k-to-l-out-of-n:G systems," *IEEE Trans. Reliab.*, vol. 44, no. 1, pp. 132-136, March 1995.

[21] P. Deheuvels, M. L. Puri, and S. S. Ralescu, "Asymptotic expansions for sums of nonidentically distributed Bernoulli random variables," *Journal of Multivariate Analysis*, vol. 28, no. 2, pp. 282-303, 1989.

[22] V. G. Mikhailov, "On a refinement of the central limit theorem for sums of independent random indicators," *Theory Probab. Appl.*, vol. 38, no. 3, pp. 479-489, 1994.

[23] A. Yu. Volkova, "A refinement of the central limit theorem for sums of independent random indicators," *Theory Probab. Appl.*, vol. 40, no. 4, pp. 791-794, 1996.

[24] W. Ehm, "Binomial approximation to the Poisson binomial distribution," *Statist. & Probab. Lett.*, vol. 11, no. 1, pp. 7-16, Jan. 1991.

[25] L. Le Cam, “An approximation theorem for the Poisson binomial distribution,” *Pacific J. of Math.*, vol. 10, no. 4, pp. 1181-1197, 1960.

[26] J. M. Steele, “Le Cam’s inequality and Poisson approximations,” *The American Mathematical Monthly*, vol. 101, no. 1, pp. 48-54, 1994.

[27] T. Hagerup, and C. Rüb, “A guided tour of Chernoff bounds,” *Information Processing Letters*, vol. 33, no. 6, pp. 305-308, Feb. 1990.

Chapter 8

General Conclusions and Open Problems

8.1 General Conclusions

In this Doctoral Thesis, setting *human* and *nature* as the key pillars, we have dealt with the environmentally-aware design of wireless networks, which definitely constitutes a new research direction. By taking into consideration environmental factors, the proposed EE-optimization algorithms aim to satisfy the traffic demand of users with the lowest energy consumption. In other words, the primary goal is to build *green communication networks* that provide high-quality services, while keeping the electromagnetic radiation at safety levels and reducing the carbon-dioxide (CO_2) emissions (*low carbon footprint*). Moreover, the operational expenditure (OPEX) of network service providers as well as the mass of satellites can be significantly reduced. In addition, the designed algorithms are able to prolong the battery lifetime of users' devices and can be used in applications with strict computation-time requirements (due to their *low complexity* and *fast convergence*). In any case, the contribution of this Dissertation is just *a small piece of the puzzle* and should be combined with further research in order to make this scientific challenge a reality.

Furthermore, we have showed that the optimum selection of GSs in RF/optical satellite networks with site diversity (under availability constraints) is an *NP-hard* problem. Also, we have developed *global optimization algorithms (B&B and DP)* as well as a *polynomial-time approximation algorithm* with provable performance guarantee. These algorithms might be useful in the initial network design, since they can provide significant cost savings in terms of the installation of GSs. Finally, we have studied in detail the performance of a *load-sharing SGD (LS-SGD)* architecture in satellite networks, which has been recently proposed in the literature. For this diversity scheme, several methods for the exact and approximate calculation of *system outage probability (SOP)* have been presented.

8.2 Open Problems

The algorithms presented in this Dissertation can be applied in several types of wireless networks as well as in other scientific fields. Subsequently, some interesting research directions stemming from this work are discussed.

- *Design of new EE-optimization algorithms:* The SCO method achieves a KKT solution (first-order optimality guarantee), which is a necessary (provided that some regularity conditions are satisfied) but not a sufficient condition for global optimality. As a result, it would be very useful to design *optimization algorithms with higher-order guarantees*, or even *global optimization algorithms* that can be used as a benchmark in order to evaluate the performance of suboptimal algorithms. Recently, the successive incumbent transcending (SIT) algorithm [1], the framework of mixed monotonic programming (MMP) which generalizes monotonic optimization [2], and the branch-and-bound (B&B) method [3] have been used to develop global optimization algorithms (with exponential complexity) for various EE-metrics in wireless networks. In summary: *Are there low-complexity algorithms with higher-order optimality guarantees or global optimization algorithms with faster convergence?*

- *Joint resource allocation for EE maximization:* Due to the fact that next-generation wireless networks require full exploitation of the available resources, an important research direction is the study of joint resource allocation problems (in this Thesis, we have only dealt with power control strategies). For example, transmit power could be optimized together with other resources, such as SC/time-slot allocation and BS/relay selection. In general, these *mixed-integer optimization problems* (i.e., with integer and continuous variables) are NP-hard and probably very difficult to solve.

- *Design of global optimization algorithms for GSs selection in SatNets, considering the spatial correlation between sites:* In this Dissertation, the weather conditions in the candidate locations are assumed independent. Nevertheless, the spatial correlation between sites is very important in practice, since it may have a significant impact on the network availability (especially when the GSs are relatively close to each other) [4]. Due to the fact that the optimization problem with independent weather conditions has been proven to be NP-hard, the general problem with spatially-correlated sites is *NP-hard* as well. Moreover, the existing methods that take into account the spatial correlation between sites are heuristic algorithms without performance guarantees [5–7]. Consequently, it would be very useful to develop global optimization algorithms for GSs selection in spatially-correlated SatNets.

- *Optimal selection of GWs in SGD-based SatNets:* In Chapter 7, we have studied the load-sharing SGD architecture, which has been recently proposed in the literature [8, 9] in order to provide very high availability and throughput with reasonable cost. An interesting research direction is *the selection of smart-GWs that minimize the total installation cost, satisfying given SOP-requirements.* This combinatorial problem might be solved using advanced algorithm design techniques and the approximation formulas of SOP presented in Chapter 7. Therefore, the following questions should be answered: 1) *What is the computational complexity of this problem?* and 2) *Are there global optimization algorithms or approximation algorithms (i.e., with provable performance guarantees) that achieve remarkable trade-offs between performance and complexity?*

Bibliography

[1] B. Matthiesen and E. A. Jorswieck, “Efficient global optimal resource allocation in non-orthogonal interference networks,” *IEEE Transactions on Signal Processing*, vol. 67, no. 21, pp. 5612-5627, Nov. 2019.

[2] B. Matthiesen, C. Hellings, E. A. Jorswieck and W. Utschick, “Mixed monotonic programming for fast global optimization,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 2529-2544, 2020.

[3] B. Matthiesen, A. Zappone, K. Besser, E. A. Jorswieck and M. Debbah, “A globally optimal energy-efficient power control framework and its efficient implementation in wireless interference networks,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 3887-3902, 2020.

[4] N. K. Lyras, C. I. Kourogiorgas and A. D. Panagopoulos, “Cloud free line of sight prediction modeling for optical satellite communication networks,” *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1537-1540, July 2017.

[5] C. Fuchs and F. Moll, “Ground station network optimization for space-to-ground optical communication links,” *IEEE/OSA J. Optical Commun. and Network.*, vol. 7, no. 12, pp. 1148-1159, Dec. 2015.

[6] N. K. Lyras, C. N. Efrem, C. I. Kourogiorgas, and A. D. Panagopoulos, “Optimum monthly-based selection of ground stations for optical satellite networks,” *IEEE Communications Letters*, vol. 22, no. 6, pp. 1192-1195, June 2018.

[7] N. K. Lyras, C. N. Efrem, C. I. Kourogiorgas, A. D. Panagopoulos, and P.-D. Arapoglou, “Optimizing the ground network of optical MEO satellite communication systems,” *IEEE Systems Journal*, vol. 14, no. 3, pp. 3968-3976, Sept. 2020.

[8] T. Rossi, M. De Sanctis and F. Maggio, “Evaluation of outage probability for satellite systems exploiting smart gateway configurations,” *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1541-1544, July 2017.

[9] T. Rossi *et al.*, “Smart gateway diversity optimization for EHF satellite networks,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 130-141, Feb. 2020.

List of Publications

International Peer-Reviewed Journals

1. **C. N. Efrem** and A. D. Panagopoulos, "On the computation and approximation of outage probability in satellite networks with smart gateway diversity," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 1, pp. 476-484, Feb. 2021.

2. **C. N. Efrem** and A. D. Panagopoulos, "Minimizing the installation cost of ground stations in satellite networks: Complexity, dynamic programming and approximation algorithm," *IEEE Wireless Communications Letters*, vol. 10, no. 2, pp. 378-382, Feb. 2021.

3. N. K. Lyras, **C. N. Efrem**, C. I. Kourogiorgas, A. D. Panagopoulos, and P.-D. Arapoglou, "Optimizing the ground network of optical MEO satellite communication systems," *IEEE Systems Journal*, vol. 14, no. 3, pp. 3968-3976, Sept. 2020.

4. **C. N. Efrem** and A. D. Panagopoulos, "Globally optimal selection of ground stations in satellite systems with site diversity," *IEEE Wireless Communications Letters*, vol. 9, no. 7, pp. 1101-1104, July 2020.

5. **C. N. Efrem** and A. D. Panagopoulos, "Dynamic energy-efficient power allocation in multibeam satellite systems," *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 228-231, Feb. 2020.

6. N. K. Lyras, **C. N. Efrem**, C. I. Kourogiorgas, and A. D. Panagopoulos, "Medium earth orbit optical satellite communication networks: Ground terminals selection optimization based on the cloud-free line-of-sight statistics," *International Journal of Satellite Communications and Networking*, vol. 37, no. 4, pp. 370-384, July/August 2019.

7. **C. N. Efrem** and A. D. Panagopoulos, "Energy efficiency optimization: A new trade-off between fairness and total system performance," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 853-856, June 2019.

8. **C. N. Efrem** and A. D. Panagopoulos, “A framework for weighted-sum energy efficiency maximization in wireless networks,” *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 153-156, Feb. 2019.

9. N. K. Lyras, **C. N. Efrem**, C. I. Kourogiorgas, and A. D. Panagopoulos, “Optimum monthly-based selection of ground stations for optical satellite networks,” *IEEE Communications Letters*, vol. 22, no. 6, pp. 1192-1195, June 2018.

International Peer-Reviewed Conferences

1. **C. N. Efrem**, N. K. Lyras, C. I. Kourogiorgas, A. D. Panagopoulos, P.-D. Arapoglou, “Low-complexity optimization algorithm for ground network design in optical satellite networks,” *International Conference on Space Optics (ICSO) 2020*, virtual conference, 2021.

2. **C. N. Efrem**, N. K. Lyras, C. I. Kourogiorgas, A. D. Panagopoulos, P.-D. Arapoglou, “Cloud free LOS probability estimation for MEO optical satellite systems and optical satellite network dimensioning,” *14th European Conference on Antennas and Propagation (EuCAP 2020)*, Copenhagen, Denmark, 2020.

3. **C. N. Efrem** and A. D. Panagopoulos, “Joint power allocation and transmission time selection in duty-cycled wireless sensor networks,” *2018 Global Information Infrastructure and Networking Symposium (GIIS)*, Thessaloniki, Greece, 2018.

Book Chapters

1. **C. N. Efrem** and A. D. Panagopoulos, “MIMO in vehicular communication networks: Channel modeling, cooperative relaying, and resource allocation,” in *Vehicle-to-Vehicle and Vehicle-to-Infrastructure Communications: A Technical Approach*, Taylor & Francis Group, CRC Press, 2018.