université PARIS-SACLAY

CentraleSupélec

N° national de thèse : 2016SACLC011

THESE DE DOCTORAT DE L'UNIVERSITÉ PARIS-SACLAY PREPAREE A CENTRALESUPÉLEC

ECOLE DOCTORALE N°580
Sciences et technologies de l'information et de la communication (STIC)

Spécialité : Réseaux, Information et Communications

Par

M. Bhanukiran Perabathini

Limites fondamentales de l'efficacité énergétique dans les réseaux sans fils

(Fundamental limits of Energy Efficiency in Wireless Networks)

Thèse présentée et soutenue à Gif-sur-Yvette, le 18 Janvier 2016 :

Composition du jury :

Mari Kobayashi	CentraleSupélec	Director of thesis
Marios Kountouris	CentraleSupélec	Co-director of thesis
Samson Lasaulce	LSS	Examiner
Mérouane Debbah	CentraleSupélec	Examiner
Luca Sanguinetti	University of Pisa	Examiner
Aris Moustakas	University of Athens	Examiner
Jean-Claude Belfiore	Telecom ParisTech	Reviewer
Jean-Marie Gorce	INSA-Lyon	Reviewer
Alberto Conte	Alcatel-Lucent Bell labs	Invited

université PARIS-SACLAY

ÉCOLE DOCTORALE

Sciences et technologies de l'information et de la communication (STIC)

Titre : Limites fondamentales de l'efficacité énergétique dans les réseaux sans fils

Mots clés : efficacité énergétique, thermodynamique, géométrie stochastique, réseaux cellulaires, 5G

Résumé : La tâche de répondre à une demande croissante pour une meilleure qualité de l'expérience utilisateur dans les communications sans fil, est contestée par la quantité d'énergie consommée par les technologies concernées et les méthodes employées. Sans surprise, le problème de la réduction de la consommation d'énergie doit être abordé à diverses couches de l'architecture de réseau et de diverses directions. Cette thèse traite de certains aspects cruciaux de la couche physique de l'architecture de réseau sans fil afin de trouver des solutions efficaces d'énergie.
Dans la première partie de cette thèse, nous explorons l'idée de l'efficacité énergétique à un niveau fondamental. À commencer par répondre aux questions telles que: - Qu'est-ce que la forme physique d'information ?, nous construisons un dispositif de communication simple afin d'isoler certaines étapes clés dans le processus physique de la communication et nous dire comment elles affectent l'efficacité énergétique d'une communication système.
Dans la deuxième partie, nous utilisons des outils de la géométrie stochastique pour modéliser théoriquement réseaux cellulaires afin d'analyser l'efficacité énergétique du système. L'exploitation de la traçabilité d'une telle modélisation mathématique, nous explorons les conditions dans lesquelles la consommation d'énergie peut être réduite. En outre, dans cette partie, nous introduisons le concept de la mise en cache des données des utilisateurs à la périphérie du réseau (à savoir le final ac BS qui est en contact avec l'utilisateur) et de montrer quantitativement comment la mise en cache peut aider à améliorer l'efficacité énergétique d'un cellulaire réseau.
Nous tenons également à ce traitement à un ac Hetnet scénario (à savoir quand il ya plus d'un type de glspl déployé BS) et étudions divers indicateurs de performance clés. Nous explorons également les conditions où l'efficacité énergétique d'un tel système peut être amélioré.
Les résultats de thèse fournissent quelques idées clés pour améliorer l'efficacité énergétique dans un réseau cellulaire sans fil contribuant ainsi à l'avancement vers la prochaine génération (5 G) les réseaux cellulaires.

université PARIS-SACLAY

ÉCOLE DOCTORALE

Sciences et technologies de l'information et de la communication (STIC)

Title: Fundamental limits of Energy Efficiency in Wireless Networks

Keywords: energy efficiency, communication systems, thermodynamics, stochastic geometry, cellular networks, 5G

Abstract: The task of meeting an ever growing demand for better quality of user experience in wireless communications is challenged by the amount of energy consumed by the technologies involved and the methods employed. Not surprisingly, the problem of reducing energy consumption needs to be addressed at various layers of the network architecture as well as from various perspectives. This thesis addresses some important energy saving possibilities at the physical layer of wireless network architecture in order to find energy efficient solutions.

In the first part of this thesis, we explore the idea of energy efficiency at a fundamental level. Starting with answering questions such as *'What is the physical form of information?'*, we build a simple communication device in order to isolate the key steps in the physical process of communication and we study how each of these steps finally influences the energy efficiency of the communication system. In the second part, we use tools from stochastic geometry to theoretically model a cellular network so as to analyze its energy consumption. Exploiting the tractability of such a mathematical modeling, we explore the conditions under which the energy consumption can be either reduced or utilized more efficiently or both.

Further, we introduce the concept of caching users' data at the edge of the network (namely the final Base Station (BS) that is in contact with the end user) and show quantitatively how incorporating caching in the network can help improve the energy efficiency. We also extend this treatment to a Heterogeneous Network (HetNet) scenario (namely when there are more than one type of BSs deployed) and study various key performance metrics. We also explore the conditions where energy efficiency of such a system can be improved.

The results in thesis provide some key ideas regarding improving energy efficiency in a wireless cellular network thereby contributing to the advancement towards the next generation (5G) cellular networks.

.. to all those curious minds who defiantly refused to be humbled by the sheer magnitude of the unknown.

Acknowledgments

My sincere thanks to my PhD advisors **Prof Mérouane Debbah** and **Prof Marios Kountouris** for their guidance throughout the period of my PhD. Also I would like to thank **Prof Mari Kobayashi** for her timely intervention and accepting to be my official PhD advisor during my defense. **Prof Debbah** will remain forever in my mind as a symbolism for energy and dedication! I would also like to thank **Mr Alberto Conte**, my industrial supervisor at Alcatel Lucent Bell-labs for his guidance and support throughout my PhD. It is also a pleasure to thank my colleague **Dr Chung Shue Chen (Calvin)** (along with the other colleagues at Bell labs) for the valuable discussion sessions we had at the beginning of my PhD.

I would also like to thank my group members **Ejder Bastug**, **Marco Maso**, **Francesca Garavello**, **Loïg Godard**, **Apostolos Destounis**, **Subhash Lakshminarayana**, **Luca Rose**, **Axel Müller**, **Matthieu de Mari**, **Sylvain Azarian**, **Raul de Lacerda**, **Romain Couillet**, **Jakob Hoydis**, **Abla Kammoun**, **Emil Björnson**, **Franck Iutzeler**, **Anthony Mays**, **Stefano Boldrini**, **Nikolaos Pappas**, **Yacine Hebbal**, **Adrien Pelletier**, **Harry Sevi**, **Stefan Mijovic**, **Kenza Hamidouche**, **Azary Abboud**, **Gil Katz**, **Matha Deghel**, **Evgeny Kusmenko**, **Hafiz Tiomoko Ali**, **Fei Shen**, **Luca Sanguinetti**, **Zheng Chen**, **Salah Eddine Hajri**, **Bakarime Diomande**, **Tanumay Datta**, **Meysam Sadeghi**, and **Apostolos Karadimitrakis** with who I have gathered numerous memories to live with.

I would also like to thank my school physics teacher **Mr Uppuluri Kameshwarao** who was instrumental in nurturing confidence in me to pursue my love for physics.

I would also like to thank my friends **Vineeth S Varma**, **Achal Agarwal**, **Pramod Padmanabhan**, **Sudheer Kolla**, **Solomon Babu**, **Arun Kumar Raju** and **'others'** for helping me during hard times of my life. I would also like to thank my wife **Desham** for being there to bank on and for creating a purpose for my life. Finally, I like to thank my parents for this life which I cherish with pleasure, pain and curiosity!

Contents

Acronyms

APC	Area Power Consumption. xii, 5, 6, 44–49, 56, 58, 59, 61–63, 65, 69–71, 80, 81
ASE	Area Spectral Efficiency. 47, 57
BPSK	Binary Phase-Shift Keying. 6, 15, 19–22, 24
BS	Base Station. ii, vi, xi, xii, 3–6, 27–37, 39–49, 51, 53, 54, 56, 57, 63
CN	Core Network. 5
D2D	Device-To-Device. 39
EE	Energy Efficiency. xii, xiii, 5, 6, 47–50, 57–59, 61–63, 66, 67, 75, 80, 81
HetNet	Heterogeneous Network. ii
i.i.d.	Independent and Identically Distributed. 52
MBS	Macro Base Station. 52, 53, 55–58, 62, 63, 69, 80
MIMO	Multiple-Input Multiple-Output. 40
PDF	Probability Distribution Function. 42, 53
PPP	Poisson Point Process. 5, 40, 43, 51, 52, 54, 55, 65
QoS	Quality-of-Service. 5, 6, 40, 41, 44, 47, 57
QPSK	Quadrature Phase-Shift Keying. 6, 20–22, 24
RAN	Radio Access Network. 5
SBS	Small Base Station. xii, xiii, 51–69, 71, 73–75, 80
SINR	Signal-To-Interference-Plus-Noise Ratio. 5, 29, 34, 40–42, 52, 55, 56, 63, 64, 68, 69
UT	User Terminal. 39, 40
ZF	Zero Forcing. 57

List of Figures

Chapter 1

Resumé (French)

1.1 Contexte et Motivation

Alimentée par l'omniprésence de nouveaux appareils sans fil et les téléphones intelligents, ainsi que la prolifération des applications gourmandes en bande passante, dernière décennie a vu une augmentation spectaculaire de l'utilisation et de la demande pour une connectivité mobile. La demande transparente pour les données sans fil par de riches applications et services multimédias souvent associée avec le streaming vidéo et les réseaux sociaux sur les smartphones et autres appareils informatiques connectés sans fil est prévu seulement pour augmenter plus loin et plus rapidement [2] dans les prochaines années. Un enjeu crucial pour les réseaux de prochaine génération est d'être en mesure de soutenir l'augmentation prévue du trafic d'une manière durable et économiquement viable. En plus d'assurer la connectivité aux utilisateurs, le suivi des dépenses d'énergie (ainsi l'empreinte carbone résultant) est également une grande importance à la fois des points économiques et environnementaux de vue. La dépense d'énergie est soit minimisée ou être utilisé aussi efficacement que possible. En dépit de la ressemblance, les deux problèmes ne sont pas les mêmes. Pour expliquer cela en termes généraux: Imaginez qu'une voiture dispose d'un réservoir de carburant de taille fixe. Le moteur peut être fabriqué soit de telle sorte qu'elle se déplace sur une longue distance, mais à une vitesse de compromis ou de telle sorte qu'elle passe à une vitesse impressionnante tandis que la distance qu'il peut parcourir est plus courte. De toute évidence, il existe un compromis entre les deux cas et, au lieu d'être simple, définissant EE dépend de l'utilité spécifique, nous cherchons d'un système. Dans le contexte des télécommunications sans fil, EE est souvent définie dans la littérature comme étant le rapport entre le débit total et la puissance totale consommée dans l'émetteur. Maximiser la EE ne garantit que l'énergie est utilement dépensé, mais il est toujours important de garder à l'esprit que la réalisation de ce maximum ne doit pas dépasser nos limites budgétaires.

Il y a eu des études portant sur ce sujet à partir de directions différentes. Dans [3], optimisation de EE d'un Multiple-Input Multiple-Output (MIMO) pointer-à-point est étudié dans les cas où imparfaite Channel State Information (CSI) est disponible uniquement à l'émetteur et est disponible uniquement à la fin récepteur. Dans [4], EE est optimisé par rapport au nombre d'antennes système MIMO, où un modèle de consommation d'énergie réaliste à chaque BS a été pris en compte. Dans [5], la densité de points d'accès (AP), le nombre d'antennes et d'utilisateurs par AP, et la puissance de transmission ont été optimisés pour maximal EE. Travaux de Li Chang et al, (voir [6], par exemple) étude EE aspects de petits réseaux cellulaires en termes de BS densités et des coûts d'énergie au BSs. Le consortium GreenTouch récemment [1] a publié un livre blanc [7] sur l'étude vert recherche compteur qui prétend qu'il est possible grâce à une combinaison de changements dans les technologies, les architectures, des composants, des algorithmes et des protocoles à réduire la consommation d'énergie nette dans les réseaux

[1] 18 Juin, 2015

de bout en bout des communications jusqu'à 98 % en 2020 par rapport au scénario de référence 2010 défini par GreenTouch. Selon cette augmentation de 10.000 fois de l'efficacité énergétique dans les réseaux d'accès mobiles est possible en intégrant l'évolution des technologies, des architectures, des composants, des algorithmes et des protocoles existants. En ce qui concerne les réseaux d'accès mobiles, parmi la liste proposée des idées sont:

- Au-delà de Cellular Generation Green (BCG^2): les données de fractionnement et de signalisation entre deux couches superposées et incorporant a été montré à la demande de connectivité pour être une solution efficace de l'énergie.
- Grande échelle Systèmes d'antennes (LSAS): Au lieu de servir des terminaux mobiles au moyen d'antennes encombrants qui rayonnent de puissance dans les secteurs, cette technique essaie de diriger le pouvoir à l'utilisateur où il est nécessaire. Ceci est réalisé en utilisant un grand nombre de personnes physiquement petit, de faible puissance, des antennes à commande individuelle pour créer une pluralité de faisceaux sélectifs par l'utilisateur de données. La puissance porteur de données est donc dirigée uniquement où elle est nécessaire et évite les interférences inutiles.

Bien que ces résultats semblent prometteurs, une analyse plus approfondie doit encore être fait pour élargir le nombre de possibilités d'atteindre les technologies de réseau sans fil de la prochaine génération efficace de l'énergie. Comme il ne peut être un remède universellement valable et viable à intensifier vers les technologies vertes, cette recherche est cruciale. La présente thèse vient dans le cadre de l'exploration diverses possibilités d'accroître l'efficacité énergétique dans les réseaux sans fil.

1.2 Les aspects physiques de la transmission sans fil des informations

D'un point de vue fondamental, il est connu depuis longtemps que le processus de *communication* a toujours un coût en termes d'énergie [8–10]. En 1948, Shannon a publié son révolutionnaire papier [11] dans lequel les conditions dans lesquelles l'information peut être communiquée de manière fiable sont mathématiquement quantifiés. Il est d'un grand intérêt pour aborder ce point précis à nouveau, mais du point de vue de la dépense énergétique prenant création et le traitement des informations en compte.

Dans son sens courant, l'information est un concept abstrait et est souvent considérée comme subjective. Cependant, dans ses informations de sens objectif signifie précisément un ensemble de réseau convenable de systèmes physiques préparés dans les états souhaités. La nature physique de l'information est traitée dans [12]. Pour donner quelques exemples:

- Un message de texte écrit sur un morceau de papier avec un crayon, mais est une parmi plusieurs agencements possibles d'un ensemble de molécules de carbone dans le plan bidimensionnel d'un document.
- Un fichier audio inscrit sur un disque compact est essentiellement une collection de fosses et des bosses gravées sur elle dans un ordre unique.
- Braille tracée sur un papier, des signaux tels que des ronds de fumée, tambours, sirènes, etc. sont toutes les entités physiques quantifiables et elles impliquent toutes l'utilisation de l'énergie lors de leur création, la destruction propagation de amd.

Si un système physique choisi a deux états distincts de l'information représentée par un tel système est appelé un *bit*. Il a été démontré qu'une grande quantité d'information peut être représentée en termes de nombre suffisant de ces bits de [11]. Préparation d'un système physique dans un état, ou changement de son état de l'un à l'autre nécessite de l'énergie. Bien qu'il puisse y avoir d'innombrables possibilités

pour les systèmes physiques qui peuvent être utilisés dans la préparation, la modification et le stockage de l'information donnée, le choix du système ne décide combien d'énergie est nécessaire pour effectuer ces opérations. Le *première* question que nous posons dans cette partie de la thèse porte sur ce point. Dans une tentative d'explorer la réponse à la question: *'Quelle est l'énergie choix efficace la plupart des systèmes physiques pour représenter et traiter l'information?*, nous commençons par un système thermodynamique simples (section 3.2.1) pour représenter un *bit d'information.* Dans le domaine de la physique classique est vraiment le meilleur moyen de stocker des bits du point de vue énergétique. Dans le chapitre 3, nous discutons en détail comment l'énergie est dépensée dans la préparation et de traitement (opérations sur) ces bits.

Le processus dans lequel les systèmes physiques sont préparés dans les états souhaités sont transportés de la *transmission* pointer vers le *réception* point dans l'espace et le temps à travers un milieu est appelé *transfert d'informations.* Ceci est un processus physique et donc nécessite de l'énergie. La corrélation entre le réseau d'états envoyé et le tableau des états reçu est ce qui détermine la qualité du transfert de l'information. Cette corrélation peut être déformée à cause d'interférences du monde physique (bruit), et une telle distorsion est quantifiée comme *erreur.* Il existe un compromis entre la quantité d'énergie dépensée au point de transmission et l'information mutuelle (entre les états reçus et les états transmis) qui peuvent être atteints. Une autre question que nous posons dans la première partie de cette thèse porte sur ce point. Dans une tentative d'explorer la réponse à la question: *" Quelle est la limite de l'efficacité d'un système de communication prenant la création de l'information en compte "?*, nous concevons un système pour convertir les informations représengtées par la thermodynamique mentionné ci-dessus . bits en ondes électromagnétiques pour atteindre la réception, et d'évaluer l'efficacité énergétique du système Nous sommes particulièrement intéressés à évaluer l'efficacité énergétique dans la communication un bit d'information comme un cas limite (section 3.4.4). Contrairement à la notion intuitive que l'efficacité peut être arbitrairement augmentée en augmentant l'investissement énergétique, nos résultats montrent que les dépenses excès d'énergie ne doit pas nécessairement être une solution efficace en communication point-à-point. Cette approche est importante non seulement pour le sans fil les fournisseurs de services, mais aussi du point de vue des utilisateurs d'appareils sans fil qui sont souvent compromises par l'ennui de charge fréquemment les batteries des appareils mobiles. En fait, cette partie du travail vient en lignes d'explorer les aspects fondamentaux du calcul ainsi que la communication.

1.3 Analyse architecturale la dépense énergétique

Stochastique modélisation basée sur la géométrie

La demande croissante de connectivité, débit de données, et la qualité de service (QoS) ne peut pas être satisfaite par la simple augmentation indéfiniment la puissance d'émission des stations de base BSs. Ceci est principalement dû au fait qu'une augmentation de la puissance d'émission, en plus d'augmenter la puissance du signal de la station de base souhaité, augmente également les interférences reçues par les stations de base de non-desserte. Cela peut effectivement diminuer le rapport signal-sur-interférence-plus-bruit (SINR) subie par le terminal utilisateur, ayant ainsi un impact négatif sur la Quality-of-Service (QoS). En outre, il est également essentiel du point de vue de l'efficacité énergétique pour résoudre le problème de minimiser les dépenses d'énergie tout en conservant certaines contraintes telles que la couverture de la cible et le débit de données minimum.

Il existe différentes approches pour résoudre les problèmes de réduction de la consommation d'énergie théoriquement. Par exemple, si un nombre fini de BSs sont déployées dans un hexagone régulier (ou grille) mode cellulaire comme dans la figure 4.2 [13], on pourrait chercher à minimiser la puissance totale consommée par l'obtention des paramètres de fonctionnement optimaux, tels que la taille de cellule hexagonale et l'amplitude de la puissance d'émission au niveau de chaque BS, tout en garantissant une certaine QoS. Néanmoins, cette approche implique une analyse lourde que l'évaluation de la répartition spatiale de la SINR dans un modèle basé sur une grille devient prohibitif complexe que la taille du système augmente

et on peut avoir recours à des simulations. Une simplification commune dans la modélisation de réseaux cellulaires qui nous permet traitons le problème analytiquement est de supposer que les emplacements des stations de base sont dispersés de façon aléatoire sur une surface à deux dimensions en plan selon un homogène PPP [14]. Plusieurs travaux ont étudié la validité de PPP modélisation des stations de base (par rapport à des modèles cellulaires réguliers,) et il est souvent illustrées afin de fournir des indications utiles sur le comportement statistique des indicateurs de performance clés [15]. Certains chercheurs, en particulier ceux avec un arrière-plan de l'industrie, expriment souvent le scepticisme sur l'utilisation de la géométrie stochastique pour l'analyse des performances des réseaux cellulaires sans fil. Leur accusation principale est que, dans un traitement à base de la modélisation géométrique stochastique néglige le déploiement réel de BSs et se concentre plutôt sur la théorie BS configurations. Il faut souligner à ce stade que le but de l'utilisation de la géométrie stochastique est pas de résoudre toute configuration réelle de la vie réelle déployée d'une collection de BSs, mais d'acquérir qualitative ainsi que des aperçus quantitatifs sur certains paramètres du système cruciales avec l'aide du formalisme mathématique traitable plutôt que selon les simulations spécifiques du système. Cela a été assez souligné dans les travaux de Andrews et al. Des résultats plus récents sur l'analyse stochastique des réseaux cellulaires sans fil peuvent être trouvés dans [16–20] et références qui y sont.

Est caching une solution efficace de l'énergie?

Les deux universités et l'industrie sont maintenant dans une envie de faire évoluer les réseaux cellulaires traditionnels vers la prochaine génération de réseaux mobiles à large bande, les réseaux forgés 5G, visant à satisfaire le tsunami de données mobiles tout en minimisant les dépenses et la consommation d'énergie. Parmi ces efforts intensifs, le contenu des utilisateurs de mise en cache localement sur le bord du réseau est considéré comme l'un des paradigmes les plus perturbateurs dans les réseaux 5G [21].

Il est intéressant mais pas surprenant, des résultats récents ont montré que distribué le cache de contenu peut décharger considérablement différentes parties du réseau, comme dans Radio Access Networks (RANs) et Core Network (CN), par intelligemment préchargement et le stockage de contenu plus proche des utilisateurs finaux [22]. En effet, les réseaux cellulaires traditionnels, qui ont été conçus pour les appareils mobiles avec des capacités de traitement et de stockage limitées, ont commencé à intégrer les capacités sensibles au contexte et proactives, alimentées par les récents progrès dans la puissance et de stockage de traitement. En conséquence, la mise en cache a récemment pris la littérature 5G par la tempête.

Caching dans les réseaux cellulaires de niveau unique

Comme mentionné précédemment, les aspects de la consommation d'énergie des réseaux cellulaires sont habituellement étudiées en plaçant BSs sur une topologie hexagonale ou grille régulière et la conduite intensive des simulations de niveau système [23]. Là, le but est de trouver les valeurs optimales des paramètres tels que la plage de cellules et BS transmettent la puissance, tout en minimisant la puissance totale consommée dans certaines QoS contraintes. En dépit de son attrait, il est assez lourde, voire impossible, d'évaluer analytiquement les indicateurs de performance clés dans les réseaux à grande échelle. Une simplification commune dans la modélisation de réseaux cellulaires est de placer le BSs sur un plan à deux dimensions selon un espace homogène PPP, ce qui permet de traiter le problème analytiquement [24]. Plusieurs travaux ont étudié la validité de PPP modélisation de BSs par rapport à des modèles cellulaires réguliers (par exemple [25]) et des idées supplémentaires peuvent être obtenus en caractérisant analytiquement les indicateurs de performance, tels que la distribution spatiale des Signal-To-Interference-Plus-Noise Ratio (SINR), la probabilité de couverture, et le taux moyen [18, 19]. Plus récemment, l'optimisation de la EE a été abordée assez bien dans les œuvres d'Emil et al, comme [26].

Dans cette thèse, nous analysons (voir chapitre 5) aspects de la consommation d'énergie des déploiements sans fil réseau à palier unique cache activé au moyen d'un modèle spatial basé sur la géométrie

stochastique. Nous considérons le problème de l'optimisation APC et EE de cache-enabled BSs dans un scénario où BSs et les utilisateurs mobiles sont répartis selon indépendante homogène PPPs. Nous fournissons des conditions dans lesquelles la consommation d'énergie de la zone est de réduire au minimum l'objet d'une certaine qualité de service (QoS) en termes de probabilité de couverture. Par ailleurs, nous apportons la puissance d'émission BS optimale qui maximise le rendement énergétique, défini comme le rapport de la surface au-dessus de l'efficacité spectrale de la puissance totale consommée. Le principal résultat de cette analyse est que la mise en cache distribuée de contenu dans les réseaux cellulaires sans fil se révèle être une solution efficace de l'énergie.

Caching dans des réseaux hétérogènes

Réseaux cellulaires sans fil de l'avenir sont susceptibles d'être hétérogènes, à savoir avoir un ou plusieurs niveaux de petites cellules superposées sur le niveau macrocellulaire. Des œuvres telles que [27] étude EE dans un HetNet scénario avec des stratégies de couchage incorporés dans. Motivé pour progresser dans cette ligne de recherche, nous étendons l'idée de la mise en cache du contenu des utilisateurs dans un (ou plusieurs) des niveaux dans un HetNet et d'étudier comment cela influe sur les aspects liés à la consommation d'énergie du système sans fil. Nous définissons et résoudre des problèmes d'optimisation APC et EE par rapport à la densité de stations de base et de leurs puissances d'émission. Dans le chapitre 6, nous analysons la consommation d'énergie aspects du réseau sans fil à deux vitesses, où les petites cellules sont cache-enabled.

1.4 Conclusions

Cette thèse a présenté des solutions au problème de l'amélioration de l'efficacité énergétique dans les réseaux sans fil. Il est clair que l'éventail des approches possibles à ce problème est énorme et cette thèse a porté principalement sur les principes fondamentaux de la physique dans la première partie et sur les questions de conception au niveau du système dans la deuxième partie.

1.5 limites fondamentales

Dans la première partie de cette thèse, nous avons conçu une expérience de pensée comme système de communication et calculé la limite supérieure de son efficacité énergétique en utilisant la thermodynamique et de la théorie électromagnétique. Avec l'aide du modèle proposé, nous avons vu que pour un montant total donné de l'énergie qui est disponible pour préparer, transmettre, et de réécrire des bits d'information, il existe une transmission de fréquence d'antenne unique qui donne une efficacité énergétique maximale. Nous avons également constaté la dépendance de l'efficacité énergétique de la bande passante et l'amplitude du signal, et qu'il a un optimum à des valeurs convenables pour les deux. Nous avons fait une comparaison de ces dépendances dans les cas de modèle de décalage de clé de phase binaire et le modèle de décalage de clé de phase en quadrature. Nos résultats montrent que, contrairement aux résultats de [28], lorsque le système de communication fonctionne à une température non nulle, la puissance optimale consommée par le système afin d'optimiser l'efficacité énergétique est non nul.

1.5.1 Orientations futures

Le traitement peut être étendu et appliqué à différents choix de systèmes physiques pour stocker et transporter l'information. Notre choix de boîtes de Szilard thermodynamiques et des ondes électromagnétiques est dans l'intérêt d'explorer les limites dans le domaine de la physique classique. Cependant,

lorsque l'on considère les systèmes d'énergie très faible, les effets de la mécanique quantique pourrait être plus prononcé et cela nécessite une étude plus approfondie. À l'avenir, il peut être intéressant d'explorer d'autres systèmes physiques fondamentaux pour représenter bits. Par exemple: états de spin quantique de particules, quantiques systèmes thermodynamiques décrits mécaniquement [29]. En outre, il est très important d'étudier les limites de l'efficacité énergétique d'un dispositif de calcul de cette ligne d'approche. Des études telles que [30] où les limites sur les coûts d'énergie de calcul sont explorées sont d'une grande importance dans le contexte actuel où la batterie untethered entraîné l'informatique mobile est omniprésent. Ces demandes de renseignements sur EE limites de calcul doivent nous conduire acquérir un aperçu des limites même à la couche MAC d'un système de communication.

1.6 Stochastic modélisation basée sur la géométrie

Dans le chapitre 4, nous avons abordé l'optimisation des APC dans les réseaux cellulaires sans fil à palier unique dans lequel les emplacements de BSs sont modélisés ont été distribués selon un homogène PPP. Sous forte BS politique de l'association, nous avons tiré des bornes sur la puissance d'émission optimale afin de garantir un certain taux de couverture et les données minimum. Sous le même QoS contraintes, nous avons dérivé la forme d'expression fermée pour optimale BS densité qui minimise la APC. A partir de ces résultats, on constate que l'existence d'un tel optimale BS densité dépend de la valeur de l'exposant d'affaiblissement de propagation.

1.6.1 Orientations futures

Dans [31], les auteurs ont exploré la convexité de la fonction couverture de probabilité pour le modèle général de pathloss unbounded. Cette analyse pourrait être reporté vers l'optimisation de la APC et de telles mesures.

Modélisation BS déploiements selon un uniforme PPP est une généralisation peu idéaliste. Dans des scénarios réels, BSs peut être vu déployé densément à un seul endroit et relativement rarement dans d'autres endroits. Prendre ce genre de regroupement des densités variées à des emplacements spatiaux variés et la réalisation des EE optimisations est un problème intéressant. Backhaul (longueur de la fibre optique etc) coûts peuvent également être incorporés dans le modèle du système et étudié comment cela va affecter les paramètres système optimaux tels que BS densité, etc.

1.7 Caching dans les réseaux cellulaires unique de niveau

Dans le chapitre 5, nous avons étudié comment intégrant des capacités de mise en cache à la BSs influe sur la consommation d'énergie dans les réseaux cellulaires sans fil. L'adoption d'un détail BS modèle de puissance et la modélisation de la BS emplacements selon un PPP, nous avons calculé les expressions pour le APC et EE, qui sont en outre simplifiée dans le régime à faible bruit. Une observation clé de ce travail est que cache-enabled BSs peut diminuer de manière significative le APC et d'améliorer la EE par rapport à traditionnelle BSs. Nous avons également observé que l'existence d'un point de consommation d'énergie optimale pour le APC dépend de l'exposant de pathloss.

1.7.1 Orientations futures

Les aspects de l'énergie et les implications de la mise en cache dans les réseaux cellulaires sans fil, en particulier pour 5G systemes, sont d'un intérêt pratique et en temps opportun et clairement besoin d'une enquête plus approfondie. Les travaux futurs peuvent inclure des scénarios de réseau hétérogènes, y compris les petites cellules, les cellules macro et accès WiFi points de déploiement. En outre, le stockage du contenu populaire exige une estimation précise de la distribution de la popularité du contenu, qui ne peut être réalisée facilement dans la pratique et peut coûter l'énergie en termes de puissance de traitement. Par conséquent, plutôt que de compter sur cette approche, les politiques de mise en cache randomisées dans un scénario stochastique [32] peut être considéré comme un moyen de fournir des indications nettes sur les avantages de l'efficacité énergétique de la mise en cache dans les réseaux sans fil denses. En outre, la distribution de popularité peut être considérée comme une fonction de localisation spatiale comme il est raisonnable de supposer que ce qui est populaire à un endroit peut ne pas être populaire ailleurs. À la lumière de cela, il est logique de prendre l'ensemble du processus de mise en cache à un détail afin de faire un commentaire rigoureuse si la mise en cache est une solution efficace de l'énergie ou non. [33], enquête sur le problème de la mise en place de cache optimale au bord sans fil pour réduire au minimum le taux de transmission dans des réseaux hétérogènes.

1.8 Caching dans les réseaux cellulaires à deux niveaux

Dans le chapitre 6 nous avons proposé un système à deux niveaux hétérogènes, un Macro Base Stations (MBSs) (sans mise en cache des capacités) et un SBSs (avec des capacités de mise en cache), nous avons défini plusieurs paramètres importants du système et provenant de leurs expressions. Nous choisissons APC et EE comme deux mesures raisonnables pour optimiser afin de minimiser la consommation d'énergie dans un réseau. Prenant λ_x et P_x (SBS densité et transmettre la puissance respectivement) soient les variables, nous avons eu quatre problèmes d'optimisation indépendants à effectuer. Nous avons imposé les contraintes C1-5 et avons découvert que APC est une fonction convexe dans λ_x et P_x. Étant donné que les équations impliquées sont compliquées à trouver les expressions analytiques exactes pour les valeurs optimales, nous découvrons les intervalles au sein de laquelle les optima peut être trouvée. Pour une valeur de système donné, cependant, un optimum peut être trouvé à l'aide de méthodes numériques. Imposer les contraintes C1-5, nous avons découvert que EE est une fonction concave par rapport à λ_x et une fonction monotone décroissante par rapport à P_x. Nous avons trouvé la limite supérieure pour optimale λ_x pour lesquels EE prend la valeur de crête.

1.8.1 Orientations futures

Dans les travaux futurs, il peut être intéressant de comparer les deux valeurs optimales de SBS densité - celui qui minimise le APC et celui qui maximise la EE, et de vérifier lequel des deux métriques est relativement plus bénéfique sur le plan économique et en termes de consommation d'énergie. Le choix de travailler avec deux niveaux (*macro* et *petite*) a été arbitraire et ce traitement peut être généralisée à un système à plusieurs niveaux (comme dans [34]) chacun avec un autre type de la capacité de mise en cache. Prenant la consommation d'énergie à chaque étape du processus de mise en cache du contenu des utilisateurs à la SBSs en compte, nous donne une prise plus complète sur les facteurs influençant l'efficacité énergétique et laissant ainsi nous optimisons judicieusement.

Chapter 2

Introduction

2.1 Background and Motivation

Fueled by the ubiquity of new wireless devices and smart phones, as well as the proliferation of bandwidth-intensive applications, past decade has seen a dramatic increase in the usage and demand for mobile connectivity. The seamless demand for wireless data by rich media applications and services often associated with video streaming and social networking on smartphones and other wirelessly connected computing devices is expected only to increase further and more rapidly [2] in coming years. One crucial challenge for next-generation networks is to be able to support the predicted increase in the traffic in a sustainable and economically viable way. In addition to ensuring connectivity to the users, monitoring energy expenditure (thereby the resulting carbon footprint) is also of great importance from both enocomic and environmental points of view. Energy expenditure shall either be minimized or be utilized as efficiently as possible. Despite the resemblance, the two problems aren't the same. To explain this in general terms: Imagine that a car has a fuel tank of fixed size. Its engine may be manufactured either such that it travels a long distance but at a compromised speed or such that it goes at an impressive speed while the distance it can travel is shorter. Clearly, there is a tradeoff between the two cases and, instead of being straightforward, defining EE depends on the specific utility we seek out of a system. In the context of wireless telecommunications, EE is often defined in the literature as the ratio between the total data rate and the total power consumed at the transmitter. Maximizing the EE does ensure that the energy is usefully spent, but it is still important to keep in mind that achieving this maximum should not surpass our budget limitations.

There have been studies addressing this subject from different directions. In [3], optimization of EE of a MIMO point-to-point link is studied in cases when imperfect CSI is available only at the transmitter and is available only at the receiver end. In [4], EE is optimized with respect to the number of antennas in MIMO system, where a realistic power consumption model at each BS has been taken into account. In [5], the density of access points (APs), number of antennas and users per AP, and transmission power have been optimized for maximal EE. Works of Li Chang et al, (see [6], for example) study EE aspects of small cell networks in terms of BS densities and power cost at the BSs. The GreenTouch consortium recently[1] released a white paper [7] on the Green Meter Research study that claims that it is possible through a combination of changes in technologies, architectures, components, algorithms and protocols to reduce the net energy consumption in end-to-end communications networks by up to 98% by 2020 compared to the 2010 reference scenario defined by GreenTouch. According to this a 10,000-fold increase of energy efficiency in mobile access networks is possible by incorporating changes in the existing technologies, architectures, components, algorithms and protocols. Concerning mobile access networks, among the

[1] June 18, 2015

proposed list of ideas are:

- Beyond Cellular Green Generation (BCG^2): Splitting data and signaling between two overlapping layers and incorporating on-demand connectivity has been shown to be an energy efficient solution.
- Large Scale Antenna Systems (LSAS): Instead of serving mobile terminals by means of bulky antennas that radiate power in sectors, this technique tries to direct power to user where it is needed. This is achieved by utilizing a large number of physically small, low-power, individually controlled antennas to create a multiplicity of user-selective beams of data. The data-bearing power is therefore only directed where it is needed and avoids unnecessary interference.

While these results seem promising, further analysis still needs to be done to broaden the number of possibilities of achieving energy efficient next generation wireless network technologies. Since there may not be a universally valid and viable remedy to step up towards green technologies, this search is crucial. The present thesis comes in the context of exploring various possibilities of increasing energy efficiency in wireless networks.

2.2 The physical aspects of wireless transfer of information

From a fundamental point of view, it has long been known that the process of *communication* always has a cost in terms of energy [8–10]. In 1948, Shannon published his revolutionary paper [11] in which the conditions under which information can be communicated reliably are mathematically quantified. It is of great interest to address this specific point again but from the perspective of energy expenditure taking information creation and processing into account.

In its colloquial sense, information is an abstract concept and is often considered subjective. However, in its objective sense information precisely means an ensemble of suitable array of physical systems prepared in desired states. The physical nature of information is discussed in [12]. To give some examples:

- A text message written on a piece of paper with a pencil is but one among several possible arrangements of an ensemble of carbon molecules in the two dimensional plane of a paper.
- An audio file inscribed on a compact disk is essentially a collection of pits and bumps engraved on it in a unique order.
- Braille scribed on a paper, signals such as smoke rings, drums, sirens etc are all quantifiable physical entities and they all involve utilization of energy during their creation, propagation amd destruction.

If a chosen physical system has two distinct states the information represented by such a system is called a *bit*. It has been shown that a large amount of information may be represented in terms of enough number of such bits [11]. Preparing a physical system in a state, or to *change* its state from one to another requires energy. While there may be countless possibilities for physical systems that can be used in preparing, changing and storing the given information, the choice of the system does decide how much energy is needed to perform these operations. The *first question* we pose in this part of thesis addresses this point. In an attempt to explore the answer to the question: *'What is the most energy efficient choice of physical systems to represent and process information?'*, we begin with a simple thermodynamic system (section 3.2.1) to represent an *information bit*. In the realm of classical physics this is indeed the cheapest way of storing bits from energy perspective. In Chapter 3, we discuss in detail how energy is spent in preparing and processing (operations on) such bits.

The process in which physical systems are prepared in desired states are transported from the *transmission* point to the *reception* point in space and time through a medium is called *information transfer*.

This is a physical process and hence requires energy. The correlation between the array of states sent and the array of states received is what determines the quality of information transfer. This correlation may be distorted due to interference from the physical world (noise), and such a distortion is quantified as *error*. There exists a trade off between the amount of energy spent at the transmission point and the mutual information (between the received states and the transmitted states) that can be achieved. Another question we pose in the first part of this thesis addresses this point. In an attempt to explore the answer to the question:*"What is the limit on the efficiency of a communication system taking information creation into account?"*, we design a system to convert the information represented by the above mentioned thermodynamic bits into electromagnetic waves to reach the reception, and evaluate the energy efficiency of the system. We are particularly interested in evaluating energy efficiency in communicating one bit of information as a limiting case (section 3.4.4). Contrary to the intuitive notion that efficiency can be arbitrarily increased by increasing energy investment, our results show that excess expenditure of energy does not necessarily have to be an efficient solution in point-to-point communication. This approach is important not just for the wireless service providers but also from the point of view of the users of wireless devices who are often compromised by the annoyance of frequently charging batteries of the mobile devices. As a matter of fact this part of the work comes in lines of exploring fundamental aspects of computation as well as communication.

2.3 Architectural analysis of energy expenditure in wireless communication

Stochastic geometry based modeling

The increasing demand for connectivity, data rate, and quality of service (QoS) cannot be satisfied merely by increasing indefinitely the transmit power of the base stations BSs. This is mainly due to the fact that an increase in transmit power, besides increasing the signal strength from the desired BS, also increases the interference received by the non-serving BSs. This may effectively decrease the signal-to-interference-plus-noise ratio (SINR) experienced by the user terminal, thus having a negative impact on the QoS. Besides that, it is also essential from an energy efficiency perspective to address the problem of minimizing the energy expenditure while maintaining certain constraints such as target coverage and minimum data rate.

There are various approaches to address the problems of reducing energy consumption theoretically. For instance, if a finite number of BSs are deployed in a regular hexagonal (or grid) cellular fashion as in Figure 4.2 [13], one might seek to minimize the total power consumed by obtaining the optimal operating parameters, such as the hexagonal cell size and the magnitude of transmit power at each BS, while guaranteeing a certain QoS. Nevertheless, this approach involves cumbersome analysis as evaluating the spatial distribution of the SINR in a grid-based model becomes prohibitively complex as the system size increases and one may have to resort to extensive simulations. A common simplification in modeling cellular networks which enables us handle the problem analytically is to assume that the locations of BSs are randomly scattered on a two dimensional plane surface according to a homogeneous PPP [14]. Several works studied the validity of PPP modeling of BSs (in comparison with regular cellular models,) and it is often shown to provide useful insights into the statistical behavior of key performance metrics [15]. Some researchers, particularly those with an industry background, often express scepticism about using stochastic geometry for performance analysis of wireless cellular networks. Their prime accusation is that in a treatment based of stochastic geometrical modeling one neglects the actual deployment of BSs and focuses instead on the theoretical BS configurations. It needs to be stressed at this point that the purpose of using stochastic geometry is not to solve any actual real-life deployed configuration of a collection of BSs, but to acquire qualitative as well as quantitative insights on some crucial system parameters with the help of tractable mathematical formalism rather than depending on system specific simulations. This

has been stressed enough in the works of Andrews et al. More recent results on stochastic analysis of wireless cellular networks can be found in [16–20] and references therein.

Is caching an energy efficient solution?

Both academia and industry are now in an urge of evolving traditional cellular networks towards the next-generation broadband mobile networks, coined as 5G networks, targeting to satisfy the mobile data tsunami while minimizing expenditures and energy consumption. Among these intensive efforts, caching users' content locally at the edge of the network is considered as one of the most disruptive paradigms in 5G networks [21].

Interestingly yet not surprisingly, recent results have shown that distributed content caching can significantly offload different parts of the network, such as in RANs and CN, by smartly prefetching and storing content closer to the end-users [22]. Indeed, traditional cellular networks, which have been designed for mobile devices with limited processing and storage capabilities, have started incorporating context-aware and proactive capabilities, fueled by recent advancements in processing power and storage. As a result, caching has recently taken the 5G literature by storm.

Caching in Single tier cellular networks

As mentioned earlier, energy consumption aspects of cellular networks are usually investigated by placing BSs on a regular hexagonal or grid topology and conducting intensive system-level simulations [23]. Therein, the aim is to find the optimal values of parameters, such as cell range and BS transmit power, while minimizing the total power consumed under certain QoS constraints. Despite its attractiveness, it is rather cumbersome or even impossible to analytically evaluate key performance metrics in large-scale networks. A common simplification in modeling cellular networks is to place the BSs on a two dimensional plane according to a homogeneous spatial PPP, enabling to handle the problem analytically [24]. Several works studied the validity of PPP modeling of BSs compared to regular cellular models (e.g. [25]) and additional insights can be obtained by analytically characterizing performance metrics, such as the spatial distribution of SINR, coverage probability, and average rate [18, 19]. More recently, optimization of EE has been addressed quite well in the works of Emil et al such as [26].

In this thesis we analyze (see Chapter 5) energy consumption aspects of cache-enabled wireless single-tier network deployments using a spatial model based on stochastic geometry. We consider the problem of optimizing APC and EE of cache-enabled BSs in a scenario where BSs and mobile users are distributed according to independent homogeneous PPPs. We provide conditions under which the area power consumption is minimized subject to a certain quality of service (QoS) in terms of coverage probability. Furthermore, we provide the optimal BS transmit power that maximizes the energy efficiency, defined as the ratio of area spectral efficiency over the total power consumed. The main result of this analysis is that distributed content caching in wireless cellular networks turns out to be an energy efficient solution.

Caching in heterogeneous networks

Wireless cellular networks of the future are likely to be heterogeneous, i.e. have one or more tiers of small cells overlaid on the macrocellular tier. Works such as [27] study EE in a HetNet scenario with sleeping strategies incorporated within. Motivated to progress in this line of research, we extend the idea of caching users' content in one (or more) of tiers in a HetNet and study how this influences the energy consumption aspects of the wireless system. We define and solve problems of optimizing APC and EE with respect to base station densities and their transmit powers. In Chapter 6, we analyze energy consumption aspects of two-tier wireless network, where the small cells are cache-enabled.

2.4 Thesis Outline

This thesis consists of two parts. The first part, comprises Chapter 3, and the second part comprises of Chapters 4, 5, and 6.

Chapter 3 focuses on fundamental aspects of energy consumption in wireless communication where we use tools from thermodynamics and electromagnetic theory to make interesting observations regarding the energy efficiency of a simple wireless telecommunication device. The chapter is organized as follows: In Section 3.2, we elaborate on the communication system that we intend to study. In Section 3.4, we discuss the energy efficiency of the proposed system that uses Binary Phase-Shift Keying (BPSK). In Section 3.5, we extend our analysis to higher constellations by comparing the present case with that of Quadrature Phase-Shift Keying (QPSK). In Section 3.6, we make plots of energy efficiency of the proposed system and comment on its behavior for cases of BPSK and QPSK signaling.

In the second part of this thesis, we study the architectural limitations on the energy efficiency. In **Chapter 4** we use tools from stochastic geometry to model wireless network systems and carry out a performance analysis. The chapter is organized as follows: In Section 4.2, we present the network model and motivate the optimization problem. In Section 4.3, we provide bounds on the optimal transmit power to minimize the area power consumption under minimum coverage and rate constraints. In Section 4.4, we evaluate the optimal BS density that minimizes the area power consumption. validation of the theoretical results by comparing with those of a numerical simulation are given in Section 4.5. Section 4.6 presents the concluding remarks.

In **Chapter 5**, we extend this analysis by studying how caching users' content at the edge of the network can influence the over all energy efficiency of the system. The chapter is organized as follows: In Section 5.2 we describe our system model and introduce a detailed power model capturing caching and QoS constraints. We then formulate the problem of minimizing the area power consumption in Section 5.3. Results on the optimal power that maximizes the EE are provided in Section 5.4. Section 5.5 presents the concluding remarks.

In **Chapter 6**, we extend the study of energy consumption aspects of a system with BSs endowed with caching capabilities to a HetNet scenario. The chapter is organized as follows: In Section 6.2 we describe the model according to which the two kinds of BSs are deployed and we define the connection policy of the User Terminals (UTs). In Subsection 6.2.1 we explain the caching model which is slightly different from the one used in Chapter 5. In Section 6.3, we define and derive anaytical expressions of key system metrics that we use later in the chapter. Subsection 6.3.1 gives details of the power model describing various components of power consumed as part of signaling and operation of a BS. In Section 6.4 we present the main results of the work where we define the problems of optimizing APC and EE with respect to spatial density and transmit power of SBSs separately. In Section 6.5, we present the numerical results to validate the theoretical results. Section 6.6 presents the concluding remarks.

2.5 Publications

Journal Articles

- B. Perabathini, E. Baştuğ, M. Kountouris, M. Debbah and A. Conte, "Energy Consumption Aspects of Cache-Enabled 5G Wireless Networks". (To be submitted)

Conference Papers

- B. Perabathini, E. Baştuğ, M. Kountouris, M. Debbah and A. Conte, "Energy Consumption Aspects of Cache-Empowered Heterogeneous Networks: Optimization and Analysis", (to be submitted

to a conference), 2015.

- B. Perabathini, E. Baştuğ, M. Kountouris, M. Debbah, and A. Conte, "Caching at the Edge: a Green Perspective for 5G Networks", IEEE International Conference on Communications (**ICC**'2015), London, UK, June 2015.
- B. Perabathini, M. Kountouris, M. Debbah, and A. Conte, "Optimal Area Power Efficiency in Cellular Networks", **Globecom** - 2nd Workshop on Green Broadband access: energy efficient wireless and wired network solutions, Austin, Texas, December 2014.
- B. Perabathini, V. S. Varma, M. Debbah, M. Kountouris, and A.. Conte, "Physical Limits of Point-To-Point Communication Systems", Workshop on Physics-Inspired Paradigms in Wireless Communications and Networks, co-located with **WiOpt** 2014, Hammamet, Tunisia, June 2014.

Part I

Physical limits

Chapter 3

Physical Limits of Point-to-Point Communication Systems

3.1 Overview

Digitization of information, by which we mean converting phone calls, music, videos, and such virtual content into digital bits of data and transiting billions of them per second, is founded upon the original work of Claude Shannon (1916-2001) published in 1948 [11]. Shannon's work laid the definition of 'information' for communication networks. Works such as [8–10] explored on the physical nature of information and that energy expenditure is inevitable in processing and communicating information. Shannon's channel capacity theorem gives us an upper bound of the rate at which information can be transmitted over a communication channel for a fixed signal strength. The converse of this result should address the question of finding the upper bound of signal strength that enables us transmit a given amount of information. The question gets much more interesting if the energy invested in creation, transfer as well as processing of the information is taken into account while finding this upper bound.

In this chapter, we explore these physical limits of successful information transfer in a point-to-point communication system. In interest of studying the upper bound on the energy usage imposed by physics on communication systems in general, we model a simple generic system that enables us to make some basic inquiries about the energy efficiency in information creation, transfer as well as processing. We use ideas from thermodynamics such as Szilard engine to represent information bits. We further use ideas from electromagnetic theory for transfer of information, and information theory to define the energy efficiency metric. We find the upper limit of this efficiency and conditions at which it can be achieved.

3.2 Proposed point-to-point communication system

Figure 3.1 is the schematic representation of a typical communication system, where it has been shown that information takes three (abstract, storable and communicable) forms during a typical process of communication. In order to model a communication system at its most basic physical form, we first look at how information itself can be physically represented. Secondly, we look at how this information can be extracted from the physical system and converted to electromagnetic waves. Finally, this signal is received by a physical antenna and stored as thermodynamic information at the receiver. What we call 'information' here is the source coded form of raw information. We do not take into account the influence of source coding on the efficiency.

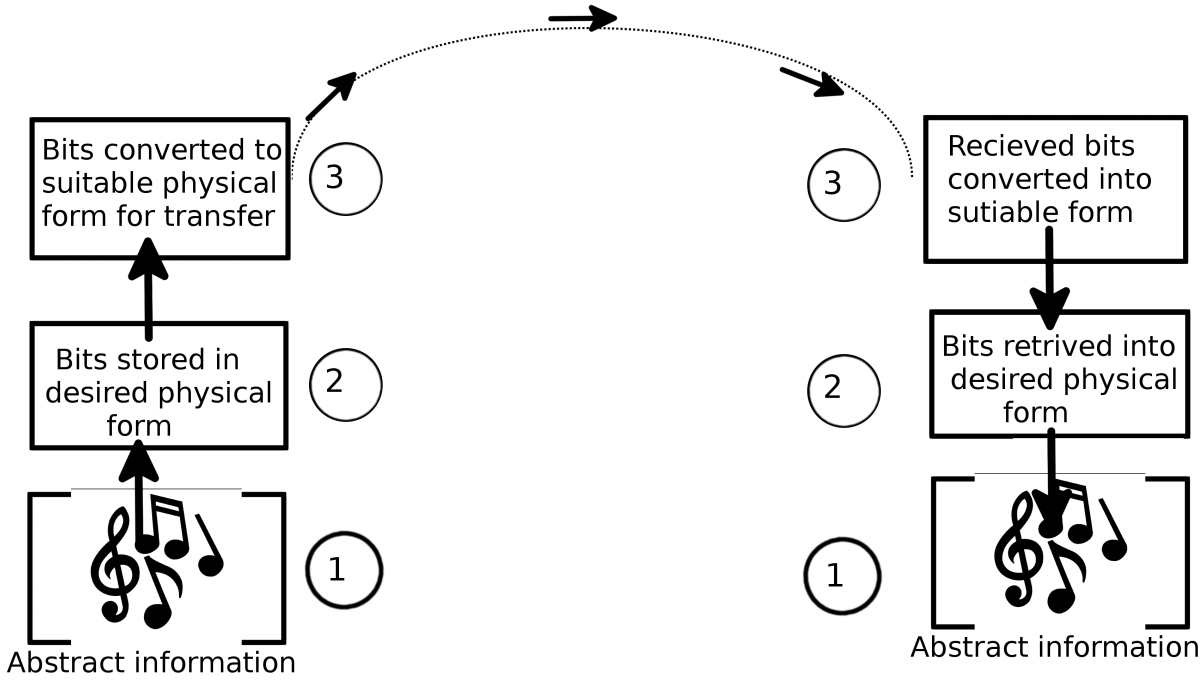


Figure 3.1: A schematic picture of a typical communication system. Abstract information, that is saved in a physical form, is converted to some other suitable physical form in order to perform the information transfer. Numbers 1,2 and 3 represent information in the three (abstract, storable and communicable) forms.

3.2.1 Thermodynamic representation of information

Consider an isolated cylindrical container of gas molecules in equilibrium. Now consider that an insulating wall with a small door is introduced in the middle. Maxwell postulated that a small demon that can operate this door can cause a violation of second law of thermodynamics by allowing gas molecules with high speeds pass in one direction and slow ones in the other direction. This is due to the fact that transferring heat between systems that are mutually at thermal equilibrium without doing any work violates second law of thermodynamics [35, 36]. Szilard resolved this paradox by arguing based on the proposition that "destruction of physical information creates entropy in the environment". For the demon to create such a temperature difference, without disturbing the system, it should to have an infinite memory at hand. Because having finite memory restricts the amount of information that can be stored at a given time and therefore excess information obtained by observing positions and velocities of several molecules would have to be destroyed continuously. This destruction of information would have dissipated energy back into the environment (container). This dissipated energy would in turn have increased the net disorder in the system contradicting Maxwell's prediction that the demon would actually decrease it [37–39].

An "engine" that Szilard used to illustrate the above idea, consists of a single molecule of an ideal gas in a cylindrical vessel with an adiabatic friction less piston locked in the middle separating it into two halves. The single molecule has two distinct physical states (left/right of the box) and therefore can represent one bit of information. To destroy this information, this system is brought in contact with a thermal reservoir at the same temperature as the molecule and the piston is unlocked. As a result the molecule jitters back and forth and repeatedly kicks the wall until the wall reached the other end of the container. Finally the molecule's position is totally randomized and the information about its location is destroyed with the help of heat drawn from the reservoir. Since the internal energy of the molecule has

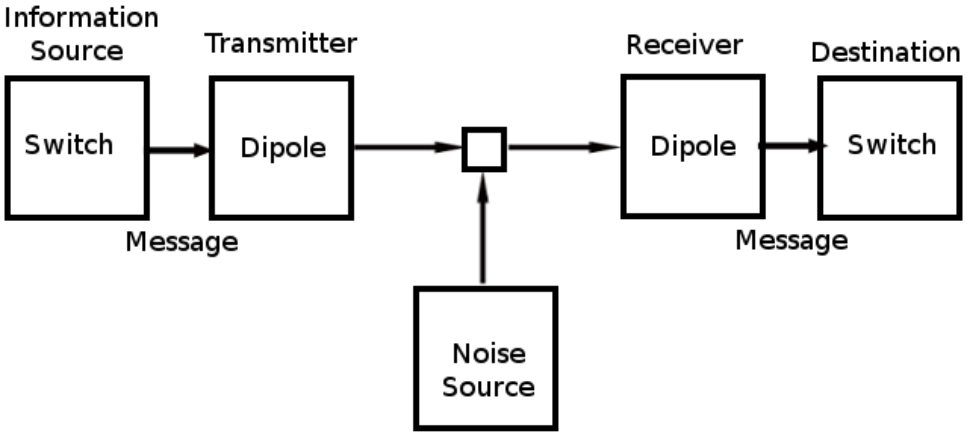


Figure 3.2: Point-to-Point communication scheme.

not changed, the excess work is dissipated into the environment.

Heat dissipated in destroying a bit ($Q_{\text{dissip.}}$) which is the work done by the molecule ($W_{\text{mol.}}$) to push the volume from $V/2$ to V in an isothermal limit, using the ideal gas law $PV = kT$ is given by [40]

$$Q_{\text{dissip.}} = W_{\text{mol.}} = \int_{V/2}^{V} P\, dV = \int_{V/2}^{V} kT\, \frac{dV}{V} = kT \ln 2, \tag{3.1}$$

where $k = 1.381 \times 10^{-23}$J/K is the Boltzmann constant. It has to be noted that the time taken to push the piston (τ) at a given temperature depends on the size of the container V. Note that τ cannot be infinitely small unless V is infinitely small in which case quantum effects have to be taken into account [29].

We use a modification of Szilard engine as shown in Figure 3.3 in our model. There are two possible states for such a system. The Bernoulli random variable corresponding to each state is $\mathcal{P}_{\text{right}} := \mathbb{1}\{\text{Probability}[X > L - \epsilon] = 0\}$, where X is the x coordinate of the position of the molecule, L is the horizontal dimension of the box, ϵ is any arbitrarily small positive number. Clearly $\mathcal{P}_{\text{right}} \in \{0, 1\}$.

1. **State 0:** The piston is pushed to the extreme end and the information about the location of the molecule is totally randomized. There is no mechanical energy that can be derived out of this system ($\mathcal{P}_{\text{right}} = 0$).

2. **State 1:** The molecule is confined in the left half and the piston is pushed to the middle from the right. A mechanical energy of at least $kT \ln 2$ can be derived out of it ($\mathcal{P}_{\text{right}} = 1$).

We say that these systems come in state 0 by default and we spend energy $kT \ln 2$ in creating state 1. The two states together represent one *bit* of information which we here on wards refer to as the *Thermal-bit*.[1] A certain amount of raw information that is sampled and converted into necessary number of bits (0's and 1's) can be physically represented using a tape (an array) of *Thermal-bits* as shown in Figure 3.4.

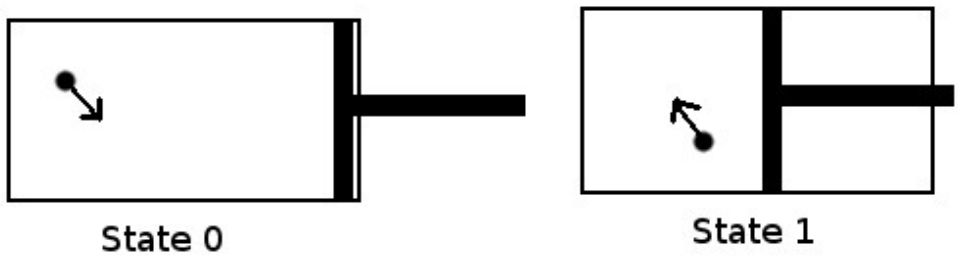


Figure 3.3: The Thermal Bit.

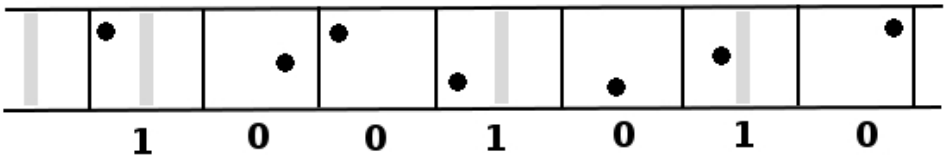


Figure 3.4: Information as a tape of thermodynamic bits.

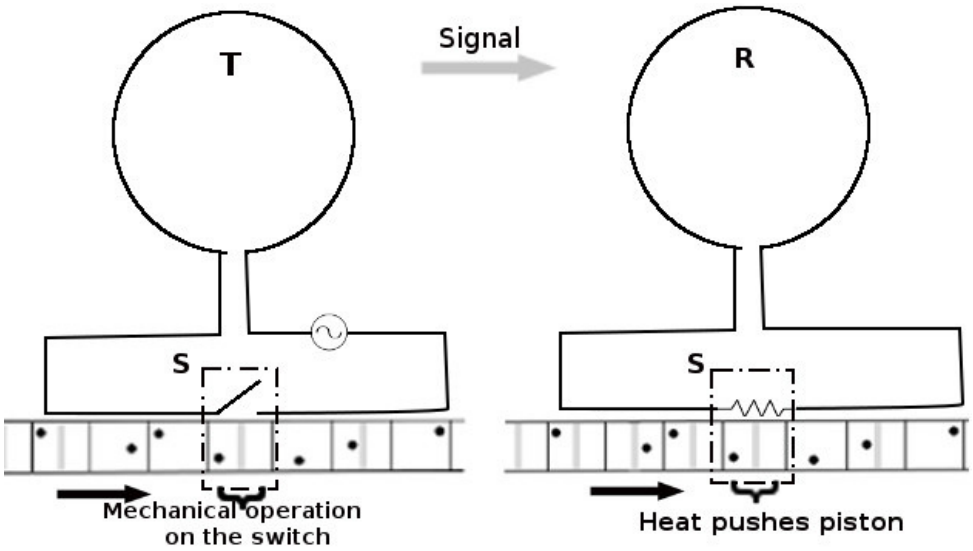


Figure 3.5: Communication device.

3.2.2 Transmission setup

Figure 3.5 depicts the communication device that we intend to study in the subsequent sections. Part **T** of the system is the transmission setup and **R** is the reception setup. We intend to send n bits of information from **T** to **R** which a distance r apart in space. For a very large n it maybe assumed that the source generates as many bits in state 0 as in state 1. That implies that the probability of having either of them is 1/2.

The Transmission Setup consists of an ideal AC power source (with no internal resistance) connected to a wire loop antenna of radius b. There is a switching mechanism **S** at the bottom and works as follows:

1. A tape of *Thermal-bits* is fed into the switching mechanism. Each time a bit enters in, it remains in **S** for τ seconds before the next bit enters.

2. If a bit enters the switch in *state 1*, mechanical energy is extracted out of it and is used to open the switch of the circuit and keeps it so for the time τ until the bit is replaced by the next one. The amount of mechanical energy extracted out of each *Thermal-bit* is given by eq. (3.1). Whereas, if it is a *Thermal-bit* in *state 2* that enters in, the switch remains closed for τ seconds until the next bit enters.

Each time the switch is closed, current from the AC source flows through the loop antenna and an electromagnetic field is radiated (magnetic dipole radiation). Let us represent the k^{th} bit to be sent by a_k, a Bernoulli random variable of probability 1/2. Therefore, depending on a_k, the switch remains either closed or open each for a period of time τ. The current 'signal' through the loop can be written as $I(t) = I_0(t)\cos(\omega_c t)$, where $I_0(t) = I_0 \sum_k a_k S(t - k\tau)$ with I_0 as the amplitude of the AC current, ω_c is the carrier frequency and S represents a square wave of period τ. We assume that the switching time $\tau \gg 2\pi/\omega_c$.

Information that has been stored in the form of *Thermal-bits* thus far is now converted into a series of electromagnetic pulses. These pulses travel at the speed of light and each pulse has a time length of τ. The phase difference between the pulses is what represents the information making it a BPSK signal.

3.2.3 Reception setup

Reception setup consists of a wire loop receiving antenna (radius b) attached in series to a resistor (Figure 3.5). The resistance of this resistor is adjusted/chosen such that the total impedance of the receiving antenna matches the free space impedance. By this, we can safely assume that all the power that *reaches* the antenna is absorbed in the form of induced electromotive force (EMF) which in turn is amplified and then completely converted into heat at the resistor.

A switching mechanism **S** is located at the bottom. It works as follows:

1. A tape of thermal bits created in *states 1* is sent in.

2. Each time the *Thermal-bit* is kept in contact with the resistor for a time τ allowing heat to transfer.

3. If the temperature of the resistor $T_R \geq T$, the piston of the *Thermal-bit* will be pushed out creating a bit in *state 0*.

[1]Similar *Thermal-bits* can be prepared with systems with any number of molecules and not necessarily just one. Nevertheless, we deliberately choose single molecule systems as a limiting case.

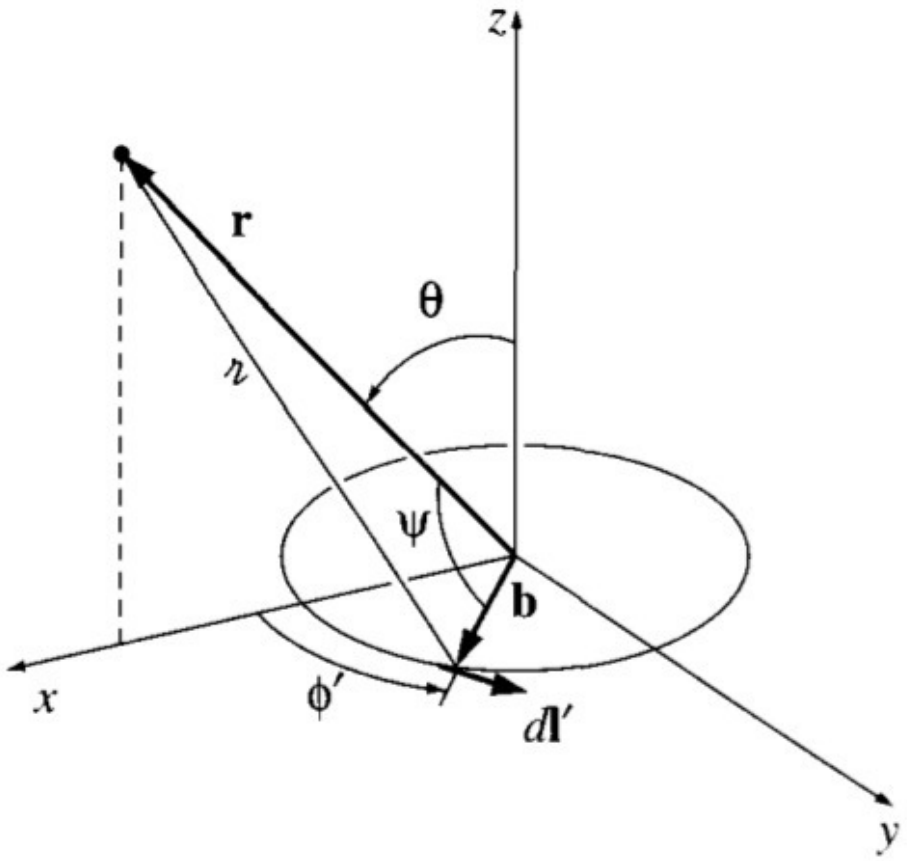


Figure 3.6: Magnetic dipole [1].

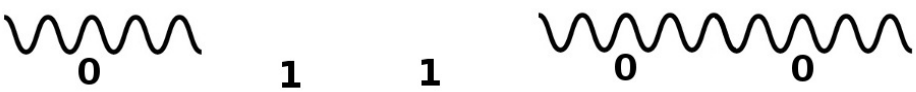


Figure 3.7: Information as a series of electromagnetic pulses.

3.3 Remarks on the limitations of using Szillard boxes

Following the transmission as well as the receiver setups described in the previous section, our model requires that the thermal-bit stays in contact with the thermal reservoir for exactly τ seconds and it has been assumed that this amount of time is enough for the particle to make repeated collisions with the piston and eventually push it significantly forward. There are two issues that need to be taken into account here:

1. The particle inside the thermodynamic box (Fig 3.3) hits the piston with a certain probability and therefore, there is a non-zero probability that it never hits the piston within the time period τ. This may actually lead to an error during the process of reading the bit. In our theoretical treatment, we assumed the probability of not hitting the piston to be negligible. But a more rigorous analysis would actually consider the Maxwell-Boltzmann distribution of the particle inside the box and evaluate the probability that the particle hits the piston within the time τ and study the significance of its effects on the energy efficiency limits calculated in the subsequent sections.

2. The particle loses some of its kinetic energy each time it hits the piston since (unlike in ideal cases) the piston also has a certain mass. The idea of keeping the thermal bit in contact with the reservoir is that the particle would redraw kinetic energy as it hits the bottom wall which is in contact with the reservoir (Fig 3.3). Here again, there is a finite chance that the particle fails to hit the bottom wall at least once within the time period τ. A more complete treatment would take this issue into account.

We defer including these aspects to future work.

3.4 The energy efficiency limit

In this section, we derive the equations for the energy consumed by the communication system and the achievable rate of the system proposed in 3.2. With the following theoretical treatment, we finally calculate the fundamental limits on energy efficiency of a physical system that communicates with electromagnetic waves.

3.4.1 The signaling system

We know that the magnitude of the Poynting vector generated by a magnetic dipole, as defined in the previous section, at a distance r from the dipole and at an angle θ with vertical axis (Figure 3.6), making the assumptions $b \ll r$, $b \ll \frac{c}{\omega_c}$ and $r \gg \frac{c}{\omega_c}$, is given by

$$|\vec{S}| = \frac{\mu_0}{c}\left[\frac{m_0(t)\omega_c^2}{4\pi c}\frac{\sin\theta}{r}\cos(\omega_c(t-r/c))\right]^2 \tag{3.2}$$

where $c = 3 \times 10^8 \mathrm{ms}^{-1}$ is velocity of light and $m_0(t) = \pi b^2 I_0(t)$ is the maximum dipole of the loop [1].

We find the average of the power density over the time period of one complete wave cycle $\tau_0 = \frac{2\pi}{\omega_c}$.

$$\begin{aligned}\langle S\rangle &= \left(\frac{\mu_0 b^4 \omega_c^4 \sin^2\theta}{16 r^2 c^3}\right) I_0^2 \\ &\quad \times \mathbb{E}_{a_k,\tau_0}[(\sum_k a_k S(t-k\tau))^2 \times \cos^2(\omega_c(t-r/c))] \\ &= \left(\frac{\mu_0 b^4 \omega_c^4 \sin^2\theta}{16 r^2 c^3}\right) I_0^2 \\ &\quad \times \mathbb{E}_{\tau_0}\Big[[\sum_{k\neq l} \mathbb{E}[a_k a_l] S(t-k\tau)^2 + \sum_k \mathbb{E}[a_k^2] S(t-k\tau)^2] \\ &\qquad\qquad \times \cos^2(\omega_c(t-r/c))\Big].\end{aligned}$$

Since $a_k \in \{0,1\}$ with probability $1/2$ each and that a_k's are not correlated, we hence have that, $\mathbb{E}[a_k]\mathbb{E}[a_l] = 1/4$ and $\mathbb{E}[a_k^2] = 1/2$, which gives

$$\begin{aligned}\langle S\rangle &= \left(\frac{\mu_0 b^4 \omega_c^4 \sin^2\theta}{16 r^2 c^3}\right) I_0^2 \\ &\quad \times \frac{3}{4}\mathbb{E}_{\tau_0}\Big[\sum_k S(t-k\tau)^2 \times \cos^2(\omega_c(t-r/c))\Big].\end{aligned}$$

By using the assumptions that $\tau \gg \tau_0$ and that $S(t-k\tau)$ is either 0 or 1, we get

$$\begin{aligned}\langle S\rangle &= \left(\frac{\mu_0 b^4 \omega_c^4 \sin^2\theta}{16 r^2 c^3}\right) I_0^2 \frac{3}{4}.\frac{1}{2} \\ &= \left(\frac{3\mu_0 b^4 \omega_c^4 \sin^2\theta}{128 r^2 c^3}\right) I_0^2. \qquad (3.3)\end{aligned}$$

The total radiating power emitted by the antenna is obtained by integrating Eq (3.3) over the surface area of any sphere surrounding the loop. Therefore the *average energy per bit* emitted by the transmission setup during one operation for time τ (that is for bit in *state 0*) is given by

$$\begin{aligned}E_T^{\text{rad}} &= \tau \times \left(\frac{3\mu_0 b^4 \omega_c^4 I_0^2}{128 c^3}\right) \int_0^{\pi} (\sin\theta)^2 \sin\theta \, d\theta \int_0^{2\pi} d\phi \\ &= \tau \times \left(\frac{3\mu_0 b^4 \omega_c^4 I_0^2}{128 c^3}\right) \times \frac{8\pi}{3} \\ &= \alpha \tau \omega_c^4 I_0^2, \qquad (3.4)\end{aligned}$$

where $\alpha = \left(\frac{\mu_0 \pi b^4}{16 c^3}\right)$.

3.4.2 Energy considerations at Transmission Setup

In the whole process of transmitting n bits of information, we like to evaluate the total amount of energy spent in the Transmission Setup. We have shown earlier that the energy required for an operation on *state 1* to switch it to *state 0* is $kT\ln 2$ and an operation on *state 0* costs zero energy. And by the assumption that the states are equally probable, we can say that an operation on a *Thermal-bit* on an average costs an energy of $\frac{1}{2}kT\ln 2$. Therefore, the average amount of energy required to read a total of n *Thermal-bits* would be $E_T^{\text{bits}} = \frac{n}{2}kT\ln 2$ joules.

To emit all the bits, the radiator must have been operated $n/2$ times and for time τ each time. Therefore the fundamental lower bound on the total energy spent is:

$$E_{\mathrm{T}}^{\mathrm{tot}}(n) := n[\alpha\tau\omega_c^4 I_0^2 + \frac{1}{2}kT\ln 2]. \tag{3.5}$$

3.4.3 Energy considerations at Receiver Setup

Equation (3.3) gives the magnitude of the Poynting vector at a distance r from the transmitting antenna. We assume that both transmitting and receiving loops are in the same plane ($\theta = 0$). Assuming that all the power absorbed into the loop is equal to the product of the magnitude of the Poynting vector and the area πb^2, the received signal energy on average per bit is given by

$$E_R^0 := \tau \times \frac{3\mu_0 b^4 \omega_c^4 I_0^2}{128 r^2 c^3} \times \pi b^2 = \beta\tau\omega_c^4 I_0^2, \tag{3.6}$$

where $\beta = \frac{3\mu_0\pi b^6}{128 r^2 c^3}$.

If we assume that the distance is large, i.e. for large enough r, the sum of received energy per bit and the channel noise can be much less than the energy needed to write a bit at the reception end, i.e.

$$E_R^0 \ll \frac{1}{2}kT\ln 2.$$

In this case, in order to write a bit into the tape at the receiver, the received non-zero signal will be amplified by an amplifier that has to supply an average total energy (for all n bits)

$$E_R^{\mathrm{tot}}(n) := \frac{n}{2}kT\ln 2. \tag{3.7}$$

Finally, we define the *total* energy cost of *preparing a Thermal-bit → transmitting it electromagnetically → constructing it back at the receiver* as (from eq. (3.5) and eq. (3.7))

$$\begin{aligned} E_{\mathrm{tot}}^0 &= E_{\mathrm{T}}^{\mathrm{tot}}(n=1) + E_{\mathrm{R}}^{\mathrm{tot}}(n=1) \\ &= \left(\alpha\tau\omega_c^4 I_0^2 + \frac{1}{2}kT\ln 2\right) + \frac{1}{2}kT\ln 2 \\ &= \alpha\tau\omega_c^4 I_0^2 + kT\ln 2. \end{aligned} \tag{3.8}$$

3.4.4 Energy Efficiency

The energy per bit to noise energy ratio snr_0^b is given as

$$snr_0^b = \frac{E_R^0}{N_0 W} = \frac{\beta\omega_c^4}{W^2}\frac{I_0^2}{N_0}, \tag{3.9}$$

where, I_0^2 which is proportional to the energy spent by the AC source at $\mathbf{T}$ per bit, N_0 is the noise spectral density, and $W := \frac{1}{\tau}$ is the bandwidth where τ is the time length of each pulse.

We know that the mutual information of BPSK channel as a function of snr_0 is given by [41]

$$\begin{aligned} I_b(snr_0^b) &= \frac{1}{\ln 2}\Big(snr_0^b - \\ &\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\frac{-y^2}{2}}\log\Big(\cosh(snr_0^b - y\sqrt{snr_0^b})\Big)\,dy\Big). \end{aligned} \tag{3.10}$$

Finally, we define energy efficiency η as the ratio

$$\eta_b = \frac{W I_b(snr_0^b)}{E_{\text{tot}}^0} \text{ bits/s/J.}$$

By substituting snr_0^b from Eq (3.9), $\eta_b(\omega_c)$ can be expressed as a function of ω_c as

$$\begin{aligned}&\eta_b(W,\omega_c,I_0,kT) \quad = \\ &\frac{W(\frac{\beta\omega_c^4 I_0^2}{W^2 N_0} - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\frac{-y^2}{2}} \log\left(\cosh(\frac{\beta\omega_c^4}{W^2}\frac{I_0^2}{N_0} - y\sqrt{\frac{\beta\omega_c^4}{W^2}\frac{I_0^2}{N_0}})\right) dy)}{(\ln 2)(\frac{\alpha\omega_c^4 I_0^2}{W} + kT\ln 2)}.\end{aligned} \tag{3.11}$$

3.5 Energy efficiency for higher constellations

So far, we only considered the case where only one bit is encoded by one symbol. However, by making appropriate modifications to the proposed setup and the scheme of phasing, one should in principle be able to encode any number of bits per symbol.

Instead of caring for the details of the engineering, we focus on the energy considerations for the case of QPSK. The main difference in the case of QPSK is that we have four symbols $\{11, 01, 10, 00\}$ and the probability that the source produces a typical symbol a_k is now assumed to be $1/4$. This will result in the following changes:

1. The average energy per symbol emitted is $\frac{1}{4}(2E_T^{\text{rad}} + E_T^{\text{rad}} + E_T^{\text{rad}} + 0) = E_T^{\text{rad}}/2$.
2. Similarly, the received signal energy on average per symbol E_R^0 is changed by a factor $1/2$.
3. A thermal operation in this case is on two bits at once. The average energy spent on the a symbol is $kT\ln 2$. Including the cost of operation at **T** and **R** makes the average energy spent per operation per symbol $2kT\ln 2$, which is twice the case compared to BPSK.

The total energy spent on a symbol in preparing, transmitting and writing (analogous to Eq (3.8)), will be

$$E_{\text{tot}}^0 = \frac{\alpha}{2}\tau\omega_c^4 I_0^2 + 2kT\ln 2 \tag{3.12}$$

The energy per bit to noise energy ratio snr_0^q is given as

$$snr_0^q = \frac{E_R^0}{2N_0 W} = \frac{\beta\omega_c^4}{2W^2}\frac{I_0^2}{N_0}. \tag{3.13}$$

We know that the mutual information of QPSK channel as a function of snr_0^q is given by [41]

$$\begin{aligned}&I_q(snr_0^q) \quad = \\ &\frac{1}{\ln 2}\Big(2snr_0^q - \sqrt{\frac{2}{\pi}}\int_{-\infty}^{\infty} e^{\frac{-y^2}{2}} \log\Big(\cosh(snr_0^q - y\sqrt{snr_0^q})\Big)\, dy\Big).\end{aligned} \tag{3.14}$$

Energy efficiency in this case is given as

$$\eta_q(W,\omega_c,I_0,kT) = \frac{W(\frac{\beta\omega_c^4 I_0^2}{2W^2N_0} - \sqrt{\frac{2}{\pi}}\int_{-\infty}^{\infty} e^{\frac{-y^2}{2}} \log\left(\cosh(\frac{\beta\omega_c^4 I_0^2}{2W^2N_0} - y\sqrt{\frac{\beta\omega_c^4 I_0^2}{2W^2N_0}})\right) dy)}{(\ln 2)(\frac{\alpha\omega_c^4 I_0^2}{2W} + 2kT\ln 2)}. \tag{3.15}$$

Energy efficiency for any M-ary constellation can be theoretically determined along similar lines. However, the engineering of the setup and the scheme of phasing would become much more complicated.

3.6 Numerical results

Equations (3.11) and (3.15) provide expressions on energy efficiency in the case of BPSK and QPSK, respectively. In this section, we numerically verify and compare the behavior of the functions $\eta_b(W,\omega_c,I_0,kT)$ and $\eta_q(W,\omega_c,I_0,kT)$.

We use the values $\mu_0 = 4\pi \times 10^{-7}\text{N A}^{-2}$, $c = 3\times 10^8\text{ms}^{-1}$, $r = 1\text{m}$ and $b = 1\text{m}$ to get the values of $\alpha = 9.25926\times 10^{-34}$ and $\beta = 3.42695\times 10^{-33}$.

In Figure 3.8 we plot η_b and η_q versus carrier frequency ω_c for $W = 1\text{Hz}$, $N_0 = 1$, $kT = 1$ and $I_0 = 1\text{A}$. It can be seen that at lower frequencies, BPSK is more efficient that QPSK, but at higher frequencies, QPSK, is more efficient than BPSK. We numerically determine the optimal carrier frequencies from the plots as $\omega_b^* = 1.57013\times 10^8\text{Hz}$ and $\omega_q^* = 2.00214\times 10^8\text{Hz}$.

Figure 3.9 shows η_b and η_q against bandwidth W for optimal frequencies $\omega_b^* = 1.57013\times 10^8\text{Hz}$, $\omega_q^* = 2.00214\times 10^8\text{Hz}$, $N_0 = 1$, $kT = 1$ and $I_0 = 1\text{A}$. The energy efficiency approaches 0 when the bandwidth is infinitely large.

Figure 3.10 shows η_b and η_q against square of the amplitude square of current I_0 for $W = 1\text{Hz}, N_0 = 1, \omega_b = \omega_q = 10^8\text{Hz}$ and $kT = 1$ For low values of current (amplitude), BPSK is seen to be more efficient than QPSK, but for higher amplitudes QPSK is much more efficient than BPSK.

While the analytical solution of the optimal frequency ω_c^* is not trivial to compute analytically from Eq (3.11), a plot of $\eta(\omega_c)$ vs ω_c in figure 3.8 suggests that there exists an optimal frequency for a set of numerical values of the parameters at which the energy efficiency has its maximum.

Figure 3.11 plots the optimal frequency ω_c^* against temperature (kT) of the *Thermal-bits* in both cases of BPSK and QPSK for $W = 1$, $N_0 = 1$ and $I_0 = 1\text{A}$. Since temperatures cannot be arbitrarily small in classical conditions, this suggests that the optimal frequency cannot be arbitrarily small either.

3.7 Concluding remarks

We devised a hypothetical communication system and calculated the upper bound on energy efficiency using electromagnetic and thermodynamic theory. We have seen that, in the proposed model, for a given total amount of energy that is at disposal to prepare, transmit and rewrite information bits, there exists a unique transmit antenna frequency that yields maximum energy efficiency. We have also seen the dependence of energy efficiency on the bandwidth and the amplitude of the signal, and that there is an optimum at operating point for the two. We have made a comparison of these dependencies in the cases of BPSK and QPSK models. Our results show that, contrary to the results in [28], when the communication system is operating at a non-zero temperature, the optimal power consumed by the system in order to maximize energy efficiency is non-zero.

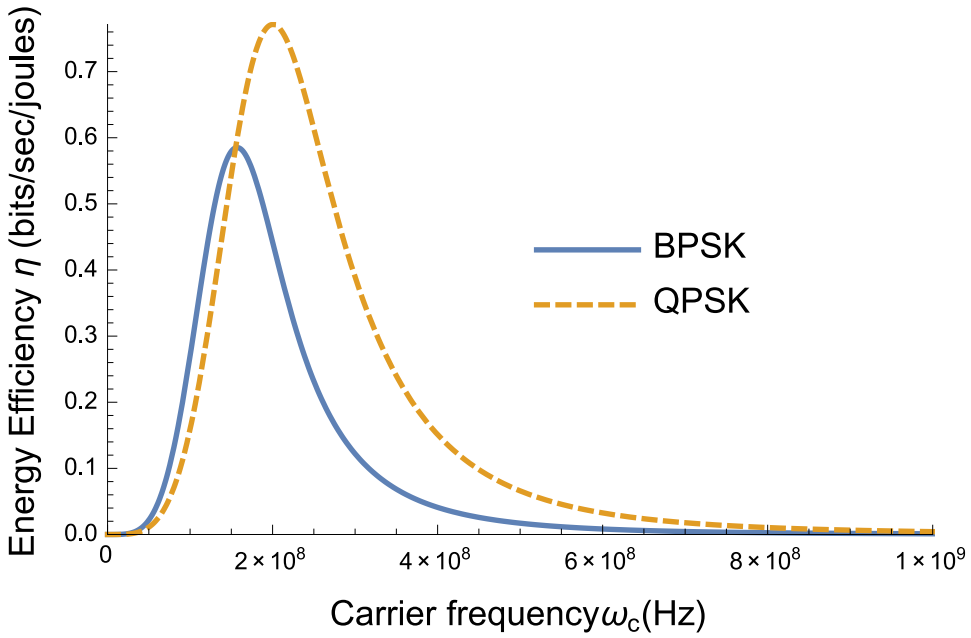


Figure 3.8: Energy efficiency vs. Frequency (ω_c).

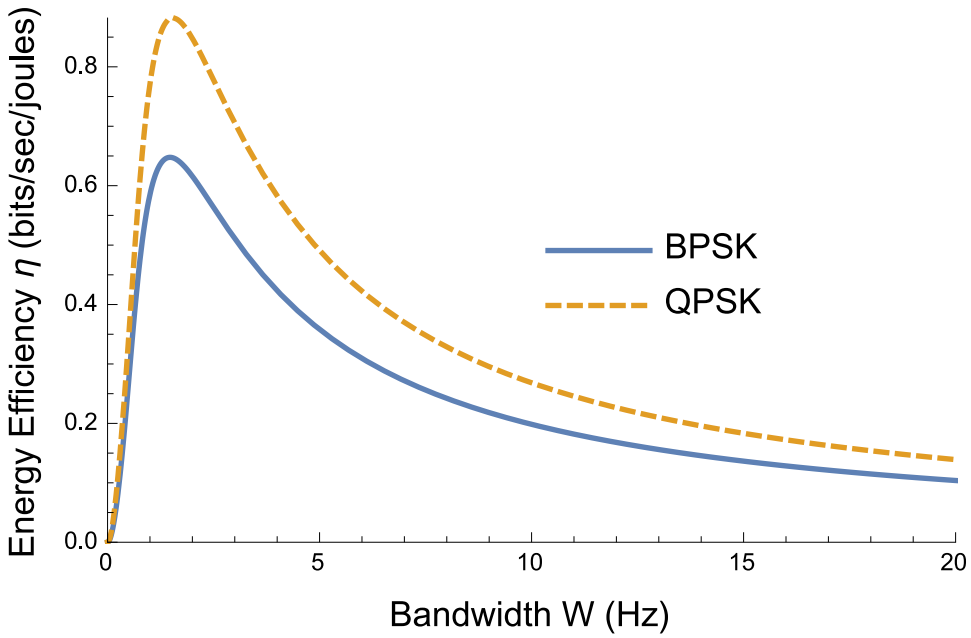


Figure 3.9: Energy efficiency vs. Bandwidth (W).

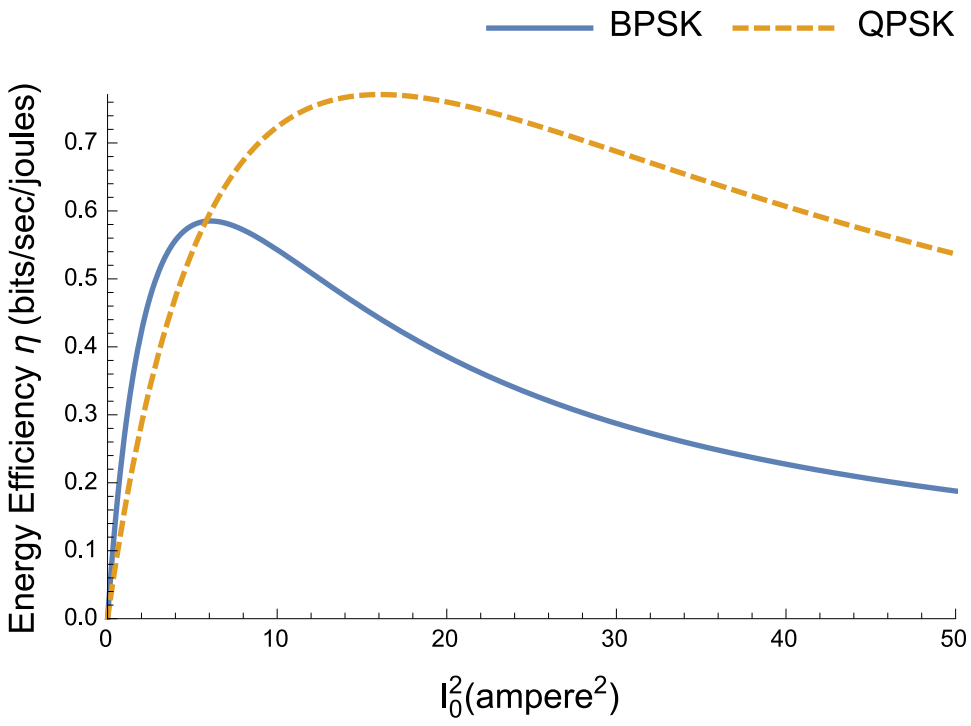


Figure 3.10: Energy efficiency vs. I_0^2.

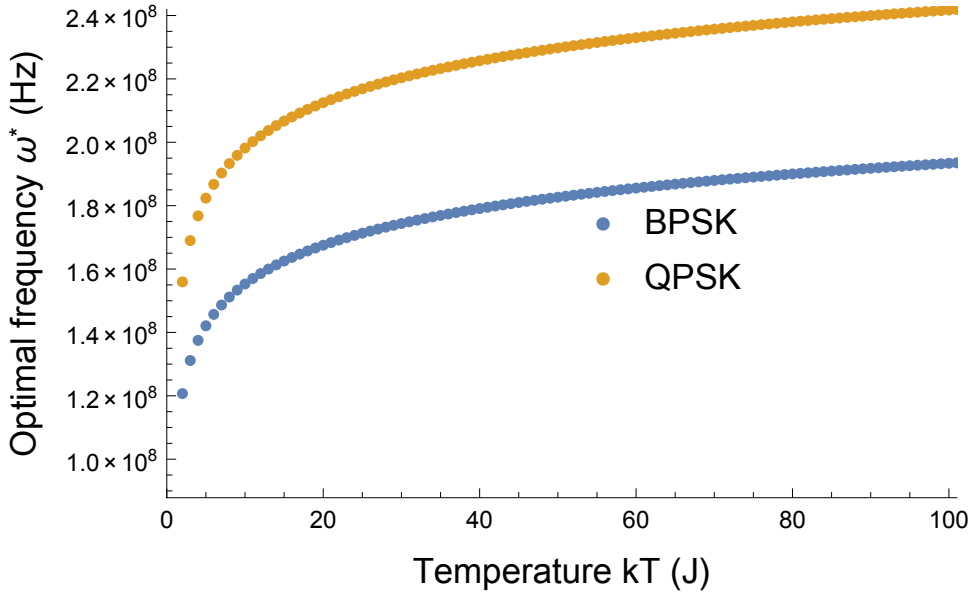


Figure 3.11: ω_c^* vs. Temperature (T).

Part II

Architectural Limits

Chapter 4

Energy Aspects of a Single-Tier Poisson Cellular Network

4.1 Overview

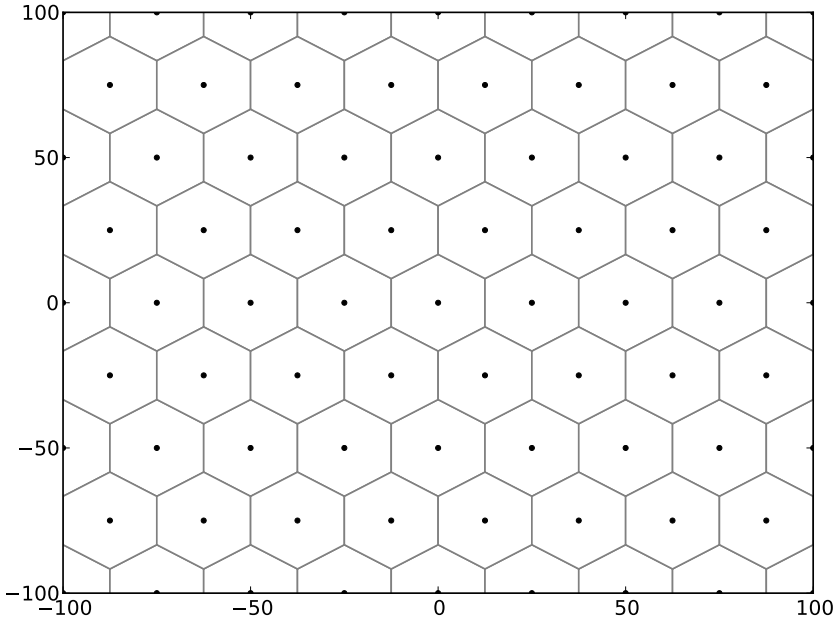
In this chapter, we investigate the problem of minimizing the area power consumption in wireless cellular networks. We focus on the downlink of a single-tier network, in which the locations of base stations (BSs) are distributed according to a homogeneous Poisson point process (PPP). Assuming that a mobile user is connected to its strongest candidate BS, we derive bounds on the optimal transmit power in order to guarantee a certain minimum coverage and data rate. Under the same QoS constraints, we find the optimal network density that minimizes the APC. Our results show that the existence of an optimal BS density for minimizing the power consumption depends on the value of the pathloss exponent.

We begin by considering a single-tier cellular network in which the BS locations are modeled according to a homogeneous spatial PPP. We assume that a mobile user is connected to the BS that provides the highest SINR and we impose two QoS constraints, namely a target coverage probability and a target minimum average rate experienced by the typical user. We aim at deriving the optimal BS density that maximizes the power efficiency, i.e. minimizes the power consumption per unit area. Evidently, a network is power efficient if the area power consumption decreases with increasing the BS density or reducing the cell size. Most prior works analyzed the performance of single-tier or heterogeneous Poisson cellular networks in terms of energy efficiency [27, 42, 43]. The most related works to our single-tier analysis is [44], in which the authors analyze the impact of transmit power reduction on the area power consumption of the network under closest BS association. In this thesis, under strongest BS association, we derive bounds on the optimal transmit power in order to guarantee a certain minimum coverage and data rate. Under the same QoS constraints, we find the optimal network density that minimizes the area power consumption, whose existence depends on the pathloss exponent and the target QoS guarantees.

4.2 System Model

4.2.1 Network model

We consider the downlink of a single-tier cellular network, in which the locations of BSs are distributed on a two-dimensional Euclidean plane $\mathbb{R}^2$ according to a homogeneous PPP $\Phi = \{r_i\}_{i\in\mathbb{N}}$ with density λ_b, where we denote by $r_i \in \mathbb{R}^2$ the location of the i-th BS. We assume that the users are also randomly



BS deployment modeled as a fixed regular lattice.

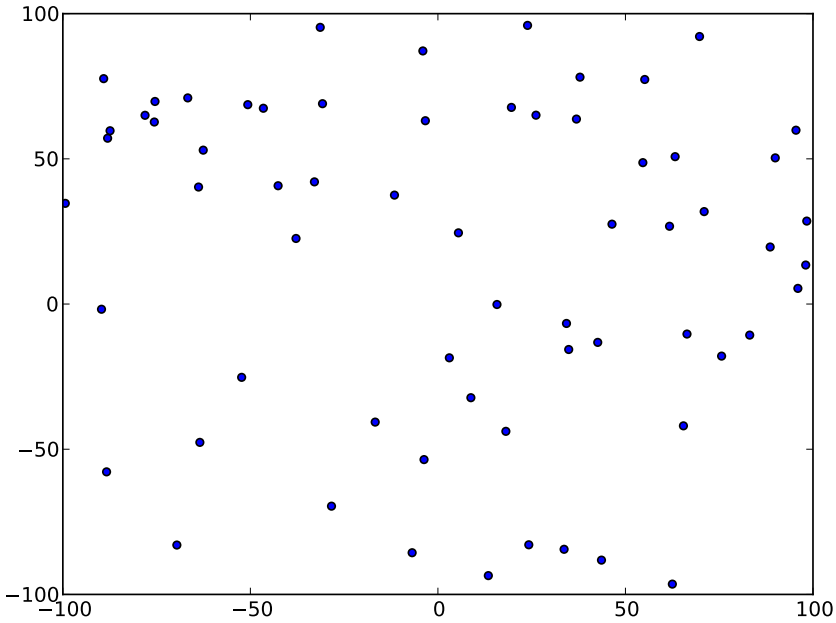


Figure 4.2: BS deployment modeled as a homogeneous Poisson point process

The above figure depicts a snapshot of a cellular network in which the same no. of BSs are deployed in a given area assuming a hexogonal lattice model (above) and a homogeneous PPP (below).

distributed according to an independent PPP of density λ_u, such that $\lambda_u \geq \lambda_b$. Without loss of generality, we focus on a mobile (typical) user at the origin for calculating the performance metrics of interest, i.e. coverage probability and average rate. The total bandwidth is denoted by B and the bandwidth per user is given by $B_u = B\frac{\lambda_b}{\lambda_u}$. We assume that all BSs transmit with the same constant power P and an additional operational power P_o (e.g. due to hardware and signaling) is consumed at each BS. In such a system, the power expenditure per unit area (in $\mathrm{W/m^2}$), coined as APC, is given as

$$\mathcal{P} = \lambda_b(P + P_o). \tag{4.1}$$

We model the system under consideration such that the transmitted signal from a given BS is subject to two propagation phenomena before it reaches the user: (i) a distance-dependent pathloss governed by the pathloss attenuation function $g(r) = br^{-\alpha}$, where b is the pathloss coefficient and α is the pathloss exponent (ii) Rayleigh fading with mean 1. According to the above assumptions and notation, the signal strength from i-th BS as received by the reference user is given as

$$p_i(r_i) = h_i P b r_i^{-\alpha}. \tag{4.2}$$

We further assume the presence of noise in the medium with power variance

$$\sigma^2 = \beta\lambda_b,$$

with $\beta = B\frac{1}{\lambda_u}\frac{FkT}{b}$, where F is the receiver noise figure, k is Boltzmann constant, and T is the ambient temperature. If the reference user is connected to the i-th BS, it receives a signal of power $p_i(r_i)$ from it. The sum of the received powers from the remaining BSs contributes to interference to this signal. As a result, the received SINR at the reference user when served by the i-th BS is given by

$$\mathrm{SINR} = \frac{h_i g(r_i) P}{\sigma^2 + I_i}, \tag{4.3}$$

where, $I_i = \sum_{r_j \in \Phi \setminus r_i} p_j(r_j)$.

In a downlink scenario, although the reference user can practically be served from any BS, a connection with a particular BS has to be established according to an association policy to ensure high SINR, which may have an impact on the APC optimization. In this thesis, we assume that a mobile user connects to the strongest BS, i.e. the BS that provides the maximum received SINR. The reference user is said to be *covered* when there is at least one BS that offers an SINR $> \gamma$. If no BS offers an SINR greater than this threshold, we say that the reference user is on coverage outage. We assume the condition $\gamma > 1$, which is needed to ensure that there is only one BS that serves the required level of SINR at a given instant [34].

4.2.2 Problem formulation

Our objective is to obtain the optimal BS density λ_b and transmit power P so that the area power consumption $\lambda_b(P + P_0)$ is minimized subject to a minimum coverage probability constraint and a minimum data rate guarantee. As stated above, the coverage probability, P_{cov}, is defined as the probability that the reference user is covered. Following the aforementioned definition, the coverage probability is given as the probability that the SINR received by the reference user on an average is greater than γ. The user data rate, R is defined as the expectation value $B\frac{\lambda_b}{\lambda_u}\mathbb{E}\left[\log_2\left(1 + \mathrm{SINR}\right)\right]$.

4.3 Optimal power for target coverage and rate

In this section we address the following optimization problem:

$$\begin{aligned} &\underset{P\in(0,\infty)}{\arg\min} \quad \mathcal{P} = \lambda_b(P+P_o) \\ &\text{s.t.} \qquad \begin{cases} (i) \quad P_{\text{cov}} \geq \epsilon P_{\text{cov}}^{\text{NN}} \\ (ii) \quad R \geq \delta R^{\text{NN}} + R_{\min} \end{cases} \end{aligned} \tag{4.4}$$

where $P_{\text{cov}}^{\text{NN}}$ and R^{NN} are the coverage probability and the per-user rate, respectively at the no noise regime, and ϵ, δ are both positive and ≤ 1. Note that when the transmit power is infinity (no noise case), the coverage probability is scale invariant, i.e. the coverage probability and also the spectral efficiency do not depend on the BS density [14].

Lemma 1. *If P_c^* is the minimum transmit power that satisfies the constraint $P_{cov} \geq \epsilon P_{cov}^{NN}$, then $P_c^* \geq \frac{A_1}{\lambda_b^{\frac{\alpha}{2}-1}}$, where $A_1 = \frac{\beta\Gamma(1+\frac{\alpha}{2})}{bC^{\frac{\alpha}{2}}(\alpha)(1-\epsilon)}$ and $C(\alpha) = \frac{2\pi^2}{\alpha}\text{cosec}\frac{2\pi}{\alpha}$.*

Proof. The coverage probability under strongest BS association for a general pathloss function $g(r)$ is given as [34, 45]

$$\begin{aligned} P_{\text{cov}}(P,\lambda_b) &= \mathbb{P}[\text{SINR} \geq \gamma] \\ &= \pi\lambda_b \int_0^\infty \exp(-q(\gamma,\lambda_b,r))\, dr \end{aligned} \tag{4.5}$$

where

$$q(P,\lambda_b,r) = \frac{\gamma\sigma^2}{Pg(\sqrt{r})} + \lambda_b \int_0^\infty \frac{\pi\gamma g(\sqrt{r_i})}{g(\sqrt{r}) + \gamma g(\sqrt{r_i})}\, dr_i. \tag{4.6}$$

We use the standard power-law pathloss model $g(r) = br^{-\alpha}$ and incorporate it in (4.6) and (4.5) to get the expression for coverage probability as

$$P_{\text{cov}}(P,\lambda_b) = \pi\lambda_b \int_0^\infty \exp\left[-\lambda_b \frac{\gamma\beta r^{\frac{\alpha}{2}}}{Pb} - \lambda_b C(\alpha)\gamma^{\frac{2}{\alpha}} r\right] dr, \tag{4.7}$$

where $C(\alpha) := \frac{2\pi^2}{\alpha}\text{cosec}\frac{2\pi}{\alpha}$. In the case of low noise ($\sigma^2 \to 0$), the above expression can be simplified by using the approximation $e^{-x} \approx 1 - x$.

$$\begin{aligned} P_{\text{cov}}(P,\lambda_b) &\approx \pi\lambda_b \int_0^\infty \left(1 - \frac{\gamma\beta\lambda_b r^{\frac{\alpha}{2}}}{Pb}\right) e^{-\lambda_b C(\alpha)\gamma^{\frac{2}{\alpha}} r}\, dr \\ &= P_{\text{cov}}^{\text{NN}} \left(1 - \frac{\beta\Gamma(1+\frac{\alpha}{2})}{Pb\lambda_b^{\frac{\alpha}{2}-1} C^{\frac{\alpha}{2}}(\alpha)}\right) \end{aligned} \tag{4.8}$$

where $P_{\text{cov}}^{\text{NN}} := \frac{\pi}{\gamma^{\frac{2}{\alpha}} C(\alpha)}$ is the coverage probability observed by the reference user in the case of negligible noise. Fig. 4.4 gives a justification to the above approximation by comparing the numerical plots of coverage probability before and after the approximation. By substituting the expression for P_{cov} from (4.8) into the constraint equation $P_{\text{cov}} \geq \epsilon P_{\text{cov}}^{\text{NN}}$ we get a condition on the range of optimal transmit power P_c^* as

$$P_c^* \geq \frac{A_1}{\lambda_b^{\frac{\alpha}{2}-1}}, \tag{4.9}$$

where $A_1 = \frac{\beta\Gamma(1+\frac{\alpha}{2})}{bC^{\frac{\alpha}{2}}(\alpha)(1-\epsilon)}$. This equation establishes the approximate minimum transmit power as a function of λ_b that satisfies the coverage constraint. □

Lemma 2. *If P_r^* is the minimum transmit power to satisfy $R \geq R_{min} + \delta R^{NN}$, then $P_r^* \geq \frac{A_2(\lambda)}{\lambda_b^{\frac{\alpha}{2}}}$, where $A_2(\lambda_b) = \frac{\sigma^2 \Gamma(1+\frac{\alpha}{2})}{bC^{\frac{\alpha}{2}}(\alpha)(1-\delta B \frac{\lambda_b}{\lambda_u})}$.*

Proof. The per-user rate is analytically given as

$$\begin{aligned} R &= B\frac{\lambda_b}{\lambda_u}\mathbb{E}[\log_2(1+\mathrm{SINR})] \\ &= R_{\min} + B\frac{\lambda_b}{\lambda_u}\int_{\ln 2}^{\infty}\mathbb{P}[\mathrm{SINR} \geq e^t - 1]\,dt, \end{aligned} \tag{4.10}$$

where $R_{\min}$ is the minimum rate [34].

As in eq (4.10), the rate per BS experienced by the reference user is given by

$$\begin{aligned} R &= R_{\min} + B\frac{\lambda_b}{\lambda_u}\mathbb{E}[\log(1+\mathrm{SINR})] \\ &= R_{\min} + B\frac{\lambda_b}{\lambda_u}\int_{t>\ln 2}\mathbb{P}[\mathrm{SINR} > e^t - 1]\,dt \\ &= R_{\min} + B\frac{\lambda_b}{\lambda_u}\pi\lambda_b\int_0^{\infty}\int_{\ln 2}^{\infty}\exp\Big[-\frac{\sigma^2 r^{\frac{\alpha}{2}}}{Pb}(e^t-1) \\ &\qquad -\lambda_b C(\alpha) r (e^t-1)^{\frac{2}{\alpha}}\Big]\,dr\,dt \\ &\approx R_{\min} + B\frac{\lambda_b}{\lambda_u}\Big(\frac{\pi}{C(\alpha)}\int_{\ln 2}^{\infty}(e^t-1)^{-\frac{2}{\alpha}}\,dt \\ &\qquad -\frac{\pi\sigma^2\lambda_b}{Pb}\int_{\ln 2}^{\infty}\frac{(e^t-1)^{-\frac{2}{\alpha}}\Gamma(1+\frac{\alpha}{2})}{C^{1+\frac{\alpha}{2}}(\alpha)\lambda_b^{\frac{\alpha}{2}}}\,dt\Big) \\ &= R_{\min} + B\frac{\lambda_b}{\lambda_u}R^{\mathrm{NN}}\left(1 - \frac{\beta\Gamma(1+\frac{\alpha}{2})}{Pb\lambda_b^{\frac{\alpha}{2}-1}C^{\frac{\alpha}{2}}(\alpha)}\right) \end{aligned} \tag{4.11}$$

where $R^{\mathrm{NN}} := \frac{\alpha\pi 2^{\frac{\alpha-2}{2}}\,{}_2F^1(\frac{2}{\alpha},\frac{2}{\alpha},\frac{2+\alpha}{\alpha},\frac{1}{2})}{C(\alpha)}$, is the rate per user when the noise is negligible. By substituting the expression for rate R in the constraint equation $R \geq \delta R^{\mathrm{NN}} + R_{\min}$ to get a condition on the optimal transmit power P_r^* as

$$P_r^* \geq \frac{A_2(\lambda_b)}{\lambda_b^{\frac{\alpha}{2}-1}}, \tag{4.12}$$

where $A_2(\lambda_b) = \frac{\beta\Gamma(1+\frac{\alpha}{2})}{bC^{\frac{\alpha}{2}}(\alpha)(1-\delta\frac{\lambda_u}{B\lambda_b})}$. This equation establishes the approximate minimum transmit power as a function of λ_b that satisfies the rate constraint. □

It follows naturally that the optimal transmit power that satisfies both the conditions (4.9) and (4.12) will therefore be $P^* = \max\{P_c^*, P_r^*\}$.

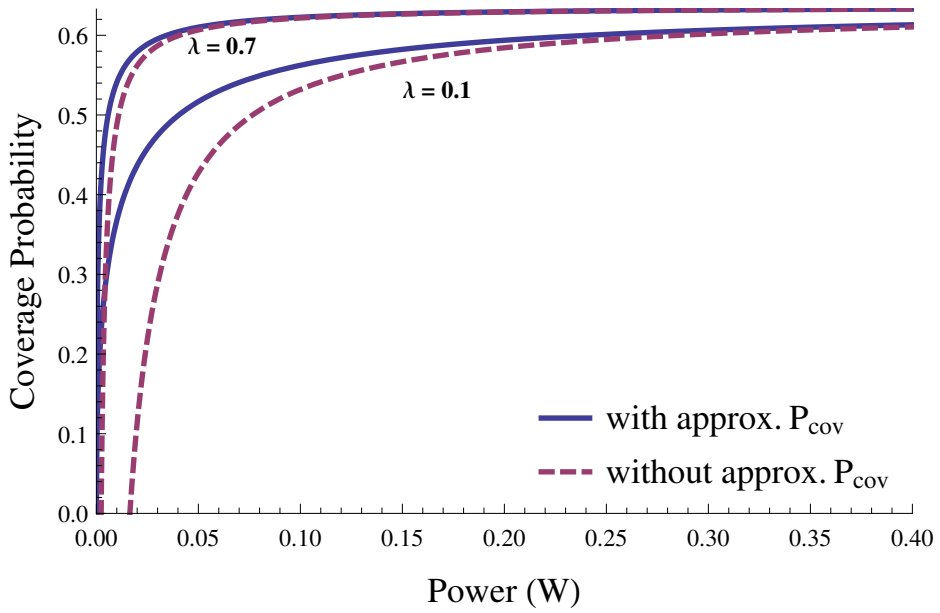


Figure 4.4: Coverage probability vs transmit power P with and without approximation (cf. (4.8)); for two different values of $\lambda_b, \beta = 10^{-3}, \lambda_u = 0.01\text{m}^{-2}, B = 20 \times 10^6\text{Hz}, \alpha = 4$. The curves almost coincide for low values of β.

4.4 Optimal BS density

Since the objective of our optimization problem is to minimize the APC ($\mathcal{P}$), we seek to minimize the function

$$\begin{aligned} \mathcal{P}(\lambda_b) &= (P^* + P_o)\lambda_b \\ &= \frac{\max\{A_1, A_2(\lambda_b)\}}{\lambda_b^{\frac{\alpha}{2}-1}} + P_o\lambda_b, \end{aligned} \tag{4.13}$$

with respect to λ_b.

Following Eq (4.13), there are two possible expressions for $\mathcal{P}$ depending on which is larger between A_1 and $A_2(\lambda_b)$, which in turn depends on the value of λ_b.

We study the two cases $A_1 > A_2(\lambda_b)$ and $A_1 < A_2(\lambda_b)$ and proceed with the optimization of $\mathcal{P}$ with respect to λ_b in each case.

Case 1

If $A_1 > A_2(\lambda_b)$ then $\lambda_b \in [\frac{\delta\lambda_u}{\epsilon B}, \infty)$ and the optimal transmit power $P^* = P_c^*$.

Therefore, the APC at optimal power follows as

$$\mathcal{P}(\lambda_b) = \frac{\beta\Gamma(1+\frac{\alpha}{2})}{bC^{\frac{\alpha}{2}}(\alpha)(1-\epsilon)} \frac{1}{\lambda_b^{\frac{\alpha}{2}-2}} + P_o\lambda_b. \tag{4.14}$$

This is clearly a convex function in λ_b for $\alpha > 2$. Therefore, we differentiate $\mathcal{P}(\lambda_b)$ with respect to λ_b and solve it for the optimum λ_b^*, i.e.

$$\frac{d\mathcal{P}(\lambda_b)}{d\lambda_b} = 0 \Rightarrow \lambda_b^* = \left[\frac{\beta\Gamma(1+\frac{\alpha}{2})}{bC^{\alpha}(\alpha)(1-\epsilon)}\frac{(\frac{\alpha}{2}-2)}{P_o}\right]^{\frac{1}{\frac{\alpha}{2}-1}}. \tag{4.15}$$

It can be noticed that $\lambda_b^* = 0$ for $\alpha = 4$, which means that when $\alpha = 4$ the optimal BS density has only the trivial solution. We comment further on the relation between the existence of optimum and the pathloss exponent α in Section 4.5, where we analyze (4.15) numerically and compare it with simulation.

Case 2

If $A_1 < A_2(\lambda_b)$ then $\lambda_b \in (0, \frac{\delta\lambda_u}{\epsilon B}]$ and the optimal transmit power $P^* = P_r^*$. Therefore, the APC at optimal power follows as

$$\mathcal{P}(\lambda_b) = \frac{\beta\Gamma(1+\frac{\alpha}{2})}{bC^{\frac{\alpha}{2}}(\alpha)(1-\frac{\delta\lambda_u}{B\lambda_b})}\frac{1}{\lambda_b^{\frac{\alpha}{2}-2}} + P_o\lambda_b. \tag{4.16}$$

We find again the optimum by equating the derivative with respect to λ_b to zero, i.e.

$$\begin{aligned} \frac{d\mathcal{P}(\lambda_b)}{d\lambda_b} &= 0 \\ \Rightarrow P_o - \frac{p\lambda_b^{-\frac{\alpha}{2}}\delta'}{\left(1-\frac{\delta'}{\lambda_b}\right)^2} + \frac{\left(1-\frac{\alpha}{2}\right)p\lambda_b^{2-\alpha/2}}{1-\frac{\delta'}{\lambda_b}} &= 0, \end{aligned} \tag{4.17}$$

where $p = \frac{\beta\Gamma(1+\frac{\alpha}{2})}{bC^{\frac{\alpha}{2}}(\alpha)}$ and $\delta' = \delta\frac{\lambda_u}{B}$.

Now, (4.17) can be simplified to the follow equation in λ_b:

$$2P_o\lambda_b^{\alpha/2}(\lambda_b-\delta)^2 - p(\alpha-4)\lambda_b^3 + p\delta(\alpha-6)\lambda_b^2) = 0, \tag{4.18}$$

which is a polynomial for $\alpha > 4$. The existence of a real solution for the polynomial depends on the value of α and the coefficients.

4.5 Simulation Results

In this section, we numerically plot the results obtained in Sections 4.3 and 4.4 and validate them with respect to simulations of our system model. A general remark is that the theoretical results match perfectly the simulated ones. We set up a square of dimension 200km × 200km and the reference user is placed at the center of the square. The number of BSs to be deployed is drawn as a Poisson random variable with its mean as the product of area of the square and the BS density (λ_b). The location coordinates of the BSs are drawn uniformly from the region of the square. Each BS is assumed to transmit at constant power and we choose the standard pathloss function $g(r_i) = r_i^{-\alpha}$. The reference user is placed at the center of the square and the received signals and interference powers are governed by (4.2), which includes the exponential random variable h_i which is the Rayleigh fading factor. In every realization, we calculate the received SINR at the reference user. We compute the coverage probability as the proportion of the number of instances (out of 2000 realizations) in which the received SINR is greater the threshold γ. Similarly the average rate per user is computed as the average value of $B\frac{\lambda_b}{\lambda_u}\log_2(1+\text{SINR})$ as experienced by the reference user.

In Figs. 4.5 and 4.6, we plot the analytical results for the coverage probability and the per-user rate (cf. 4.5) and compare it with simulations. The two plots demonstrate that both these performance metrics asymptotically saturate to a constant value rather than increasing with BS transmit power increasing. This asserts that increasing transmit power of BSs may not always be the best solution to increase the

QoS. This further motivates us to search for the minimum amount of transmit power, which ensures a minimum level of QoS.

In Fig. 4.7, we compare

- the theoretically derived expression for the approximate optimum power $P_c^*(\lambda_b)$ given in (4.9),
- the exact optimum P_c^*, numerically evaluated through exhaustive search for the least value of P that satisfies the coverage constraint of (4.4), and
- the optimum P_c^* evaluated using simulations,

as functions of λ_b. It can be noticed that the curves corresponding to theoretical exact minimum and the simulation coincide, whereas expectedly, the approximate theoretical result has a small gap with respect to the simulated exact result.

Fig. 4.8 depicts a similar treatment as described in Fig. 4.7, but for the rate constraint of (4.4). It can be noticed that the curves corresponding to theoretical exact minimum and the simulation fairly coincide while the approximate theoretical result is slightly different, as expected. This again validates the correctness of our theoretical analysis.

In Fig. 4.9, we plot for different values of α, the theoretical expression for APC ($\mathcal{P}$) (4.14) versus BS density (λ_b). We compare this with the simulation result where $\mathcal{P}$ is plotted against λ_b for values of transmit power (P) that satisfy the constraints in (4.13). Since the bandwidth is reasonably large (of the order of 10^6 Hz), the region $\lambda_b \in (0, \frac{\delta\lambda_u}{\epsilon B}]$ is very narrow and it does not have much of importance. Therefore, we only consider the region $\lambda_b \in [\frac{\epsilon\lambda_u}{\delta B}, \infty)$ in our plots. We verify that $\mathcal{P}(\lambda_b)$ has no minima for $\alpha = 4$ and it is negligibly small for $\alpha = 3$. This is a key message of our work, which dictates a relation between the pathloss exponent and the existence of a minima for the APC $\mathcal{P}$. Furthermore, we observe that in the cases of $\alpha = 5$ and 6 deploying too few BSs is not an energy efficient solution.

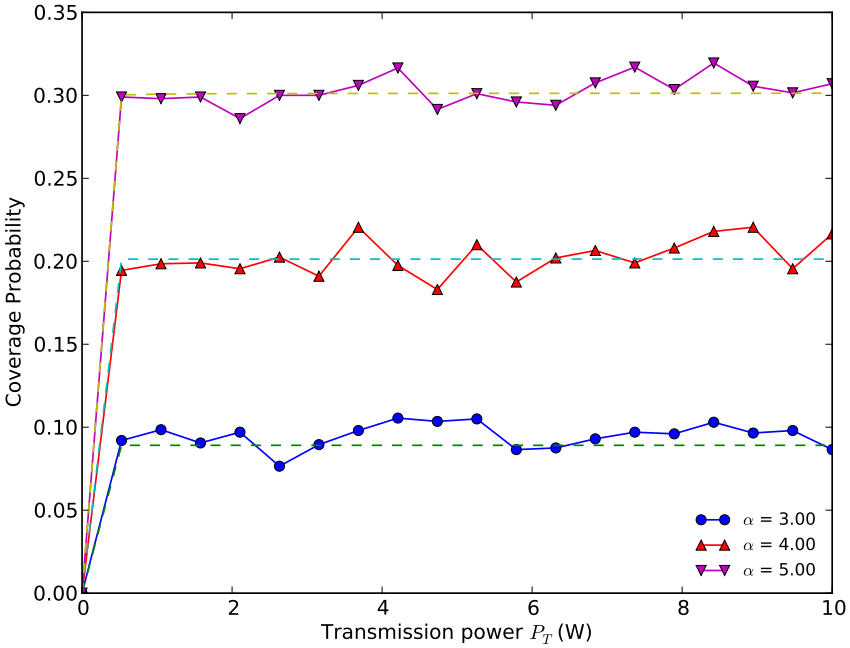


Figure 4.5: Coverage probability P_{cov} vs. transmit power P(W) for different values of α and $\beta = 2 \times 10^{-7}$. P_{cov} asymptotically saturates to a constant value with indefinite increase in P.

4.6 Concluding remarks

We have studied the problem of minimizing the power consumption in single-tier cellular wireless networks with best BS association policy. Using a low-noise approximation, we derived bounds on the minimum transmit power for achieving certain QoS constraints in terms of coverage and user rate. Based on these optimal transmit power values, we derived the optimal BS density that minimizes the area power consumption subject to minimum coverage probability and per-user rate guarantees. A takeaway message of this chapter is that the existence of an optimal BS density for optimizing area power efficiency depends on the specific value of the pathloss exponent. When the pathloss exponent $\alpha < 4$, APC has no minimum value with respect to λ_b and when $\alpha > 4$ there's always an optimal λ_b^* at which the APC assumes a minimum.

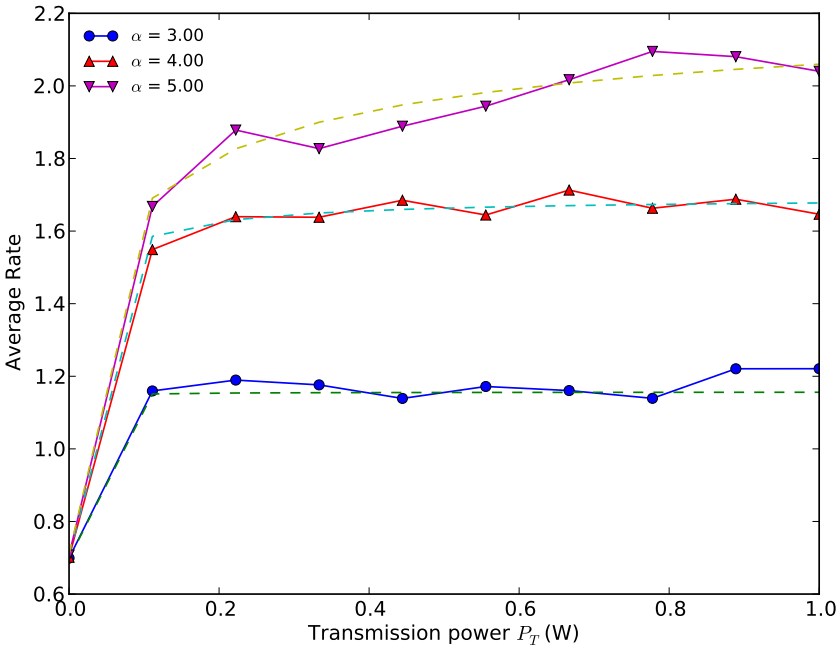


Figure 4.6: Rate (R) (bits/Hz) vs. transmit power P(W) for different values of α. R asymptotically saturates to a constant value with indefinite increase in P.

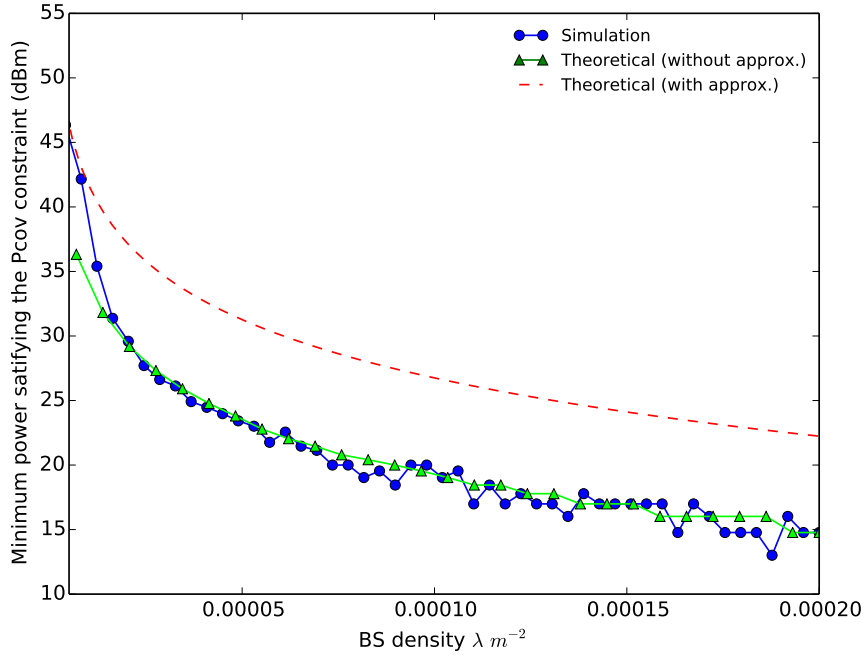


Figure 4.7: Optimal transmit power satisfying the coverage constraint (P_c^*) vs. BS density (λ_b) for $\alpha = 5$, $\epsilon = 0.6$, and $\beta = 2 \times 10^{-7}$.

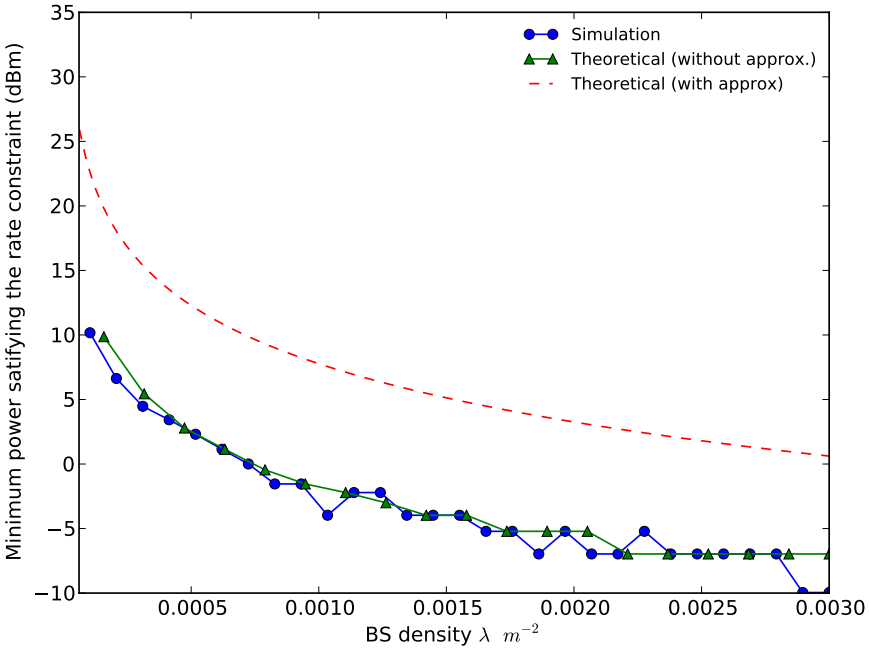


Figure 4.8: Optimal transmit power satisfying the rate constraint P_r^* vs. BS density λ_b for $\alpha = 5$, $\delta = 0.6$, $P_o = 1$W, and $\beta = 2 \times 10^{-7}$.

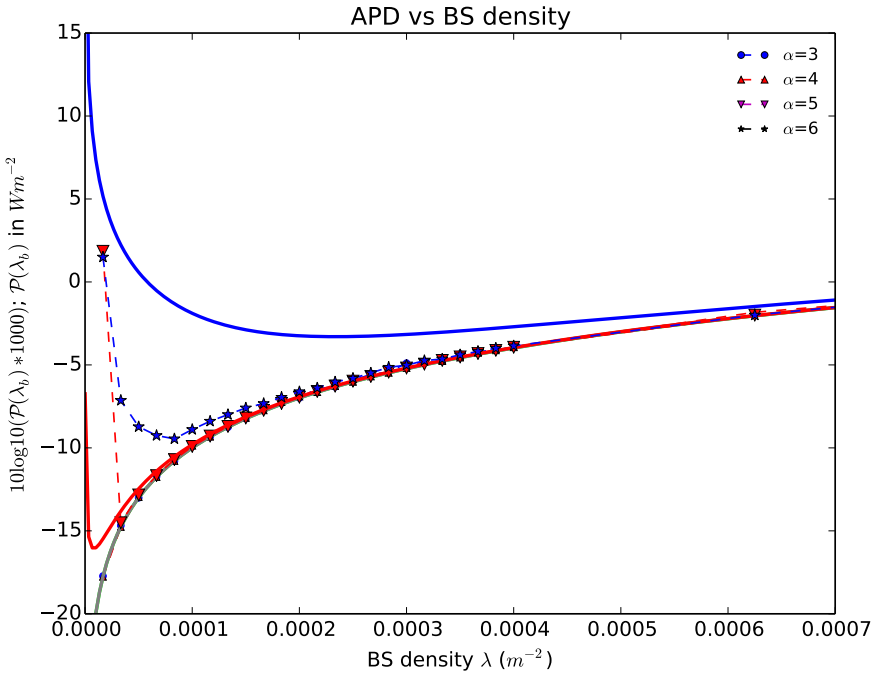


Figure 4.9: APC vs. λ_b: Continuous curves represent the theoretical result in (4.14) and broken curves represent the simulation result of APC for different values of α. $\lambda_b \in [\frac{\delta\lambda_u}{\epsilon B}, \infty)$, $\lambda_u = 1$ m^{-2}, $\gamma = 10$, $\epsilon = 0.6$, $\delta = 0.6$, $\beta = 2 \times 10^{-7}$, and $B = 20 \times 10^6$M Hz.

Chapter 5

Energy Aspects of a Single-Tier Poisson Cellular Network with Caching Capabilities

5.1 Overview

Endowed with context-awareness and proactive capabilities, caching users' content locally at the edge of the network is able to cope with increasing data traffic demand in 5G wireless networks. In this chapter, we focus on the energy consumption aspects of cache-enabled wireless cellular networks, specifically in terms of area power consumption (APC) and energy efficiency (EE). We assume that both base stations (BSs) and mobile users are distributed according to homogeneous Poisson point processes (PPPs) and we introduce a detailed power model that takes into account caching. We study the conditions under which the area power consumption is minimized with respect to BS transmit power, while ensuring a certain quality of service (QoS) in terms of coverage probability. Furthermore, we provide the optimal BS transmit power that maximizes the area spectral efficiency per unit total power spent. The main takeaway of this analysis is that caching seems to be an energy efficient solution.

The idea of caching at the edge of the network, namely at the level of BSs and UTs, have been highlighted in various works, including edge caching [46], FemtoCaching [47], and proactive caching [48]. Spatially random distributed cache-enabled BSs are modeled in [49], and expressions for the outage probability and average delivery rate are derived therein. Another stochastic framework but in the context of cache-enabled Device-To-Device (D2D) communications is presented in [50], studying performance metrics that quantify the local and global fraction of served content requests. Additionally, multi-armed bandit problem under unknown content popularity [51], an approximation framework based on the facility location problem [52], a many-to-many matching game formulation [53] and information-theoretic local and global caching gains [54] are studied. The performance of simple caching, replication and regenerating codes for D2D scenario [55], a joint content-aware user clustering and content caching [56] and joint optimization of power and cache control for cooperative MIMO [57] are investigated. Although most prior works (see [58] for a recent survey) deals with different aspects of caching, i.e. performance characterization and approximate algorithms, the energy consumption behavior of cache-enabled BSs in densely deployed scenarios has not been investigated. This is precisely the focus of this analysis.

5.2 System Model

We consider the downlink scenario of a single-tier cellular network in which the BSs are distributed on the two-dimensional Euclidean plane $\mathbb{R}^2$ according to a homogeneous spatial PPP with density λ_b denoted by $\Phi_b = \{r_i\}_{i\in\mathbb{N}}$, where $r_i \in \mathbb{R}^2$ is the location of the i-th BS. The mobile UTs (or users) are modeled to be distributed on the same plane according to an independent homogeneous PPP with density $\lambda_u > \lambda_b$, which is denoted by $\Phi_u = \{s_j\}_{j\in\mathbb{N}}$ with $s_j \in \mathbb{R}^2$ being the location of the j-th user. Without loss of generality, we focus on a *typical* UT placed at the origin of the coordinate system for calculating the key performance metrics of interest.

In this system model, we assume that the signal transmitted from a given BS is subject to two propagation phenomena before reaching a user: (i) a distance dependent pathloss governed by the pathloss function $g(r) = br^{-\alpha}$, where b is the pathloss coefficient and α is the pathloss exponent, and (ii) Rayleigh fading with mean 1. Therefore, the signal strength from the i-th BS as received by the typical user can be expressed as

$$p_i(r_i) = h_i P b r_i^{-\alpha}, \tag{5.1}$$

where the random variable h_i denotes the power of Rayleigh fading and P is the transmitted power. In addition, we assume that background noise is present in the system, with variance $\sigma^2 = \beta\lambda_b$, with $\beta = B\frac{1}{\lambda_u}\frac{FkT}{b}$, where B is the total available bandwidth, F is the receiver noise figure, k is Boltzmann constant, and T is the ambient temperature.

As alluded earlier, the typical user connected to the i-th BS receives a signal of power $p_i(r_i)$. This in turn means that the sum of the received powers from the rest of the BSs contributes to the interference to this signal. As a result, the received SINR at the typical user is given by

$$\text{SINR} = \frac{h_i g(r_i) P}{\sigma^2 + I_i}, \tag{5.2}$$

where $I_i = \sum_{r_j\in\Phi_b\setminus r_i} p_j(r_j)$ is the cumulative interference experienced from all the BSs except the i-th BS.

In a downlink scenario, although the typical user may technically be served from any BS, a connection with a particular BS has to be established according to a association policy that satisfies a certain performance metric, e.g. ensuring QoS. In this analysis, we assume that a mobile user connects to the BS that provides the maximum SINR. This is formally expressed as

$$\max_{r_i\in\Phi_b} \text{SINR}(i) > \gamma,$$

where γ is the target SINR. Given the above definition, the typical user is said to be *covered* when there is at least one BS that offers an SINR $> \gamma$. If not, we say that the typical user is not covered. We assume that $\gamma > 1$, which is needed to ensure that there is at maximum one BS that provides the highest SINR for a user at a given instant [59].

5.2.1 Coverage

Among the main objectives of wireless networks is to guarantee a certain QoS for the users. The choice of a specific QoS metric influences both the complexity and the operation of the network. In this chapter, we consider the fact that coverage probability is one measure of QoS as it is related with user experience and satisfaction. We impose a constraint to assure its value to be above a threshold limit. The coverage probability for strongest BS association for a general pathloss function $g(r)$ is given as [31, 59]

$$\begin{aligned} P_{\text{cov}}(P,\lambda_b) &= \mathbb{P}[\text{SINR} \geq \gamma] \\ &= \pi\lambda_b \int_0^\infty \exp(-q(P,\lambda_b,r))\, dr, \end{aligned} \tag{5.3}$$

where

$$q(P,\lambda_b,r)=\frac{\gamma\sigma^2}{Pg(\sqrt{r})}+\lambda_b\int_0^{\infty}\frac{\pi\gamma g(\sqrt{r_i})}{g(\sqrt{r})+\gamma g(\sqrt{r_i})}\,dr_i. \tag{5.4}$$

Denoting by $P_{\text{cov}}^{\text{NN}}$ the coverage probability in the case of no noise ($\sigma^2\to 0$), the expression for the coverage probability in (5.3) may be approximated in a low noise regime as [60]

$$P_{\text{cov}}(P,\lambda_b)=P_{\text{cov}}^{\text{NN}}\left(1-\frac{A'}{\lambda_b^{\frac{\alpha}{2}-1}}\frac{1}{P}\right), \tag{5.5}$$

where $A'=\frac{\beta\Gamma(1+\frac{\alpha}{2})}{b[C(\alpha)]^{\frac{\alpha}{2}}}$ and $C(\alpha)=\frac{2\pi^2}{\alpha}\csc\left(\frac{2\pi}{\alpha}\right)$.

The specific constraint we consider here is that the coverage probability experienced by a typical user has to be always greater than the coverage probability achieved in the absence of noise, i.e.

$$P_{\text{cov}}\geq P_{\text{cov}}^{\text{NN}}. \tag{5.6}$$

As a consequence, it can be shown (from (5.5) and (5.6)) that the optimal BS density λ_b^* that guarantees the above QoS constraint is

$$\lambda_b^*\geq\frac{A}{P^{\frac{2}{\alpha-2}}}, \tag{5.7}$$

with $A=A'^{\frac{2}{\alpha-2}}$. We use this expression in the optimization problem we formulate in Section 5.3.

Once the typical user is covered with a certain SINR, the traffic requests of this typical user have to be satisfied from its BS, by bringing the content from its source on the Internet via the backhaul. In practice, even though the typical user is covered and benefits from high SINR, it is obvious that any kind of bottleneck in the backhaul may result into long delays to the content, degrading the overall quality of experience (QoE). As part of dealing with this bottleneck, we assume that the BSs are able to store the users' (popular) content in their caches, so that requests can be satisfied locally, without passing over the limited backhaul. This is detailed in the following section.

5.2.2 Cache-enabled Base Stations

Several studies have shown that multiple users actually access the same content very frequently. Take for instance some popular TV shows, the case of viral videos with over a billion viewings, news blogs, online streaming, etc. In this context, the network will be inundated with requests for the same content that might largely increase the latency or, eventually, congest the network itself. Otherwise stated, certain types of content (or information) are relatively more popular than others and are requested/accessed more often by the users [61]. Therefore, it is reasonable to assume that a user's choice distribution matches with the global content popularity distribution. As mentioned before, the logic behind having cache-enabled BSs is to exploit this likelihood and store locally at the BSs serving a typical user the content with highest demand (popular demand) so that both users and service providers get an incentive when a popular request is made.

Let us assume that each BS is equipped with a storage unit (hard disk) which caches popular content. Since the storage capacity cannot be infinite, we assume that at each BS a set of content up to f_0 (the catalog) is stored on the hard disk. Rather than caching uniformly at random, a smarter approach would be to store the most popular content according to the given global content popularity statistics. We model the content popularity distribution at a typical user to be a right continuous and monotonically decreasing Probability Distribution Function (PDF), denoted as [62]

$$f_{\text{pop}}(f,\eta)=\begin{cases}(\eta-1)\,f^{-\eta}, & f\geq 1,\\ 0, & f<1,\end{cases} \tag{5.8}$$

where f is a point in the support of the corresponding content, and η represents the steepness of the popularity distribution curve. We define the steepness factor to be the (average) number of users per BS, that is $\eta = \frac{\lambda_u}{\lambda_b}$. The justification for the above model is that the higher the number of users attached to a BS, the more accurately the trend is sampled, hence the more content is sorted towards the left of the distribution thereby making it steeper. Moreover, since $\lambda_u > \lambda_b$, we have that $\eta > 1$.

Now, given the fact that BSs cache the catalog according to the content popularity distribution in (5.8), the probability that a content demanded by a connected user falls within the range $[0, f_0]$ is given by

$$\begin{aligned} \mathbb{P}_{\text{hit}} &= \int_0^{f_0} f_{\text{pop}}(f, \eta)\, df \\ &= \int_0^{f_0} (\eta - 1) f^{-\eta}\, df \\ &= 1 - f_0^{1-\eta}. \end{aligned} \tag{5.9}$$

It can be verified that $\mathbb{P}_{\text{hit}}$ converges to 1 when $f_0 \to \infty$, namely when the catalog stored in the BSs goes to infinity. Consequently, the probability that a request is missing from the catalog can be expressed as $\mathbb{P}_{\text{miss}} = f_0^{1-\eta}$. An illustration of the system model is given in Fig. 5.1, including snapshots of PPPs and visualization of the content popularity distribution. In the following, we introduce a power model which takes into account the caching capabilities at BSs, and will be used for investigating the energy aspects of cache-enabled BS deployment.

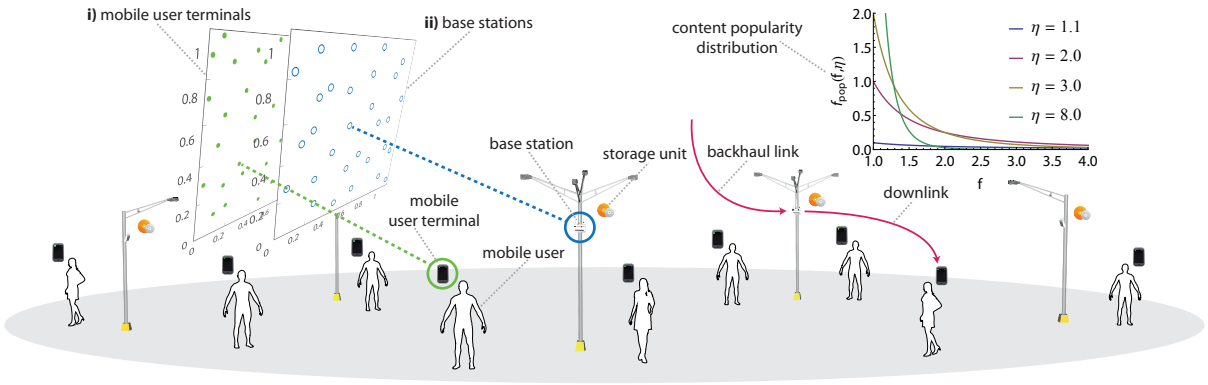


Figure 5.1: An illustration of the considered system model. The snapshots of PPPs for i) mobile users and ii) base stations on unit area are given on the left side. The content popularity distribution for different values of η is given on the top right side, showing that lower values of η corresponds to a more uniform behavior.

5.2.3 Power Consumption Model

In the context of studying caching at the edge, there is a large scope for adopting an extensive power model that takes into account various detailed factors. In this analysis, we deliberately restrict ourselves to a basic power model as a first attempt to relate energy efficiency to a cache enabled wireless network. We address a more complete model in an extension of this analysis presented in Chapter 6. We consider two different power models depending on whether BSs have caching capabilities or not.

With caching

The components of the total power consumed at an operating BS is given as follows:

i) A constant transmit power, P.

ii) An operational charge at each BS, P_o.

iii) Power needed to retrieve data from the local hard disk when a content from the catalog is requested, P_{hd}.

iv) Power needed to retrieve data from the backhaul when a content outside the catalog is requested, P_{bh}.

We assume that $P_{\text{bh}} > P_{\text{hd}}$ motivated by the realistic constraint that it is more power consuming to utilize the backhaul connection than to retrieve stored information from the local caching/storage entity.

Therefore, the total power consumed at a given BS is sum of all the components such as

$$\begin{aligned} P_{\text{tot}}^{(\text{c})} &= P + P_o + P_{\text{bh}} \times \mathbb{P}_{\text{miss}} + P_{\text{hd}} \times \mathbb{P}_{\text{hit}} \\ &= P + P_o + P_{\text{hd}} + (P_{\text{bh}} - P_{\text{hd}}) f_0^{1-\eta} \\ &= P + P_{\text{s}} + P_{\text{d}} f_0^{1-\eta}, \end{aligned} \tag{5.10}$$

where $P_{\text{s}} = P_o + P_{\text{hd}}$ and $P_{\text{d}} = P_{\text{bh}} - P_{\text{hd}}$.

Without caching

In the absence of caching, the BS has to retrieve the requested content from the backhaul every service timeslot. This is equivalent to the case where $\mathbb{P}_{\text{hit}} = 0$ (or $\mathbb{P}_{\text{miss}} = 1$). The total power consumed at a given BS is given as

$$\begin{aligned} P_{\text{tot}} &= P + P_o + P_{\text{bh}} \\ &= P + P_s + P_{\text{d}}. \end{aligned} \tag{5.11}$$

5.3 Area Power Consumption

The power expenditure per unit area, also termed as APC, is an important metric to characterize the deployment and operating costs of BSs, also indicating the compatibility of the system with the legal regulations. In our system model, the APC for cache-enabled BSs is defined as

$$\mathcal{P}^{(\text{c})} = \lambda_{\text{b}} P_{\text{tot}}^{(\text{c})}. \tag{5.12}$$

In the same way, the APC of BSs with no caching capabilities is defined as $\mathcal{P} = \lambda_{\text{b}} P_{\text{tot}}$. In this section, we aim at minimizing separately the APC for both cases (caching and no caching), while satisfying a certain QoS. This can be formally written as

$$\begin{aligned} &\underset{P \in [0,\infty)}{\text{minimize}} \quad \mathcal{P}(P) \text{ or } \mathcal{P}^{(\text{c})}(P) \\ &\text{subject to} \quad P_{\text{cov}}(P, \lambda_b) \geq P_{\text{cov}}^{\text{NN}}. \end{aligned} \tag{5.13}$$

Consider first the case where BSs have caching capabilities. The following result can be obtained for the solution of the optimization problem in (5.13).

Proposition 1. *Suppose that BSs have caching capabilities, thus $\mathcal{P}^{(c)}(P)$ is the objective (utility) function in (5.13). Then, for $\alpha = 4+\epsilon$, $\epsilon > 0$, the optimal power allocation $P^\star$ that solves (5.13) is lower bounded as*

$$P^* > \frac{2P_s}{\epsilon}. \tag{5.14}$$

Proof. Using the expression for optimum $\lambda_b^\star$ from (5.7) and incorporating it in (5.12), we get the expression for APC as

$$\mathcal{P}^{(c)} = \frac{A}{P^{\frac{2}{\alpha-2}}}(P + P_{\mathrm{s}} + P_{\mathrm{d}} f_0^{1-\frac{\lambda_u}{A}P^{\frac{2}{\alpha-2}}}). \tag{5.15}$$

Let ϵ be a real number and write $\alpha = 4+\epsilon$.

$$\begin{aligned}\mathcal{P}^{(c)} &= \frac{A}{P^{\frac{2}{2+\epsilon}}}(P + P_{\mathrm{s}} + P_{\mathrm{d}} f_0^{1-\frac{\lambda_u}{A}P^{\frac{2}{2+\epsilon}}}) \\ &= A(P^{1-\frac{2}{2+\epsilon}} + P_{\mathrm{s}}P^{-\frac{2}{2+\epsilon}} + P_{\mathrm{d}}P^{-\frac{2}{2+\epsilon}} f_0^{1-\frac{\lambda_u}{A}P^{\frac{2}{2+\epsilon}}}).\end{aligned} \tag{5.16}$$

For $\epsilon \leq 0$, $\mathcal{P}^{(c)}$ is a monotonically decreasing function and no minimum point exists. However, for $\epsilon > 0$ (i.e. $\alpha > 4$), the first term in (5.16) dominates as $P \to \infty$, indicating that there exists a minimum where the derivative of the curve changes its sign from negative to positive. We set $\epsilon > 0$ for what follows.

Differentiating $\mathcal{P}^{(c)}$ with respect to P we get

$$\frac{d\,\mathcal{P}^{(c)}}{d\,P} = \frac{P^{-\frac{\epsilon+4}{\epsilon+2}} f_0^{-\frac{\lambda_u P^{\frac{2}{\epsilon+2}}}{A}}}{\epsilon+2} \times A(P\epsilon - 2P_{\mathrm{s}}) f_0^{\frac{\lambda_u P^{\frac{2}{\epsilon+2}}}{A}} - 2\lambda_u f_0 P_{\mathrm{d}} \log(f_0) P^{\frac{2}{\epsilon+2}} - 2A f_0 P_{\mathrm{d}}. \tag{5.17}$$

Given the fact that P is always positive, the derivative in (5.17) remains negative as P increases from 0 until it is sufficiently greater than $\frac{2P_{\mathrm{s}}}{\epsilon}$, after which the derivative can change its sign to positive. This indicates that there exists a minimum value for $\mathcal{P}^{(c)}$ when

$$P^* > \frac{2P_{\mathrm{s}}}{\epsilon}, \tag{5.18}$$

which concludes the proof. □

For the case where BSs have no caching capabilities, the following result is derived.

Proposition 2. *Suppose that the BSs have no caching capabilities, thus $\mathcal{P}(P)$ is the objective function in (5.13). Then, for $\alpha = 4+\epsilon$, $\epsilon > 0$, the optimal power allocation $P^\star$ that solves (5.13) is given by*

$$P^* = \frac{2(P_s + P_d)}{\epsilon}. \tag{5.19}$$

Proof. In the case without caching, by similar treatment as in the proof of Proposition 1, we write the expression for APC as

$$\mathcal{P} = \frac{A}{P^{\frac{2}{2+\epsilon}}}(P + P_{\mathrm{s}} + P_{\mathrm{d}}). \tag{5.20}$$

It can be noticed that $\mathcal{P}$ is a monotonically decreasing function and has no minimum except when $\epsilon > 0$ (or $\alpha > 4$).

Differentiating $\mathcal{P}$ with respect to P we get

$$\frac{d\,\mathcal{P}}{d\,P} = \frac{AP^{-\frac{\epsilon+4}{\epsilon+2}}\left(-2P_{\mathrm{d}} - 2P_{\mathrm{s}} + P\epsilon\right)}{\epsilon+2}. \tag{5.21}$$

By equating the derivative to zero, the optimum power

$$\frac{d\,\mathcal{P}}{d\,P} = 0 \Rightarrow P^{*} = \frac{2(P_{\mathrm{s}} + P_{\mathrm{d}})}{\epsilon}. \tag{5.22}$$

It can easily be verified that the second derivative $\frac{d^2\,\mathcal{P}}{d\,P^2} > 0$ for $P = P^{*}$. □

5.3.1 Remarks

For a given finite value of transmit power P, from (5.16) and (5.20), we observe that, in all sensible cases,

$$\mathcal{P}^{(c)} < \mathcal{P}. \tag{5.23}$$

This indicates that BSs with caching capabilities always outperform those without caching. Additionally, $\mathcal{P}^{(c)}$ can be made smaller by increasing the catalog size f_0 in BSs. This is indeed intuitively correct, as shown in Fig. 5.2 where some realistic power values from [63] are considered.

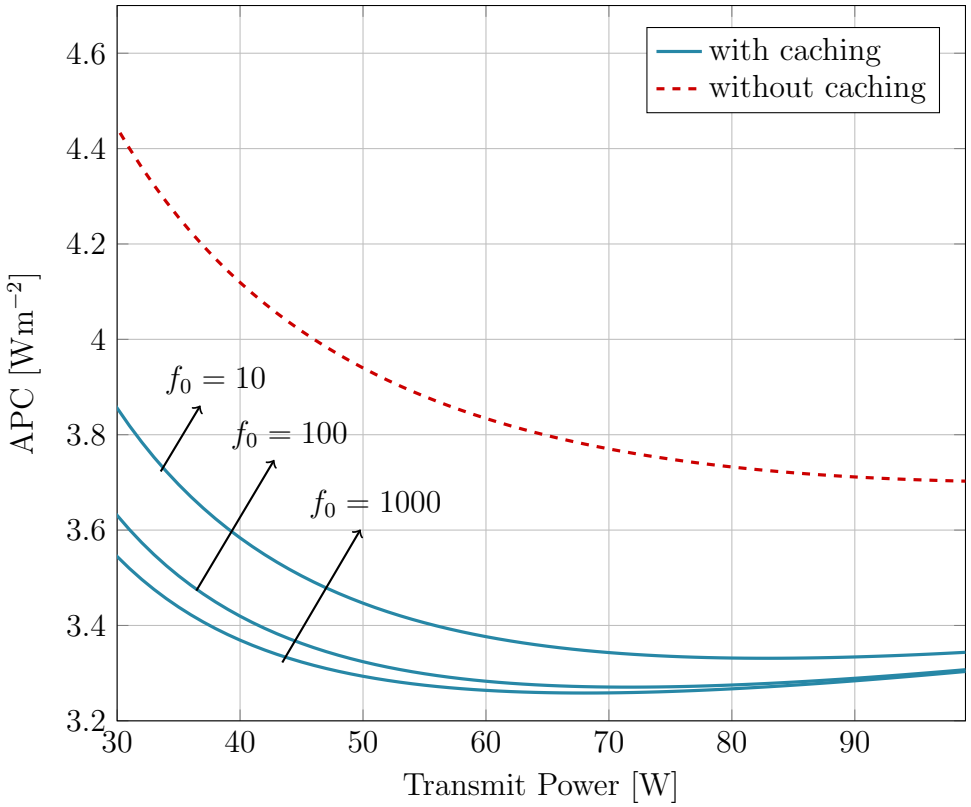


Figure 5.2: APC vs. transmit power with and without caching for values: $P_{\mathrm{s}} = 25\mathrm{W}$, $P_{\mathrm{d}} = 10\mathrm{W}$, $\beta = 1$, and $\alpha = 4.75$ in (5.15) and (5.20).

Fig. 5.3 illustrates the variation of APC with respect to the transmit power in the cases with and without caching, for different values of pathloss exponent α. It can be noticed that the APC has a

minimum for a certain power value only when $\alpha > 4$. However, APC can be significantly reduced with caching in all cases, and the performance gap between caching and no caching cases is increased for α increasing.

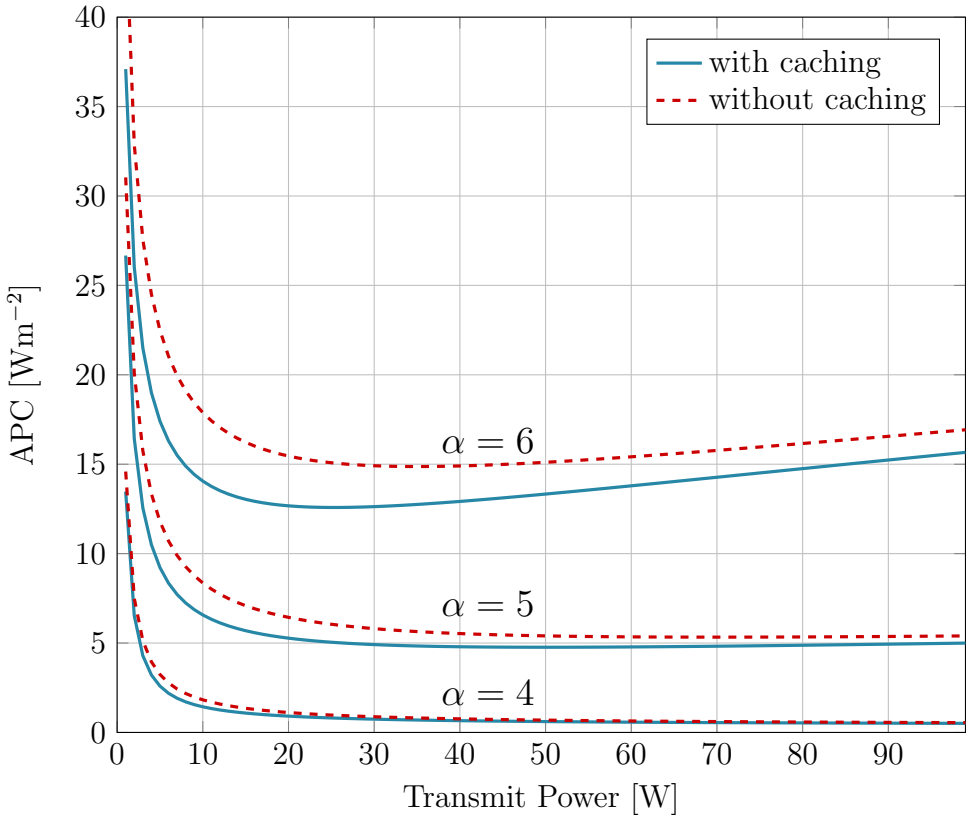


Figure 5.3: APC vs. Transmit power with and without caching for values: $P_{\mathrm{s}} = 25\mathrm{W}, P_{\mathrm{d}} = 10\mathrm{W}, f_0 = 10, A = 2$ in (5.15) and (5.20)

5.4 Energy Efficiency

Another key performance metric that should be studied is the energy efficiency, which indicates the amount of utility (throughput) that is extracted out of a unit power invested for network operation. The standard QoS factor chosen as the utility in literature is the Area Spectral Efficiency (ASE). In (5.24), we define the EE as the ratio between ASE and the total power spent at the BS, i.e.

$$\mathcal{E} = \frac{\lambda_b \log(1+\gamma) P_{\mathrm{cov}}(P, \lambda_b)}{P_{\mathrm{tot}}}. \tag{5.24}$$

In order to define more precisely the EE metric for both cases, we use the expression for coverage probability established in (5.5). With caching, the expression for EE in (5.24) therefore becomes

$$\mathcal{E}^{(c)}(P, \lambda_b) = \frac{\lambda_b \log(1+\gamma) P_{\mathrm{cov}}^{\mathrm{NN}} (1 - \frac{A}{\lambda_b^{\frac{\alpha}{2}-1}} \frac{1}{P})}{P + P_{\mathrm{s}} + P_{\mathrm{d}} f_0^{1-\frac{\lambda_u}{\lambda_b}}}. \tag{5.25}$$

In the case without caching, EE is given as

$$\mathcal{E}(P,\lambda_b)=\frac{\lambda_b \log(1+\gamma)P_{\text{cov}}^{\text{NN}}(1-\frac{A}{\lambda_b^{\frac{\alpha}{2}-1}}\frac{1}{P})}{P+P_{\text{s}}+P_{\text{d}}}. \tag{5.26}$$

The following results are given for the maximization of EE and the discussions are carried out afterwards.

Proposition 3. *Suppose that the BSs have caching capabilities and let $P_1^\star$ denote the optimal power allocation that maximizes $\mathcal{E}^{(c)}(P,\lambda_b)$. Then*

$$P_1^*=1+\sqrt{1+P_s+P_d f_0^{1-\frac{\lambda_u}{\lambda_b}}}. \tag{5.27}$$

Proof. We differentiate the expression (5.25) with respect to P and solve for the optimum power P_1^*

$$\begin{aligned}\frac{d\,\mathcal{E}^{(c)}}{d\,P}=0 \quad &\Rightarrow \quad 2P-P^2+P_{\text{s}}+P_{\text{d}}f_0^{1-\frac{\lambda_u}{\lambda_b}}=0\\ &\Rightarrow \quad P_1^*=1+\sqrt{1+P_{\text{s}}+P_{\text{d}}f_0^{1-\frac{\lambda_u}{\lambda_b}}}.\end{aligned} \tag{5.28}$$

It can be verified that the second derivative $\frac{d^2\,\mathcal{E}^{(c)}}{d\,P^2}$ is negative for the positive solution of P_1^*. □

Proposition 4. *Suppose that the BSs have no caching capabilities and let $P_2^\star$ denote the optimal power allocation that maximizes $\mathcal{E}^{(c)}(P,\lambda_b)$. Then*

$$P_2^*=1+\sqrt{1+P_s+P_d}. \tag{5.29}$$

Proof. Similar to (5.28), we can show that the optimum power $P_2^*=1+\sqrt{1+P_{\text{s}}+P_{\text{d}}}$. It can be verified that the second derivative $\frac{d^2\,\mathcal{E}}{d\,P^2}$ is negative for the positive solution of P_2^*. □

5.4.1 Remarks

Based on the above results, we can make the following observations:

1. For given positive values of λ_b and P, $\mathcal{E}^{(c)}(P,\lambda_b)$ is always higher than $\mathcal{E}(P,\lambda_b)$ making it apparent that implementing BSs with caching capabilities is an energy-efficient solution.
2. For a fixed value of λ_b, $\mathcal{E}^{(c)}$ has a maximum at $P_1^*=1+\sqrt{1+P_{\text{s}}+P_{\text{d}}f_0^{1-\frac{\lambda_u}{\lambda_b}}}$ and $\mathcal{E}(P,\lambda_b)$ has a maximum at $P_2^*=1+\sqrt{1+P_{\text{s}}+P_{\text{d}}}$. Noting that $P_1^*<P_2^*$, we observe that the optimal EE may be attained at a smaller value of transmit power in the case of cache-enabled BSs.

In Fig. 5.4 we plot the variation of EE as a function of the transmit power P. It can be seen that EE can be significantly increased (in the case of caching) by increasing the size of the catalog in BSs, namely f_0.

5.5 Concluding remarks

In this chapter, we studied how incorporating caching capabilities at the BSs affects the energy consumption in wireless cellular networks. Adopting a detailed BS power model and modeling the BS locations according to a PPP, we derived expressions for the APC and the EE, which are further simplified in the low noise regime. A key observation of this chapter is that cache-enabled BSs can significantly decrease the APC and improve the EE as compared to traditional BSs. We also observed that the existence of an optimum power consumption point for the APC depends on the pathloss exponent.

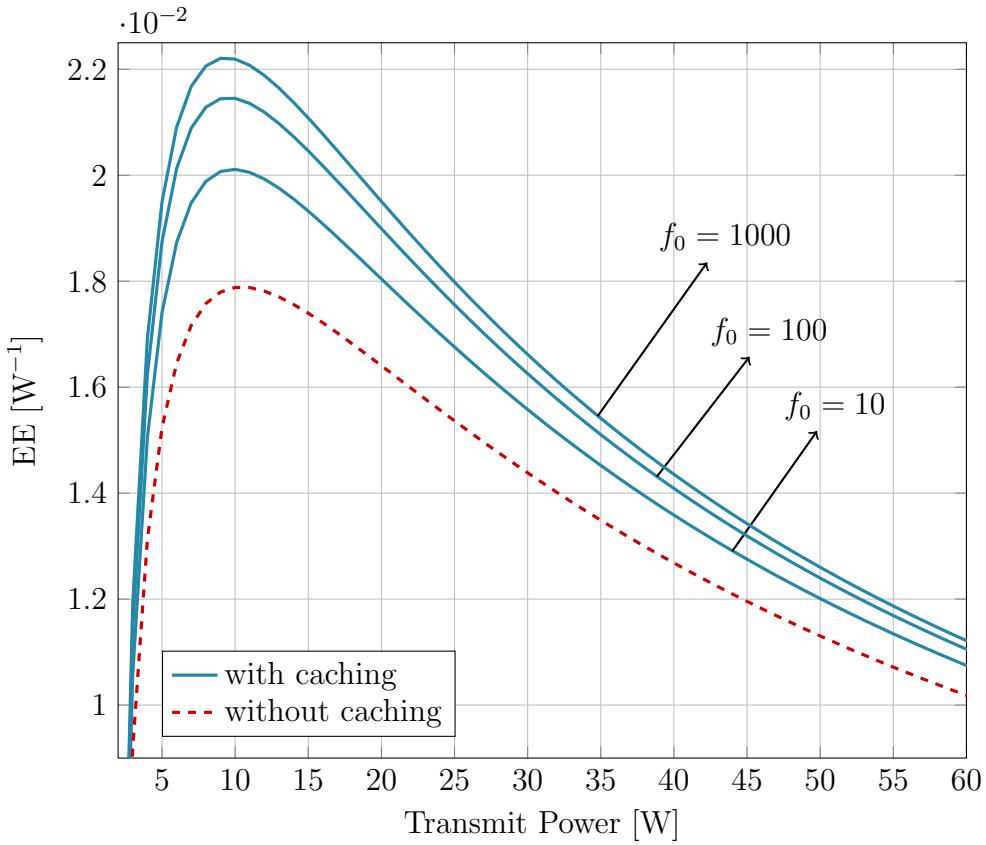


Figure 5.4: EE vs. transmit power with and without caching for values: $P_{\mathrm{s}} = 25\mathrm{W}, P_{\mathrm{d}} = 10\mathrm{W}, \alpha = 4.75, \beta = 1, P_{\mathrm{cov}}^{\mathrm{NN}} = 1, \lambda_b = 0.5, \lambda_u = 0.6,$ and $\gamma = 2$ in (5.25) and (5.26).

Chapter 6

Energy Aspects of a Two-Tier Poisson Cellular Network with Caching Capabilities

6.1 Overview

Future cellular networks are very likely to be heterogeneous with each tier handling a different task in providing services to the end user. Recently, there has been an increased focus on the study of energy saving aspects of caching users' content at the edge of a heterogeneous cellular networks. For instance, [64] presents an energy efficient power control scheme for a network with energy harvesting SBSs equipped with wireless backhaul and local storage. In [65], a joint caching and base station activation policy for saving energy in a heterogeneous cellular networks has been proposed. In [66], the proposed network design is such that the most popular content is cached at SBSs and an optimized transmission policy at the MBS and the caching policy at the SBS has been proposed in order to minimize the energy consumption in the backhaul link. [67] shows that introducing cache can improve the EE if the file catalog size is not indefinitely large and placing the cached content at the pico BSs is more energy efficient. In this part of the thesis, we explore the energy saving aspects of caching from the point of view of BS densities and transmission powers. Both these are significant in terms of deployment as well as maintenance costs.

Keeping in view the prospects of improving energy efficiency of a wireless network by incorporating caching end users' data at the edge of the network, we propose a system with two tier cellular network - one with caching capabilities and the other without. We model the deployment of these two tiers (one with small base stations and one with macro base stations) according to two independent Poisson point processes. We derive expressions for quantities such as coverage probability of the users and activation probability of base stations, and take these into account while defining key two system performance metrics, namely the Area Power Consumption (APC) and the Energy Efficiency (EE). Imposing some sensible constraints we define and solve two problems of optimizing APC and EE with respect to base stations densities and their transmit powers separately. We study how caching affects these optimization problems and show that it is indeed an energy efficient solution.

6.2 Network Model

In order to make a performance analysis of a heterogeneous cellular network, we consider the following system model:

Topology: In our network model, we have two layers of BSs each randomly deployed according to an independent PPP. In particular, the first tier is made of SBSs where they are deployed according to a PPP $\Phi_x = \{x_1, x_2 \ldots\}$ of intensity λ_x. The second tier is composed of MBSs and are also deployed according to an independent PPP $\Phi_y = \{y_1, y_2 \ldots\}$ of intensity λ_y. Additionally, users (mobile terminals) are modelled according to an independent PPP $\Phi_z = \{z_1, z_2 \ldots\}$ of intensity λ_z. Without loss of generality, we consider a *typical* mobile user at Cartesian origin $(0, 0)$ for obtaining the performance metrics of this heterogeneous network. The results derived for typical node apply to any mobile user due to the Slivnyak-Mecke theorem and the stationary and isotropy properties of the PPP [16].

Signal Model: As a worst case scenario, we suppose that small and macro cells operates in the same frequency band, and thus interfere with each other. Each small and macro cell has constant signalling transmit power of P_x and P_y respectively. For the sake of simplicity, we suppose that a base station (macro or small cell) is represented by its position. The received power experienced at the typical mobile user due to the base station x (or y in case of macro cell) is given by $P_x h_x g(x)$, where h_x is the fading power coefficient (or square of the fading amplitude) of the channel between the base station x and the typical user, and $g(x) = \|x\|^{-\alpha}$ is the singular path-loss function with $\alpha > 2$ (see [68] for details of this path-loss function). Note that the fading power coefficients are Independent and Identically Distributed (i.i.d.) exponential random variables (namely Rayleigh fading) with mean μ. Hence, received SINR of the typical user from the signaling base station x is expressed by

$$\mathrm{SINR}_x = \frac{P_x h_x g(x)}{\sigma^2 + I_x + I_y} \tag{6.1}$$

where $I_x = \sum_{x_i \in \Phi_x / x} P_x h_{x_i} g(x_i)$ and $I_y = \sum_{y_j \in \Phi_y} P_y h_{y_j} g(y_j)$ are the cumulative interference experienced from non-signalling small and macro cells respectively, and σ^2 denotes the power of white additive Gaussian noise given as:

$$\sigma^2 = W \left(\frac{\lambda_x + \lambda_y}{\lambda_z} \right) FKT \tag{6.2}$$

where W is the bandwidth, FKT is the noise figure. Finally, the target SINR, that is the minimum SINR that an SBS is expected to provide in order to establish a connection, is denoted by γ. We impose the constraint that $\gamma > 1$ as this facilitates the analysis by making it more tractable as it makes sure that only one SBS provides an SINR value above the desired threshold[59].

Connectivity Policy: In our network model, we define the connectivity policy by which the typical user initiates contact with the network. The user is either connected to a SBS (call event 1) or a MBS (call event 2). In particular, the typical user connects to 1) an SBS that provides the strongest received SINR *if* it is greater than the threshold γ , or 2) to the closest MBS if it is within a distance R_0 from the given mobile terminal irrespective of how much SINR it provides. Therefore, total coverage probability may be approximately given as the sum of probability of these two events, that is

$$\begin{aligned} \mathbb{P}_{\mathrm{cov}}(\lambda_x, \lambda_y, P_x, P_y) &\approx \mathbb{P}_x + (1 - \mathbb{P}_x)\mathbb{P}_y \\ &= \mathbb{P}_x(1 - \mathbb{P}_y) + \mathbb{P}_y \end{aligned} \tag{6.3}$$

where $\mathbb{P}_x$ is the probability of the occurrence of the first event (namely SBS connection) and $\mathbb{P}_y$ is the probability of the occurrence of second event (namely MBS connection). It has to be noted at this point that the above two events are not necessarily mutually exclusive and we only adopt an approximate treatment in this work. A more rigorous approach would take into consideration the correlation between these two events.

6.2.1 Caching Model

Many studies have pointed out that people tend to access the same content very frequently. Take for example some popular TV shows, viral videos with millions of views, news blogs and online streaming. In such conditions, the underlying network has to deliver the same content multiple times which in turn cause high latency and congestion. In cellular networks, we also know that specific types of content (or information) are highly popular than the others [61], thus it is reasonable to assume a global content popularity distribution which highly matches with users' choice distribution. In this context, as alluded to earlier, cache-enabled BSs can exploit this information by storing the popular contents locally, thus yielding higher satisfactions of mobile users and offloading the communication infrastructure of the service provides.

Suppose that each BS is assembled with a storage unit (i.e., solid-state-drive or hard disk) in order to cache popular content. Since these storage capacities are bounded in practice, we let each BS to store a set of content up to $\frac{f_0}{\lambda_x}$ (the catalog). Here f_0 is a constant representing the total catalog size of SBS deployment, and λ_x is the density of SBS. Instead of caching content uniformly at random, as mentioned before, we focus on storing the most popular content according to the given global popularity statistics. To show this, we assume that the content popularity distribution at a typical user is a right continuous and monotonically decreasing PDF, defined as [62]

$$f_{\text{pop}}\left(f,\eta\right) = \begin{cases} \left(\eta-1\right) f^{-\eta}, & f \geq 1, \\ 0, & f < 1, \end{cases} \tag{6.4}$$

where f indicates a point in the support of the corresponding content, and η characterize the steepness of the content popularity distribution. Lower values of η results in almost uniform behaviour, whereas higher η values means that more content is sorted toward the left of the distribution, thereby yielding a steeper behaviour.

Now, as we know that BSs store the catalog according to the content popularity distribution in (6.4), the *cache hit probability* becomes paramount of interest and needs be calculated. In our case, the probability that a content demanded by a connected user to be in the cache of BS (within the range $[0, f_0]$), namely cache hit probability, is given by

$$\begin{aligned} \mathbb{P}_{\text{hit}} &= \int_0^{\frac{f_0}{\lambda_x}} f_{\text{pop}}(f,\eta)\, df \\ &= \int_0^{\frac{f_0}{\lambda_x}} (\eta-1) f^{-\eta}\, df \\ &= 1 - \left(\frac{f_0}{\lambda_x}\right)^{1-\eta}. \end{aligned} \tag{6.5}$$

It can be shown that $\mathbb{P}_{\text{hit}}$ converges to 1 as $\frac{f_0}{\lambda_x} \to \infty$, namely when the catalog stored in the BSs grows to infinity. As a consequence, the probability that a demanded content is missing from the catalog is written as $\mathbb{P}_{\text{miss}} = \left(\frac{f_0}{\lambda_x}\right)^{1-\eta}$. The sketch of the system model is shown in Fig. 6.1, including realizations of PPPs for small/macro cells and users. In what follows, we present our main results on performance metrics introduce a detailed area power consumption model and an energy efficiency metric with both of them taking into consideration the caching capabilities at BSs as well as other load dependent and independent components of the network. This extensive model will be used to quantify the benefits of cache-enabled BS deployments from point of energy aspects.

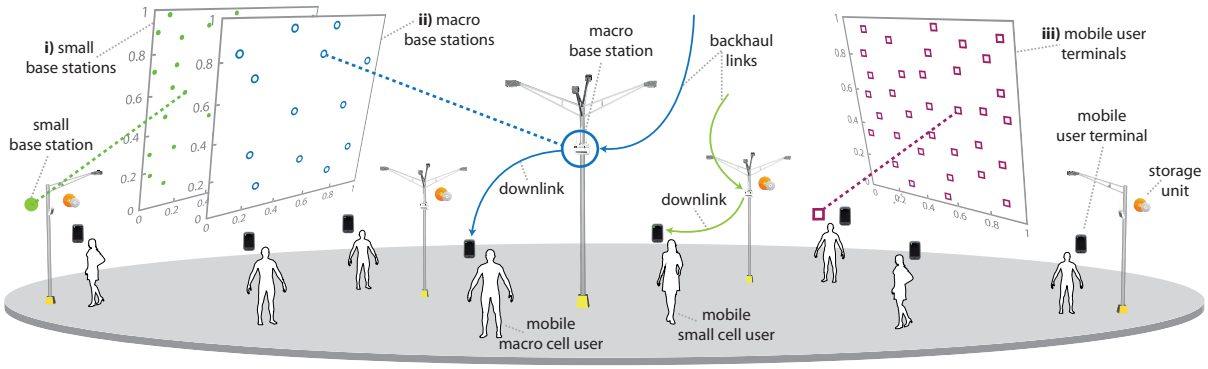


Figure 6.1: An illustration of the considered network model. The snapshots of PPPs for i) cache-enabled small base stations, ii) macro base stations, and iii) mobile user terminals.

6.3 Performance Metrics and Quality of Service

In this section, we state and prove four important results in order to establish the analytical expressions for 1) the coverage probabilities of typical user and 2) activation probabilities of small and macro BSs. We start by focusing on coverage probabilities as follows.

Lemma 3. *In a low noise regime, probability that a typical user is covered by a SBS is given by*

$$\mathbb{P}_x = \frac{\pi\lambda_x}{\mu}\left(\frac{1}{P\lambda_x + Q\lambda_y} - R(P\lambda_x + Q\lambda_y)^{-(\frac{\alpha+2}{2})}\Gamma(\frac{\alpha+2}{2})\right) \tag{6.6}$$

where $P = \frac{2\pi^2\gamma\,{}_2F_1\left(1,\frac{\alpha-2}{\alpha};2-\frac{2}{\alpha};-\gamma\right)}{\mu(\alpha-2)}$, $Q = \frac{2\pi^2(\frac{P_x}{P_y\gamma})^{-\frac{2}{\alpha}}\csc[\frac{2\pi}{\alpha}]}{\alpha\mu}$, $R = \frac{b\gamma\sigma^2}{P_x}$, *and* $\alpha > 2$ *is a sufficiently small number less than* 1. *Therein,* $F_1(.,.;.;.)$ *and* $\Gamma(.)$ *are the hyper-geometric and gamma functions respectively [69]. On the other hand, in case of no noise regime* ($R \to 0$)*, it holds that*

$$\mathbb{P}_x^{NN} = \frac{\pi\lambda_x}{P\lambda_x + Q\lambda_y}. \tag{6.7}$$

Proof. See Appendix 6.8. □

Lemma 4. *Probability that a typical user is covered by a macro cell base station is given by*

$$\mathbb{P}_y = 1 - e^{-\lambda_y\pi R_0^2}. \tag{6.8}$$

Proof. The proof can be easily shown from void probability of PPP [16]. □

Using Lemmas 3 and 4, the coverage probability under no noise regime can be straightforwardly expressed as follows.

Corollary 1. *Probability that a typical user is either covered by a small cell or macro cell is approximately given by*

$$\mathbb{P}_{\text{cov}}(\lambda_x,\lambda_y,P_x,P_y) \approx 1 + \frac{-\pi\lambda_x e^{-\lambda_y\pi R_0^2}}{P\lambda_x + Q\lambda_y} - e^{-\lambda_y\pi R_0^2} \tag{6.9}$$

where P, Q *and* Q *are defined as in Lemma 3.*

Now, we are ready to state and prove two results in order to establish the analytical expressions of the activation probabilities of a given SBS and an MBS. Note that the activation probabilities together with coverage probabilities shall be extensively used in our area power consumption and energy efficiency metrics given in the following sections. First, in the following lemma, we define the activation probability of an SBS as the probability that it offers an SINR above the threshold to at least one user and we derive its analytical expression.

Lemma 5. *The probability that a given SBS is active is given as*

$$\mathbb{P}_x^{\text{act}} \approx \frac{\mu-1}{\mu} + \frac{\pi\lambda_z}{\mu}\left(\frac{1}{P'\lambda_x + Q\lambda_y} - R(P'\lambda_x + Q\lambda_y)^{-(\frac{\alpha+2}{2})}\Gamma(\frac{\alpha+2}{2})\right) \tag{6.10}$$

where $P' = \frac{2\pi(\frac{1}{\gamma})^{-\frac{2}{\alpha}}\csc[\frac{2\pi}{\alpha}]}{\alpha\mu}$, *$Q$ and R are defined as in Lemma 3. On the other hand, in case of no noise regime ($R \to 0$), it holds that*

$$\mathbb{P}_x^{act\ NN} = \frac{\pi\lambda_z}{P'\lambda_x + Q\lambda_y}. \tag{6.11}$$

Proof. See Appendix 6.10. □

In the following lemma, we define the activation probability of an MBS as the probability that a typical user is not covered by an SBS *and* an MBS is at a distance R_0 from the considered user.

Lemma 6. *The probability that a given MBS is providing an SINR above the threshold γ to at least one user at given instance is approximately given by*

$$\mathbb{P}_y^{\text{act}} \approx (1 - \frac{\pi\lambda_x}{\mu}\left(\frac{1}{P\lambda_x + Q\lambda_y} - R(P\lambda_x + Q\lambda_y)^{-(\frac{\alpha+2}{2})}\Gamma(\frac{\alpha+2}{2})\right)) \times (1 - \exp(-\lambda_z\pi R_0^2)), \tag{6.12}$$

where P, Q and R are defined as in Lemma 3.

Proof. See Appendix 6.11. □

Having these analytical expressions above, we now focus on the definition of our main performance metrics that are related to green aspects of caching in cellular networks.

6.3.1 Area Power Consumption

The power expenditure per unit area, also coined as APC, is an crucial metric to quantize the deployment and operating costs of BSs, also required for the compatibility of the network with the legal regulations. The total APC is the sum of the power consumed by different analog components and digital signal processing units of the BS in a given area. That is to say

$$\mathcal{P} = \underbrace{\lambda_x P_x^{\text{fix}} + \lambda_y P_y^{\text{fix}}}_{\text{(i)}} + \underbrace{\lambda_x \mathbb{P}_x^{\text{act}} P_x^{\text{act}} + \lambda_y \mathbb{P}_y^{\text{act}} P_y^{\text{act}}}_{\text{(ii)}} \tag{6.13}$$

where component (i) captures *load-independent* area power consumption per unit area, *i.e* the power consumed per unit area irrespective of weather a BS is serving a user or not. It consists of:

- P_x^{fix}: The fixed power consumed at each SBS regardless of being active or non-active. We define $P_x^{\text{fix}} = P_x^o + P_x$, where P_x^o is the maintenance cost of a SBS and P_x is the load independent wireless transmission cost. This term in overall accounts for load-independent power consumption required for site-cooling, control signalling, baseband processors and backhaul usage [63]. Power consumption of load-independent caching operation is also incorporated into the term P_x^{fix} (i.e., regular/periodic estimation of content popularity).

- P_y^{fix}: The fixed power consumed at each MBS regardless of being active and non-active. We again define $P_y^{\text{fix}} = P_y^o + P_y$, where P_y^o is the maintenance cost and P_y is the load independent wireless transmission cost. The elements that contributes to P_y^{fix} are similar to the elements mentioned for P_x^{fix}, except some hardware and software differences between MBSs and SBSs (i.e., no caching capabilities at the MBSs).

Component (ii) in (6.13) accounts for *load-dependent* area power consumption with

- $\mathbb{P}_x^{\text{act}}$: Probability that a randomly chosen SBS serves at least on user. This quantifies the activity of an SBS, thus representing the amount of load on a given SBS in general. The expressions for $\mathbb{P}_x^{\text{act}}$ are provided in Lemma 5.

- P_x^{act}: Total power consumed when an SBS is active. This term accounts for the load-dependent power consumed by the circuit components in transceiver chains of each SBS (i.e., amplifiers, converters, mixers and filters). Moreover, the consumed power due to channel estimation via standard linear algebra operations [70], coding and decoding with some suboptimal fixed-complexity algorithms [71], operations for linear processing (i.e., Cholesky factorization for Zero Forcing (ZF) processing [70]), backhaul usage in case of cache misses [72] and access to the cache are considered in this term. When caching is incorporated into the network architecture, P_x^{act} accounts for the average power consumed at an SBS considering the cases that the desired content is and is not cached in the local hard disk. Therefore, from (6.5), it follow that $P_x^{\text{act}} = P_{\text{bh}}(\frac{f_0}{\lambda_x})^{1-\eta} + P_{\text{hd}}(1-(\frac{f_0}{\lambda_x})^{1-\eta})$, where P_{bh} includes the power needed to retrieve data from the backhaul when a content outside the catalog is requested, and P_{hd} includes the power needed to retrieve data from the local hard disk when a content from the catalog is requested. In the absence of caching, $P_x^{\text{act}} = P_{\text{bh}}$.

- $\mathbb{P}_y^{\text{act}}$: Probability that an MBS is active and captures the load of a MBS. The expressions for $\mathbb{P}_y^{\text{act}}$ are given in Lemma 6.

- P_y^{act}: Total power consumed when an MBS is active. Similar power components as in $\mathbb{P}_x^{\text{act}}$ are included in P_y^{act} with slight differences in the setup (i.e., no caching capability at the MBS).

6.3.2 Energy Efficiency

Yet another key performance metric that should be investigated is the energy efficiency, which quantize the amount of utility that is gathered out of a unit power invested for network operation. Some standard QoS factors chosen as the utility in literature are the ASE, throughput and coverage probability [73]. In our case, we define EE of the system as the ratio between the average coverage probability offered by a BS and the power consumed at an average active BS. The expression of EE is given as

$$\begin{aligned}\mathcal{E} &= \frac{\lambda_x \mathbb{P}_x + \lambda_y(1-\mathbb{P}_x)\mathbb{P}_y}{\text{power consumed at an average active BS}} \\ &= \frac{\lambda_x \mathbb{P}_x + \lambda_y(1-\mathbb{P}_x)\mathbb{P}_y}{\mathbb{P}_x^{\text{act}} P_x^{\text{act}} + \mathbb{P}_y^{\text{act}} P_y^{\text{act}}}. \end{aligned} \tag{6.14}$$

Definitions of the terms involved hold same as in Section 6.3.1.

6.4 Suitable System Optimizations towards Cache-empowered Green Networks

In order to reduce energy consumption in the considered network model, it is important to optimize right metrics. In this section, we explore the conditions under which APC can be minimized and EE can be maximized when the density λ_x and transmit power P_x of SBSs are varied separately. This results in four different optimization problems and are elaborated in the subsequent sections. In fact, another important factor while performing optimizations on energy consumption related metrics is to have a right set of justifiable constraints which can reduce the overall cost of network deployment. These constraints are given as follows.

C1: The constraint $\frac{P_x}{P_y} \leq k_p; 0 < k_p < 1$ fixes the the ratio between the transmit powers of SBSs and MBSs to be always less than k_p.

C2: The constraint $\frac{\lambda_y}{\lambda_x} \leq k_{yx}; 0 < k_{yx} < 1$ fixes the ratio between the number of MBSs and SBSs to be always less than k_{yx} in a given patch of area. k_{yx} here indicates the constraint on the infra structural costs.

C3: The constraint $\frac{\lambda_z}{\lambda_x} \geq k_{zx}; k_{zx} > 1$ fixes the ratio between the number of users and number of SBSs to be always greater than k_{zx} in a given patch of area. Typically we keep this as one SBS per every $10 - 15$ users.

C4: The constraint $\mathbb{P}_x \geq \xi \mathbb{P}_x^{\mathrm{NN}}$ makes sure that the coverage probability an SBS is always positive and is always greater that the coverage probability offered during a no noise regime by a factor ξ such that $0 < \xi < 1$.

C5: The constraint $\mathbb{P}_x^{\mathrm{act}} \geq \delta \mathbb{P}_x^{\mathrm{act\ NN}}$ makes sure that the activation probability of an SBS is always positive and is always greater that the activation probability during a no noise regime by a factor δ such that $0 < \delta < 1$.

The following result holds for optimal minimum SBS transmit power.

Proposition 5. *Minimum SBS transmit power (P_x^{min}) for the constraint $\mathbb{P}_x \geq \xi \mathbb{P}_x^{NN}$ to be satisfied is given as $P_x^{min} = \frac{\max\{A_1, A_2\}}{\lambda_x^{\frac{\alpha}{2}}}$, where, $A_1 = \frac{\gamma \sigma^2 \Gamma(\frac{\alpha+2}{2})}{(1-\xi)(P+Qk_{yx})^{\frac{\alpha}{2}}}$ and $A_2 = \frac{\gamma \sigma^2 \Gamma(\frac{\alpha+2}{2})}{(1-\delta)(P'+Qk_{yx})^{\frac{\alpha}{2}}}$.*

Proof. See Appendix 6.9. □

Conversely, it follows from the above proposition that the minimum SBS density at which the constraint on coverage probability is satisfied is given as

$$\lambda_x^{\min} = \frac{\max\{A_1^{\frac{2}{\alpha}}, A_2^{\frac{2}{\alpha}}\}}{P_x^{\frac{2}{\alpha}}}. \tag{6.15}$$

Now, we focus on aforementioned four optimization problems which we have bounds and observations for optimal values of λ_x and P_x.

6.4.1 Optimization of APC with respect to λ_x

In this section, we formulate the problem of optimizing APC ($\mathcal{P}$) with respect to the SBS density λ_x

Theorem 7. *If λ_x^* is the optimal SBS density that minimizes the APC, $\mathcal{P}(\lambda_x)$, subject to the constraints C1-5, then*

$$\lambda_x^* \in \left(0, \left((\alpha-2)\frac{b}{2c\eta}\right)^{\frac{1}{1+\eta+\frac{\alpha-4}{2}}}\right] \tag{6.16}$$

where $b = A(k_{yx}+k_p)$, $c = P_d f_0^{1-\eta}\left[\pi k_{zx}\left(\frac{1}{P'+Qk_{yx}} - (1-\xi)\frac{(P'+Qk_{yx})^{-(\frac{\alpha+2}{2})}}{P+\pi+Qk_{yx}}\right)\right]$.

Proof. See Appendix 6.12. □

6.4.2 Optimization of APC with respect to P_x

Theorem 8. *If P_x^* is the optimal SBS transmit power that minimizes the APC, $\mathcal{P}(\lambda_x)$, subject to the constraints C1-5, then*

$$P_x^* \in \left(0, \left(\frac{\tilde{c}\eta}{\tilde{b}(\alpha-2)}\right)^{\frac{\alpha}{\alpha-2+\eta}}\right] \tag{6.17}$$

where $\tilde{b} = (k_{yx}+k_p)A^{\frac{2}{\alpha}}$; $\tilde{c} = P_d f_0^{1-\eta} A^{\frac{2\eta}{\alpha}}\left[\pi k_{zx}\left(\frac{1}{P'+Qk_{yx}} - (1-\xi)\frac{(P+\pi+Qk_{yx})^{\frac{\alpha}{2}}}{(P'+Qk_{yx})^{\frac{\alpha+2}{2}}}\right)\right]$.

Proof. See Appendix 6.13. □

6.4.3 Optimizing EE with respect to λ_x

Theorem 9. *If λ_x^* is the optimal SBS density that maximizes the EE, $\mathcal{E}(\lambda_x)$, subject to the constraints C1-5, then*

$$\lambda_x^* \in \left(0, \frac{1}{g}\ln\left(\frac{s}{r+s}\right)\right] \tag{6.18}$$

where $r = \xi\frac{\pi}{P+Qk_{yx}}$, $s = k_{yx}\left(1-\xi\frac{\pi}{P+Qk_{yx}}\right)$ *and* $g = k_{yx}\pi R_0^2$.

Proof. See Appendix 6.14. □

6.4.4 Optimizing energy efficiency with repect to P_x

Proposition 6. *The expression for EE as a function of P_x is monotonically decreasing.*

Proof. See Appendix 6.15. □

6.5 Numerical Results and Validation

In this section, we present numerical results to confirm our theoretical expressions. We also discuss the behavior of various quantities of interest and compare them with our theoretical conclusions.

Fig 6.2 illustrates the variation of the SBS coverage probability $\mathbb{P}_x$ with respect to the SBS density λ_x. It can be noticed the coverage probability asymptotically attains a constant value as $\lambda_x \to \infty$ just as expected from Eq (6.28).

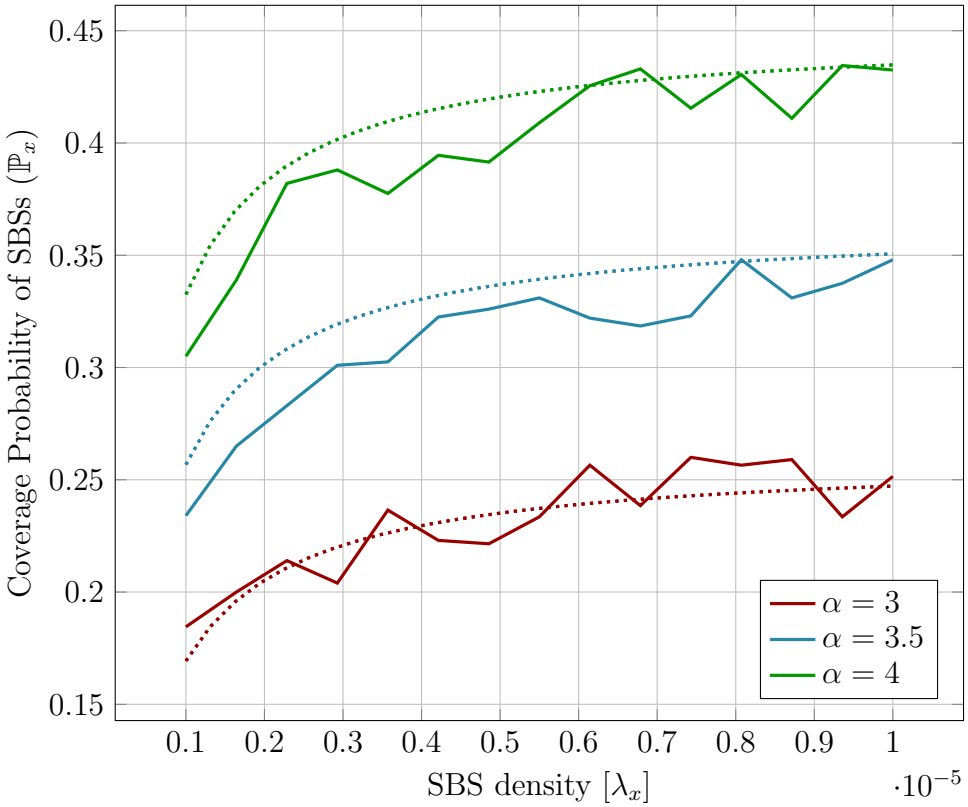


Figure 6.2: Coverage probability offered by SBSs $\mathbb{P}_x$ vs the SBS density λ_x for values $\lambda_y = 10^{-6}\text{m}^{-2}$, $\lambda_z = 10\text{m}^{-2}$, bandwidth $W = 20\text{M Hz}$, Noise figure $FKT = 10^{-15}$ and $\gamma = 1$ in Eq (6.28). The dotted curves represent theoretical values and the continuous lines represent values form numerical simulations.

Fig 6.3 illustrates the variation of the SBS coverage probability $\mathbb{P}_x$ with respect to the SBS transmit power P_x. It can be noticed the coverage probability asymptotically attains a constant value as $P_x \to \infty$ just as expected from Eq (6.28).

Fig 6.4 illustrates the variation of the SBS activation probability $\mathbb{P}_x^{\text{act}}$ with respect to the SBS density λ_x. It can be noticed the activation probability asymptotically attains a constant value as $\lambda_x \to \infty$ just as expected from Eq (6.28).

Fig 6.5 illustrates the variation of the SBS activation probability $\mathbb{P}_x^{\text{act}}$ with respect to the SBS transmit power P_x. When there are very few SBSs, the optimal power that satisfies the constraints is seen to be high.

Fig 6.6 illustrates the variation of the optimal SBS transmit power $P_x^{\min}$ that satisfies the constraints C1-5. When the SBS transmit power is very low, the optimal SBS density that satisfies the constraints is seen to be high.

Fig 6.7 illustrates the variation of the optimal SBS density $\lambda_x^{\min}$ that satisfies the constraints C1-5. When the SBS density is very low, the optimal SBS transmit power that satisfies the constraints is seen to be high.

Fig 6.8 illustrates the variation of APC with respect to the SBS density λ_x. It can be noticed that APC is a convex function and assumes its lowest value at optimum λ_x^*.

Fig 6.9 illustrates the variation of APC with respect to the SBS transmit power P_x. It can be noticed

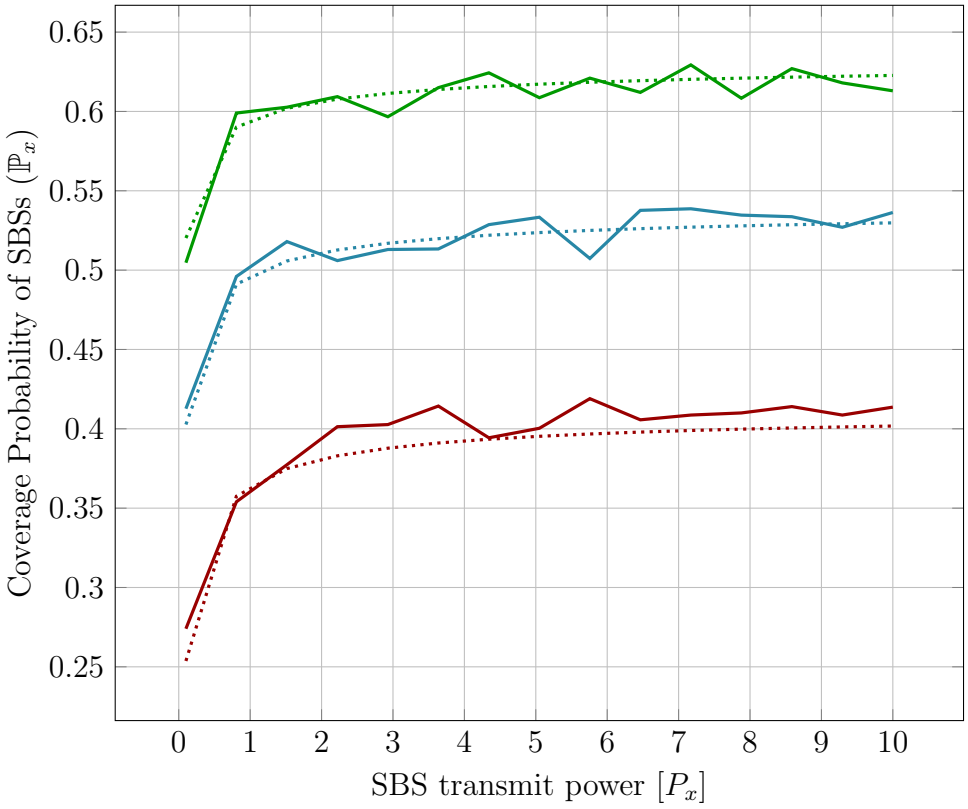


Figure 6.3: Coverage probability offered by SBSs $\mathbb{P}_x$ vs the SBS transmit power P_x for values $\lambda_x = 10^{-6}\text{m}^{-2}$, $\lambda_z = 10\text{m}^{-2}$, bandwidth $W = 20\text{M Hz}$, Noise figure $FKT = 10^{-15}$ and $\gamma = 1$ in Eq (6.28). The dotted curves represent theoretical values and the continuous lines represent values form numerical simulations.

that APC is a convex function and assumes its lowest value at optimum P_x^*.

Fig 6.10 illustrates the variation of EE with respect to the SBS density λ_x. It can be noticed that EE is a concave function and assumes its highest value at optimum λ_x^*.

Fig 6.11 illustrates the variation of EE with respect to the SBS density P_x. It can be noticed that EE is a decreasing function of P_x^*. Therefore, as far as transmit power is considered we can conclude that the lower the better.

6.6 Conclusions

1. Choosing to work with a two tier heterogeneous system, one MBSs (without caching abilities) and one SBSs (with caching abilities), we defined several important system parameters and derived their expressions.
2. APC and EE are two sensible metrics to optimize in order to minimize energy consumption in a network. Taking λ_x and P_x (SBS density and transmit power respectively) to be the variables, we had four independent optimization problems to perform.
3. Imposing the constraints C1-5, we found out that APC is a convex function in both λ_x and P_x. Although the equations involved are complicated to find out the exact analytical expressions for

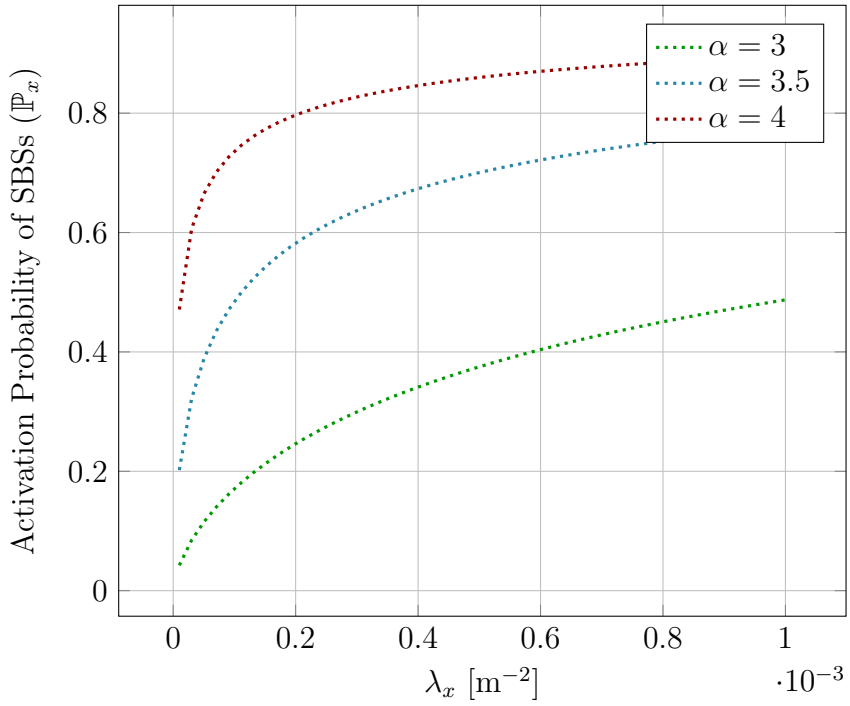


Figure 6.4: Activation probability vs SBS density for values $\lambda_y = 10^{-6}\text{m}^{-2}$, $\lambda_z = 10\text{m}^{-2}$, bandwidth $W = 20\text{M Hz}$, Noise figure $FKT = 10^{-15}$ and $\gamma = 1$ in Eq (6.10).

the optimal values, in this work we find out the intervals within which the optima can be found. For a given system values, though, an optimum can be found out using numerical methods.

4. Imposing the constraints C1-5, we found out that EE is a concave function with respect to λ_x and a monotonically decreasing function with respect to P_x. We found out the upper limit for optimal λ_x for which EE takes the peak value.

5. In future work, it may be of interest to compare the two optimum values of λ_x - one that minimizes the APC and one that maximizes the EE, and verify which of the two metrics is more beneficial economically and in terms of energy consumption.

6. The choice of working with two tiers (*macro* and *small*) is arbitrary and this treatment can be generalized to a system with multiple tiers each with a different kind of caching capability.

6.7 Appendices

6.8 Proof of Lemma 3

We define $\mathbb{P}_x$ as the probability that the best received SINR from the small BSs is greater than a threshold γ. Note that the interference is contributed by both MBSs and SBSs.

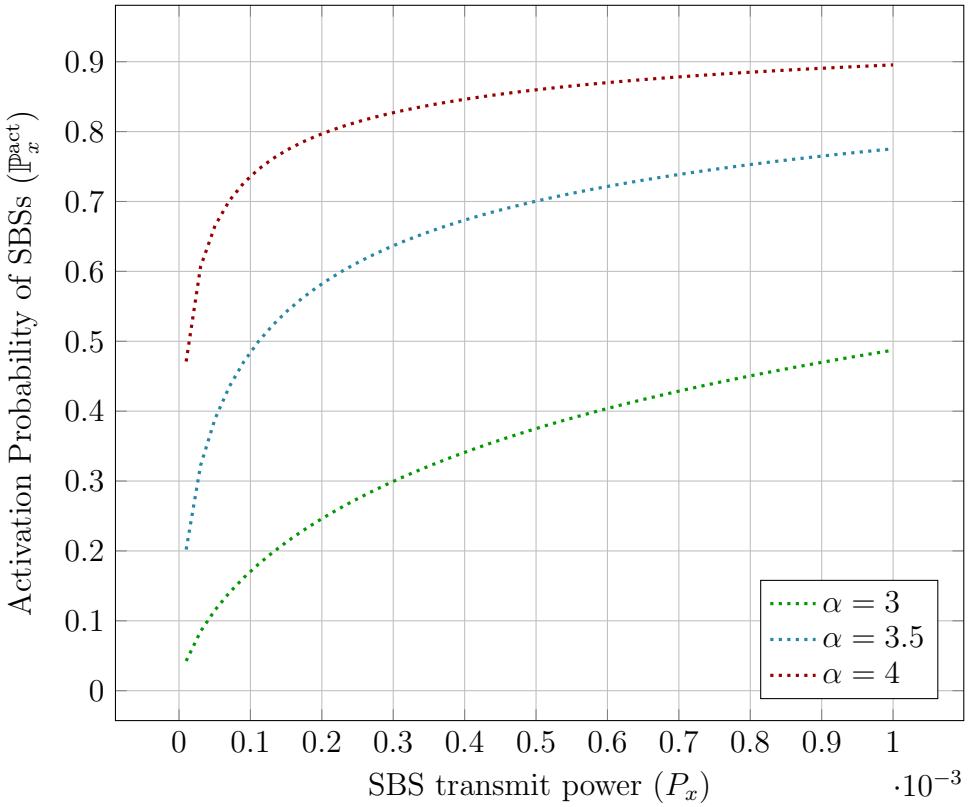


Figure 6.5: Activation probability offered by SBSs $\mathbb{P}_x^{\text{act}}$ vs the SBS transmit power P_x for values $\lambda_x = 10^{-6}\text{m}^{-2}$, $\lambda_z = 10\text{m}^{-2}$, bandwidth $W = 20\text{M}$ Hz, Noise figure $FKT = 10^{-15}$ and $\gamma = 1$ in Eq (6.10).

The expression for received SINR can therefore be written as:

$$\begin{aligned} \text{SINR}_x &= \frac{P_x h_x g(x)}{\sigma^2 + I_x + I_y} \\ &= \frac{P_x h_x g(x)}{\sigma^2 + \sum\limits_{x_i \in \Phi_x / x} P_x h_{x_i} g(x_i) + \sum\limits_{y_j \in \Phi_y} P_y h_{y_j} g(y_j)} \end{aligned} \tag{6.19}$$

For this SINR to be greater than the target value,

$$\text{SINR}_x > \gamma \Rightarrow h_x > \frac{\gamma(\sigma^2 + I_x + I_y)}{P_x g(x)}. \tag{6.20}$$

Probability that the received SINR is greater than the target is given by

$$\begin{aligned} \mathbb{P}_x &= 2\pi\lambda_x \int_0^\infty \mathbb{P}[h_x > \frac{\gamma(\sigma^2 + I_x + I_y)}{P_x g(x)}] d\,x \\ &= \mathbb{E}_{I_x, I_y}\left[2\pi\lambda_x \int_0^\infty \mathbb{P}[h_x > \frac{\gamma(\sigma^2 + I_x + I_y)}{P_x g(x)}] x d\,x | I_x, I_y \right] \\ &\overset{(*)}{=} \frac{2\pi\lambda_x}{\mu} \mathbb{E}_{I_x, I_y}\left[\int_0^\infty \exp(-\frac{\mu\gamma(\sigma^2 + I_x + I_y)}{P_x g(x)}) x\, dx | I_x, I_y \right] \end{aligned} \tag{6.21}$$

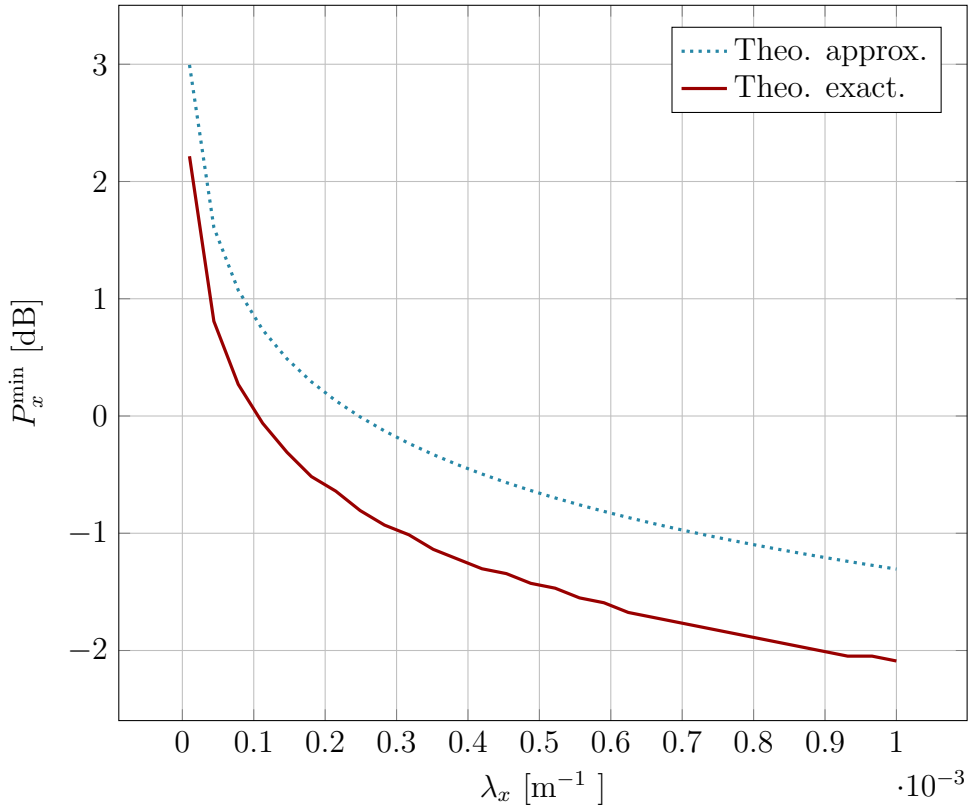


Figure 6.6: Minimum SBS transmit power satisfying the constraints C1-4, vs SBS density for values $k_{yx} = 0.1$, $k_{zx} = 10$, $k_p = 0.5$, $\alpha = 4.3$, $\gamma = 1$,$\xi = 0.1$, $\delta = 0.1$, bandwidth $W = 20$M Hz and Noise figure $FKT = 10^{-15}$.

$$\mathbb{P}_x = \frac{2\pi\lambda_x}{\mu}\int_0^\infty x e^{-\frac{\mu\gamma\sigma^2}{P_x g(x)}}\mathcal{L}_{I_x}\left(\frac{\mu\gamma}{P_x g(x)}\right)\mathcal{L}_{I_y}\left(\frac{\mu\gamma}{P_x g(x)}\right)\mathrm{d}x \tag{6.22}$$

where, $\mathcal{L}_{I_x}$ and $\mathcal{L}_{I_y}$ represent the Laplace transforms with respect to the random variables I_x and I_y respectively.

The Laplace transform, $\mathcal{L}_{I_x}\left(\frac{\mu\gamma}{P_x g(x)}\right)$ is given as

$$\begin{aligned}\mathcal{L}_{I_x}\left(\frac{\mu\gamma}{P_x g(x)}\right) &= \mathbb{E}_{\Phi_x h_i}\left[e^{-s\sum\limits_{x_i\in\Phi_x/x} h_i g(x_i)}\right] \\ &= \mathbb{E}_{\Phi_x h_i}\left[\prod_{x_i\in\Phi_x/x} e^{-s h_i g(x_i)}\right]\end{aligned} \tag{6.23}$$

where, $s = \frac{\cancel{P_x}\mu\gamma}{\cancel{P_x}\, g(\sqrt{x})}$.

The probability generating functional (PGFL) of [16] the PPP states that for some function $f(x)$

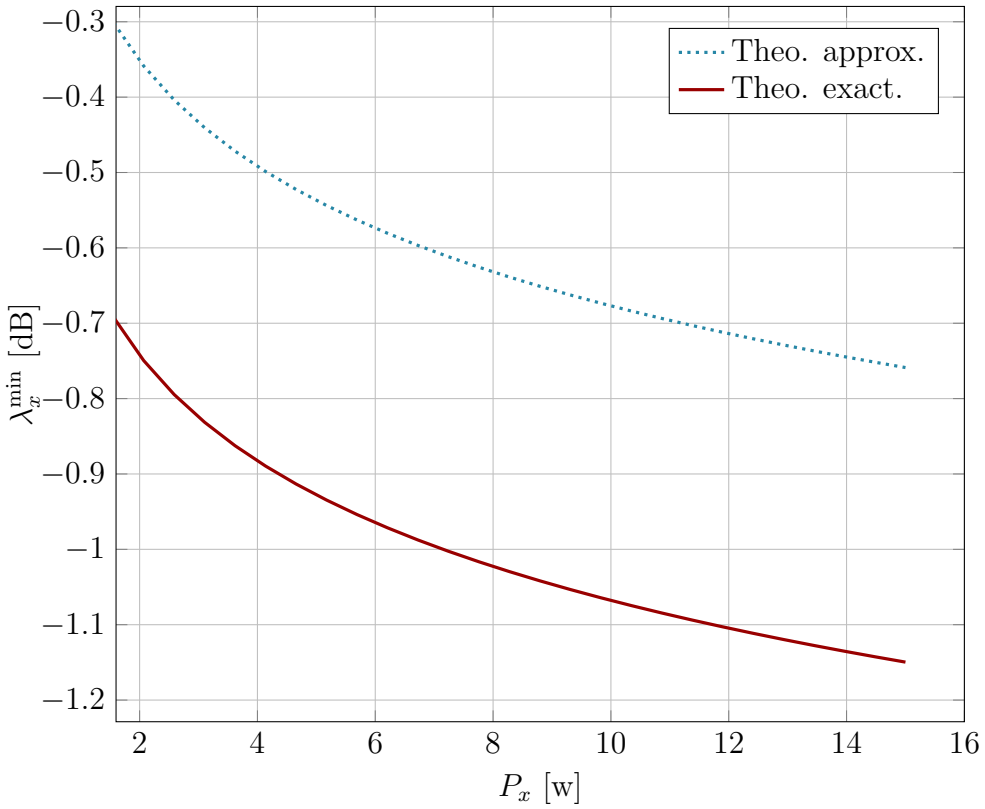


Figure 6.7: Minimum SBS density satisfying the constraints C1-4, vs SBS transmit power for values $k_{yx} = 0.1$, $k_{zx} = 10$, $k_p = 0.5$, $\alpha = 4.3$, $\gamma = 1$, $\xi = 0.1$, $\delta = 0.1$, bandwidth $W = 20$M Hz and Noise figure $FKT = 10^{-15}$.

that $\mathbb{E}[\prod_{x\in\Phi} f(x)] = \exp(-\lambda \int_{\mathbb{R}^2}(1 - f(x))\, dx)$

$$\begin{aligned}
\mathcal{L}_{I_x}\left(\frac{\mu\gamma}{P_x g(x)}\right) &= \exp\left(-2\pi\lambda_x \int_0^\infty v\mathrm{d}v\, \mathbb{E}_h[1 - e^{-\frac{\mu\gamma h g(v)}{g(x)}}]\right) \\
&= \exp\left(-2\pi\lambda_x \int_0^\infty v\mathrm{d}v\, (\int_0^\infty dh\, e^{-\mu h}(1 - e^{-\frac{\mu\gamma h g(v)}{g(x)}}))\right) \\
&= \exp\left(-2\pi\lambda_x \int_0^\infty v\mathrm{d}v\, \frac{\gamma g(v)}{\gamma\mu g(v) + \mu g(x)}\right) \qquad (6.24)
\end{aligned}$$

Similarly,

$$\mathcal{L}_{I_y}\left(\frac{\mu\gamma}{P_x g(x)}\right) = \exp\left(-2\pi\lambda_y \int_0^\infty v\mathrm{d}v\, \frac{\gamma P_y g(v)}{\gamma\mu P_y g(v) + \mu P_x g(x)}\right) \qquad (6.25)$$

By taking path loss function to be $g(x) = bx^{-\alpha}$, and making appropriate substitutions we get

$$\begin{aligned}
\mathcal{L}_{I_x}\left(\frac{\mu\gamma}{P_x g(x)}\right) &= \exp\left(-\lambda_x \frac{2\pi^2\gamma^{2/\alpha}\csc\left(\frac{2\pi}{\alpha}\right)}{\alpha\mu} x^2\right) \\
&= \exp\left(-\lambda_x P x^2\right) \qquad (6.26)
\end{aligned}$$

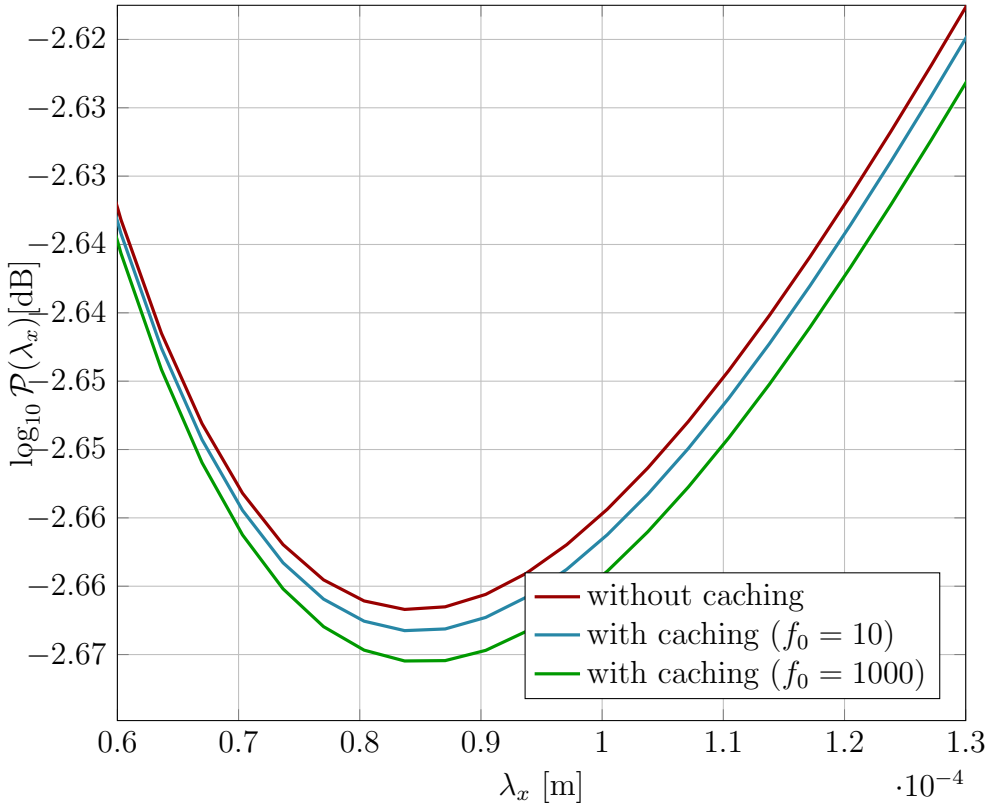


Figure 6.8: APC vs SBS density, with constraints C1-5 for values $k_{yx} = 0.1$, $k_{zx} = 10$, $k_p = 0.5$, $\alpha = 4.3$, $\gamma = 1$, $P_x^{\text{fix}} = 10$, $P_y^{\text{fix}} = 20$, $\xi = 0.1$, $\delta = 0.1$, bandwidth $W = 20$M Hz and Noise figure $FKT = 10^{-15}$ in Eq (6.13).

where, $P = \frac{2\pi^2\gamma^{2/\alpha}\csc\left(\frac{2\pi}{\alpha}\right)}{\alpha\mu}$.

$$\begin{aligned}\mathcal{L}_{I_y}\left(\frac{\mu\gamma}{P_x g(x)}\right) &= \exp\left(-\lambda_y \frac{2\pi(\frac{P_x}{P_y\gamma})^{-\frac{2}{\alpha}}\csc[\frac{2\pi}{\alpha}]}{\alpha\mu}x^2\right)\\ &= \exp\left(-\lambda_y Q x^2\right)\end{aligned} \tag{6.27}$$

where, $Q = \frac{2\pi^2(\frac{P_x}{P_y\gamma})^{-\frac{2}{\alpha}}\csc[\frac{2\pi}{\alpha}]}{\alpha\mu}$.

Substituting (6.26) and (6.27) in (6.22), we can write

$$\mathbb{P}_x = \frac{2\pi\lambda_x}{\mu}\int_0^\infty x e^{-Rx^\alpha} e^{-\lambda_x P x^2} e^{-\lambda_y Q x^2}\,\mathrm{d}x$$

where, $R = \frac{b\gamma\sigma^2}{P_x}$.

Using the low noise approximation $e^{-Rx^\alpha} \approx (1 - Rx^\alpha)$,

$$\begin{aligned}\mathbb{P}_x &= \frac{\pi\lambda_x}{\mu}\int_0^\infty 2x dx\,(1 - Rx^\alpha)e^{-(\lambda_x P + \lambda_y Q)x^2}\\ &= \frac{\pi\lambda_x}{\mu}\left(\frac{1}{P\lambda_x + Q\lambda_y} - R(P\lambda_x + Q\lambda_y)^{-(\frac{\alpha+2}{2})}\Gamma(\frac{\alpha+2}{2})\right)\end{aligned} \tag{6.28}$$

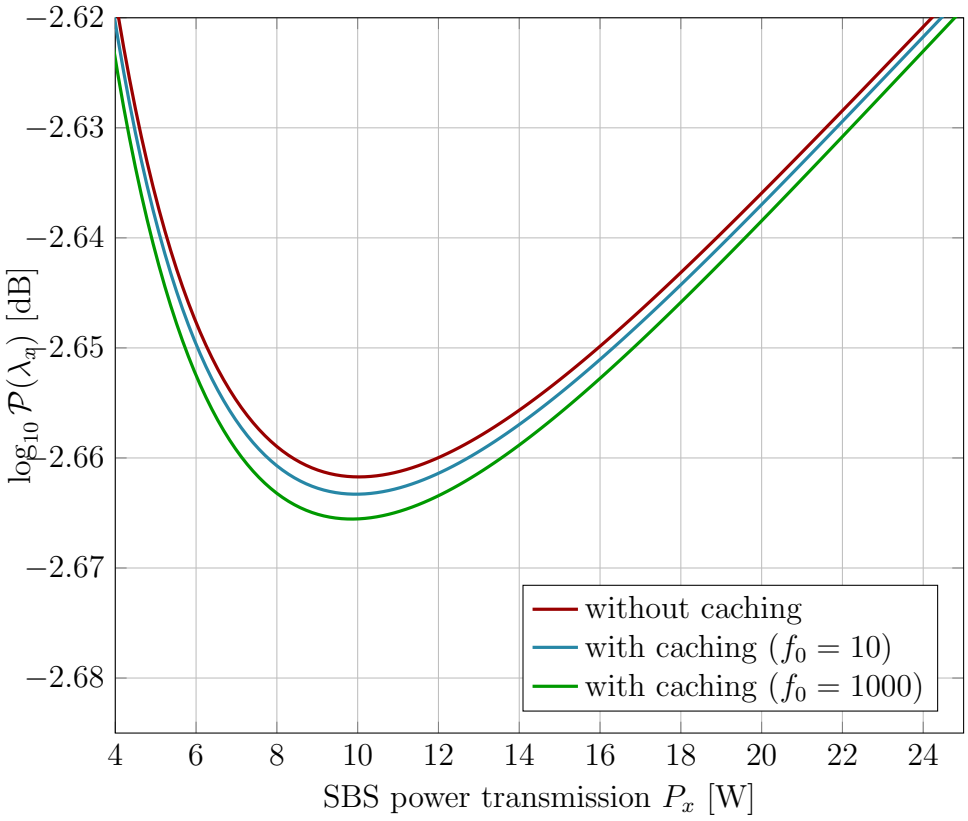


Figure 6.9: APC vs SBS transmit power, with constraints C1-5 for values $k_{yx} = 0.1$, $k_{zx} = 10$, $k_p = 0.5$, $\alpha = 4.3$, $\gamma = 1$, $P_x^{\text{fix}} = 10$, $P_y^{\text{fix}} = 20$, $\xi = 0.1$, $\delta = 0.1$, bandwidth $W = 20$M Hz and Noise figure $FKT = 10^{-15}$ in Eq (6.13).

under no noise regime, $(R \to 0)$,

$$\mathbb{P}_x^{\text{NN}} = \frac{\pi \lambda_x}{\mu} \frac{1}{P\lambda_x + Q\lambda_y}. \tag{6.29}$$

■

6.9 Proof of Proposition 5

The constraint on the SBS coverage probability (C-4) requires that

$$\begin{aligned} & \mathbb{P}_x \geq \xi \mathbb{P}_x^{\text{NN}} \\ \Rightarrow \quad & \frac{1}{P\lambda_x + Q\lambda_y} - R(P\lambda_x + Q\lambda_y)^{-(\frac{\alpha+2}{2})}\Gamma(\frac{\alpha+2}{2}) \geq \xi \frac{1}{P\lambda_x + Q\lambda_y} \\ \Rightarrow \quad & \frac{1-\xi}{P\lambda_x + Q\lambda_y} \geq R'(P\lambda_x + Q\lambda_y)^{-(\frac{\alpha+2}{2})} \\ \Rightarrow \quad & P\lambda_x + Q\lambda_y \leq \left(\frac{1-\xi}{R'}\right)^{\frac{-2}{\alpha}} \end{aligned} \tag{6.30}$$

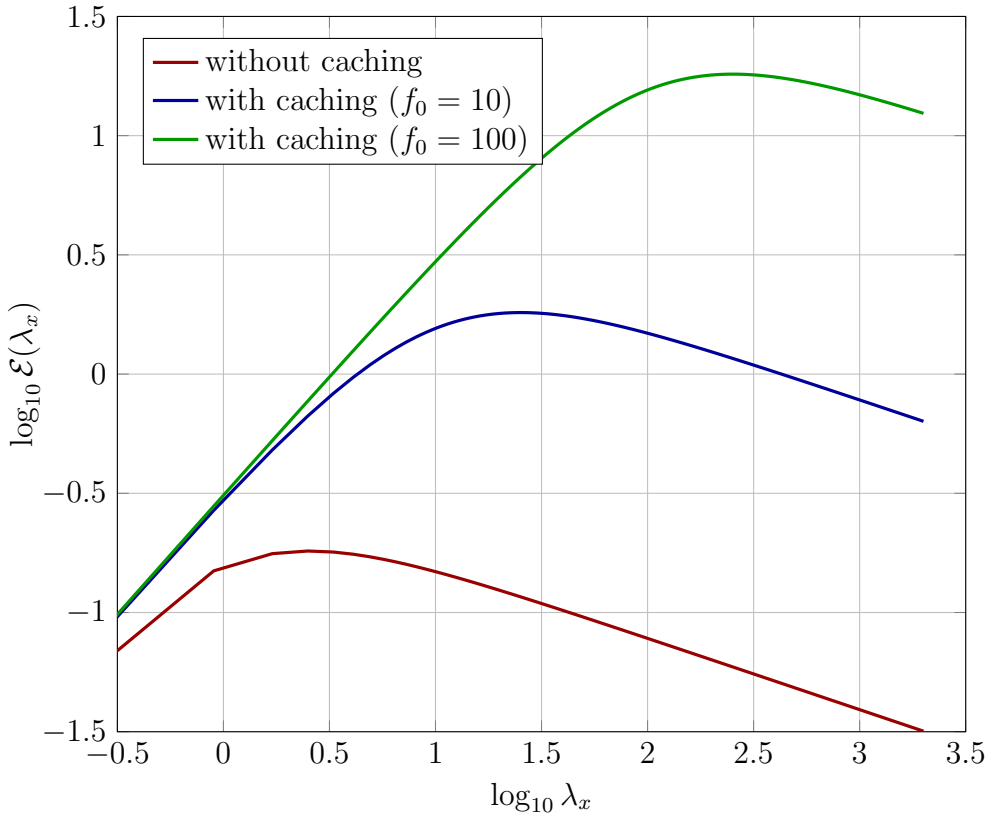


Figure 6.10: EE vs the SBS density, with constraints C1-5 for values $k_{yx} = 0.1$, $k_{zx} = 10$, $k_p = 0.5$, $\alpha = 4.3$, $\gamma = 1$, $P_x^{\text{fix}} = 10$, $P_y^{\text{fix}} = 20$, $\xi = 0.1$, $\delta = 0.1$, bandwidth $W = 20$M Hz and Noise figure $FKT = 10^{-15}$ in Eq (6.14).

Including the constraints $\frac{P_x}{P_y} = k_p$ and $\frac{\lambda_y}{\lambda_x} = k_{yx}$, Eq (6.30) becomes

$$\lambda_x(P + Qk_{yx}) \leq \left(\frac{1-\xi}{R'}\right)^{\frac{-2}{\alpha}} \tag{6.31}$$

substituting $R' = \frac{b\gamma\sigma^2}{P_x}\Gamma(\frac{\alpha+2}{2})$, we see that this condition is valid when

$$(1-\xi)\lambda_x^{\frac{\alpha}{2}}(P + Qk_{yx})^{\frac{\alpha}{2}} \leq \frac{\gamma\sigma^2\Gamma(\frac{\alpha+2}{2})}{P_x} \tag{6.32}$$

$$\Rightarrow P_x \geq \frac{A_1}{\lambda_x^{\frac{\alpha}{2}}} \tag{6.33}$$

where, $A_1 = \frac{\gamma\sigma^2\Gamma(\frac{\alpha+2}{2})}{(1-\xi)(P+Qk_{yx})^{\frac{\alpha}{2}}}$.

By similar analysis of the constraint on activation probability (C-5) ($\mathbb{P}_x^{\text{act}} \geq \delta\mathbb{P}_x^{\text{act NN}}$), it can be easily shown that $P_x \geq \frac{A_2}{\lambda_x^{\frac{\alpha}{2}}}$, where $A_2 = \frac{\gamma\sigma^2\Gamma(\frac{\alpha+2}{2})}{(1-\delta)(P'+Qk_{yx})^{\frac{\alpha}{2}}}$. Combining this with Eq (6.32), the minimum transmit power that satisfies the constraint C-4 and C-5 can be obtained as

$$P_x^{\min} = \frac{\max\{A_1^{\frac{2}{\alpha}}, A_2^{\frac{2}{\alpha}}\}}{P_x^{\frac{2}{\alpha}}}. \tag{6.34}$$

■

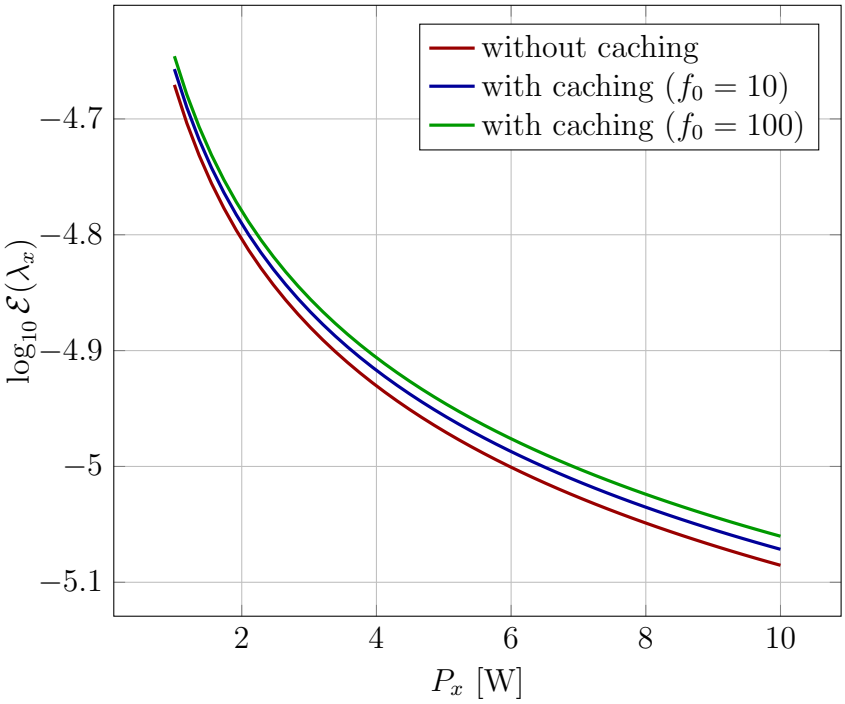


Figure 6.11: EE vs the SBS transmit power, with constraints C1-5 for values $k_{yx} = 0.1$, $k_{zx} = 10$, $k_p = 0.5$, $\alpha = 4.3$, $\gamma = 1$, $P_x^{\text{fix}} = 10$, $P_y^{\text{fix}} = 20$, $\xi = 0.1$, $\delta = 0.1$, bandwidth $W = 20$M Hz and Noise figure $FKT = 10^{-15}$ in Eq (6.14).

6.10 Proof of Lemma 5

We define $\mathbb{P}_x^{\text{act}}$ as the probability that a given SBS is active. By this we mean that a given SBS is providing an SINR above the threshold γ to at least one user at a given instance. To calculate this we consider the probability that the SINR emitted by an SBS is below the threshold everywhere in $\mathbb{R}^2$ and subtract it from 1. z in the below expression represents the position coordinate of a user.

$$\begin{aligned}
\mathbb{P}_x^{\text{act}} &= 1 - \mathbb{P}[\text{SINR}_x < \gamma] \\
&= 1 - \mathbb{P}[h_x < \frac{\gamma(\sigma^2 + I_x + I_y)}{P_x g(z)}] \\
&= 1 - \mathbb{E}_{I_x, I_y}\left[\mathbb{P}[h_x < \frac{\gamma(\sigma^2 + I_x + I_y)}{P_x g(z)}] | I_x, I_y, z\right] \\
&= 1 - \frac{1}{\mu}\mathbb{E}_{I_x, I_y}\left[\int_0^\infty \left(1 - \exp(-\frac{\mu\gamma(\sigma^2 + I_x + I_y)}{P_x g(z)})\right) f(z)\, \mathrm{d}z | I_x, I_y\right] \\
&= \frac{\mu - 1}{\mu} + \frac{2\pi\lambda_z}{\mu}\int_0^\infty z e^{-\frac{\mu\gamma\sigma^2}{P_x g(z)}} \mathcal{L}_{I_x}\left(\frac{\mu\gamma}{P_x g(z)}\right) \mathcal{L}_{I_y}\left(\frac{\mu\gamma}{P_x g(z)}\right) \mathrm{d}z
\end{aligned}$$

Under low noise approximation ($\sigma^2 \to 0$),

$$\mathbb{P}_x^{\text{act}} \approx \frac{\mu - 1}{\mu} + \frac{\pi\lambda_z}{\mu}\left(\frac{1}{P'\lambda_x + Q\lambda_y} - R(P'\lambda_x + Q\lambda_y)^{-(\frac{\alpha+2}{2})}\Gamma(\frac{\alpha+2}{2})\right) \tag{6.35}$$

where $P' = \frac{2\pi^2(\frac{1}{\gamma})^{-\frac{2}{\alpha}}\csc[\frac{2\pi}{\alpha}]}{\alpha\mu}$, $Q = \frac{2\pi^2(\frac{P_x}{P_y\gamma})^{-\frac{2}{\alpha}}\csc[\frac{2\pi}{\alpha}]}{\alpha\mu}$ and $R = \frac{b\gamma\sigma^2}{P_x}$. ■

6.11 Proof of Lemma 6

Probability that an MBS is active $\mathbb{P}_y^{\text{act}}$ is given as,

$$\mathbb{P}_y^{\text{act}} \approx (1 - \mathbb{P}_x) \times \mathbb{P}_{z,R_0}, \tag{6.36}$$

where, $\mathbb{P}_{z,R_0} = 1 - \exp(-\lambda_z \pi R_0^2)$ is probability that a randomly chosen MBS has at least one user at a distance R_0 and $1 - \mathbb{P}_x$ probability that the user is not covered by any SBS. Using Lemma 3 for $\mathbb{P}_x$ gives the desired result. ■

6.12 Proof of Theorem 7

As shown in the equation (6.13), APC consists of two components namely the load-independent component and the load-dependent component. We express them both in terms of λ_x by incorporate constraints C1-5 into the expressions and eliminating other system variables as shown below:

6.12.1 Load independent part

$$\lambda_x P_x^{\text{fix}} + \lambda_y P_y^{\text{fix}} = \lambda_x (P_x^o + P_x) + \lambda_y (P_y^o + P_y) \tag{6.37}$$

From constraints C-1 and C-2 we substitute in the above expression $\frac{P_x}{P_y} = k_p$ and $\frac{\lambda_y}{\lambda_x} = k_{yx}$. Further, from constraint C-4 and proposition 5, we substitute P_x with $P_x^{\min} = \frac{A}{\lambda^{\frac{\alpha}{2}}}$, where $A = \frac{\gamma \sigma^2 \Gamma(\frac{\alpha+2}{2})}{(1-\xi)(P+Qk_{yx})^{\frac{\alpha}{2}}}$. We get,

$$\begin{aligned} \lambda_x P_x^{\text{fix}} + \lambda_y P_y^{\text{fix}} &= \lambda_x (P_x^o + P_x^{\min}) + k_{yx} \lambda_x (P_y^o + k_p P_x^{\min}) \\ &= \lambda_x (P_x^o + k_{yx} P_y^o) + \lambda_x (k_{yx} + k_p) P_x^{\min} \\ &= \lambda_x (P_x^o + k_{yx} P_y^o) + (k_{yx} + k_p) \frac{A}{\lambda_x^{\frac{\alpha-2}{2}}} \end{aligned} \tag{6.38}$$

6.12.2 Load dependent part

Taking $\mu = 1$, we can expand the load-dependent part of APC as

$$\begin{aligned} & \lambda_x \mathbb{P}_x^{\text{act}} P_x^{\text{act}} + \lambda_y \mathbb{P}_y^{\text{act}} P_y^{\text{act}} = \\ & \lambda_x \left[\pi \lambda_z \left(\frac{1}{P' \lambda_x + Q \lambda_y} - R' (P' \lambda_x + Q \lambda_y)^{-(\frac{\alpha+2}{2})} \right) \right] \\ & \times (P_{\text{bh}} (\frac{f_0}{\lambda_x})^{1-\eta} + P_{\text{hd}} (1 - (\frac{f_0}{\lambda_x})^{1-\eta})) \\ + \quad & \lambda_y \left[(1 - e^{-\lambda_z \pi R_0^2}) \left(1 - \pi \lambda_x \left(\frac{1}{P \lambda_x + Q \lambda_y} - R' (P \lambda_x + Q \lambda_y)^{-(\frac{\alpha+2}{2})} \right) \right) \right] \times P_{bh} \end{aligned} \tag{6.39}$$

Including the constraint $\mathbb{P}_x \geq \xi \mathbb{P}_x^{\mathrm{NN}}$, we get

$$\begin{aligned} & \lambda_x \mathbb{P}_x^{\mathrm{act}} P_x^{\mathrm{act}} + \lambda_y \mathbb{P}_y^{\mathrm{act}} P_y^{\mathrm{act}} = \\ & \lambda_x \left[\pi \lambda_z \left(\frac{1}{P' \lambda_x + Q \lambda_y} - \frac{\gamma \sigma^2 \Gamma(\frac{\alpha+2}{2})}{P_x} (P' \lambda_x + Q \lambda_y)^{-(\frac{\alpha+2}{2})} \right) \right] \\ & \times (P_{\mathrm{bh}} \left(\frac{f_0}{\lambda_x} \right)^{1-\eta} + P_{\mathrm{hd}} (1 - \left(\frac{f_0}{\lambda_x} \right)^{1-\eta})) \\ + \quad & \lambda_y \left[(1 - e^{-\lambda_z \pi R_0^2}) \left(1 - \xi \frac{\pi \lambda_x}{P \lambda_x + Q \lambda_y} \right) \right] \times P_{bh} \end{aligned} \tag{6.40}$$

Including the constraints $\frac{P_x}{P_y} = k_p$, $\frac{\lambda_y}{\lambda_x} = k_{yx}$ and $\frac{\lambda_z}{\lambda_x} = k_{zx}$, and by substituting P_x with $P_x^{\min} = \frac{A}{\lambda_x^{\frac{\alpha}{2}}}$, we get

$$\begin{aligned} & \lambda_x \mathbb{P}_x^{\mathrm{act}} P_x^{\mathrm{act}} + \lambda_y \mathbb{P}_y^{\mathrm{act}} P_y^{\mathrm{act}} = \\ & \lambda_x \left[\pi k_{zx} \left(\frac{1}{P' + Q k_{yx}} - (1 - \xi) \frac{(P' + Q k_{yx})^{-(\frac{\alpha+2}{2})}}{P + \pi + Q k_{yx}} \right) \right] \\ & \times (P_{\mathrm{hd}} + P_{\mathrm{d}} \left(\frac{f_0}{\lambda_x} \right)^{1-\eta}) \\ + \quad & k_{yx} \lambda_x \left[(1 - e^{-k_{zx} \lambda_x \pi R_0^2}) \left(1 - \xi \frac{\pi}{P + Q k_{yx}} \right) \right] \times P_{bh} \end{aligned}$$

where $P_{\mathrm{d}} = P_{\mathrm{bh}} - P_{\mathrm{hd}}$. Since $k_{zx} > 1$, it can be noted that the above expression is always positive.

The total expression of APC can be written as

$$\begin{aligned} \mathcal{P}(\lambda_x) &= \lambda_x P_x^{\mathrm{fix}} + \lambda_y P_y^{\mathrm{fix}} + \lambda_x \mathbb{P}_x^{\mathrm{act}} P_x^{\mathrm{act}} + \lambda_y \mathbb{P}_y^{\mathrm{act}} P_y^{\mathrm{act}} \\ &= a \lambda_x + b \lambda_x^{-\frac{\alpha-2}{2}} + c \lambda_x^{\eta} + d \lambda_x (1 - e^{-f \lambda_x}) \end{aligned} \tag{6.41}$$

$a = (P_x^o + k_{yx} P_y^o) + P_{\mathrm{hd}} \left[\pi k_{zx} \left(\frac{1}{P' + Q k_{yx}} - (1 - \xi) \frac{(P' + Q k_{yx})^{-(\frac{\alpha+2}{2})}}{P + Q k_{yx}} \right) \right]$, $b = A(k_{yx} + k_p)$, $c = P_{\mathrm{d}} f_0^{1-\eta} \left[\pi k_{zx} \left(\frac{1}{P' + Q k_{yx}} - (1 -$
$d = k_{yx} P_{\mathrm{bh}} \left[\left(1 - \xi \frac{\pi}{P + Q k_{yx}} \right) \right]$, $f = k_{zx} \pi R_0^2$.

We differentiate $\mathcal{P}(\lambda_x)$ in Eq (6.41) with respect to λ_x, we get

$$\frac{d\,\mathcal{P}(\lambda_x)}{d\,\lambda_x} = a + d\left(1 - e^{-f\lambda_x}\right) + df e^{-f\lambda_x} \lambda_x - \frac{(\alpha - 2)}{2} b \lambda_x^{-\frac{\alpha}{2}} + c\eta \lambda_x^{-1+\eta} \tag{6.42}$$

It can be examined at this point that the limit $\lim_{\lambda_x \to 0} \frac{d\,\mathcal{P}(\lambda_x)}{d\,\lambda_x} = -\infty$ (since $\eta > 1$ and $\alpha > 2$), and $\frac{d\,\mathcal{P}(\lambda_x)}{d\,\lambda_x}$ is always positive if

$$\begin{aligned} & c\eta \lambda_x^{\eta-1} - \frac{(\alpha - 2)}{2} b \lambda_x^{-\frac{\alpha}{2}} > 0 \\ \Rightarrow \quad & \lambda_x > \left[(\alpha - 2) \frac{b}{2c\eta} \right]^{\frac{1}{\eta + \frac{\alpha-2}{2}}} \end{aligned} \tag{6.43}$$

Therefore, the optimal SBS density is located within the interval where the derivative $\frac{d\,\mathcal{P}(\lambda_x)}{d\,\lambda_x}$ changes its sign $\lambda_x^* \in \left(0, \left((\alpha - 2) \frac{b}{2c\eta} \right)^{\frac{1}{\eta + \frac{\alpha-2}{2}}} \right)$. ■

6.13 Proof of Theorem 8

As shown in the equation (6.13), APC consists of two components namely the load-independent component and the load-dependent component. We express them both in terms of P_x by incorporate constraints C1-5 into the expressions and eliminating other system variables as shown below:

6.13.1 Load independent part

$$\lambda_x P_x^{\text{fix}} + \lambda_y P_y^{\text{fix}} = \lambda_x (P_x^o + P_x) + \lambda_y (P_y^o + P_y) \tag{6.44}$$

Including the constraints $\frac{P_x}{P_y} = k_p$ and $\frac{\lambda_y}{\lambda_x} = k_{yx}$, and by substituting $\lambda_x = \frac{A^{\frac{2}{\alpha}}}{P_x^{\frac{2}{\alpha}}}$, we get

$$\begin{aligned} \lambda_x P_x^{\text{fix}} + \lambda_y P_y^{\text{fix}} &= \lambda_x (P_x^o + P_x) + k_{yx}\lambda_x (P_y^o + k_p P_x) \\ &= \lambda_x (P_x^o + k_{yx} P_y^o) + \lambda_x (k_{yx} + k_p) P_x \\ &= \frac{A^{\frac{2}{\alpha}}}{P_x^{\frac{2}{\alpha}}} (P_x^o + k_{yx} P_y^o) + (k_{yx} + k_p) A^{\frac{2}{\alpha}} P_x^{\frac{\alpha-2}{\alpha}} \end{aligned} \tag{6.45}$$

6.13.2 Load dependent part

With $\mathbb{P}_x \geq \xi \mathbb{P}_x^{\text{NN}}$, Eq (6.39) becomes

$$\begin{aligned} & \lambda_x \mathbb{P}_x^{\text{act}} P_x^{\text{act}} + \lambda_y \mathbb{P}_y^{\text{act}} P_y^{\text{act}} = \\ & \lambda_x \left[\pi \lambda_z \left(\frac{1}{P'\lambda_x + Q\lambda_y} - \frac{\gamma \sigma^2 \Gamma(\frac{\alpha+2}{2})}{P_x} (P'\lambda_x + Q\lambda_y)^{-(\frac{\alpha+2}{2})} \right) \right] \\ & \times (P_{\text{bh}} \left(\frac{f_0}{\lambda_x} \right)^{1-\eta} + P_{\text{hd}} (1 - \left(\frac{f_0}{\lambda_x} \right)^{1-\eta})) \\ + \quad & \lambda_y \left[(1 - e^{-\lambda_z \pi R_0^2}) \left(1 - \xi \frac{\pi \lambda_x}{P\lambda_x + Q\lambda_y} \right) \right] \times P_{bh} \end{aligned} \tag{6.46}$$

Including the constraints $\frac{P_x}{P_y} = k_p$, $\frac{\lambda_y}{\lambda_x} = k_{yx}$ and $\frac{\lambda_z}{\lambda_x} = k_{zx}$, we get

$$\begin{aligned} & \lambda_x \mathbb{P}_x^{\text{act}} P_x^{\text{act}} + \lambda_y \mathbb{P}_y^{\text{act}} P_y^{\text{act}} = \\ & \lambda_x \left[\pi k_{zx} \left(\frac{1}{P' + Qk_{yx}} - \lambda_x^{-\frac{\alpha}{2}} \frac{\gamma \sigma^2 \Gamma(\frac{\alpha+2}{2})}{P_x} (P' + Qk_{yx})^{-(\frac{\alpha+2}{2})} \right) \right] \\ & \times (P_{\text{hd}} + P_{\text{d}} \left(\frac{f_0}{\lambda_x} \right)^{1-\eta}) \\ + \quad & k_{yx}\lambda_x \left[(1 - e^{-k_{zx}\lambda_x \pi R_0^2}) \left(1 - \xi \frac{\pi}{P + Qk_{yx}} \right) \right] \times P_{bh} \end{aligned} \tag{6.47}$$

By substituting $\lambda_x = \frac{A^{\frac{2}{\alpha}}}{P_x^{\frac{2}{\alpha}}}$, we get

$$\begin{aligned}
&\lambda_x \mathbb{P}_x^{\text{act}} P_x^{\text{act}} + \lambda_y \mathbb{P}_y^{\text{act}} P_y^{\text{act}} = \\
&\left(\frac{A}{P_x}\right)^{\frac{2}{\alpha}} \left[\pi k_{zx} \left(\frac{1}{P' + Qk_{yx}} - \left(\frac{A}{P_x}\right)^{-1} \frac{\gamma \sigma^2 \Gamma(\frac{\alpha+2}{2})}{P_x} (P' + Qk_{yx})^{-(\frac{\alpha+2}{2})}\right)\right] \\
&\times (P_{\text{hd}} + P_{\text{d}} f_0^{1-\eta} A^{\frac{2(\eta-1)}{\alpha}} P_x^{\frac{2(1-\eta)}{\alpha}}) \\
+ \quad & k_{yx} \left(\frac{A}{P_x}\right)^{\frac{2}{\alpha}} \left[(1 - e^{-k_{zx}(\frac{A}{P_x})^{\frac{2}{\alpha}} \pi R_0^2}) \left(1 - \xi \frac{\pi}{P + Qk_{yx}}\right)\right] \times P_{bh}
\end{aligned} \tag{6.48}$$

From Eq (6.45) and Eq (6.48),

$$\mathcal{P}(P_x) = \tilde{a} P_x^{-\frac{2}{\alpha}} + \tilde{b} P_x^{\frac{\alpha-2}{\alpha}} + \tilde{c} P^{-\frac{\eta}{\alpha}} + \tilde{d} P_x^{-\frac{2}{\alpha}} (1 - e^{-\tilde{f} P_x^{-\frac{2}{\alpha}}}) \tag{6.49}$$

where, $\tilde{a} = A^{\frac{2}{\alpha}}(P_x^o + k_{yx} P_y^o) + A^{\frac{2}{\alpha}} \left[\pi k_{zx} \left(\frac{1}{P' + Qk_{yx}} - (1-\xi)\frac{(P+Qk_{yx})^{\frac{\alpha}{2}}}{(P'+Qk_{yx})^{\frac{\alpha+2}{2}}}\right)\right]$; $\tilde{b} = (k_{yx} + k_p) A^{\frac{2}{\alpha}}$; $\tilde{c} = P_{\text{d}} f_0^{1-\eta} A^{\frac{2\eta}{\alpha}} \left[\pi k_{zx} \left(\frac{1}{P' + Qk_{yx}} - (1-\xi)\frac{(P+Qk_{yx})^{\frac{\alpha}{2}}}{(P'+Qk_{yx})^{\frac{\alpha+2}{2}}}\right)\right]$;
$\tilde{d} = k_{yx} P_{bh} A^{\frac{2}{\alpha}} \left[\left(1 - \xi \frac{\pi}{P + Qk_{yx}}\right)\right]$; $\tilde{f} = k_{zx} A^{\frac{2}{\alpha}} \pi R_0^2$.

We differentiate $\mathcal{P}(P_x)$ with respect to P_x to get

$$\begin{aligned}
\frac{d\,\mathcal{P}(P_x)}{d\,P_x} \quad = \quad & -\frac{2\tilde{d} e^{-\tilde{f} P_x^{-\frac{2}{\alpha}}} \tilde{f} P_x^{-1-\frac{4}{\alpha}}}{\alpha} - \frac{2\tilde{a} P_x^{-1-\frac{2}{\alpha}}}{\alpha} - \frac{2\tilde{d}\left(1 - e^{-\tilde{f} P_x^{-\frac{2}{\alpha}}}\right) P_x^{-1-\frac{2}{\alpha}}}{\alpha} \\
& + \frac{\tilde{b} P_x^{-1+\frac{\alpha-2}{\alpha}} (\alpha - 2)}{\alpha} - \frac{\tilde{c} P_x^{-1-\frac{\eta}{\alpha}} \eta}{\alpha}
\end{aligned} \tag{6.50}$$

This expression is always negative when

$$\begin{aligned}
& \frac{\tilde{b} P_x^{-1+\frac{\alpha-2}{\alpha}} (\alpha - 2)}{\alpha} - \frac{\tilde{c} P_x^{-1-\frac{\eta}{\alpha}} \eta}{\alpha} \leq 0 \\
& P_x \leq \left(\frac{\tilde{c}\eta}{\tilde{b}(\alpha - 2)}\right)^{\frac{\alpha}{\alpha-2+\eta}}
\end{aligned} \tag{6.51}$$

The optimal transmit power of SBSs $P_x^* \in [0, \left(\frac{\tilde{c}\eta}{\tilde{b}(\alpha-2)}\right)^{\frac{\alpha}{\alpha-2+\eta}}]$. ■

6.14 Proof of Theorem 9

The numerator in the expression for $\mathcal{E}(\lambda_x)$ in Eq (6.14)

$$\begin{aligned}
\lambda_x \mathbb{P}_x + \lambda_y (1 - \mathbb{P}_x) \mathbb{P}_y \quad &= \quad \lambda_x \xi \frac{\pi}{P + Qk_{yx}} + k_{yx} \lambda_x \left(1 - \xi \frac{\pi}{P + Qk_{yx}}\right) (1 - e^{-k_{yx} \lambda_x \pi R_0^2}) \\
&= \quad r\lambda_x + s\lambda_x (1 - e^{-g\lambda_x}),
\end{aligned} \tag{6.52}$$

where $r = \xi \frac{\pi}{P + Qk_{yx}}$, $s = k_{yx}\left(1 - \xi \frac{\pi}{P + Qk_{yx}}\right)$ and $g = k_{yx} \pi R_0^2$.

The denominator in the expression for $\mathcal{E}(\lambda_x)$ in Eq (6.14)

$$\begin{aligned}
&\mathbb{P}_x^{\text{act}} P_x^{\text{act}} + \mathbb{P}_y^{\text{act}} P_y^{\text{act}} = \\
&\left[\pi k_{zx}\left(\frac{1}{P' + Qk_{yx}} - (1-\xi)\frac{(P' + Qk_{yx})^{-(\frac{\alpha+2}{2})}}{P + Qk_{yx}}\right)\right] \\
&\times (P_{\text{hd}} + P_{\text{d}}\left(\frac{f_0}{\lambda_x}\right)^{1-\eta}) \\
+ \quad &\left[(1 - e^{-k_{zx}\lambda_x \pi R_0^2})\left(1 - \xi\frac{\pi}{P + Qk_{yx}}\right)\right] \times P_{bh}
\end{aligned} \tag{6.53}$$

where $P_{\text{d}} = P_{\text{bh}} - P_{\text{hd}}$. Since $k_{zx} > 1$, it can be noted that the above expression is always positive.

$$\mathbb{P}_x^{\text{act}} P_x^{\text{act}} + \mathbb{P}_y^{\text{act}} P_y^{\text{act}} = l + m\lambda_x^{\eta-1} + n(1 - e^{-f\lambda_x}) \tag{6.54}$$

where, $l = P_{\text{hd}}\left[\pi k_{zx}\left(\frac{1}{P'+Qk_{yx}} - (1-\xi)\frac{(P'+Qk_{yx})^{-(\frac{\alpha+2}{2})}}{P+Qk_{yx}}\right)\right]$,
$m = P_{\text{d}} f_0^{1-\eta}\left[\pi k_{zx}\left(\frac{1}{P'+Qk_{yx}} - (1-\xi)\frac{(P'+Qk_{yx})^{-(\frac{\alpha+2}{2})}}{P+Qk_{yx}}\right)\right]$,
$n = P_{\text{bh}}\left[\left(1 - \xi\frac{\pi}{P+Qk_{yx}}\right)\right]$, and $f = k_{zx}\pi R_0^2$.

From Eq (6.52) and Eq (6.54),

$$\mathcal{E}(\lambda_x) = \frac{r\lambda_x + s\lambda_x(1 - e^{-g\lambda_x})}{l + m\lambda_x^{\eta-1} + n(1 - e^{-f\lambda_x})} \tag{6.55}$$

Differentiating $\mathcal{E}(\lambda_x)$ with respect to λ_x

$$\frac{d\,\mathcal{E}(\lambda_x)}{d\,\lambda_x} = \frac{r + \left(1 - e^{-g\lambda_x}\right)s + e^{-g\lambda_x} g s \lambda_x}{l + \left(1 - e^{-f\lambda_x}\right)n + m\lambda_x^{\eta-1}} - \frac{\left(r\lambda_x + \left(1 - e^{-g\lambda_x}\right)s\lambda_x\right)\left(e^{-f\lambda_x} f n + m\lambda_x^{\eta-2}(\eta - 1)\right)}{\left(l + \left(1 - e^{-f\lambda_x}\right)n + m\lambda_x^{\eta-1}\right)^2} \tag{6.56}$$

The derivative is positive when

$$\begin{aligned}
&r\left(l + \left(1 - e^{-f\lambda_x}\right)n + m\lambda_x^{\eta-1}\right) + e^{-g\lambda_x} g s \lambda_x \left(l + \left(1 - e^{-f\lambda_x}\right)n + m\lambda_x^{\eta-1}\right) \\
&\quad + \left(1 - e^{-g\lambda_x}\right)\left(s\left(l + \left(1 - e^{-f\lambda_x}\right)n + m\lambda_x^{\eta-1}\right) + s\lambda_x\left(-e^{-f\lambda_x} f n - m\lambda_x^{-2+\eta}(\eta - 1)\right)\right) \\
&\quad + r\lambda_x\left(-e^{-f\lambda_x} f n - m\lambda_x^{-2+\eta}(\eta - 1)\right) \geq 0
\end{aligned} \tag{6.57}$$

Noting that the terms $\eta - 1$, $1 - e^{-g\lambda_x}$, and $1 - e^{-f\lambda_x}$ are positive, we can conclude that the above expression is always positive when

$$\begin{aligned}
& -e^{-(f+g)\lambda_x}\left(-s + e^{g\lambda_x}(r+s)\right)\left(f n\lambda_x + e^{f\lambda_x} m\lambda_x^{\eta-1}(-1+\eta)\right) \geq 0 \\
\Rightarrow \quad & s - e^{g\lambda_x}(r+s) \geq 0 \\
\Rightarrow \quad & \lambda_x^* \leq \frac{1}{g}\ln\left(\frac{s}{r+s}\right).
\end{aligned} \tag{6.58}$$

This sets an upper bound for the optimal SBS density. ■

6.15 Proof of Proposition 6

We write the expression for energy efficiency, defined in Eq (6.14), in terms of P_x.

The numerator in the expression for $\mathcal{E}(P_x)$ becomes

$$\begin{aligned} & \lambda_x \mathbb{P}_x + \lambda_y (1-\mathbb{P}_x)\mathbb{P}_y \\ = \; & \lambda_x \xi \frac{\pi}{P+Qk_{yx}} + k_{yx}\lambda_x \left(1 - \xi \frac{\pi}{P+Qk_{yx}}\right)(1 - e^{-k_{yx}\lambda_x \pi R_0^2}) \\ = \; & r\lambda_x + s\lambda_x (1 - e^{-g\lambda_x}), \\ = \; & (\frac{A}{P_x})^{\frac{2}{\alpha}} (r + s(1 - e^{-g(\frac{A}{P_x})^{\frac{2}{\alpha}}})), \end{aligned} \tag{6.59}$$

where $r = \xi \frac{\pi}{P+Qk_{yx}}$, $s = k_{yx}\left(1 - \xi \frac{\pi}{P+Qk_{yx}}\right)$ and $g = k_{yx}\pi R_0^2$.

Using Eq (6.54), denominator in the expression for $\mathcal{E}(P_x)$ in Eq (6.14)

$$\begin{aligned} & \mathbb{P}_x^{\text{act}} P_x^{\text{act}} + \mathbb{P}_y^{\text{act}} P_y^{\text{act}} = \\ & \left[\pi k_{zx}\left(\frac{1}{P' + Qk_{yx}} - (1-\xi)\frac{(P+Qk_{yx})^{\frac{\alpha}{2}}}{(P'+Qk_{yx})^{\frac{\alpha+2}{2}}}\right)\right] \\ & \times (P_{\text{hd}} + P_{\text{d}} f_0^{1-\eta} A^{\frac{2(\eta-1)}{\alpha}} P_x^{\frac{2(1-\eta)}{\alpha}}) \\ + \; & \left[(1 - e^{-k_{zx}(\frac{A}{P_x})^{\frac{2}{\alpha}}\pi R_0^2})\left(1 - \xi \frac{\pi}{P+Qk_{yx}}\right)\right] \times P_{bh} \end{aligned} \tag{6.60}$$

Substituting the numerator and denominator from Eq (6.59) and Eq (6.60) in Eq (6.14), it can be observed that the function $\mathcal{E}(P_x)$ is monotonically decreasing with P_x. By which we deduce that there is no optimum value for EE with respect to SBS transmit power P_x. In this case, one can choose the least P_x possible and fix the value of λ_x to λ_x^* in Eq (6.58). ■

Part III

Conclusions

Chapter 7

Conclusions and future directions

This thesis presented solutions to the problem of bettering energy efficiency in wireless networks. Clearly the range of possible approaches to this problem is huge and this thesis focused mainly on the fundamentals of physics in the first part and on the system level design issues in the second part.

7.1 Fundamental limits

In the first part of this thesis, we devised a thought-experiment like communication system and calculated the upper bound on its energy efficiency using thermodynamics and electromagnetic theory. With the help of the proposed model, we have seen that for a given total amount of energy that is available to prepare, transmit, and rewrite information bits, there exists a unique transmitting antenna frequency that yields maximum energy efficiency. We have also seen the dependence of energy efficiency on the bandwidth and the amplitude of the signal, and that it has an optimum at suitable values for the two. We have made a comparison of these dependencies in the cases of binary phase key shifting model and quadrature phase key shifting model. Our results show that, contrary to the results from [28], when the communication system is working at a non-zero temperature, the optimal power consumed by the system in order to maximize energy efficiency is non-zero.

7.1.1 Future directions

The treatment can be extended and applied on various choices of physical systems to store and transport information. Our choice of thermodynamic Szilard boxes and electromagnetic waves is in the interest of exploring the limits in the realm of classical physics. However, when we consider very low energy systems, the quantum mechanical effects could be more pronounced and this requires further study. In future work it may be interesting to explore further fundamental physical systems to represent bits. For example: quantum spin states of particles, quantum mechanically described thermodynamic systems [29]. Also it is very important to study the limits on energy efficiency of a computation device from this line of approach. Studies such as [30] where limits on the energy costs of computation are explored are of great importance in the current era where untethered battery driven mobile computing is ubiquitous. Such enquiries on EE limits of computation shall lead us acquire insights into the limits even at the MAC layer of a communication system.

7.2 Stochastic geometry based modeling

In chapter 4, we addressed optimization of APC in single-tier wireless cellular networks in which the locations of BSs are modeled to have been distributed according to a homogeneous PPP. Under strongest BS association policy, we derived bounds on the optimal transmit power in order to guarantee a certain minimum coverage and data rate. Under the same QoS constraints, we derived the closed form expression for optimal BS density that minimizes the APC. From these results it is seen that the existence of such optimal BS density depends on the value of the pathloss exponent.

7.2.1 Future directions

In [31], the authors explored the convexity of the coverage probability function for general unbounded pathloss model. This analysis could be carried forward towards optimization of APC and such metrics.

Modeling BS deployments according a uniform PPP is a slightly idealistic generalization. In real scenarios, BSs may be seen deployed densely at one place and relatively rarely in other places. Taking this sort of clustering of varied densities at varied spatial locations and carrying out the EE optimizations is an interesting problem. Backhaul (length of the optical fiber etc) costs may also be incorporated into the system model and studied how this is going to affect the optimal system parameters such as BS density etc.

7.3 Caching in single tier cellular networks

In Chapter 5, we studied how incorporating caching capabilities at the BSs influences the energy consumption in wireless cellular networks. Adopting a detailed BS power model and modeling the BS locations according to a PPP, we derived expressions for the APC and the EE, which are further simplified in the low noise regime. A key observation of this work is that cache-enabled BSs can significantly decrease the APC and improve the EE as compared to traditional BSs. We also observed that the existence of an optimum power consumption point for the APC depends on the pathloss exponent.

7.3.1 Future directions

The energy aspects and implications of caching in wireless cellular networks, especially for 5G systems, are of practical and timely interest and clearly require further investigation. Future work may include heterogeneous network scenarios, including small cells, macro cells and WiFi access points deployment. Furthermore, storing the popular content requires accurate estimation of content popularity distribution, which cannot be easily performed in practice and may cost energy in terms of processing power. Therefore, rather than relying on this approach, randomized caching policies in a stochastic scenario [32] can be considered as a means to provide crisp insights on the energy efficiency benefits of caching in dense wireless networks. Also, popularity distribution may be considered a function of spatial location as it is sensible to assume that what is popular at one place may not be popular elsewhere. In light of this, it makes sense to take the whole process of caching at a great detail in order to make a rigorous comment on whether caching is an energy efficient solution or not. [33], investigates the problem of optimal cache placement at the wireless edge to minimize the backhaul rate in heterogeneous networks.

7.4 Caching in two-tier cellular networks

In Chapter 6 we proposed a two tier heterogeneous system, one MBSs (without caching abilities) and one SBSs (with caching abilities), we defined several important system parameters and derived their expressions. We choose APC and EE as two sensible metrics to optimize in order to minimize energy consumption in a network. Taking λ_x and P_x (SBS density and transmit power respectively) to be the variables, we had four independent optimization problems to perform. We imposed the constraints C1-5 and found out that APC is a convex function in both λ_x and P_x. Since the equations involved are complicated to find out the exact analytical expressions for the optimal values, we find out the intervals within which the optima can be found. For a given system values, though, an optimum can be found out using numerical methods. Imposing the constraints C1-5, we found out that EE is a concave function with respect to λ_x and a monotonically decreasing function with respect to P_x. We found out the upper limit for optimal λ_x for which EE takes the peak value.

7.4.1 Future directions

In future works, it may be of interest to compare the two optimum values of SBS density - the one that minimizes the APC and the one that maximizes the EE, and verify which of the two metrics is relatively more beneficial economically and in terms of energy consumption. The choice of working with two tiers (*macro* and *small*) has been arbitrary and this treatment can be generalized to a system with multiple tiers (as in [34]) each with a different kind of caching capability. Taking energy consumption at each step of the process of caching users' content at the SBSs into account, gives us a more comprehensive hold on factors influencing energy efficiency and thereby letting us optimize it sensibly.

Bibliography

[1] D. J. Griffiths, *Introduction to electrodynamics*. Prentice Hall, 1999.

[2] Cisco. (2014) Cisco visual networking index: Global mobile data traffic forecast update, 2013-2018. [Online]. Available: http://goo.gl/l77HAJ

[3] V. S. Varma, "Energy efficiency optimization in wireless networks," Ph.D. dissertation, Université Paris Sud-Paris XI, 2014.

[4] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user mimo systems: Is massive mimo the answer?" *arXiv preprint arXiv:1403.6150*, 2014.

[5] E. Björnson, L. Sanguinetti, and M. Kountouris, "Designing wireless broadband access for energy efficiency: Are small cells the only answer?" *CoRR*, vol. abs/1504.07566, 2015. [Online]. Available: http://arxiv.org/abs/1504.07566

[6] C. Li, J. Zhang, and K. Letaief, "Energy efficiency analysis of small cell networks," in *Communications (ICC), 2013 IEEE International Conference on*. IEEE, 2013, pp. 4404–4408.

[7] G. Foundation, "Greentouch final results from green meter research study," 2015. [Online]. Available: http://www.greentouch.org/index.php?page=greentouch-green-meter-research-study

[8] J. D. Bekenstein, "Energy cost of information transfer," *Physical Review Letters, vol. 46 - 10.*, Mar 1981.

[9] S. Lloyd, "Physical limits to communication," *Physical Review Letters, vol. 93 - 10.*, 2004.

[10] L. Brillouin, "Science and information theory," *Courier Dover Publications.*, 2004.

[11] C. E. Shannon, ""a mathematical theory of communication"," *Bell System Technical Journal 27 (3): 379–423, 1948.*, 1948.

[12] R. Landauer, "Information is a physical entity," *Physica A 263 63-67.*, 1999.

[13] P. Gonzalez-Brevis, J. Gondzio, Y. Fan, H. Poor, J. Thompson, I. Krikidis, and P.-J. Chung, "Base station location optimization for minimal energy consumption in wireless networks," in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, May 2011, pp. 1–5.

[14] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, 2011.

[15] C. S. Chen, V. M. Nguyen, and L. Thomas, "On small cell network deployment: A comparative study of random and grid topologies," in *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, Sept 2012, pp. 1–5.

[16] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks: Volume 1: THEORY*. Now Publishers Inc, 2009, vol. 1.

[17] ——, *Stochastic Geometry and Wireless Networks Volume II: APPLICATION*. Now Publishers Inc, 2009.

[18] C.-H. Lee, C.-Y. Shih, and Y.-S. Chen, "Stochastic geometry based models for modeling cellular networks in urban areas," *Wireless Networks*, vol. 19, no. 6, pp. 1063–1072, 2013. [Online]. Available: http://dx.doi.org/10.1007/s11276-012-0518-0

[19] S. Mukherjee, *Analytical Modeling of Heterogeneous Cellular Networks*. Cambridge University Press, 2014.

[20] M. Haenggi, *Stochastic geometry for wireless networks*. Cambridge University Press, 2012.

[21] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[22] E. Baştuğ, M. Bennis, and M. Debbah, "Living on the Edge: The role of proactive caching in 5G wireless networks," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 82–89, August 2014.

[23] R. Jain, C. So-In, and A.-k. Al Tamimi, "System-level modeling of ieee 802.16 e mobile wimax networks: key issues," *IEEE Wireless Communications*, vol. 15, no. 5, pp. 73–79, October 2008.

[24] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, November 2011.

[25] C. S. Chen, V. M. Nguyen, and L. Thomas, "On small cell network deployment: A comparative study of random and grid topologies," in *IEEE Vehicular Technology Conference (VTC Fall'12)*, September 2012.

[26] E. Bjornson, L. Sanguinetti, and M. Kountouris, "Energy-efficient future wireless networks: A marriage between massive mimo and small cells," in *Signal Processing Advances in Wireless Communications (SPAWC), 2015 IEEE 16th International Workshop on*. IEEE, 2015, pp. 211–215.

[27] Y. S. Soh, T. Quek, M. Kountouris, and H. Shin, "Energy efficient heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 5, pp. 840–850, May 2013.

[28] S. Verdu, "On channel capacity per unit cost"," *IEEE Trans. on Inf. Theory, vol. 36, no. 5, pp. 1019-1030.*, Sep 1990.

[29] S. W. Kim, "Quantum szilard engine," *Phys. Rev. Lett. 106, 070401.*, Feb 2011.

[30] I. L. Markov, "Limits on fundamental limits to computation," *Nature*, vol. 512, no. 7513, pp. 147–154, 08 2014. [Online]. Available: http://dx.doi.org/10.1038/nature13570

[31] T. Samarasinghe, H. Inaltekin, and J. Evans, "Optimal SINR-based coverage in poisson cellular networks with power density constraints," in *IEEE 78th Vehicular Technology Conference (VTC Fall'13)*, September 2013.

[32] B. Blaszczyszyn and A. Giovanidis, "Optimal geographic caching in cellular networks," *arXiv preprint: 1409.7626*, 2014.

[33] V. Bioglio, F. Gabry, and I. Land, "Optimizing mds codes for caching at the edge," *arXiv preprint arXiv:1508.05753*, 2015.

[34] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of K-tier downlink heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550 – 560, Apr. 2012.

[35] C. H. Bennett, "Demons, engines and the second law," *Scientific American 257, no. 5, 108–116.*, 1987.

[36] ——, "Notes on landauer's principle, reversible computation, and maxwell's demon," *Studies in History and Philosophy of Modern Physics 34, 501–510.*, 2003.

[37] L. Szilard, "On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings," *Behavioral Science 9, no. 4, 301–310.*, 1964.

[38] M. Gopalkrishnan, "The hot bit i: The szilard-landauer correspondence." [Online]. Available: http://arxiv.org/abs/1311.3533v2

[39] L. Brillouin, "Maxwell's demon cannot operate:information and entropy. i," *J. Appl. Phys. 22, 334*, 1951.

[40] R. P. F. et al, *Feynman Lectures On Computation.* Westview Press, 2000.

[41] S. A. M. Ghanem, "Mutual information for generalized arbitrary binary input constellations," in *MAP-Tele Workshop.*, 2010.

[42] F. Richter, A. Fehske, and G. Fettweis, "Energy efficiency aspects of base station deployment strategies for cellular networks," in *Proc. IEEE Vehicular Techno. Conf.*, Anchorage, AL, Sep. 2009, pp. 1–5.

[43] D. Cao, S. Zhou, and Z. Niu, "Optimal base station density for energy-efficient heterogeneous cellular networks," in *2012 IEEE International Conference on Communications (ICC)*, June 2012, pp. 4379–4383.

[44] S. Sarkar, R. K. Ganti, and M. Haenggi, "Optimal base station density for power efficiency in cellular networks," in *2014 IEEE International Conference on Communications (ICC)*, June 2014.

[45] T. Samarasinghe, H. Inaltekin, and J. Evans, "Optimal sinr-based coverage in poisson cellular networks with power density constraints," in *Vehicular Technology Conference (VTC Fall), 2013 IEEE 78th*, Sept 2013, pp. 1–5.

[46] S. Göbbels, "Disruption tolerant networking by smart caching," *IEEE International Journal of Communication Systems*, pp. 569–595, Apr 2010.

[47] N. Golrezaei, K. Shanmugam, A. G. Dimakis, A. F. Molisch, and G. Caire, "Femtocaching: Wireless video content delivery through distributed caching helpers," in *IEEE INFOCOM*, 2012, pp. 1107–1115.

[48] E. Baştuğ, J.-L. Guénégo, and M. Debbah, "Proactive small cell networks," in *20th International Conference on Telecommunications (ICT'13)*, Casablanca, Morocco, May 2013.

[49] E. Baştuğ, M. Bennis, M. Kountouris, and M. Debbah, "Cache-enabled small cell networks: Modeling and tradeoffs," *EURASIP Journal on Wireless Communications and Networking*, Under minor revision (2014).

[50] A. Altieri, P. Piantanida, L. R. Vega, and C. Galarza, "On fundamental trade-offs of device-to-device communications in large wireless networks," *arXiv preprint arXiv:1405.2295*, 2014.

[51] P. Blasco and D. Gunduz, "Learning-based optimization of cache content in a small cell base station," *arXiv preprint arXiv:1402.3247*, 2014.

[52] K. Poularakis, G. Iosifidis, and L. Tassiulas, "Approximation algorithms for mobile data caching in small cell networks," *IEEE Transactions on Communications*, vol. 62, no. 10, pp. 3665–3677, October 2014.

[53] K. Hamidouche, W. Saad, and M. Debbah, "Many-to-many matching games for proactive social-caching in wireless small cell networks," in *12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, May 2014, pp. 569–574.

[54] M. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.

[55] J. Paakkonen, C. Hollanti, and O. Tirkkonen, "Device-to-device data storage for mobile cellular systems," in *IEEE Globecom Workshops (GC Workshops)*, 2013, pp. 671–676.

[56] M. S. ElBamby, M. Bennis, W. Saad, and M. Latva-aho, "Content-aware user clustering and caching in wireless small cell networks," in *11th International Symposium on Wireless Communication Systems (ISWCS)*, Barcelona, Spain, August 2014.

[57] A. Liu and V. Lau, "Cache-enabled opportunistic cooperative mimo for video streaming in wireless systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 390–402, January 2014.

[58] E. Baştuğ, M. Bennis, and M. Debbah, *Proactive Caching in 5G Small Cell Networks.* Wiley, [In Minor Revision] 2015. [Online]. Available: http://goo.gl/vxKNz1

[59] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of K-tier downlink heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550–560, April 2012.

[60] B. Perabathini, M. Kountouris, M. Debbah, and A. Conte, "Optimal area power efficiency in cellular networks," in *IEEE Globecom Workshops (GC Wrokshops)*, Austin, Texas, December 2014. [Online]. Available: http://goo.gl/ZP865y

[61] M. Z. Shafiq, L. Ji, A. X. Liu, and J. Wang, "Characterizing and modeling internet traffic dynamics of cellular devices," in *ACM SIGMETRICS joint international conference on Measurement and modeling of computer systems*, 2011, pp. 305–316.

[62] M. E. Newman, "Power laws, Pareto distributions and Zipf's law," *Contemporary physics*, vol. 46, no. 5, pp. 323–351, 2005.

[63] M. Imran, E. Katranaras, G. Auer, O. Blume, V. Giannini, I. Godor, Y. Jading, M. Olsson, D. Sabella, P. Skillermark *et al.*, "Energy efficiency analysis of the reference systems, areas of improvements and target breakdown," Tech. Rep. ICT-EARTH deliverable, Tech. Rep., 2011. [Online]. Available: http://www.ict-earth.eu

[64] A. Kumar and W. Saad, "On the tradeoff between energy harvesting and caching in wireless networks," in *IEEE International Conference on Communications (ICC'2015), Workshop on Green Communications and Networks*, London, UK, June 2015.

[65] K. Poularakis, G. Iosifidis, and L. Tassiulas, "Joint caching and base station activation for green heterogeneous cellular networks," in *IEEE International Conference on Communications (ICC'15)*, London, UK, June 2015.

[66] M. Gregori, J. Gomez-Vilardebo, J. Matamoros, and D. Gunduz, "Joint transmission and caching policy design for energy minimization in the wireless backhaul link," in *IEEE International Symposium on Information Theory (ISIT).* IEEE, 2015, pp. 1004–1008.

[67] D. Liu and C. Yang, "Will caching at base station improve energy efficiency of downlink transmission?" in *EEE Global Conference on Signal and Information Processing (GlobalSIP'14).* Atlanta, Georgia, USA: IEEE, December 2014.

[68] H. Inaltekin, M. Chiang, H. V. Poor, and S. B. Wicker, "On unbounded path-loss models: effects of singularity on wireless network performance," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1078–1092, September 2009.

[69] D. Zwillinger, *Table of integrals, series, and products.* Elsevier, 2014.

[70] S. Boyd and L. Vandenberghe, "Numerical linear algebra background," Tech. Rep., 2013. [Online]. Available: http://www.ee.ucla.edu/ee236b/lectures/num-lin-alg.pdf

[71] A. Mezghani and J. A. Nossek, "Power efficiency in communication systems from a circuit perspective," in *IEEE International Symposium on Circuits and Systems (ISCAS'11)*, May 2011, pp. 1896–1899.

[72] S. Tombaz, A. Vastberg, and J. Zander, "Energy-and cost-efficient ultra-high-capacity wireless access," *IEEE Wireless Communications*, vol. 18, no. 5, pp. 18–24, October 2011.

[73] B. Perabathini, E. Baştuğ, M. Kountouris, M. Debbah, and A. Conte, "Caching on the edge: a green perspective for 5G networks," in *IEEE International Conference on Communications (ICC'15)*, London, UK, June 2015.