Universidad
Rey Juan Carlos

TESIS DOCTORAL

Stochastic Schemes for Dynamic Network Resource Allocation

Autor:

Luis Miguel López Ramos

Directores:

Prof. Dr. Antonio García Marqués

Prof. Dr. Francisco Javier Ramos López

Programa de Doctorado Interuniversitario en Multimedia y Comunicaciones

Escuela Internacional de Doctorado

2016

Resumen

Las redes de comunicaciones inalámbricas y las redes de distribución de electricidad experimentan crecientes demandas en lo referente a eficiencia y fiabilidad. La necesidad de desarrollar métodos para asignar unos recursos escasos, como son la potencia y el ancho de banda, motivan la incorporación de herramientas de optimización no lineal y procesamiento de señal para su diseño óptimo. Esta tesis desarrolla esquemas para asignación eficiente de recursos en redes dinámicas, haciendo hincapié en el carácter estocástico tanto de la formulación de los problemas como del funcionamiento de los algoritmos. Se investiga el uso de optimización estocástica y de herramientas de descomposición para el desarrollo de algoritmos de baja complejidad con aplicaciones específicas en el diseño a través de capas (*cross-layer*) de comunicaciones inalámbricas, redes de radio cognitiva, y redes de distribución de electricidad (*smart grids*).

El carácter limitado de los recursos necesarios para el funcionamiento de las redes, junto con la inclusión de diversos requisitos de calidad de servicio (*Quality of Service, QoS*), hace que el diseño, la gestión y la operación de las redes planteen una serie de retos. Además, cuando la información de estado de la red es variante en el tiempo, los nodos deben adaptar sus parámetros de operación a dicha variabilidad. Se hace necesario el diseño mediante esquemas novedosos que no sólo utilicen información local sino que integren la información de los nodos a lo largo de la red, y que tengan en cuenta la incertidumbre referente los parámetros del sistema y en los eventos futuros.

Metodología

En esta tesis proponemos el diseño de esquemas óptimos mediante la solución de problemas matemáticos rigurosamente formulados teniendo en cuenta las variables aleatorias presentes en los modelos de red. Los tres elementos básicos de la formulación de un problema de optimización son las variables de diseño (recursos a asignar así como información de monitorización y de control), las restricciones de sistema y de calidad de servicio), y la función a optimizar u objetivo (usualmente formulada como suma ponderada de varias funciones de utilidad).

Una vez formulado el problema de optimización, se identifica el tipo de problema que es y se procede a su solución, ya sea mediante la aplicación de algoritmos existentes para el tipo de problema en cuestión, o el desarrollo de nuevos algoritmos mediante la adaptación o combinación de algoritmos existentes para problemas relacionados. La validación experimental de los algoritmos propuestos es imprescindible, ya que permite validar el modelado del problema, el funcionamiento del algoritmo propuesto y el análisis de aspectos técnicos que trascienden la notación matemática.

Antecedentes

Existen trabajos previos en asignación dinámica de recursos en redes inalámbricas que integran información de diversas capas (física, enlace, red...). Algunos de estos trabajos están basados en optimización convexa y descomposición en el dominio dual, a la vez que utilizan información sobre la ganancia instantánea de los canales inalámbricos (*fading*); otros trabajos establecen políticas de

backpressure basadas en técnicas de control adaptativo y que tienen como objetivo estabilizar las colas de transmisión de la red.

En los últimos años, las redes de *cognitive radio* (CR) se han posicionado como la próxima generación de comunicaciones inalámbricas y móviles para mitigar el problema de la infrautilización del espectro, gracias a su capacidad para limitar la interferencia que se causa a los usuarios coexistentes (usuarios*primarios*). Para hacer esto, las CRs deben adquirir información sobre la presencia de señales en el espectro y adaptar sus parámetros de transmisión de acuerdo a la información percibida y a los recursos de transmisión disponibles. Los trabajos existentes en este contexto limitan la interferencia causada, bien mediante la imposición de restricciones (a corto o largo plazo) sobre la potencia de transmisión; o bien controlando la probabilidad de interferir con las transmisiones primarias. La información captada por los sensores es generalmente imperfecta, debido ya sea a errores en el sensado, a la cuantificación (baja resolución) de la información codificada por los sensores, o a que la información percibida quede desactualizada. Ante este problema, los algoritmos estocásticos de asignación de recursos permiten adaptarse a condiciones variables en los canales o a cambios en el patrón de comportamiento de los usuarios primarios. En el mismo contexto, existen también trabajos que ponen de manifiesto el compromiso (*tradeoff*) entre la adquisición de información (*sensing*) y la asignación de recursos en CRs y que investigan su diseño conjunto.

En el contexto del despacho (asignación de recursos) óptimo de energía en *smart grids*, la creciente penetración de las fuentes de energía renovables plantea nuevos retos relacionados con la variabilidad temporal y la incertidumbre. En este contexto, existen trabajos que formulan problemas de asignación de recursos con restricciones de seguridad y de fiabilidad basados en herramientas de optimización estocástica, mientras que otros trabajos más recientes demuestran que los *inverters* (presentes en los sistemas de generación fotovoltaica) se pueden controlar para apoyar la regulación del voltaje de la red. Algunos trabajos más recientes desarrollan esquemas de operación para *smart grids* en diferentes escalas de tiempo, o se basan en aproximación estocástica para explotar un cierto margen de sobrecarga esporádica permitido por los estándares para facilitar la integración de fuentes de energía renovable.

Objetivos

El primer objetivo de la tesis es resolver un problema estocástico de asignación de recursos a través de capas que utiliza información instantánea de *fading* y de longitud de colas para asignar de forma óptima recursos en las capas física, de enlace y de transporte.

El segundo objetivo de la tesis es optimizar el rendimiento (en términos de tasa de transmisión) de una red de CR, limitando la probabilidad de interferir a los usuarios primarios. Se pretende además analizar en qué medida la formulación de algoritmos estocásticos y el tratamiento mediante estos de información de estado ruidosa y/o desactualizada suponen una mejora con respecto a algoritmos previos.

Como objetivo adicional de la investigación en CR, se propone la resolución de la optimización conjunta de los procesos de adquisición de información (*sensing*) y asignación de recursos para CR y el análisis del impacto de tener en cuenta la correlación temporal de la información de estado sobre la actividad de los usuarios primarios utilizando programación dinámica estocástica (SDP).

Finalmente, en el contexto de *smart grids*, se aborda el objetivo de asignar recursos de una red de distribución en dos escalas de tiempo diferentes, en presencia de restricciones en promedio o en probabilidad sobre las decisiones en la escala de tiempo más rápida. Se pretende analizar la mejora en términos económicos de utilizar un modelo de red aproximado para minimizar el coste de operación esperado mediante la resolución de un problema de punto de silla estocástico.

Resultados

Los resultados teóricos sobre asignación de recursos a través de capas permiten establecer una relación entre las longitudes de las colas de transmisión y ciertos multiplicadores de Lagrange del problema planteado; esta relación puede ser utilizada para estimar y controlar los retardos de la red, y además permite establecer prioridades de retardo entre usuarios.

El diseño conjunto de la adquisición de información (*sensing*) y la asignación de recursos en CR permite alcanzar un rendimiento óptimo de la red y tiene la capacidad de adaptarse a condiciones del canal variantes. Una estrategia de diseño en dos pasos ha permitido reducir significativamente la complejidad computacional de la solución con alternativas de baja complejidad; además, a partir de dichos esquemas y los resultados experimentales se generaron mapas que permiten visualizar la decisión óptima en función de la información de estado.

Los esquemas desarrollados para *smart grids* optimizan decisiones de gestión energética en dos escalas de tiempo diferentes y manejan variables acopladas en el tiempo de forma estocástica. Las restricciones en promedio y en probabilidad se han tratado mediante descomposición dual y optimización convexa. Dichos esquemas funcionan basándose en datos de generación y demanda, y convergen a decisiones óptimas de asignación de potencia.

Los resultados obtenidos a lo largo de la tesis confirman que los esquemas estocásticos diseñados alcanzan un rendimiento óptimo o cercano al óptimo, son robustos ante procesos no estacionarios en la información de red, cumplen las restricciones instantáneas y además cumplen asintóticamente las restricciones impuestas a largo plazo.

Conclusiones

La investigación realizada ha permitido diseñar, de forma sistemática, esquemas estocásticos que alcanzan un rendimiento óptimo o cercano al óptimo, son robustos ante procesos no estacionarios en la información de red, cumplen las restricciones instantáneas y además cumplen asintóticamente las restricciones impuestas a largo plazo.

En relación con el objetivo de asignación de recursos a través de capas, se ha realizado la optimización mediante un esquema estocástico tanto para el enlace de subida (*uplink*) como de bajada (*downlink*) y se ha establecido una relación matemática entre ciertos multiplicadores de Lagrange y la longitud de las colas de transmisión, lo que adicionalmente ha permitido analizar y controlar los retardos que afectan a distintos usuarios de la red celular.

El objetivo de optimización de la red de CR se ha llevado a cabo con éxito y además se han obtenido diversas variantes del esquema de asignación de recursos, en los que la manera en que se imponen las restricciones de probabilidad de interferencia varía, pudiendo ser esta una restricción instantánea o a largo plazo.

Asimismo, la optimización conjunta de la adquisición de información y asignación de recursos se ha resuelto y se ha mejorado la eficiencia computacional del algoritmo desarrollado gracias al análisis detallado de la estructura matemática del problema. Se ha comprobado, además, que la formulación utilizada permite programar versiones estocásticas con carácter adaptativo.

En lo referente al despacho óptimo en *smart grids*, se ha verificado que la asignación conjunta recursos de generación convencional y alternativa, en dos escalas de tiempo y bajo restricciones a largo plazo sobre el voltaje, puede formularse como un problema cóncavo-convexo. La solución óptima de dicho problema ha sido adaptada para aproximar la solución de un problema análogo con restricciones probabilísticas de voltage, que ha sido validada de forma numérica.

La presencia de restricciones de potencia media ha sido un factor común en los diseños de redes inalámbricas acometidos, así como su tratamiento utilizando descomposición dual y esquemas de subgradiente dual estocástico. Se ha seguido una estrategia similar para tratar con restricciones probabilísticas (no convexas) en CRs y en *smart grids*, dando lugar a esquemas efectivos con una carga computacional asequible.

Abstract

Wireless networks and power distribution grids are experiencing increasing demands on their efficiency and reliability. Judicious methods for allocating scarce resources such as power and bandwidth are of paramount importance. As a result, nonlinear optimization and signal processing tools have been incorporated into the design of contemporary networks. This thesis develops schemes for efficient resource allocation (RA) in such dynamic networks, with an emphasis in stochasticity, which is accounted for in the problem formulation as well as in the algorithms and schemes to solve those problems. Stochastic optimization and decomposition techniques are investigated to develop low-complexity algorithms with specific applications in cross-layer design of wireless communications, cognitive radio (CR) networks and smart power distribution systems.

The costs and constraints on the availability of network resources, together with diverse quality of service (QoS) requirements, render network design, management and operation challenging tasks. Moreover, when the available network state information (NSI) is time-varying, nodes must be capable of adapting their operation parameters to the those variations. While traditional schemes only exploited local information, or did not take into account uncertainty in system parameters and future events, the methodology proposed in this thesis aims at designing schemes through the solution of rigorously formulated mathematical problems involving random variables. The three basic elements when formulating an optimization problem are the design variables (mainly corresponding to resources to be allocated, and sensing and control information), system and QoS constraints, and the objective function (usually formulated as weighted sums of utility functions). Equilibrium between problem tractability and realistic modeling is also sought.

Previous works on dynamic RA for wireless networks that take into account cross-layer information rely on convex optimization and dual decomposition techniques, while exploiting instantaneous fading; whereas others build upon dynamic backpressure policies based on adaptive control tools and aim to stabilize queues of the network. Based on such a background, our first contribution is the design of a stochastic scheme that uses instantaneous fading and queue length information to optimally allocate resources at the transport, link and physical layers. Theoretical results allow to establish links between the transmit queue lengths and Lagrange multipliers; these links can be used to estimate and control queuing delays and establish delay priorities among users. Experimental results confirm that RA feasibility and optimality are preserved when the scheme employs window averages of multipliers and queue lengths.

In the last years, CRs have emerged as the next-generation solution to the perceived spectrum underutilization, thanks to their capability to limit the interference inflicted to coexisting primary users (PUs). To do so CRs must sense the spectrum and dynamically adapt their transmission parameters according to the available resources and sensed information. Existing works limit CR-inflicted interference either through short- and long-term transmit-power constraints; or, by controlling the probability of interfering with PU transmissions. As the sensed information may be imperfect (due to errors or quantization) and get outdated, implementation of stochastic RA algorithms offers the possibility of adapting the operation of the network to varying channel conditions and PU behaviour. Motivated by these facts, we develop stochastic RA algorithms that

optimize sum-rate performance of a CR network, limit the probability of interfering with PUs, and jointly account for outdated and noisy NSI. Additional works in this context point out the tradeoff between the sensing and RA and deal with their joint design. Hence, we also develop a jointly optimal sensing and RA algorithm that additionally takes into account the temporal correlation of the primary NSI and the sensing cost by means of stochastic dynamic programming (SDP). The proposed schemes obtain optimal performance, account for imperfections in the acquired state information, and are able to adapt to varying channel conditions. A two-step strategy significantly reduces the computational solution complexity without loss of optimality, and its formulation allows developing adaptive stochastic schemes. Numerical experiments confirm a significant performance improvement with respect to low-complexity alternatives and allow to trace sensing-decision maps.

The last part of the thesis addresses the design of RA algorithms for smart grids, where increasing penetration of renewable energy sources (RESs) pose new challenges related to variability and uncertainty. Previous works formulate security-constrained and risk-limiting dispatch schemes based on stochastic optimization tools, whereas recent works demonstrate that power inverters can be controlled to effect voltage regulation. More recent works deal with smart grid operation schemes in different timescales, or capitalize on stochastic approximation to limit the magnitude of sporadic component overloads. Based on this background, we consider joint dispatch of slow- and fast- timescale distribution grid resources under average or probabilistic constraints over fast-timescale decisions. Using an approximate grid model, the expected network operation cost is minimized under inverter and voltage constraints, and the two-stage dispatch is formulated as a stochastic saddle-point problem. The developed dispatch algorithms account for conventional and alternative energy resources at different timescales, which are coupled across time in a stochastic manner. Average and probabilistic voltage constraints are tackled using dual decomposition and convex optimization. The resulting schemes rely on random samples of stochastic generation and demand, and converge to optimal dispatch decisions.

The experimental results throughout the thesis confirm that the proposed stochastic RA schemes achieve optimal or near-optimal performance, are robust against non-stationary NSI, fulfill instantaneous constraints and asymptotically fulfill long-term constraints. Average power consumption constraints and their treatment by means of dual stochastic gradients are common ground for the cross-layer and CRs designs undertaken. A similar strategy is followed to deal with (nonconvex) probability of interference constraints in CRs and voltage probabilistic constraints in smart grids with manageable complexity.

Acknowledgments

I would like to start by expressing my special appreciation and thanks to my advisor Professor Dr. Antonio G. Marques, who has been mentoring me during the last years. He has encouraged my research and supervised my doctoral work, always through constructive and respectful critique. The research in wireless communications and cognitive radios in this thesis has become reality thanks to his close collaboration, laying strong foundations for my future research.

My sincere and deep gratitude goes also to Professor Dr. Vassilis Kekatos, who has devoted time and effort to help me to expand my professional horizons in the interesting topic of smart grids. This has been possible thanks to his active motivational effort, insightful suggestions, and patient and attentive supervision of my research.

I also want to acknowledge Professor Dr. Javier Ramos not only for serving as co-advisor and helping to get the funding necessary to develop my doctoral work, but also for his everyday work to forge ahead the young but promising Fuenlabrada School of Telecommunication Engineering. This acknowledgment also extends to the professors, researchers, and PhD students in the Signal Theory and Communications research group who have supported and helped me during these years making these offices a really welcoming place to work.

I would also like to thank Professor Georgios B. Giannakis for his attention and advice, and his kind invitation for me to visit the SPiNCOM research group in the University of Minnesota School of Science and Engineering, where he leads a wonderful team of people with brilliant careers and interesting personalities. My participation in the group activities, especially for the group colloquiums, has greatly contributed to improve my research and communication skills, and has been an unforgettable experience.

The material of this thesis was also benefited from discussions with Dr. Yu Zhang, Dr. Emiliano Dall'Anese, and Professors Antonio Caamaño, Jerónimo Arenas, Seung-Jun Kim, Antonio Conejo, Eduardo del Arco, and Inmaculada Mora.

Due thanks go to Professors José Luis Rojo, Geert Leus, Luis Cadarso, Carles Antón-Haro, Milan Prodanovic, Sancho Salcedo, and Paolo Banelli for agreeing to serve in my PhD examination committee; and Professors Nikolaos Gatsis and Vassilis Kekatos for serving as reviewers of this thesis.

This research work has been financially supported by the Spanish Ministry of Education FPU grant AP2010-1050, the Spanish Ministry of Economy and Competitiveness (MINECO) Grant TEC2013-41604-R, Comunidad de Madrid (CAM) Grant S2013/ICE-2933, and the EU-FP7 (ICT-2011-9-TUCAN3G).

I would also like to acknowledge those people who have collaborated with me in research work not included in this thesis, and have undoubtedly contributed to my doctoral education: Eduardo Perdices, Prof. Jose María Cañas, Dr. Santiago Segarra, Dr. Sergio Muñoz, and Prof. José Luis Rojo.

Last but not least, I want to thank the constant and unconditional support that I have received from my family and close friends who, in many occasions, have helped me to gain and retain the self confidence I needed to complete this PhD.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Societies in an extremely connected world depend heavily on services provided by engineered networks – emblematic examples being telecommunications and electricity. The widespread availability of internet access and the maturity of wireless communication technology allow to create new personalized communication applications and evolved business models. At the same time, the increasing penetration of renewable energy sources in the power grid brings hope to circumvent the scarcity of fossil fuels. However, an ever-growing human population, together with the urge to reduce the power consumed by communication devices and the need to cope with aging infrastructures and overcrowded spectrum bands, are putting increasing demands on the required efficiency, reliability, and sustainability of such networks. To satisfy such demands, careful network planning and judicious management of the available resources are crucial. Network size, complexity, and time-variability render these tasks challenging and call for systematic design approaches.

To satisfy the demand for fast and error-resilient wireless communications providing quality of service (QoS) for diverse applications, judicious methods for allocating the available (scarce) power and bandwidth resources are of paramount importance [Gat12]. To this end, nonlinear optimization has been fruitfully brought into the design of wireless networks —see e.g., [Lin06b, Geo06, Chi07, Sha08]— and is instrumental in devising efficient link- and network-adaptive resource allocation (RA) algorithms. Similarly, the project of a single, smart electricity grid [Far10] that efficiently integrates distributed and renewable energy sources (RESs) without sacrificing reliability requires intelligent control and *dispatch* schemes. Apart from the incorporation of state-of-the-art power electronics, sensor and communication infrastructure, and computational intelligence [Fan12], nonlinear optimization and signal processing tools are also key to implement smart control strategies and optimal resource management algorithms.

This thesis develops schemes for efficient RA in dynamic (time-varying) networks, with an emphasis in stochasticity, which is incorporated in problem formulations and solution schemes. Optimization methodologies and decomposition techniques are investigated to develop algorithms with specific application to both wireless communication and electrical power networks.

The outline of this first chapter is as follows. Section 1.1 provides a more detailed explanation about network RA, starting with the general concepts and then moving to more application-specific concepts. Section 1.2 describes the mathematical tools that will be used throughout the thesis. Section 1.3 contains the motivation, methodology and research objectives pursued during the development of the thesis. The chapter ends in Section 1.4 with a presentation of the outline of this document, and the enumeration of the main contributions.

1.1 Resource Allocation in Networks

Since this thesis deals with various types of networks, we start by reviewing general network aspects. Then, we discuss the role of dynamic network RA schemes, and follow with a description of the practical scenarios for which such schemes will be developed.

A network is a system of interconnected agents (nodes) where a global behavior emerges from the aggregation of local interactions (links), defining a notion of proximity or dependence between entities with an underlying graph structure. Sometimes the network is defined by the structure underlying a set of magnitudes. In other occasions, it is the physical structure of the network what influences the relationships between such magnitudes.

This thesis focuses on networks that are engineered to have resources flowing through their links, and injected (introduced or extracted) at their nodes. The combination of injections and flows permit the transportation of commodities (such as data) or resources (such as energy) from one point of the network to another. It is desirable to do this process efficiently, and resources necessary for competitive network function are generally limited, so that the main challenge of network design and operation is to balance the associated tradeoffs. As increasingly stringent performance requirements are demanded from the aforementioned networks, their design and operation experience an intensive increase in complexity, which requires a detailed analysis.

In a *dynamic* network its state is a random process that varies in time and space, and usually there is interest in monitoring and controlling such a process. In general terms, the performance of a network depends on the availability of network state information (NSI). In wireless networks where NSI varies, nodes must be capable of adapting their transmission and reception parameters to these changes, while adhering to constraints on the usage of resources and QoS requirements.

A network will rarely perform optimally if its design and operation are based on traditional schemes that only exploit local information, or if their operation is planned based only in information about the present (myopic schemes) without having into account their impact on future events. A holistic, integrated, and multi-stage network design is necessary instead. Information on future events is often stochastic and usually more uncertain than that regarding the present. Similarly, from the point of view of a node, information on nodes that are far in the network is generally more uncertain than local information.

Optimal network RA comes from the solution of a judiciously and rigorously formulated mathematical problem. The practical implementation of the aforementioned schemes requires signaling schemes and communication protocols, and problems should be formulated in a way that facilitates the design of such protocols. The developed models should be sufficiently realistic (complex) as to grasp the coupling between variables relevant to the network performance. At the same time, simplicity is a useful property since algorithms need to be designed with implementability and scalability in mind. Knowledge of the practical application helps to keep this balance between realism and implementability. In the remainder of this section we will describe the technologies for which RA schemes are developed.

1.1.1 Wireless communications and cross-layer design

Contemporary wireless communication networks serve traffic with diverse QoS requirements for different applications including voice, data and real time, or streaming video/audio. Early wireless communications standards defined several levels of QoS requirements and established fields in protocols for terminals to request them, but did not specify the RA schemes that would ultimately provide QoS support. For this reason, the design of optimal *dynamic* RA algorithms that take into account *cross-layer* information has attracted the attention of information theory, signal processing, communications, adaptive control, and networking communities. Existing works in this area can be mainly classified in one of the following two categories. In the first category, approaches rely

on convex optimization and dual decomposition techniques; see e.g., [Lin06a, Lin06b, Che06, Lin05, Sol06]. In those, the solution of a properly formulated optimization problem dictates how resources are allocated, while the structure of the solution typically suggests the design of signaling protocols. In the second category, approaches build upon dynamic backpressure policies, see e.g., [Tas92, Nee05, Ery06, Gia06, Sto05, Geo06]. These works rely on adaptive control tools (the so-called Lyapunov optimization), and aim at stabilizing queues of the wireless network. Such queues correspond not only to actual queues where packets are buffered before transmission, but also to virtual queues whose role is to account for QoS requirements. Most existing works have focused on the impact of the cross-layer features on the higher layers, and fewer have addressed the design of cross-layer networking algorithms that take into account *instantaneous fading*; see [Lin06a, Che06, Lin05, Wan07, Rib10a, Geo06, Gat10, Mar13] for some of the exceptions.

More information about stochastic cross-layer RA based on fading and queue information can be found in Chapter 2.

1.1.2 Cognitive Radio Networks

Cognitive radios (CRs) are next-generation wireless radios that use the spectrum more efficiently by intelligently adapting their transmission parameters to a varying environment [Hay05, Zha07a] while avoiding harmful interference to other users. The perceived spectrum under-utilization has motivated intensive research on dynamic spectrum management, for which CRs are the enabling technology.

In most CR designs, users holding the license of spectrum are called primary users (PUs), whereas secondary users (SUs) are those in charge of dynamically accessing the spectrum that PUs do not utilize. In many practical scenarios, PUs are assumed not to collaborate with the SUs, mainly because this allows coexistence of non-cognitive PUs (complying with older standards) with SUs that implement more advanced processing capabilities.

Depending on the operating conditions and the requirements of the primary users, CRs can be deployed under different paradigms [Zha07a, Gol09]. Three of the most extensively studied in the literature are the interweave, overlay and underlay paradigms [Gol09, Zha10][1]. Overlay and underlay CRs transmit at the same time and band with PUs while keeping interference below a certain level. Overlay CRs align their transmissions with those of the PUs, which requires knowledge about the codebook and messages that the PUs send [DD12]. Underlay CRs adjust their power so that interference at active PUs remains below a pre-specified threshold [Zha07a, Gol09, Zha10], requiring state information about the primary-to-secondary channel gains. Differently, interweave CRs are allowed to use a frequency band only if no PU is active in it. To do so, spectrum sensing is employed to detect empty portions of the radio spectrum at a certain time. Upon detection of a transmit opportunity, the CR uses the channel but is required to vacate it when accessed again by a PU.

Effective operation of either type of CR requires the implementation of two critical tasks: i) sensing the spectrum seeking transmit opportunities and ii) dynamically adapting their transmission power and modulation according to the available resources and the sensed information [Hay05, Mar14]. Harmful interference can be avoided by adding appropriate constraints in the RA formulation. It is also worth stressing that the state information collected by sensing is generally heterogeneous because there is more uncertainty on the information related to the primary network than that of the secondary network.

More detailed information on the specific CR-related challenges that are faced in this thesis will be described in Chapters 3 and 4.

[1] Initially, most works referred to interweave CRs as overlay CRs (in contrast to underlay CRs). Nowadays, the difference between the three paradigms is clear and the term interweave CRs to refer to the scenario where the SUs exploit spectrum holes to transmit opportunistically is widely accepted; see e.g., [Gol09].

1.1.3 Electric Power Networks

Convergence of power networks and information technologies give rise to smart grid technology [Fan12], which is called upon minimizing the environmental impact of the energy production, fostering the utilization of RESs, integrating flexible energy storage, and establishing a novel social contract with network agents, among other innovations. To address them, industry capitalizes on state-of-the-art information technologies such as sensing, control, communications, and machine learning [Ami05, Gia13].

Transmission (high voltage, long distance) and distribution (medium voltage, shorter distances) networks are modeled and operated in quite different ways and face different challenges. Recent advances in network optimization and control give rise to microgrids, which are distribution systems containing loads and distributed energy resources (DERs) (such as distributed generators, storage devices, or controllable loads) that can be operated in a controlled, coordinated way either while connected to the main power network or while islanded. Microgrids act as single controllable entities with respect to the grid and are envisioned to provide resiliency against time-variability [Gia13].

Signal processing already plays an important role in various grid monitoring tasks such as power system state estimation, line outage identification [Zhu12], and load and electricity price forecasting. There are also research efforts in solving new emergent and intriguing inference and learning problems such as blind topology identification [Kek14] and cyberattack detection [Kun10].

Leveraging the state information captured by the aforementioned monitoring, the smart grid will be operated with significantly improved efficiency. Optimization tools have been successfully employed in tasks such as *economic dispatch* (setting generators power output so that the load is served at the minimum operation cost). *Optimal power flow (OPF)* formulations extend the economic dispatch by including the effects of the transmission/distribution network in the problem formulation, so that line losses are minimized and line thermal limits are preserved. As a matter of fact, optimal Lagrange multipliers of such problems are used to calculate wholesale electricity prices [Woo12] and research efforts are being placed on determining more advanced pricing schemes such as real-time prices based on more sophisticated optimal dispatch formulations.

Due to the increasing renewable generation and its stochastic nature, the growth of variability and uncertainty pose new challenges in energy management. For example, solar energy from photovoltaic (PV) units can change significantly over one-minute intervals. The power inverters found in PV units can be commanded to curtail active power generation or adjust their power factor within seconds [Liu08], [Car08], but conventional generation cannot follow such fast-varying generation. Operation in such different timescales calls for *multistage* dispatch solutions.

All these challenges motivate the incorporation of stochastic optimization (see Sec. 1.2.2) to formulate efficient, security-constrained dispatch schemes [Bou08] and risk-limiting approaches [Var11]. These tools are also envisioned to ensure efficient and reliable operation under normal and emergency conditions. While in most traditional designs reliability came at the cost of excessive conservativeness, stochastic optimization can better leverage the available (statistical) information to provide reliability at a minimum sacrifice in efficiency.

Optimal formulations aim at minimizing expected operation costs including cost penalties upon deviating from energy market schedules [Var11]. Apart from ensuring user satisfaction and representing physical limitations of network components, constraints can also take care of security issues such as thermal limits of transmission/distribution lines, and control over under/over-voltages; and reliability guarantees such as limiting the probability of blackouts and cascading failures.

More detailed information on dispatch and reliability issues in distribution grids in different timescales will be provided in Chapter 5.

1.1.4 Common aspects

Albeit different in nature, communication and power networks share a number of properties and challenges that motivate a coordinated investigation of their optimal operation and design.

Network topology influences RA formulations because of capacity limits and flow conservation laws. The amount of data that can be transmitted through a wireless link is limited, as well as the power that can flow through a power line without overheating. Routing nodes are required to balance their incoming and outbound traffic to prevent large delays, and the power injections at a bus (nodes in power networks) always must equal the power consumed by the loads located in it.

Another noticeable parallelism is the presence of time-varying magnitudes. Wireless networks face time-varying conditions such as local data-flow injections, channel gains (fading), and PU activity in the case of CRs. The main varying conditions in power networks are essentially demand and renewable available energy. As a result, both call for the implementation of adaptive schemes that monitor the NSI and adapt the allocation of resources accordingly.

In both cases, there is interest in developing RA algorithms that limit the probability of some specific undesirable events, such as outages in mobile communications, and interference to PUs in CRs (see Sec. 1.1.2). In power networks, the probability of over-/under-voltages or load shedding should also be maintained under a prespecified limit.

The mathematical formulation of optimization problems in these networks will share also common aspects, such as aggregate-utility maximization (or, equivalently, cost minimization) objectives through the (possibly weighted) sum of nodal objectives, and the expression of capacity limits and flow conservation laws as constraints. The following section elaborates on the optimization techniques that will be used to formulate such problems.

1.2 Optimization algorithms

The use of optimization theory has been key to develop cross-layer communication protocols and CR networks. It is also playing a prominent role in the design of contemporary schemes to operate smart grids [Far13b], where novel optimization techniques allow to operate with more accurate network models. This thesis is devoted to extend the currently well-studied use of optimization methods for network design through the incorporation of stochastic tools. Specifically, this implies the incorporation of stochastic programming[2] [Kal94] and stochastic approximation [Kus03], or even combinations of both [Nem09], whereas the proposed algorithms will deal with stochastic variables / processes, and the incorporation of stochastic approximation grants algorithms with a functioning of stochastic nature.

When using optimization to design practical control schemes, apart from being able to find the optimal solution, the structure of the solution itself is very important since it can be leveraged to suggest signaling or message passing algorithms. Also algorithm byproducts (such as Lagrange multipliers or value functions) contain relevant information about the solution (e.g. economic interpretations or congestion indicators).

The rest of the section provides a description of the specific mathematical tools that have been utilized in this research.

1.2.1 Convex optimization and decomposition techniques

Early works that use optimization as a mathematical tool to analyze network protocols are [Kel98] and [Low99]. The gist of these works is that congestion control protocols can be viewed as distributed implementations of algorithms that solve utility maximization problems. Source rates

[2]The term *programming* is often used to refer to optimization techniques. As well, stochastic optimization problems are commonly referred to as *stochastic programs*.

are regarded as primal variables and congestion parameters constitute variables of the corresponding Lagrange dual problems. Recursive schemes updating these variables boil down to subgradient descent iterations on the dual function—the kind of optimization algorithm also known as dual decomposition [Rib10b].

Although some works advocate alternative decomposition methods [Joh06, Pal06], most of the network optimization literature relies on dual decomposition. This is because the associated Lagrangian function exhibits a separable structure reminiscent of layered network designs [Rib10b].

On this issue, though, it is important to stress that since networking problems can be nonconvex, the duality gap is generally nonzero. As a result, the dual optimum is generally different from the primal optimum and for this reason layering is believed to come at the price of optimality loss. However, recent advances [Rib10b] show that in scenarios dealing with stochastic NSI, under certain conditions imposed on the distribution of the random state variables, the duality gap of such problems is zero. Prior to elaborating on this fact, let us describe the techniques that allow to solve optimization problems in the presence of random variables.

1.2.2 Stochastic optimization

Solving optimization problems under uncertainty (affecting model parameters or possible future events) has long been a focus of mathematical programming research, and of high relevance for engineering applications. In this context, robust optimization [Ber11] and stochastic optimization [Kal94] allow a more efficient modeling than deterministic optimization problems. When unknown variables are known to belong to a certain set and their statistical distribution is unknown, robust approaches are useful. On the other hand, information about the statistics of uncertain variables can be exploited, whether it is in the form of a probability density function (pdf) or a set of samples. In these cases, the stochastic approach is less conservative and can lead to more efficient solutions; in exchange, solving this kind of problems is usually more computationally demanding.

Stochastic optimization is frequently applied in setups where decisions must be made in several stages in order to integrate information that is revealed at different time instants. Two-stage stochastic problems have been widely studied using previously established strategies such as Benders decomposition [Bir85] or dual decomposition [Rus97] , whereas multi-stage problems call for stochastic dynamic programming (SDP) (see Sec. 1.2.4).

1.2.3 Stochastic approximation

A basic difficulty of solving (stochastic optimization) problems that often involve expectations over multidimensional random variables is that the integrals involved cannot be computed with a high accuracy. Two computational approaches to approximate these expectations based on Monte Carlo sampling techniques are the stochastic approximation (SA) and the sample average approximation (SAA) methods. The SA method is going back to the pioneering paper by Robbins and Monro [Rob51]. Since then, SA algorithms became widely used in stochastic optimization (see, e.g., [Pfl12] and references therein) and, due to especially low demand for computer memory, in new and diverse areas such as signal processing and communications.

Indeed, whether or not they are called stochastic approximations, such algorithms occur frequently in practical systems for the purposes of noise or interference cancellation, the optimization of "post processing" or "equalization" filters in time varying communication channels, adaptive antenna systems, adaptive power control in wireless communications, and many related applications [Kus03].

In several RA approaches that rely on convex optimization and dual decomposition techniques, see e.g., [Lin06a, Lin06b, Che06, Lin05, Sol06], the solution of a properly formulated optimization problem dictates how resources are allocated, while the structure of the solution typically suggests the design of signaling protocols. In the last years, works such as [Mar08b, Mar09, Gao09, Rib10a]

developed adaptive *dual* stochastic algorithms to estimate the optimal Lagrange multiplier values. The developed algorithms are robust to non-stationarities, do not require knowledge of the channel fading gain distribution, and incur arbitrarily minimal loss of performance (optimality). Motivated by these advances, recent works develop stochastic dual methods for power networks dispatch [Wanar], load shedding [Gat14] and reactive power compensation [Kek15a], which are also robust to non-stationary and unknowingly-distributed renewable energy injections.

1.2.4 Stochastic dynamic programming

Stochastic dynamic programming (SDP) is a technique used to optimize the operation of discrete-time complex systems, where decisions have to be made sequentially and there is a dependency among decisions that are made at different time instants. In this framework, presence is assumed of an agent which has the ability to take certain actions that will modify the (stochastic) way in which a controlled system evolves. To be more precise, the system to be controlled is modeled as a state-space model composed of: a set of state variables ; a set of actions which are available to the controller and which can depend on the state ; a random function that describes the transitions of the system depending on the actions taken; and a function that defines the reward associated with a state transition.

In general, finding the optimal solution of a SDP is computationally demanding. Unless the structure of the specific problem can be exploited, complexity grows exponentially with the size of the state space, the size of the action space, and the length of the temporal horizon. This is commonly referred to as the *triple curse of dimensionality* [Pow07]. Two classical strategies to mitigate such a problem are: i) framing the problem into a specific, previously studied model and ii) find approximate solutions that reduce the computational cost in exchange for a small loss of optimality.

Key to solving a SDP problem is the definition of the so-called *value function* which is a mathematical tool that allows to efficiently quantify the effects of current actions in future time instants. SDP-solving algorithms are generally based on estimating the value function by iterative means. Statistical signal processing can also be helpful in this context as it is commonly required to approximate value functions through parametric estimation, and adaptively modifying those parameters as new information on the system is learnt (see also [Bus10] and the last part of Chapter 4).

(Partially observable) Markov decision processes

Markov decision processes (MDPs) are a relevant subclass within SDP problems. For such problems, the state transition and reward function depend only on the state and the action taken in the current time instant. MDPs with finite state-action spaces can be solved exactly for finite-horizon problems. For infinite-horizon problems, the solution can be approximated with arbitrary precision.

A partially observable Markov decision process (POMDP) can be viewed as an MDP for which the state is not always known perfectly. Only an observation of the state (which may be affected by errors, missing data or ambiguity) is available. To deal with these challenging properties, it is assumed that a (possibly random) *observation function* is known. The POMDP framework provides a systematic method for using the system history of actions and observations to aid in the disambiguation of the current observation to the effect of taking optimal actions. The key point is the definition of an internal *belief state* accounting for previous actions and observations. The belief state is a probabilistic description of the system state.

A nice property of POMDPs is that, under certain conditions, they can be reformulated as a continuous-space MDPs, making possible to use MDP-solving methods to solve POMDPs as well. More details on this topic can be found in [Kae98].

1.3 Motivation and objectives

Having described the practical applications that will be addressed throughout this thesis, and the main mathematical tools that will be deployed to solve the proposed problems, we discuss here the motivation, problem-solving methodology and specific objectives of the thesis.

1.3.1 Motivation

To guarantee that dynamic networks attain (near) optimal performance while guaranteeing safe and satisfactory operation, we are interested in developing:

- Stochastic schemes that allow to modify their operation as a function of the time-varying state variables, and entail a low computational burden;
- Sensing strategies to extract relevant information from the network environment;
- Adaptive signal processing algorithms that allow to estimate unknown information and predict future state variables.

To achieve these goals, this thesis proposes the systematic and rigorous utilization of optimization and signal processing tools for network design. We consider (and will demonstrate throughout the thesis) that the rigorous mathematical formulation and analysis of the engineering problems, and the accurate incorporation of stochasticity in the models and algorithms are sources of timely and insightful research contributions.

1.3.2 Methodology

Proper formulation of mathematical problems will be critical to obtain optimal solutions and develop efficient algorithms that allow to properly integrate the distinct nature of variables and constraints under study. The three basic elements in formulating an optimization problem are the design variables, the constraints that the solution must satisfy, and the objective function (metric to be optimized). Let us further describe these elements in the context of the thesis:

1. Optimization (design) variables correspond essentially to:
   (a) resources that the network operator can modify, such as transmit power and rates, user access, user-flow allocation (scheduling), routing, power injections at power grid nodes, or available renewable energy;
   (b) sensing and control-information interchange between nodes of different nature such as user equipment, base stations, network controllers or power buses;
   (c) undesired network events, such as interferences, overvoltages, and load shedding, which can be tolerated if their probability is limited;
   (d) auxiliary parameters introduced to facilitate/simplify the previously described tasks.

   The behavior in time of these variables is potentially diverse, because they can be deterministic or random; global or local (at a node or neighborhood level); static (constant or representing averages/ probabilities/statistic moments of time-varying parameters) or dynamic (being able to vary in different timescales).

2. Regarding constraints, these can be classified into two main groups:
   (a) Quality-of-service (QoS) constraints, which will be used to guarantee different performance metrics that users and applications require (demanding certain average transmit rate, limiting the bit error rate (BER) or average delay, or allowing a maximum interfering power).

(b) System constraints related to the operating conditions that regulate the system function that can be imposed either by the physical system properties, or by the applicable standards. These will guarantee that the obtained solution can indeed be implemented in the system under study. For instances, several users cannot simultaneously access a channel, transmit power cannot exceed the established limits, power flows across power lines must not exceed thermal limits, and PV systems cannot deliver energy without sunlight.

3. The objective function must be carefully designed because there are different measures of quality that depend on the system under study (in wiereless sensor networks the power consumption is a reasonable objective, but a function of the transmit rates is better suited for a mobile data network). Formulations that use a weighted sum of several merit figures (known as aggregate utility maximization) has been successfully used in communications as well as operations research, transportation engineering, or econometrics [Lin06b, Rib10b, Mar11a].

All the previous elements depend on state variables that vary across time and space (nodes and links across the network). Such state variables are oftentimes not perfectly known, hence must be modeled as stochastic quantities. Stochasticity can be incorporated into the problem formulation in several different ways. When the costs of RA decisions are affected by stochasticity, common practice is to include the expectation of a random cost function in the objective. Similarly, when varying conditions affect variables which are involved in constraints, the constraints can also be rewritten using the expectation operators, giving rise to a so-called expected or long-term constraint. Alternatively, one can formulate probabilistic constraints guaranteeing that the original constraints are satisfied with a certain probability [cf. Chapters 3 and 5]. A more complex model of stochasticity appears in state-space models where state variables depend on previous states, decisions, and random variables, calling for SDP. The formulation of such models is based on the definition of random function ssuch as the transition, observation and reward function [cf. Chapter 4].

This methodology will be useful for the modeling of time-varying parameters, interactions between the resources to be allocated, and the acquisition and interchange of control information. Such a systematic approach allows to focus on exploiting the problem structures with the aim of reducing complexity and providing robustness against imperfect, non-stationary or incomplete information.

1.3.3 Objectives

The previous points reflect the methodology and general targets of the thesis. In the following paragraphs, some of the specific goals are briefly described:

1. Design of optimal schemes for RA in mobile communication networks with diverse QoS requirements and operation conditions, with emphasis in providing guarantees of maximum interference (deterministic and probabilistic) to other networks and users.

2. Design of convergent stochastic algorithms that adapt existing resources in the network as a function of the available state information.

3. Design of adaptive processing schemes that will allow to estimate/predict state variables that are not precisely known. These algorithms will leverage of the statistical dependencies between the different state variables (e.g. time correlation). Such schemes must be used in an integrated way using the utility-cost criteria that govern the overall network design.

4. Design of algorithms to manage heterogeneous state variables of different nature and diverse levels of uncertainty. Schemes will be designed having into account that the information

about the channel is statistical or deterministic, perfect or imperfect, stationary or non-stationary.

5. Incorporation in the RA problem formulation of the information acquisition (sensing) and optimization of the latter process.

6. Promotion of network robustness against unexpected events such as sensing errors, imprecise state information, power surges or voltage sags.

7. Characterization of the relationships that can be established between dual variables and different parameters of a network, such as queue lengths, marginal cost of resources, or the sensitivity of the optimal cost with respect to QoS parameters.

8. Development of schemes to allocate resources under probabilistic constraints [cf. Chapters 3 and 5] paying special attention to obtaining tractable solutions.

9. Design of algorithms that allow to optimize the interaction between resources that have to be allocated in different timescales.

To overcome these challenges a number of tools will be used, including: convex optimization, stochastic approximation, dual methods for convex and non-convex problems, dual decomposition, SDP, adaptive (Bayesian) estimation, and queuing theory. In certain cases it will be enough to apply existing tools; in other cases we will need modifying or tailoring them to the problem at hand.

1.4 Outline of the dissertation and contributions

This thesis comprises four technical chapters, three of them dealing with the optimization of wireless communication and cognitive networks and the last one addressing a problem in the context of power networks. Since each chapter deals with a different system setup with particular challenges, the optimization tools used in each of the chapters are different and exhibit an increasing level of elaboration.

Chapter 2 is devoted to cross-layer RA in cellular networks based on channel- and queue-state information. An optimization problem with long-term average power and rate constraints is formulated, and solved using a stochastic approximation dual descent algorithm. Links will be established between the optimal Lagrange multipliers and the *window-averaged* length of the queues. Capitalizing on those links, queue stability and average queuing delay of the developed algorithms are characterized

Chapter 3 addresses the RA in CRs with heterogeneous and imperfect knowledge of the channel state information, under probability-of-interference constraints. A stochastic dual descent algorithm (similar to that used in the previous chapter) is used, not only to solve the RA, which is formulated as convex, but also to deal with the non-convex probability-of-interference constraints.

Chapter 4 further elaborates on the model presented in Chapter 3 and deals with the design of jointly optimal sensing and RA schemes. This implies the combination of opportunistic spectrum access and a sequential bayesian estimator. When the time correlation in the NSI is incorporated in the joint design problem, it gives rise to a SDP problem, which is split into two levels, one dealing with the RA (resulting in a convex problem), and the other one being a (much more tractable) dynamic program. Finally, a stochastic approximation-based scheme is proposed to make the solution adaptive to channel non-stationarities.

Chapter 5 is dedicated to the RA in electric power networks. This goes beyond mere application of the previously exposed techniques into a power dispatch problem, and includes a novel two-stage stochastic program formulation and its application into the dispatch of a smart *distribution* power

grid in two timescales, including constraints on the long-term average voltages and over-/under-voltage probabilities. The dispatch is formulated as a stochastic saddle point problem and solved using a primal-dual stochastic algorithm.

In Chapter 6, general conclusions of the thesis and lines of future work wrap up the document.

1.4.1 Contributions

The research developed during this PhD has been published in several journals and conferences, which are listed below.

Journal articles

- López-Ramos, L.M., Marques, A.G., and Ramos, F.J., Jointly Optimal Sensing and Resource Allocation for Multiuser Interweave Cognitive Radios, *IEEE Transactions on Wireless Communications*, vol. 13,11, pp. 5954–5967, 2014 (ref. [LR14b]).
- Marques, A.G., López-Ramos, L.M., Giannakis, G.B., and Ramos, F.J., Resource allocation for interweave and underlay CRs under probability-of-interference constraints, *IEEE Journal on Selected Areas in Communications*, vol. 30, 10, pp. 1922–1933, 2012 (ref. [Mar12a]).
- Marques, A.G., Lopez-Ramos, L.M., Giannakis, G.B., Ramos, F.J., and Caamano, Antonio J., Optimal cross-layer resource allocation in cellular networks using channel-and queue-state information, *IEEE Transactions on Vehicular Technology*, vol. 61,6, pp. 2789–2807, 2012 (ref. [Mar12b]).
- Lopez-Ramos, L.M., Kekatos, V., Marques, A.G., and Giannakis, G.B., Two-timescale stochastic dispatch for power distribution networks, *IEEE Transactions on Smart Grid*, 2016 (submitted, ref. [LR16]).

International conferences

- Lopez-Ramos, L.M., Kekatos, V., Marques, A.G., and Giannakis, G.B., Microgrid Dispatch and Price of Reliability using Stochastic Approximation, *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1131–1135, 2015 (ref. [LR15]).
- López-Ramos, L.M., Marques, A.G., and Ramos, F.J., Joint sensing and resource allocation for underlay cognitive radios, *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7283–7287, 2014 (ref. [LR14a]).
- López-Ramos, L.M., Marques, A.G., and Ramos, F.J., Soft-decision sequential sensing for optimization of interweave Cognitive Radio networks, *2013 IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 235–239, 2013 (ref. [LR13]).
- López-Ramos, L.M., Marques, A.G., Ramos, F.J., and Caamano, Antonio J., Cross-layer resource allocation for downlink access using instantaneous fading and queue length information, *2010 IEEE Globecom Workshops,* pp. 1212–1216, 2010 (ref. [LR10]).
- Marques, A.G., Lopez-Ramos, L.M., Giannakis, G.B., and Ramos, F.J., Adaptive underlay cognitive radios with imperfect CSI and probabilistic interference constraints, *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2011 4th IEEE International Workshop on,* pp. 161–164, 2011 (ref. [Mar11d]).

- MARQUES, A.G., GIANNAKIS, G.B., LOPEZ-RAMOS, L.M., and RAMOS, F.J., Stochastic resource allocation for cognitive radio networks based on imperfect state information, *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3196–3199 2011 (ref. [Mar11b]).

Chapter 2

Optimal Cross-Layer Resource Allocation Using Channel and Queue State Information

Recent results in cross-layer RA for wireless systems have revealed that channel and queue state information are two of the most important parameters critically affecting the resultant designs. Motivated by these results, we rely here on stochastic convex optimization to develop optimal algorithms that use instantaneous fading and queue length information to allocate resources at the transport (flow-control), link and physical layers. The allocation strategies are obtained as the solution of a constrained utility maximization problem involving average performance metrics. It turns out that the optimal allocation at a given instant depends on the instantaneous channel state information and Lagrange multipliers, which are associated with the QoS requirements and the operating conditions of the system. The multipliers are estimated online using a dual stochastic subgradient approach, and are linked with the *window-averaged* length of the queues. Capitalizing on those links, queue stability and average queuing delay of the developed algorithms are characterized, and a simple mechanism is devised to effect delay priorities among users.

2.1 Introduction

Existing works in cross-layer RA can be mainly classified in one of the following two categories. In the first category, approaches rely on convex optimization and dual decomposition techniques; see e.g., [Lin06a, Lin06b, Che06, Lin05, Sol06]. In those, the solution of a properly formulated optimization problem dictates how resources are allocated, while the structure of the solution typically suggests the design of signaling protocols. In the second category, approaches build upon dynamic backpressure policies, pioneered by [Tas92]; see e.g., [Nee05, Ery06, Gia06, Sto05, Geo06]. These works rely on adaptive control tools (the so-called Lyapunov optimization), and aim to stabilize queues of the wireless network. Such queues correspond not only to actual queues where packets are buffered before transmission, but also to virtual queues whose role is to account for QoS requirements. Most existing works have focused on the impact of the cross-layer features on the higher layers, and fewer have addressed the design of cross-layer networking algorithms that take into account *instantaneous fading*; see [Lin06a, Che06, Lin05, Rib10a, Geo06, Gat10] for some of the exceptions.

In this context, this chapter aims to optimally design cross-layer RA algorithms for *fading* wireless *cellular* networks under the following operating conditions. At the transport layer, every node receives packets from higher layers. The packets of each user entail different utility levels and nodes implement simple flow control mechanisms to keep the network stable. At the link layer, users share orthogonally a set of parallel flat fading channels. At the physical layer, nodes can

adapt their power and rate loadings per channel realization. Both uplink (multiple access) and downlink (broadcast) channels will be considered. The optimal cross-layer RA (flows, powers, rates, and channels) is obtained as the solution of a constrained optimization problem, which naturally takes into account flow-specific utility functions, individual QoS requirements, and the network's operating conditions. The resultant optimum dynamic RA is found in closed form, and it is shown to depend only on the current channel realization, and user-specific prices (Lagrange multipliers). An adaptive dual stochastic algorithm is developed to estimate the optimal Lagrange multiplier values. The developed algorithm is robust to non-stationarities, does not require knowledge of the channel distribution, and incurs arbitrarily minimal loss of performance (optimality). Furthermore, it is shown that under very mild conditions, stochastic estimates of some of the Lagrange multipliers correspond to scaled versions of *windowed averages* of queue lengths. In other words, the developed schemes reveal one way in which queue state information (QSI) metrics related to the queue length can be used to optimize network performance. Building on the relationship between queues and multipliers, both stability and average queueing delay of the novel cross-layer RA algorithms are characterized.

The main differences relative to the state of the art and the most relevant contributions of the work reflected in this chapter are listed next. a) The specific cross-layer RA algorithms designed in this work were not considered before. b) The convergence results are stronger than those available is the QSI literature. The convergence results are similar to those in [Rib10a], which neither deals with queues nor considers windowed averages. c) Approximations for the average queuing delay are provided and a simple mechanism to effect delay priorities is presented, which is attractive from a practical perspective. d) The QSI corresponds to windowed averages of the instantaneous queue lengths. e) The channel state information (CSI)/QSI adaptive schemes are designed relying purely on convex optimization and dual stochastic algorithms, without including *explicitly* the queue lengths in the formulation. Contributions d) and e) deserve further elaboration. From a practical perspective, d) is interesting because QSI variability will be smaller and its time correlation stronger, so that, for example, more aggressive coding schemes can be employed for QSI feedback. Indeed, several actual congestion control and queue management protocols rely on window-averaged queue lengths [Low03]. From an analytical perspective, d) entails that the convergence (feasibility and optimality) proofs in the chapter have to account for QSI updates being biased and outdated versions of the actual instantaneous queue lengths. As a result, the proofs can be easily adapted to deal with additional sources of QSI imperfections which are relevant from a practical perspective (e.g., noise or delay in the signaling channels). The approach in e) opens the door to apply several known results to the problem at hand. For example, sensitivity analysis can be used to quantify trade-offs between the considered constraints and the average queuing delay. The approach also allows for different versions of the dual stochastic updates, each potentially giving rise to different forms (metrics) of QSI. See [Mar09, Gao09, Rib10a] for works that do not deal with QSI but use different stochastic updates to allocate resources in wireless networks.

The rest of the manuscript is organized as follows. The system setup and operating conditions are described in Section II. The constrained optimization problem is formulated in Section III. Its solution is presented in Section IV, where a stochastic method to estimate the optimum Lagrange multipliers is also developed, along with the connection between Lagrange multipliers and queue lengths. Queue stability and the average delay are characterized in Section V, where a method to effect delay priorities among users is also presented. Sections III-V deal with the uplink, while Section VI briefly presents their counterparts for the downlink. The signaling overhead required to implement the developed schemes is discussed in Section VII. Numerical results and conclusions in Sections VIII and IX wrap-up this manuscript.[1]

[1]Specific notation for Chapter 2: $W(x[n], L) := \frac{1}{L}\sum_{l=n-L+1}^{n} x[l]$ denotes the windowed average of length L of

Table 2.1: Summary of most significant notation in Chapter 2.

Symbol	Meaning
n	Time slot index
M , m	Number of users / User index
K , k	Number of channels / Channel index
$\mathbf{h}$, h_m^k	Channel gain vector / channel gain for user m on channel k
w_m^k	Scheduling variable (fraction of time user m occupies channel k)
$p_m^k(\mathbf{h})$	Power transmitted by user m over channel k
$r_m^k(\mathbf{h})$	Rate transmitted by user m over channel k
$\check{p}^k$	Maximum instantaneous power level in channel k
$\check{p}_m$	Maximum average power consumed by user m (Uplink)
$\check{p}$	Maximum average power consumed by the BS (Downlink)
$C_m^k(\mathbf{h}, p_m^k(\mathbf{h}))$	Capacity (rate-power) function for user m over channel k
a_m	Instantaneous arrival rate of user m
$\bar{a}_m$	Average arrival rate of user m
q_m	Length of queue m
$U_m(\bar{a}_m)$	Utility function of user m's service rate
P*	Sum-utility achieved by the optimal resource allocation [cf. (3.6)]
π_m	Lagrange multiplier for user m's average power constraint
ρ_m	Lagrange multiplier for user m's flow conservation constraint
α_m	Lagrange multiplier for user m's average arrival rate constraint
$\boldsymbol{\lambda}$	Vector gathering all Lagrange multipliers
$\pi_m^*, \rho_m^*, \alpha_m^*, \boldsymbol{\lambda}^*$	(Optimal) Value of the *non-stochastic* Lagrange multipliers
$\boldsymbol{\lambda}_L[n]$	Stochastic multiplier estimates using an averaging window of length L
μ, μ_m	Stepsize / User-selective stepsize (for delay priorities)
$f(m, k, h, \boldsymbol{\lambda})$	Link Quality Indicator of channel k for user m

2.2 Modeling preliminaries

In this section, the system setup and channel model are introduced first. Then, operation of the link and physical layers is presented. Finally, the flow-control mechanism and queue dynamics are described. To facilitate the readability of the manuscript, the most relevant notation introduced in this and the following sections is summarized in Table 2.1.

Consider M wireless terminals (users) connected to an access point (AP). The overall bandwidth B is divided into K orthogonal channels, each with bandwidth B/K small enough to ensure that the fading per channel is flat, i.e., non-selective. The wireless link between the AP and user m on channel k is characterized by its random square magnitude h_m^k, which is assumed normalized by the receiver noise variance. The overall $MK \times 1$ gain vector is denoted by $\mathbf{h} := \{h_m^k,\ m = 1, \ldots, M,\ k = 1, \ldots, K\}$. The channel fading process is assumed ergodic, stationary and statistically independent across users and time. Moreover, it is assumed that $\mathbf{h}$ remains invariant over a block of symbols, but can vary from block-to-block (block fading channel model). In other words, if n denotes the current block index (whose duration is dictated by the

variable x at time $n \geq L$.

channel coherence interval), then $\mathbf{h}[n]$ remains constant and $\mathbf{h}[n] \neq \mathbf{h}[n+1]$ with probability one (w.p.1). Notation $\mathbf{h}[n]$ will be used whenever the block time-variance of the channel needs to be emphasized.

Regarding the link layer operation, links at the outset are scheduled to access simultaneously but orthogonally in time (or frequency) any of the channels; see e.g., [Won99, Mar11c]. To describe access quantitatively, let $w_m^k(\mathbf{h}) \in [0,1]$ denote the nonnegative fraction of time (or channel bandwidth) that user m is scheduled to transmit over channel k during the channel realization $\mathbf{h}$. Since individual user transmissions interfere each other, it must hold that

$$\sum_{m=1}^{M} w_m^k(\mathbf{h}) \leq 1, \quad \forall k, \ \forall \mathbf{h}. \tag{2.1}$$

This way, if $w_m^k(\mathbf{h}) = 0.9$ and $w_{m'}^k(\mathbf{h}) = 0.1$, the traffic of user m is assigned to channel k during 90% of the duration of realization $\mathbf{h}$, to user m' during the remaining 10%, and traffic of no other user is scheduled. Clearly, if one restricts $w_m^k(\mathbf{h}) \in \{0,1\}$, then $w_m^k(\mathbf{h})$ can be readily viewed as an indicator variable; meaning that if channel k is assigned to user m, then it holds that $w_m^k(\mathbf{h}) = 1$ while $w_{m'}^k(\mathbf{h}) = 0$ for all $m' \neq m$. No restriction is placed on the maximum number of channels a user can access.

The resources adapted at the physical layer will be power and rate per user and channel. Specifically, let $p_m^k(\mathbf{h})$ and $r_m^k(\mathbf{h})$ denote, respectively, the instantaneous power and rate user m transmits over channel k during the channel realization $\mathbf{h}$, if $w_m^k(\mathbf{h}) = 1$. Two types of power constraints are considered. Spectrum mask constraints are imposed to ensure that the *instantaneous* $p_m^k(\mathbf{h})$ does not exceed a maximum prespecified level $\check{p}^k$; that is,

$$p_m^k(\mathbf{h}) \leq \check{p}^k, \quad \forall k, \ \forall m, \ \forall \mathbf{h}. \tag{2.2}$$

On the other hand, the maximum *average* power terminal m can transmit is also bounded by $\check{p}_m$; hence,

$$\mathbb{E}\Big[\sum_{k=1}^{K} p_m^k(\mathbf{h}) w_m^k(\mathbf{h})\Big] \leq \check{p}_m, \ \forall m \tag{2.3}$$

where the expectation is taken over the fading channel distribution. Since an uplink setup is considered here, users act as transmitters, and the power of *each individual user* has to be bounded [cf. (2.3)]. In the downlink, the AP acts as a transmitter and thus only the AP power must be constrained (see Section 2.6 for details).

Under bit error rate or capacity constraints, power $p_m^k(\mathbf{h})$ is coupled with the corresponding rate $r_m^k(\mathbf{h})$. This rate-power coupling will be represented by the function $C_m^k(\mathbf{h}, p_m^k(\mathbf{h}))$. It is assumed throughout that the rate-power function $C_m^k(\mathbf{h}, p_m^k(\mathbf{h}))$ is increasing and strictly concave. For example, if sufficiently strong coding schemes are used, $C_m^k(\mathbf{h}, p_m^k(\mathbf{h}))$ approaches Shannon's capacity formula $\log(1 + h_m^k p_m^k(\mathbf{h}))$, which is increasing and strictly concave. The one-to-one mapping between $p_m^k(\mathbf{h})$ and $r_m^k(\mathbf{h})$ implies that when the optimization problem is formulated, it suffices to optimize over one of them.

Regarding the transport and network layers, the operation is as follows. Packets are generated exogenously at higher layers. Packet streams will be referred to as flows, and there will be as many flows as users, each user having one flow. The packet arrival rate of flow m at a given instant n is a random variable denoted by $a_m[n]$. The average arrival rate of exogenous information of flow m is denoted by $\bar{a}_m$. Terminals are equipped with queues (buffers) capable of storing the incoming packets. Let $q_m[n]$ denote the queue size corresponding to flow m at time slot n. Then,

the queue obeys the recursion

$$q_m[n+1] = \Big[q_m[n] + a_m[n] - \sum_{k=1}^{K} r_m^k(\mathbf{h}[n]) w_m^k(\mathbf{h}[n])\Big]_0^{\infty}, \ \forall m. \tag{2.4}$$

In practice, packet arrivals and departures vary at a time scale smaller than n. This implies that definitions slightly different from the one in (2.4) are also possible [Geo06]. Such differences, however, are inconsequential for the subsequent analysis, and (2.4) has been chosen for mathematical simplicity.

Different definitions of queue stability are available. In this work, queues are deemed stable if $\lim_{n\to\infty} \frac{1}{n} \sum_{l=1}^{n} \mathbb{E}[q_m[l]] < \infty$. This definition is referred to as "strong stability" in [Geo06]. For queues to be stable, the following necessary condition needs to be satisfied

$$\bar{a}_m \leq \sum_{k=1}^{K} \mathbb{E}\left[r_m^k(\mathbf{h}) w_m^k(\mathbf{h})\right], \ \forall m. \tag{2.5}$$

The latter is typically known as necessary average flow conservation condition. Equation (2.5), together with (2.1), (2.2), and (2.3) are accounted for in the optimization problem presented in the next section.

2.3 Problem formulation and Design approach

The optimal RA will be viewed in this section as the solution of a constrained optimization problem. The *objective* will entail concave and increasing so-called *utility functions* $U_m(\cdot)$, that are commonly used in RA tasks (not only restricted to communication systems), and account for the "social" utility (reward) that a specific resource (here $\bar{a}_m$) gives rise to. Sum-utility maximization has been frequently employed by scheduling, MAC layer, and networking algorithms; see e.g., [Liu06, Jia05, Geo06], and references therein. Utility functions $U_m(\bar{a}_m)$ are chosen to be increasing (so that solutions which allow for higher arrival rates are promoted), and concave (so that the marginal utility for each user terminal diminishes as its rate increases) which offers a means to effect fairness among different users. Typical utility functions $U_m(x)$ include $\alpha_m \log(\beta_m + x)$ and $\alpha_m(1-\beta_m)^{-1}x^{1-\beta_m}$, where α_m, β_m are user-dependent positive constants. On the other hand, to effect QoS, a *minimum* average arrival rate $\check{a}_m$ is guaranteed for certain users.

The optimal allocation is obtained by solving the following constrained sum-utility maximization:

$$\mathsf{P}^* := \max_{\substack{\bar{a}_m \\ w_m^k(\mathbf{h}), p_m^k(\mathbf{h})}} \ \sum_{m=1}^{M} U_m(\bar{a}_m) \tag{2.6a}$$

$$\text{subject to}: \ \bar{a}_m \geq \check{a}_m, \ m = 1, ..., M \tag{2.6b}$$

$$(2.1), \ (2.2), \ (2.3), \ (2.5). \tag{2.6c}$$

Recall that there is no need to optimize over $r_m^k(\mathbf{h})$, because $r_m^k(\mathbf{h})$ in (2.5) can be replaced with $C_m^k(\mathbf{h}, p_m^k(\mathbf{h}))$; thus, the optimum value of $p_m^k(\mathbf{h})$ will readily yield the optimum value of $r_m^k(\mathbf{h})$. Moreover, for 'best effort' flows the corresponding $\check{a}_m$ in (2.6b) is set to zero. Both the cross-layer and the channel-adaptive attributes of the RA problem are apparent since variables of different layers are jointly optimized in (3.6), and several optimization variables as well as constraints are functions of $\mathbf{h}$. Finally, note that in (4.8a) we have defined P^* as the sum-rate utility value achieved by the optimum solution. This notation will be useful in the upcoming sections to assess the potential loss of performance of the algorithms to be developed.

The solution of (3.6) will be pursued in the next section, but several remarks are due before

that. First, if the minimum rate requirements in (2.6b) are too high, and the power budgets in (2.3) too small, the optimization in (3.6) could be infeasible. This however is readily detectable, since the Lagrange multipliers associated with some of the infeasible users would grow unbounded. In such a case, the only option to stabilize the system is to simply drop users. Unfortunately, optimally selecting which users to drop (also known as admission control) often boils down to an NP-hard problem, and goes beyond the scope of this work. Second, the average rate constraints in (2.6b) guarantee that the *average* arrival rate of a flow remains above a given requirement (elastic traffic). However, there is no guarantee on the *instantaneous* transmit-rate of real-time traffic. Third and last, it is important to stress that queue dynamics were not explicitly taken into account in (3.6). Only the necessary condition for stability in (2.5) has been explicitly accounted for. However, it will be shown in the next section that queues can in fact be used as stochastic estimates of Lagrange multiplier values associated with constraint (2.5). This result implies that the solution of (3.6) will implicitly depend on QSI too.

2.4 Optimal cross-layer allocation

The problem in (3.6) can be readily transformed to a convex one (see [Won99] or [Mar11c] for details), which can be solved using the dual approach. The optimal solution will be presented first as a function of the optimum Lagrange multipliers (dual variables). A stochastic scheme will be developed next for estimating the multipliers per time instant n. The last part of this section will be devoted to the relationship between stochastic multiplier estimates and queue lengths.

2.4.1 Optimal allocation as a function of the multipliers

Let π_m, ρ_m, and α_m denote the Lagrange multipliers associated with the *average* constraints in (2.3), (2.5), and (2.6b), respectively. Collect all these multipliers in a vector $\boldsymbol{\lambda}$. There is no need for dualizing the *instantaneous* constraints (2.1) and (2.2) because the solution will turn out to satisfy them automatically. Furthermore, let $(\dot{U}_m)^{-1}(\cdot)$ and $(\dot{C}_m^k)^{-1}(\mathbf{h},\cdot)$ denote the inverse function of the derivative of $U_m(\cdot)$ and $C_m^k(\mathbf{h},\cdot)$, respectively; and remember that x^* stands for the optimum value of a given variable x. Using these notational conventions and the Karush-Kuhn-Tucker (KKT) conditions [Ber99] associated with (3.6), it is shown in Appendix A that the optimal cross-layer RA is

$$\bar{a}_m^*(\boldsymbol{\lambda}^*) = \left[(\dot{U}_m)^{-1}(\rho_m^* - \alpha_m^*)\right]_0^{\infty} \tag{2.7}$$

$$p_m^{k*}(\mathbf{h},\boldsymbol{\lambda}^*) = \left[(\dot{C}_m^k)^{-1}(\mathbf{h}, \pi_m^*/\rho_m^*)\right]_0^{\check{p}^k} \tag{2.8}$$

$$r_m^{k*}(\mathbf{h},\boldsymbol{\lambda}^*) = C_m^k(\mathbf{h}, p_m^{k*}(\mathbf{h},\boldsymbol{\lambda}^*)). \tag{2.9}$$

For notational convenience, define $f(m,k,\mathbf{h},\boldsymbol{\lambda}^*) := \rho_m^* r_m^{k*}(\mathbf{h},\boldsymbol{\lambda}^*) - \pi_m^* p_m^{k*}(\mathbf{h},\boldsymbol{\lambda}^*)$. Then, Appendix A also shows that for $\mathbf{h}$ and $\boldsymbol{\lambda}^*$ given, only users m for which $f(m,k,\mathbf{h},\boldsymbol{\lambda}^*)$ is maximum can transmit over channel k; i.e., users satisfying $m_k = \arg\max_{m'}\{f(m',k,\mathbf{h},\boldsymbol{\lambda}^*)\}$ [Li01, Mar11c]. If the user attaining the maximum is unique, it follows readily that the optimum scheduling is

$$w_m^{k*}(\mathbf{h},\boldsymbol{\lambda}^*) = \mathbb{1}_{\{m=\arg\max_{m'}\{f(m',k,\mathbf{h},\boldsymbol{\lambda}^*)\}}. \tag{2.10}$$

It will be shown later that for the schemes developed here, the user attaining the maximum is always unique, and thus the optimum scheduling is always given by (2.10).

A close look at (2.7)-(2.10) reveals that the optimal resource management depends only on the current channel realization $\mathbf{h}$, and on the optimal Lagrange multipliers $\boldsymbol{\lambda}^*$. While (2.7)-(2.9) are easy to derive and interpret, the scheduling in (2.10) deserves further elaboration. Equation (2.10) asserts that *per channel realization* $\mathbf{h}$, each channel k is uniquely assigned to the user index

maximizing the functional $f(m,k,\mathbf{h},\boldsymbol{\lambda}^*)$; i.e., the scheduling follows a winner(s)-take-all strategy. Regarding ρ_m^* and π_m^* as prices for the rate and power, $f(m,k,\mathbf{h},\boldsymbol{\lambda}^*)$ determines the net utility (rate reward minus power cost) of the mth user on channel k. This means that users with large ρ_m^* (or small π_m^*) are promoted for selection. Although for a given $\mathbf{h}$ the winner is unique, the user accessing the channel will vary as $\mathbf{h}$ varies. Hence, long-term fairness is accommodated too. Last but not least, in several practical systems (like multicarrier systems) the channel gains in consecutive channel are expected to be very similar. This implies that the value of f would be similar for k and $k+1$ and therefore, it is likely that the winner user in both channels is the same. Additional details on this issue are provided in Section 2.7, where different alternatives to reduce the signaling overhead required to implement our schemes are discussed. The winner-takes-all strategy (also known as max-weight scheduling) has been shown optimal for additional setups dealing with orthogonal sharing of resources among users [Li01, Mar11c], [Lee06], [Lin06b], [Sto05]. It is also worth noting that (2.7) only dictates the optimal value of the average arrival rate, but not the specific distribution of the packet arrivals. In fact, from the point of view of (3.6) any arrival distribution with average given by (2.7) is equally optimal. Nevertheless, the distribution of packet arrivals does affect practical issues such as queue lengths, or the delay packets experience. These issues as well as further conditions on the distribution of packet arrivals will be re-visited when the stochastic algorithms are presented.

2.4.2 Estimating the multipliers

In the previous section, the optimal RA was characterized by two variables: the current channel state information $\mathbf{h}$, and the optimum Lagrange multipliers $\boldsymbol{\lambda}^*$. Various options are available to find $\boldsymbol{\lambda}^*$. Most are iterative numerical methods seeking $\boldsymbol{\lambda}^*$ off-line by capitalizing on knowledge of the channel distribution [Ber99, Ch. 4 and 6]. An online approach is pursued here under which the exact value of $\boldsymbol{\lambda}^*$ is never found. Instead, an estimate of $\boldsymbol{\lambda}^*$ is obtained using stochastic approximation iterations per time index n. This estimate, call it $\boldsymbol{\lambda}_L[n]$ (the reason for using the subscript L will be apparent soon), remains sufficiently close to $\boldsymbol{\lambda}^*$; see e.g., [Wan07]. The motivation behind this estimate is threefold: (i) computational complexity of the stochastic schemes is lower than that of their off-line counterparts, (ii) stochastic schemes do not need to know the channel distribution and can cope with channel non-stationarities; and (iii) connections between these multiplier estimates and queue lengths can be established as will be seen in the next section.

Toward estimating the multipliers, let $a_m^*[\boldsymbol{\lambda}_L[n]]$ denote the instantaneous arrival of flow m during block n. The latter is a random variable drawn from a distribution with mean $\bar{a}_m^*(\boldsymbol{\lambda}_L[n])$ given by (2.7). Moreover, with $x[n]$ denoting the value of a given variable x at time n, the sliding window average of length L of variable x at time $n \geq L$ is[2] $W(x[n],L) := \frac{1}{L}\sum_{l=n-L+1}^{n} x[l]$. Based on this average and with μ denoting a *constant* stepsize, the Lagrange multipliers are updated as follows:

$$\pi_{L,m}[n+1] := \Big[\pi_{L,m}[n] - \mu\Big(\check{p}_m \;\; W\Big(\textstyle\sum_k p_m^{k*}(\mathbf{h}[n],\boldsymbol{\lambda}_L[n]) w_m^{k*}(\mathbf{h}[n],\boldsymbol{\lambda}_L[n]),L\Big)\Big)\Big]_0^{\infty} \quad (2.11)$$

$$\rho_{L,m}[n+1] := \Big[\rho_{L,m}[n] + \mu\Big(W\Big(a_m^*[\boldsymbol{\lambda}_L[n]] - \textstyle\sum_k r_m^{k*}(\mathbf{h}[n],\boldsymbol{\lambda}_L[n]) w_m^{k*}(\mathbf{h}[n],\boldsymbol{\lambda}_L[n]),L\Big)\Big)\Big]_0^{\infty} \quad (2.12)$$

$$\alpha_{L,m}[n+1] := \Big[\alpha_{L,m}[n] + \mu\Big(\check{a}_m - W(a_m^*[\boldsymbol{\lambda}_L[n]],L)\Big)\Big]_0^{\infty} \quad (2.13)$$

[2]For $n < L$, the sliding window average is defined as $W(x[n],L) := \frac{1}{n}\sum_{l=1}^{n} x[l]$. Although the definition for $n < L$ is needed from completeness, it will not be critical for the performance analysis. This holds because we will be interested in the long-term performance of our allocation schemes (more specifically, in the cumulative running average of the variables as $n \to \infty$). Clearly, if $x[n]$ is bounded, the behavior during the first $L-1$ time slots (which is an interval of finite length) is not relevant for that purpose.

where the primal variables in (3.20)-(2.13), namely, powers, rates, scheduling percentages and flow rates, are found by substituting $\pi_{L,m}[n]$, $\rho_{L,m}[n]$, and $\alpha_{L,m}[n]$ into (2.7)-(2.10). Those will be referred to as stochastic primal variables.

Having described the stochastic schemes, it is appropriate to revisit the optimality of the winner-takes-all strategy in (2.10). It was stated in Section 2.4.1 that the scheduling in (2.10) is indeed optimum because for the schemes in this chapter the user m^* maximizing $f(m,k,\mathbf{h},\boldsymbol{\lambda}^*)$ is unique. To verify that this is true for the stochastic allocation, the user $m^*[n]$ maximizing $f(m,k,\mathbf{h}[n],\boldsymbol{\lambda}_L[n])$ must be unique. Clearly, $f(m,k,\mathbf{h}[n],\boldsymbol{\lambda}_L[n])$ is a continuous random variable because all terms involved in its definition are continuous random variables. Moreover $f(m,k,\mathbf{h}[n],\boldsymbol{\lambda}_L[n])$ is uncorrelated across users, because in practice both arrivals and channel gains are independent across users. As a result, the event of $f(m,k,\mathbf{h}[n],\boldsymbol{\lambda}_L[n]) = f(m',k,\mathbf{h}[n],\boldsymbol{\lambda}_L[n])$ has probability (Lebesgue measure) zero. Since for the winner not to be unique two users must attain the same value of f, the previous finding readily implies that the winner is unique w.p.1.

Convergence, feasibility and optimality of the stochastic schemes will rely on the fact that (3.20) and (2.13) are finite-length averages of stochastic subgradients of the dual function in (3.6), see e.g. [Ber99, Ch. 6]. Employing unbiased stochastic subgradients of the dual function for allocating resources in wireless fading networks has received some attention in recent years [Wan07, Gao09, Mar09, Mar10, Rib10a]. The focus and novelty here are on algorithms that use constant stepsize and combine (average) the L most recent stochastic subgradients. These two features will allow us to establish connections between some of the iterations in (3.20), and the (windowed average) length of the system queues. To be specific, assuming that the updates in (3.20)-(2.13) are bounded, the following result guarantees feasibility and near optimality of the stochastic allocation iterates.

Proposition 1 *The sample average of the stochastic RA: (i) is feasible and (ii) incurs arbitrarily small loss in performance relative to the average non-stochastic solution of the problem in* (3.6)*. Specifically, as $n\to\infty$, it holds w.p.1 that*

$$(i)\ \frac{1}{n}\sum_{l=1}^{n}\left(\sum_{m=1}^{M} r_m^*(\mathbf{h}[l],\boldsymbol{\lambda}_L[l])w_m^{k*}(\mathbf{h}[l],\boldsymbol{\lambda}_L[l])\right) \geq \frac{1}{n}\sum_{l=1}^{n} a_m^*[\boldsymbol{\lambda}_L[l]] \geq \check{a}_m$$
$$\frac{1}{n}\sum_{l=1}^{n}\left(\sum_{k=1}^{K} p_m^{k*}(\mathbf{h}[l],\boldsymbol{\lambda}_L[l])w_m^{k*}(\mathbf{h}[l],\boldsymbol{\lambda}_L[l])\right) \leq \check{p}_m; \tag{2.14a}$$

(ii) *And with $\delta_{\mathrm{P}}(\mu)$ denoting a positive number proportional to the stepsize μ, it holds that [cf.* (3.6)*]*

$$\sum_m U_m\left(\frac{1}{n}\sum_{l=1}^{n} a_m^*(\boldsymbol{\lambda}_L[l])\right) \geq \mathrm{P}^* - \delta_{\mathrm{P}}(\mu). \tag{2.15}$$

In words, Proposition 1 guarantees asymptotic optimality of the stochastic iterates because they are feasible and achieve a value (performance) arbitrarily close to P^*, which is the optimal value of the original (non-stochastic) solution of (3.6). The proof of the proposition is given in Appendix B.1 and relies on the convergence of stochastic (epsilon) subgradient methods. For simplicity, the proofs are given for convergence in probability, but the extension to convergence w.p.1 is straightforward using the arguments in [Boy06, Rib10a], [Ber03, Ch. 8]. Moreover, the proof also shows that under very mild assumptions, the first inequality in (2.14) and the inequality in (2.14a) hold with equality.

Equally relevant, convergence results can also be obtained when the cumulative running average in Proposition 1 is replaced by a finite-size sliding window averaging or with an exponentially-decaying window averaging. For those cases, convergence in distribution of the modified left hand

side of (2.15) to a Gaussian whose mean is larger than the right hand side of (2.15) can be shown. A rigorous proof can be obtained following the methodology in [Kus03, Chapter 11].

2.4.3 Relating multipliers with queue lengths

On top of being optimal, the stochastic schemes are also meaningful because they reveal the cross-layer attribute of our algorithm. To see how, set temporarily $L = 1$ and compare (2.12) to (2.4). It is clear that $\rho_m[n]$ and $q_m[n]$ are related in a way such that *the stochastic Lagrange multipliers can be interpreted as a scaled version of the queue sizes*. A very similar result is also true for $L > 1$. For this case, the following holds.

Proposition 2 *If there exists n_0 such that $\rho_{L,m}[n] > 0$ and $q_m[n] > 0$ for $n \geq n_0$ then $\rho_{L,m}[n] = \mu W(q_m[n], L) + \delta_m(n_0)$, where $\delta_m(n_0) := \rho_{L,m}[n_0] - \mu q_m[n_0]$.*

Proof: The proof for $L = 1$ is straightforward because the only difference between (2.12) and (2.4) is the stepsize μ. To prove the equivalence for $L > 1$, we will view $\rho_{L,m}[n]$ as the output $y[n]$ of a discrete linear time-invariant filter with input $x[n] = a_m^*[n] - \sum_k r_m^{k*}(\mathbf{h}[n]) w_m^{k*}(\mathbf{h}[n])$, where dependence of variables on $\boldsymbol{\lambda}_L[n]$ has been dropped for notational brevity. Upon examining (2.12) and with $\star$ denoting convolution, $y[n]$ can be written as

$$y[n] = h_{IIR}[n] \star (h_{FIR}[n] \star (\mu x[n])). \tag{2.16}$$

In the latter, $h_{FIR}[n]$ implements the window averaging of length L, while $h_{IIR}[n]$ implements an autoregressive filter of order one; i.e., if $y_{IIR}[n]$ and $x_{IIR}[n]$ denote, correspondingly, the output and the input of filter $h_{IIR}[n]$, then $y_{IIR}[n] = y_{IIR}[n-1] + x_{IIR}[n]$. Using properties of the convolution operator, (2.16) can be rewritten as

$$y[n] = \mu(h_{FIR}[n] \star (h_{IIR}[n] \star x[n])). \tag{2.17}$$

Relying on the expression in (2.4) and given that $x[n] = a_m^*[n] - \sum_k r_m^{k*}(\mathbf{h}[n]) w_m^{k*}(\mathbf{h}[n])$, it readily follows that $q_m[n] = (h_{IIR}[n] \star x[n])$. For this to be true, the projection operator in (2.4) needs to be transparent, which is the reason for requiring the queues not to be empty[3]. Substituting the former into (2.17) yields $y[n] = \mu(h_{FIR}[n] \star q_m[n])$. Since $h_{FIR}[n]$ implements the window averaging and $y[n] = \rho_{L,m}[n]$, it follows that $\rho_{L,m}[n] = \mu W(q_m[n], L)$, which is the claim of the proposition.

As mentioned earlier, the result of Proposition 2 is meaningful because it shows one way in which queues can be used to allocate resources in the network. In fact, Propositions 1 and 2 establish that if ρ_m^* in (2.7)-(2.10) is approximated by $\mu W(q_m[n], L)$, the resultant stochastic RA is optimum so long as μ is sufficiently small [cf. (2.15)]. Although the term $\delta_m(n_0)$ introduces a discrepancy between $\rho_{L,m}[n]$ and $\mu W(q_m[n], L)$, this is not a concern because: (i) in practice $\delta_m(n_0) \approx 0$; and (ii) from an implementation perspective, the value of $\rho_{L,m}[n]$ at instant n_0 can always be redefined as $\rho_{L,m}[n_0] := \rho_{L,m}[n_0] - \delta_m(n_0)$ without affecting the long-term performance of the algorithm. Similarly, the presence of n_0 is not critical from a practical perspective either; and has been assumed to simplify the proof. Nevertheless exhaustive simulations have shown that a nonzero n_0 is present for small-medium values of L and μ. As an alternative, the iterations in (3.20)-(2.13) can be slightly modified so that the correspondence $\rho_{L,m}[n] = \mu W(q_m[n], L)$ holds $\forall n$. Although the proof is a bit more tedious, it can be shown that Proposition 1 holds in this case too.

[3] Intuitively, this assumption is reasonable because the optimization in (3.6) aims to obtain as high arrival rates as possible, which require the system to work close to its saturation point.

In addition to revealing how QSI can be optimally used for allocating resources, Proposition 2 is also relevant for extra reasons too: (i) to analyze stability of RA algorithms; (ii) to estimate the queueing delay that packets will suffer from; and (iii) to establish connections with other well-known cross-layer RA algorithms such as the dynamic backpressure algorithm generalizations reported in [Geo06]. The next section deals with i) and ii). Regarding (iii), the stochastic iterates with $L = 1$ allocate resources in a manner related to those in [Geo06, Hua11] in the following way: the role of $\rho_L[n]$ is played by the actual queues; the role of $\pi_L[n]$ and $\alpha_L[n]$ is played by the so-called "virtual" queues; and some tuning parameters in [Geo06] correspond to the inverse of the stepsize μ. Nevertheless, the results presented so far are distinct from [Geo06] not only because different channel realizations are used to update the multipliers (since $L > 1$), but also because the analytical framework (Lyapunov optimization versus stochastic convex optimization), the CSI model, as well as the convergence claims are all different. Interestingly, works as [Hua11] also established relationships between the queue lengths and the optimal value of the multipliers of a related optimization problem. More specific details on this issue will be given in the next section after presenting the results on the queue stability of our schemes.

2.5 Queue stability and delay analysis

The previous section established that although queue dynamics have not been explicitly accounted for in the formulation of (3.6), they emerge naturally as *scaled* stochastic estimates of the Lagrange multipliers associated with (2.5). In this section, queue stability of the developed RA algorithms is characterized first. Next, estimates of the average queueing delay that the flows experience are obtained, and used to outline a mechanism to effect delay priorities among users.

Stability analysis of the stochastic RA in (2.7)-(2.13) relies on Proposition 2 to establish the following result regarding convergence of queue lengths.

Proposition 3 *If* $\bar{q}_m[n] := \frac{1}{n}\sum_{l=1}^{n} q_m[l]$ *denotes the sample average of the queue length* $q_m[n]$*, and* $\delta_q(\mu)$ *is a finite number satisfying* $\delta_q(\mu) \to 0$ *as* $\mu \to 0$*, it then holds that*

$$|\mu \bar{q}_m[n] - \rho_m^*| < \delta_q(\mu) \text{ as } n \to \infty \qquad \text{w.p. } 1. \tag{2.18}$$

As a corollary, (2.18) yields $\lim_{n\to\infty} \bar{q}_m[n] < \infty$ provided that $\rho_m^* < \infty$; which implies that the iterations are stable so long as the original problem in (3.6) is feasible. The proof is given in Appendix B.2. Note that Proposition 3 focuses on the convergence of the sample-average of the queues and not on the convergence of the instantaneous multipliers. That is the case because we are interested in the long-term queue stability and not in issues like queue length (or delay) moments. As simulations will confirm, the time trajectory of the stochastic estimates consists of two phases. During the first phase, the multipliers will move from the initialization point towards the optimum value. Grossly speaking, the speed of convergence is expected to be linear because our algorithms are modified versions of a stochastic first-order (gradient) iteration. During the second phase, the estimates will hover around their optimum value. The estimates will not converge to a fixed point (because they are continuously updated based on a continuous random variable) and the "hovering noise" will be proportional to the stepsize considered.

Before moving to the next section, it has to be mentioned that for an algorithm related to the one proposed here, a relationship between the *expected* queue length and the optimum value of a Lagrange multiplier has been also established in [Hua11]. Specifically, [Hua11] relies on adaptive control tools to propose a cross-layer algorithm which relies on CSI and QSI to allocate resources. The authors analyze then the performance that such an algorithm achieves and show that the expected queue length approximately corresponds to the value of a Lagrange multiplier of a deterministic optimization problem.

2.5.1 Average delay and delay priorities

The relationship between queues and Lagrange multipliers can also be leveraged to estimate the average queueing delay of the proposed stochastic RA. To this end, Little's result asserts that with stable queues the long-term average delay is given by the long-term average queue length divided by the long-term average arrival rate, meaning that the delay of a specific flow is $\bar{d}_m = \bar{q}_m/\bar{a}_m$. Using Proposition 3, it follows that $|\mu\bar{d}_m - \rho_m^*/(\bar{a}_m)| < \delta_q(\mu)/\bar{a}_m$ w.p.1. This readily implies that the average delay for the stochastic resource algorithms developed in this thesis can be *approximated* as

$$\bar{d}_m \approx \rho_m^*/(\mu\bar{a}_m^*), \quad \forall m. \tag{2.19}$$

In words, the average delay of our stochastic algorithm can be estimated based on the optimal solution of (3.6), and the stepsize of the proposed iterations [Mar10]. The KKT conditions associated with (3.6) can be readily used to show that $\rho_m^* = \dot{U}_m(\bar{a}_m^*) + \alpha_m^*$. Substituting the latter into (2.19), it follows that the delay can be written as the sum of two nonnegative terms $\bar{d}_m \approx \dot{U}_m(\bar{a}_m^*)/(\mu\bar{a}_m^*) + \alpha_m^*/(\mu\bar{a}_m^*)$. Clearly, the first term depends only on the exogenous arrival rates $\bar{a}_m^*$. This implies that the arrival rates can be used to estimate the value of the average delay for a given flow (when the minimum rate constraint of such a flow is not active, and thus $\alpha_m^* = 0$) or a lower bound on that value (when the minimum rate constraint of the flow is active, and thus $\alpha_m^* > 0$).

Upon examining (2.19), it is also apparent that changes in the stepsize induce changes in the average delay. Specifically, the larger the stepsize, the smaller the average queueing delay. The intuition is that large stepsizes accelerate convergence and improve the ability of the iterates to react against events that otherwise would increase queuing delay. However, large stepsize values also lead to more severe hovering in the dual domain, and thus more pronounced loss of optimality. Equally interesting, (2.19) can also be used to effect different delay priorities. Key for this purpose is the fact that the iterations in (3.20)-(2.13) converge not only if the stepsize is the same for all entries of $\boldsymbol{\lambda}$, but also if it is different for each entry. This way, flows (users) with stricter delay constraints can employ a larger stepsize. In other words, by allowing the stepsize to be user-dependent it is possible to control the average delay performance of individual users.

2.6 Downlink setup

In this section, we briefly elaborate on the generalization of the developed schemes for use in a downlink setup (fading broadcast channel). The first step to address the optimal design is to reinterpret some of the notation. Specifically, in this section the entries of $\mathbf{h}$ represent the channel from the AP (transmitter) to the users (receivers); $\bar{a}_m$ is the average exogenous rate at the AP destined for user m; and $p_m^k(\mathbf{h})$ and $r_m^k(\mathbf{h})$ are the power and rate transmitted by the AP to user m over channel k. With $\check{p}$ denoting the average power budget of the AP, the second step is to replace the M individual power constraints in (2.3) with the single constraint

$$\sum_{k,m} p_m^k(\mathbf{h}) w_m^k(\mathbf{h}) \leq \check{p}. \tag{2.20}$$

Now, we are ready to formulate the new optimization problem which gives rise to the optimal

cross-layer allocation for the downlink setup:

$$\mathsf{P}^* := \max_{\substack{\bar{a}_m \\ w_m^k(\mathbf{h}), p_m^k(\mathbf{h})}} \sum_{m=1}^{M} U_m(\bar{a}_m) \tag{2.21a}$$

$$\text{subject to}: \ \bar{a}_m \geq \check{a}_m, \ \forall m \tag{2.21b}$$

$$(2.1), \ (2.2), \ (2.5), \ (2.20). \tag{2.21c}$$

Similar to (3.6), the problem in (2.21) can be transformed into a convex one and solved in the dual domain. The main difference is that instead of M multipliers $\{\pi_m\}_{m=1}^M$, now only a single multiplier π is involved. Clearly, $\{\pi_m\}_{m=1}^M$ were associated with the M constraints in (2.3) while π is associated with the single constraint in (2.20). With this change, the optimal allocation of resources for the broadcast channel is

$$\bar{a}_m^*(\boldsymbol{\lambda}^*) = \left[(\dot{U}_m)^{-1}(\rho_m^* - \alpha_m^*)\right]_0^{\infty} \tag{2.22}$$

$$p_m^{k*}(\mathbf{h}, \boldsymbol{\lambda}^*) = \left[(\dot{C}_m^k)^{-1}(\mathbf{h}, \pi^*/\rho_m^*)\right]_0^{\check{p}^k} \tag{2.23}$$

$$r_m^{k*}(\mathbf{h}, \boldsymbol{\lambda}^*) = C_m^k(\mathbf{h}, p_m^{k*}(\mathbf{h}, \boldsymbol{\lambda}^*)) \tag{2.24}$$

$$w_m^{k*}(\mathbf{h}, \boldsymbol{\lambda}^*) = \mathbb{1}_{\{m = \arg\max_{m'}\{f_{DL}(m', k, \mathbf{h}, \boldsymbol{\lambda}^*)\}} \tag{2.25}$$

where $f_{DL}(m, k, \mathbf{h}, \boldsymbol{\lambda}^*) := \rho_m^* r_m^{k*}(\mathbf{h}, \boldsymbol{\lambda}^*) - \pi^* p_m^{k*}(\mathbf{h}, \boldsymbol{\lambda}^*)$. Comparing (2.22)-(2.25) to (2.7)-(2.10), we indeed observe changes in (2.8)-(2.10), which are the RA variables dependent on π_m^*.

Regarding the stochastic iterates, the only modification required is to replace the M updates in (3.20) with the following single update

$$\pi_L[n+1] = \left[\pi_L[n] - \mu W\Big(\check{p} - \sum_{k,m} p_m^{k*}(\mathbf{h}[n], \boldsymbol{\lambda}_L[n]) w_m^{k*}(\mathbf{h}[n], \boldsymbol{\lambda}_L[n]), L\Big)\right]_0^{\infty}. \tag{2.26}$$

The results in Propositions 1, 2 and 3, as well as that in (2.19) hold also for this case. Together with the changes in the formulation, consideration of a downlink setup entails changes in the signaling schemes and feedback overhead. This issue is discussed in further detail in the next section.

2.7 Signaling and computational overhead

So far, algorithms yielding the optimal allocation of resources as a function of several variables have been developed and their performance has been analyzed. This section briefly discusses the signaling (and computational) overhead required to implement the developed algorithms and points out means to alleviate it. We assume that the algorithms are run in a centralized fashion so that the AP (which acts as a scheduler) carries out most of the required tasks. The operation of the central scheduler proceeds into four steps (phases): (s1) gathering of the information required to find the optimum RA; (s2) calculation of the optimum RA; (s3) notification of the resulting allocation to the transmitters; and (s4) update of the corresponding multipliers. Carrying out these tasks is challenging because the variables involved are available at different locations. For example, the channel gains $\mathbf{h}[n]$ are available at the receiver side via training, but not at the transmitter side (unless channels are reciprocal). On the other hand, the instantaneous values of the queues $\{q_m[n]\}_{m=1}^M$ are only known at the transmitter side because they depend on the instantaneous arrivals. If convenient for reducing the signaling overhead, the multipliers can become available at both sides provided that the corresponding stochastic iterations are run both at the transmitter and the receiver. The overhead associated with each of the aforementioned steps depends on the

system setup (downlink vs. uplink), for this reason the two setups are analyzed separately.

Consider first the downlink case. To find the optimum RA, the AP needs to know $(q_m[n], \alpha_{L,m}[n])$ $\forall m$, $\mu_L[n]$, and $\mathbf{h}[n]$. In this setup the AP acts as transmitter, so that $q_m[n]$, $\alpha_{L,m}[n]$ and $\pi_L[n]$ are available at the AP, and the only problem is to get $\mathbf{h}[n]$. In Time Division Duplexing (TDD) systems, uplink and downlink channels are reciprocal. Hence, the AP can acquire $\mathbf{h}[n]$ by estimating the channel in the reverse link. However, in Frequency Division Duplexing (FDD) systems the forward and reverse channels are non-reciprocal, and therefore $\mathbf{h}[n]$ is not available at the AP. In that case, during step (s1), every user has to send the estimation of its own channels to the AP. If the channel varies sufficiently slow and the number of channels is not too high, the users could send and analog estimation of their channels. However, if it varies to fast or the number of channels is too high, the feedback rate needs to be reduced. Effective methods to feedback the values of $\mathbf{h}[n]$ are to rely on quantized CSI [Mar08a, Lov08], and exploit correlation among channel gains (e.g., by grouping channels [LR10]). Step (s2) is run locally at the AP and only entails computational overhead. For the schemes developed in this chapter the required complexity to carry out such an operation is affordable. Specifically, finding the optimum arrival flows, powers and rates requires evaluation of closed-form expressions [cf. (2.7)-(2.9)]. Finding the optimum scheduling requires the evaluation of the closed-form functional $f(m, k, \mathbf{h}[n], \boldsymbol{\lambda}_L[n])$ for every user m and channel k and computing a maximum per channel [cf. (2.10)]. During step (s3), the scheduler needs to notify the optimum allocation to the terminals. This requires broadcasting the index of the winner user in each of the (groups of) channels. It the users can detect the modulation/coding mode that the AP is using, no further action is required. If not, the AP needs to notify the modulation/coding scheme to be used. If only a finite set of adaptive modulation and coding modes is used (which is the typical case in quantized CSI systems), then only the index of the corresponding mode has to be identified. For this setup, the users do not need to know any multipliers, so that step (s4) is only run at the AP and no feedback is required. Regarding the computational complexity, step (s4) only requires run one iteration of the updates in (3.20)-(2.13) (which is trivial because the optimum allocation has been already found).

The uplink case is more challenging. In this case, we can safely assume that $\mathbf{h}[n]$, $\{\alpha_{L,m}[n]\}_{m=1}^{M}$ and $\{\pi_{L,m}[n]\}_{m=1}^{M}$ are available at the AP (the AP acts as receiver, so that the channel values are known via training). However the values of $\{q_m[n]\}_{m=1}^{M}$ are only known by the users, so that during step (s1) such values have to be feedback to the AP. If the duration of a slot n is long enough, it is reasonable to assume that every user can send its own $q_m[n]$ to the AP through a signaling (feedback) control channel. If not, schemes to reduce the signaling are required. As in the case of the CSI, worth exploring alternatives are to quantize the queue length or exploit the correlation that the queue lengths exhibit across time. A simpler alternative consists in sending the (windowed average) queue length only one out of L slots. Clearly, this would reduce the feedback rate by a factor of L. The proofs in Appendix B can be adapted to show that the optimality/convergence results in Proposition 1 would hold also for such a case. The computational complexity to run step (s2) is the same than that in the downlink setup. During step (s3), the AP broadcasts the index of the winner user in each of the channels. Moreover the power and rate that the winner users need to load have to be broadcasted too. For fast fading channels, the most reasonable alternative to reduce this overhead is to use quantized CSI. In such a case, the AP only needs to identify the index of the rate-power pair to be used.

We close this section by pointing out that worth looking lines of research are developing schemes that: i) work in a fully distributed scenario; ii) are robust against communication errors; and iii) are suboptimal (imperfect), but require a much lower signaling overhead [Lin06b]. Although all these lines are certainly of interest, they go beyond the scope of this thesis.

Table 2.2: Sum utility (here $P^* := \sum_m U_m(\bar{a}_m)$), and sum rate, individual average rates and powers for different test cases (transients have been removed when forming averages). Average power $\bar{p}$ is $[40.0, 20.0, 12.0, 7.0]$ for all five test cases.

	(AS1)	(AS2)	(AS3)	(AS4)	(AS5)
P^*	17.354	17.632	17.097	14.057	8.968
$\sum_m \bar{a}_m$	467.06	442.76	451.96	157.72	82.49
$\bar{a}_1$	282.46	239.81	300.00	76.14	53.39
$\bar{a}_2$	115.54	119.16	85.82	43.16	21.74
$\bar{a}_3$	49.36	58.31	45.00	25.09	6.68
$\bar{a}_4$	19.70	25.48	21.14	13.33	0.68
$\bar{p}$	[40.0, 20.0, 12.0, 7.0]				

2.8 Numerical results

In this section, numerical results are presented to assess the analytical claims. Unless explicitly mentioned, the default simulation setup will be an OFDM uplink system with $M = 4$ users and $K = 64$ parallel channels. Channels are time and frequency selective and the corresponding SNRs follow an exponential distribution (Rayleigh fading channel model). The average power gains for users 1, 2, 3 and 4 are 6 dB, 4.5 dB, 3 dB and 1.5 dB, respectively. Arrivals are randomly drawn from a binomial distribution. Simulations are run assuming that the signaling channels were ideal, so that both CSI and QCI are assumed to be perfectly known.

Test Case 1: optimality and feasibility. First, we illustrate that the developed algorithms are optimal and feasible, and compare their performance with that of existing alternatives. Values of the average transmit-power, transmit-rate, sum-rate and sum-utility for the five allocation schemes (AS) tested, are as follows: (AS1) utilities are linear (so that the objective corresponds to sum-rate), and there are no rate restrictions; (AS2) the utility functions are $U_m(a) = \log(1 + a)\ \forall m$, and there are no rate restrictions; (AS3) as AS2 but with $\check{a}_1 = 300$; (AS4) a simplified version of AS2 with fixed subcarrier allocation; and (AS5) an algorithm compatible with the WiMAX standard, where the subcarrier and power allocation are fixed and the SNR per subcarrier is used to select the transmit-rate from a codebook with 8 pre-specified values. In all five schemes the average power constraints are $\check{\mathbf{p}} := [\check{p}_1, \check{p}_2, \check{p}_3, \check{p}_4]^T = [40, 20, 12, 7]^T$. To reduce the number of schemes simulated, in this test case we only simulate the non-stochastic version of our algorithms. Extensive simulations assessing the optimality loss of the stochastic algorithms relative to their stochastic counterparts will be carried out in Test Case 3.

The results listed in Table 2.2 illustrate that the proposed algorithm is able to provide fairness and guarantees on minimum average rate requirements, while it clearly outperforms other suboptimal schemes which fail to satisfy these requirements. Regarding optimality, we see how indeed AS1 outperforms AS2 in terms of sum-rate (sum-rate is the metric maximized by AS1) and how AS2 does better than AS1 in terms of sum-utility (sum-utility is the metric maximized by AS2). Moreover, we also see that to satisfy the minimum rate requirement for user 1, AS3 needs to sacrifice some utility. (This is because the unconstrained problem in AS2 does not fulfill the requirement $\check{a}_1$, and therefore the constraint needs to be activated.) Recall that optimality of the non-stochastic schemes has been proved theoretically and the results here try to illustrate the gains with respect to suboptimal schemes. Finally, it should be stressed that the small differences in terms of sum-utility values among AS1, AS2 and AS3 is due to the specific set-up selected. Had the utilities been different or the rates smaller, the difference in terms of sum-utility would have been larger. However, for the simulated configuration, the transmit-rates are large enough so that

the selected utility is basically flat.

Test Case 2: dynamic behavior of the stochastic schemes. To gain further insight on the convergence of our algorithms, we analyze the dynamic behavior of the stochastic iterates. We focus on the test case in AS3, with $L = 1$ and $\mu = 10^{-4}$. Figure 3.1 comprises six subplots, each depicting the evolution over time of a different subset of variables. Subplot (a) is the evolution of the average transmit rate $\bar{a}_m[n] := \frac{1}{n}\sum_{l=1}^{n} a_m^*(\boldsymbol{\lambda}_L[l])$, while (b) is the evolution of the average power consumption $\bar{p}_m[n] := \frac{1}{n}\sum_{l=1}^{n}(\sum_k p_m^{k*}(\mathbf{h}[l], \boldsymbol{\lambda}_L[l]) w_m^{k*}(\mathbf{h}[l], \boldsymbol{\lambda}_L[l])$. In these subplots, dashed lines correspond to the performance when the optimal multipliers are known, while continuous lines correspond to the stochastic algorithms proposed in this chapter. Subplots (c), (d) and (f) plot the instantaneous value of the Lagrange multipliers: $\rho_{L,m}[n]$, $\alpha_{L,m}[n]$ and $\pi_{L,m}[n]$, respectively. Finally, subplot (e) depicts the time-trajectory of the instantaneous value of the queue lengths $q_m[n]$.

Starting with the trajectories of the running averages $\bar{a}_m[n]$ and $\bar{p}_m[n]$ in subplots (a) and (b), we observe that the considered constraints are satisfied (with equality), and that the stochastic allocation converges in some hundred (correspondingly, a couple of thousand) iterations. The behavior of the Lagrange multipliers plotted in (c), (d), and (e) is slightly different. As in (a) and (b), there is an initial phase during which the multipliers go from the initial point (which was chosen randomly) towards the optimal value. However, once the multipliers approach their optimal value, they do not strictly converge but hover around it. This in not unreasonable, because $\rho_{L,m}[n]$, $\alpha_{L,m}[n]$ and $\pi_{L,m}[n]$ are instantaneous variables while $\bar{a}_m[n]$ and $\bar{p}_m[n]$ are running averages. The numerical results also reveal that for the simulated test case, $\alpha_{L,1}$ is non-zero. This means that the rate constraint for $m = 1$ needs to be enforced, as mentioned in the Test Case 1. We also observe that all $\pi_{L,m}$ and $\rho_{L,m}$ are also non-zero $\forall m$, confirming that constraints in (2.3) and (2.5) are always active. Finally, the trajectories of queue lengths are depicted in subplot (e). Note that the queues are stable, and the relationship between $\rho_{L,m}[n]$ and $q_m[n]$ holds in practice.

Test Case 3: changing the stepsize and the window length. The numerical results for this test case are listed in Table 2.3. The results correspond to a setup where the power gains for users 1, 2, 3 and 4 are 6, 5, 4 and 3 dB, respectively; the maximum average power constraints are $\check{\mathbf{p}} = [40, 20, 16, 10]^T$. The minimum average rate constraints in the upper half of the table are $\check{\mathbf{a}} := [\check{a}_1, \check{a}_2, \check{a}_3, \check{a}_4]^T = [0, 0, 0, 55]^T$, while the average rate constraints in the lower half of the table are $\check{\mathbf{a}} = [150, 0, 0, 0]^T$. All other parameters are set as in the default case. The values for μ and L vary as indicated in the entries of the table. The numerical results reveal that larger stepsizes give rise to larger gaps from the optimum. This trend is reasonable, because the bounds on the loss of optimality in Proposition 1 are proportional to μ. It must be noted, however, that the bounds were derived using a worst-case approach, so that the actual loss does not have always to follow such a trend. The variation of the losses with respect to L is more difficult to explain. On the one hand, the bound on the optimality loss grows if the value of L increases (this was not established in the statement of Proposition 1, but it was revealed in its proof). On the other hand, larger values of L, render the stochastic updates of the multipliers closer to the original (non-stochastic) updates. This contradictory behavior may be the cause for the results in Table 2.3. For $\mu = 10^{-5}$, larger values of L incur more sizeable losses; for $\mu = 10^{-4}$, the trend is the opposite; and for $\mu = 3 \cdot 10^{-5}$, the variation of the losses is not monotonic.

In order to illustrate the behavior of our algorithm in a scenario with larger number of users, additional numerical results (confirming the previous findings) are presented in Table 2.4. The simulation setup is similar to the previous one, but in this case there are $M = 8$ users instead of 4. The average power gains are $[6, 6, 4.5, 4.5, 3, 3, 1.5, 1.5]^T$ and the maximum average power constraints are $\check{\mathbf{p}} = [80, 40, 40, 20, 30, 15, 20, 10]^T$. Two different rate requirements are considered: $\check{\mathbf{a}} = [100, 0, 0, 0, 0, 0, 0, 0]^T$ for the configurations listed in the upper half of the table and $\check{\mathbf{a}} =$

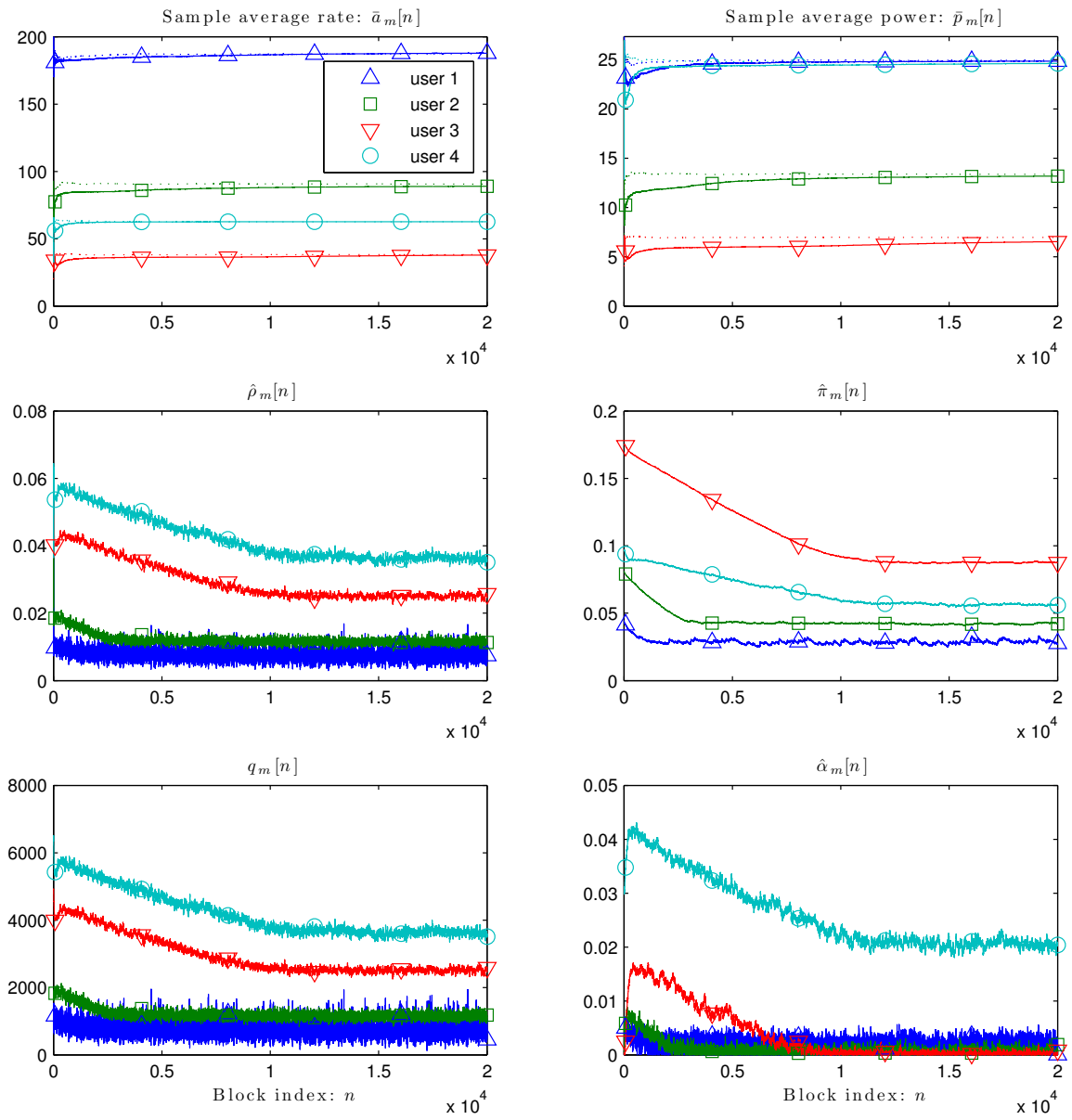

Figure 2.1: Trajectories of different primal and dual variables in test case 2: (a) sample average rate $\bar{a}_m[n]$; (b) sample average power $\bar{p}_m[n]$; (c) flow conservation multiplier $\hat{\rho}_m[n]$; (d) estimate of the average power multiplier $\hat{\pi}_m[n]$; (e) queue length $q_m[n]$; (f) estimate of the average arrival rate multiplier $\hat{\alpha}_m[n]$.

Table 2.3: Sum utility (P^*) and sum rate for different values of μ and window sizes.

$\check{\boldsymbol{a}} = [0, 0, 0, 55]$				
CASE	Stepsize μ	Window size L	P^*	$\sum_m \bar{a}_m$
OFF-LINE	N/A	N/A	18.765	490.38
ONLINE	1.0E-5	1	18.722	483.87
		3	18.712	482.44
		8	18.571	464.98
	3.0E-5	1	18.424	435.13
		3	18.474	452.21
		8	18.432	453.49
	1.0E-4	1	17.592	337.50
		3	18.085	399.57
		8	18.371	449.02
$\check{\boldsymbol{a}} = [250, 0, 0, 0]$				
CASE	Stepsize μ	Window size L	P^*	$\sum_m \bar{a}_m$
OFF-LINE	N/A	N/A	18.784	513.21
ONLINE	1.0E-5	1	18.736	508.85
		3	18.662	503.79
		8	18.389	483.58
	3.0E-5	1	18.354	483.53
		3	18.271	477.65
		8	18.062	462.59
	1.0E-4	1	17.551	424.10
		3	17.871	451.05
		8	17.735	445.18

Table 2.4: Sum utility (P^*) and sum rate for different values of μ and window sizes. M = 8 users; K = 10 channels

$\check{\boldsymbol{a}}_1 = 100$				
CASE	Stepsize μ	Window size L	P^*	$\sum_m \bar{a}_m$
OFF-LINE	N/A	N/A	24.095	211.20
ONLINE	3.0E-5	1	23.956	208.69
		4	23.951	208.64
	3.0E-6	4	24.046	210.44
		10	24.042	210.37
$\check{\boldsymbol{a}}_2 = 70; \check{\boldsymbol{a}}_4 = 40$				
CASE	Stepsize μ	Window size L	P^*	$\sum_m \bar{a}_m$
OFF-LINE	N/A	N/A	24.010	197.09
ONLINE	3.0E-5	1	23.845	193.74
		4	23.846	193.81
	3.0E-6	4	23.974	195.98
		10	23.967	195.85

Table 2.5: Queue lengths, average delays [cf. (2.19)], and sum utility (P^*) for the allocation schemes in Test Case 4.

	(ASa)	(ASb)	(ASc)	(ASd)	(ASe)	(ASf)
$\bar{q}_1$	960	7791	115	726	143608	463
$\bar{q}_2$	1051	9436	119	974	70832	333
$\bar{q}_3$	1480	13382	104	1449	46116	217
$\bar{q}_4$	2162	19543	220	2077	32603	203
$\bar{d}_1$	4.5	39.0	0.6	3.6	717.8	2.4
$\bar{d}_2$	10.6	89.8	1.1	9.3	673.6	3.3
$\bar{d}_3$	21.8	181.2	1.4	19.6	623.4	3.0
$\bar{d}_4$	47.0	389.5	4.5	41.3	649.1	4.2
P^*	18.050	18.218	18.153	18.220	18.222	18.101

$[0, 70, 0, 40, 0, 0, 0, 0]^T$ for the configurations listed in the lower half of the table.

Test Case 4: evaluating the average delay. Finally, we analyze the queuing delay associated with the developed schemes. We consider a downlink setup with $M = 4$, $L = 1$, $K = 64$, $\check{\mathbf{a}} = [200, 80, 40, 30]^T$, $\check{p} = 80$. Six test cases are considered: (ASa) our stochastic scheme with $\mu = 10^{-5}$ and $\boldsymbol{\lambda}_L[0] = 0$; (ASb) our stochastic scheme with $\mu = 10^{-6}$ and $\boldsymbol{\lambda}_L[0] = 0$; (ASc) the scheme in ASa but with $\boldsymbol{\lambda}_L[0] = 0.9\boldsymbol{\lambda}^*$; (ASd) the counterpart of ASc for ASb; (ASe) the non-stochastic scheme which assumes that $\boldsymbol{\lambda}^*$ is known and uses these values to adapt resources; and (ASf) a modified version of the non-stochastic scheme in ASe, where the average arrival rate $\bar{a}^*(\boldsymbol{\lambda}^*)$ is slightly backed-off so that the delay is decreased (specifically, simulations are run with $\bar{a} = 0.97\bar{a}^*(\boldsymbol{\lambda}^*)$). In all six cases, packet arrival rates follow a binomial distribution. The corresponding average queue length, average delay and utility loss are listed in Table 2.5.

The main observation from the numerical results in Table 2.5 is that the average delay is indeed inversely proportional to the stepsize. Moreover, the results also show that the delay of the stochastic schemes ASa, ASb, ASc and ASd is considerably lower than that of the non-stochastic optimal solution in ASe. In addition, comparison of the simulated average delays for the schemes ASf and ASe reveals that the delay of the non-stochastic schemes can be considerably reduced at the expense of a small loss in terms of optimality (sum-utility). Differently, stochastic schemes ASc and ASd also achieve a delay reduction relative to their counterparts ASa and ASb, but without incurring a sum-utility loss (relative to ASa and ASb). The delay improvement observed in the stochastic schemes that use a "smart" initialization is in agreement with the results in [Hua11]. Specifically, [Hua11] develops different algorithms that explicitly consider the instantaneous queue length and shows that if the actual queue lengths are offset by a positive number (referred to as place-holder bits), the delay optimality trade-offs improve considerably. The effect of such place-holder bits is very similar to the non-zero initialization considered in ASc and ASd. Characterizing the optimality versus delay trade-off for the nonstochastic schemes, their stochastic counterparts (relying on the results in [Hua11]), and stochastic versus nonstochastic alternatives, are interesting research lines, but go beyond the scope of this thesis.

To gain insights on the sensitivity of the average delay with respect to the stepsize, we run a last experiment. In this experiment, the value of the stepsize is modified online and the effect on the size of the queues and the average delay is analyzed. Specifically, the stepsize remains constant during a period of 10,000 time iterations, and then it is reduced by a factor of $1/4$. The simulation comprises 5 periods so that the total simulated time is 50,000 instants. The corresponding results are plotted in Figure 2.2, where subplot (a) represents the queue lengths of all users; (b) depicts

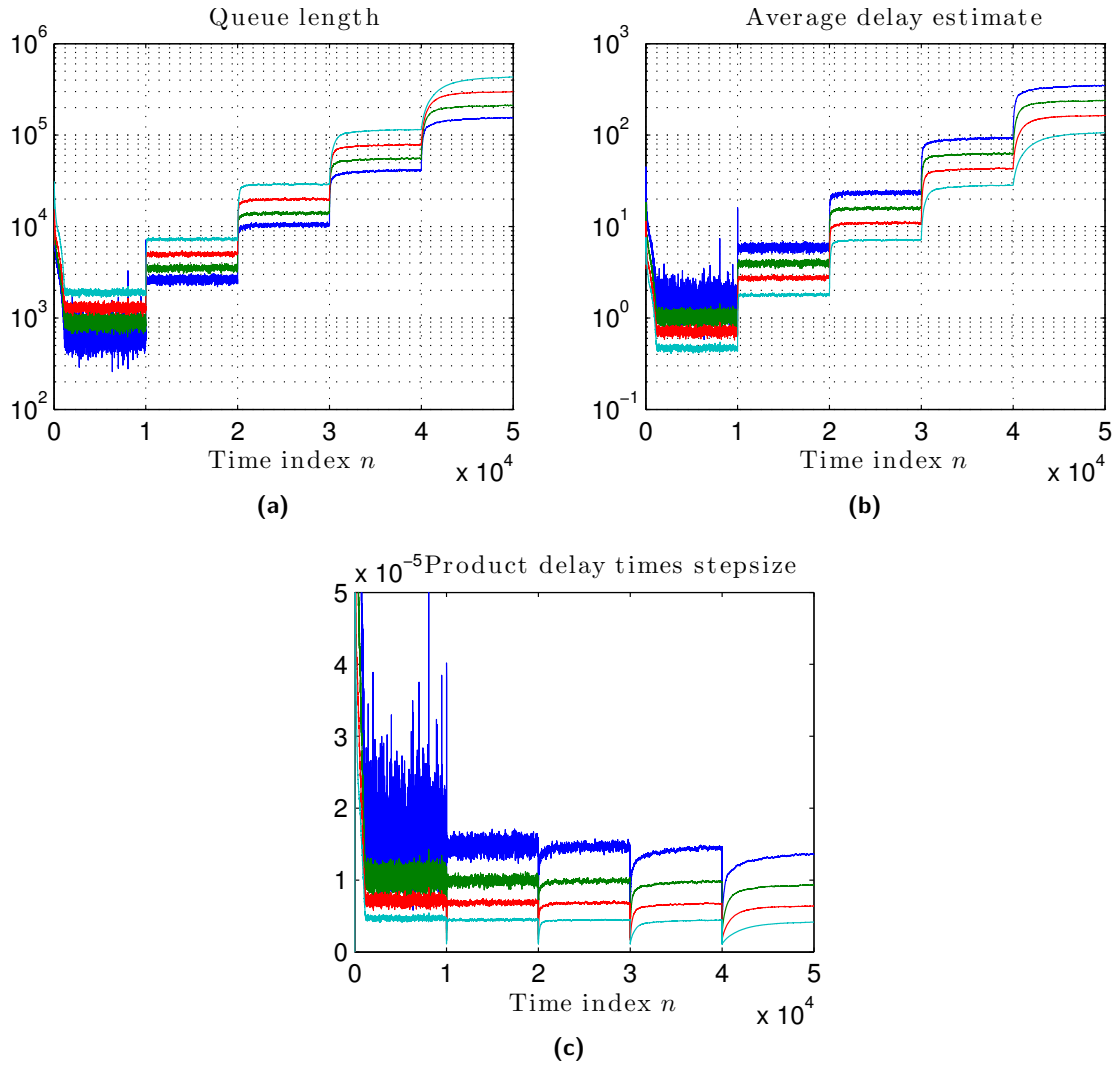

Figure 2.2: Trajectories of (a) queue length, (b) average delay and (c) product delay stepsize for a slightly varying stepsize μ. The variation of the stepsize corresponds to a (decreasing) piecewise constant function.

Table 2.6: Sum utility (P^*), sum rate, and average delays [cf. (2.19)] for the allocation schemes in Test Case 5.

	(A1)	(A2)	(A3)
P^*	23.954	24.077	24.055
$\sum_m \bar{a}_m$	208.65	210.89	210.78
$\sum_m \bar{d}$	1531	12314	1526

the delay experienced by the users; and, (c) represents the product delay times stepsize. Results in subplots (a) and (b) show that every time the stepsize changes, the queue length and delay increase notably. However, after a transient period the stochastic schemes are able to drive the queues to a stable level (note also that the duration of this transient seems also to increase as the stepsize decreases). Moreover, the product delay times stepsize, is stabilized around a steady state value confirming that the average delay is inversely proportional to the stepsize (as predicted by Proposition 3).

Test Case 5: comparison with other optimal RA algorithms. The last test case is devoted to compare the performance of our algorithms with that of existing state-of-the art alternatives. The metrics to be compared are sum average delay and sum-utility. Three algorithms are compared: (A1) our stochastic optimization algorithm with $L = 1$ and $\mu = 10^{-5}$; (A2) an stochastic algorithm with time diminishing stepsize similar to the proposed in [Wan07] (the stepsize considered is $\mu[n] = \mu[0]/(1 + 100n/N)$, where $N = 50000$ is the total simulation time); and (A3) the optimal queue control algorithm proposed in [Hua11]. Neither [Wan07] nor [Hua11] consider the exact same operating conditions than the ones in this chapter, so that the solutions proposed in those papers require some adaptation. Since [Hua11] considers an unconstrained optimization problem, we will assume that the value of the *optimum* values of the multipliers associated with the average rate and power constraints (i.e., α_m^* and π_m^* $\forall m$) are perfectly known. To ensure a fair comparison, the initialization of the queues (multipliers) in both algorithms is the same and the parameter V in [Hua11] is set to $V = 1/\mu$.

The scenario considered is the same than one of the considered in Test Case 3 ($M = 8$ and $\check{\mathbf{a}} = [100, 0, 0, 0, 0, 0, 0, 0]^T$). Results are listed in Table 2.6. The first observation is that A2 achieves the best sum-utility value. This was expected because the schemes in [Wan07] are optimal as $n \to \infty$. On the other hand, A2 also exhibits the highest delay. This was expected too, because the steady state of the algorithms in [Wan07] is in fact the same than that of the optimal non-stochastic solution (recall that the results in Test Case 4 confirmed that the queues for the non-stochastic algorithm can grow unboundedly). Comparing A3 with A1, the results show that the sum-utility loss is slightly higher for our scheme while the achieved delays are similar. This is also reasonable, because A3 capitalizes on the perfect knowledge of α_m^* and π_m^* (which amounts to have knowledge of the channel distribution) while A1 is a purely stochastic (online) algorithm.

2.9 Concluding summary

Cross-layer algorithms were designed in this chapter to allocate resources (flows, channel access, power and rates) in a cellular system, where users transmit over a set of orthogonal channels. Both uplink and downlink setups were addressed. The developed RA strategy depends on the instantaneous fading, the queue lengths, and user-specific weights. The latter correspond to Lagrange multipliers and are estimated online using stochastic approximation iterations. Capitalizing on a relationship between multiplier estimates and the windowed average queue lengths, stability and average delay performance were analyzed. Finally, a mechanism to effect delay priorities among

users by tuning the stepsize of individual users was also discussed.

Appendices

Appendix A: Optimal solution for (3.6)

To solve for the optimal RA the constraints in (3.6) will be split into two groups. The first group is formed by (2.3), (2.5) and (2.6b). Those average constraints are generally difficult to handle and dual techniques will be used to deal with them. The second group is formed by (2.1) and (2.2) plus the nonnegative constraints that were not explicitly written in (3.6) but confine all $w_m^k(\mathbf{h}), p_m^k(\mathbf{h})$ and $\bar{a}_m$ to be nonnegative. The constraints in the second group are easy to handle, and there is no need to dualize them.

To find the optimal solution for (3.6) we will: a) form the corresponding Lagrangian; b) minimize the Lagrangian while guaranteeing that the constraints in the second group are satisfied and c) substitute for the optimum value of the Lagrange multipliers to recover the primal variables.

We start by introducing notation to form the Lagrangian associated with the problem in (3.6). Let $\mathbf{x}$ be a vector containing all primal variables $(\bar{a}_m, w_m^k(\mathbf{h}), p_m^k(\mathbf{h}),\ \forall m,k,\mathbf{h})$. Note that $\mathbf{x}$ has infinite length because $\mathbf{h}$ takes infinite values. Using these notational conventions, the Lagrangian is:

$$\mathcal{L}(\mathbf{x},\boldsymbol{\lambda}) = \sum_{m=1}^{M}\Big(-U_m(\bar{a}_m) + \alpha_m(\bar{a}_m - \check{a}_m) \quad + \rho_m\big(\mathbb{E}\Big[\bar{a}_m - \sum_{k=1}^{K} w_m^k(\mathbf{h})C_m^k(\mathbf{h},p_m^k(\mathbf{h}))\Big]\big) + \pi_m\Big(\mathbb{E}\Big[\sum_{k=1}^{K} w_m^k(\mathbf{h})p_m^k(\mathbf{h}) - \check{p}_m\Big]\Big)\Big). \tag{2.27}$$

For a given $\boldsymbol{\lambda}$, we need to minimize $\mathcal{L}(\mathbf{x},\boldsymbol{\lambda})$ with respect to (w.r.t.) $\mathbf{x}$. As it will be apparent next, the structure of $\mathcal{L}(\mathbf{x},\boldsymbol{\lambda})$ allows the minimization w.r.t. $\bar{a}_m$, $p_m^k(\mathbf{h})$ and $\{w_m^k(\mathbf{h})\}_{m=1}^M$ to be performed separately. Since $\mathcal{L}(\mathbf{x},\boldsymbol{\lambda})$ is strictly convex and differentiable w.r.t. $\bar{a}_m$ and $p_m^k(\mathbf{h})$, minimization w.r.t. those variables amounts to equating the corresponding partial derivatives of $\mathcal{L}(\mathbf{x},\boldsymbol{\lambda})$ to zero. Differently, the minimization w.r.t. $w_m^k(\mathbf{h})$ is slightly more complicated, mainly because $\mathcal{L}(\mathbf{x},\boldsymbol{\lambda})$ is linear in $w_m^k(\mathbf{h})$.

Consider first the optimum arrival flows. Differentiating $\mathcal{L}(\mathbf{x},\boldsymbol{\lambda})$ w.r.t $\bar{a}_m$ and equating the derivative to zero yields $\dot{U}_m(\bar{a}_m^*) + \alpha_m - \rho_m = 0$. Solving the latter w.r.t. $\bar{a}_m$ yields

$$\bar{a}_m^*(\boldsymbol{\lambda}) = (\dot{U}_m)^{-1}(\rho_m - \alpha_m). \tag{2.28}$$

To guarantee that the latter satisfies the nonnegative constraints, we just need to project $\bar{a}_m^*$ onto the feasible set (nonnegative orthant). Thus, the solution is

$$\bar{a}_m^*(\boldsymbol{\lambda}) = \big[(\dot{U}_m)^{-1}(\rho_m - \alpha_m)\big]_0^{\infty}. \tag{2.29}$$

Proceeding similarly with the optimum power allocation, we set the partial derivative of $\mathcal{L}(\mathbf{x},\boldsymbol{\lambda})$ w.r.t. $p_m^k(\mathbf{h})$ to zero. This yields $\rho_m w_m^k(\mathbf{h})\dot{C}_m^k(\mathbf{h},p_m^k(\mathbf{h})) - \pi_m w_m^k(\mathbf{h}) = 0$, which can be rewritten as $[\rho_m \dot{C}_m^k(\mathbf{h},p_m^k(\mathbf{h})) - \pi_m]w_m^k(\mathbf{h}) = 0$. Clearly, the equality is satisfied if: i) $w_m^k(\mathbf{h}) = 0$ or ii) $\rho_m \dot{C}_m^k(\mathbf{h},p_m^k(\mathbf{h})) - \pi_m = 0$. Supposing that $w_m^k(\mathbf{h}) \neq 0$, then ii) needs to hold. The root of ii) is $p_m^{k\star}(\mathbf{h},\boldsymbol{\lambda}) = (\dot{C}_m^k)^{-1}(\mathbf{h},\pi_m/\rho_m)$. As before, we must guarantee that $p_m^{k\star}(\mathbf{h},\boldsymbol{\lambda})$ is nonnegative and satisfies (2.3); hence, the optimal power allocation is

$$p_m^{k\star}(\mathbf{h},\boldsymbol{\lambda}) = \big[(\dot{C}_m^k)^{-1}(\mathbf{h},\pi_m/\rho_m)\big]_0^{\check{p}^k}. \tag{2.30}$$

If now $w_m^k(\mathbf{h}) = 0$, then any finite value of $p_m^k(\mathbf{h})$ is equally optimum. In fact, $w_m^k(\mathbf{h})$ being zero means that channel k is not assigned to user m, so that the effective transmit power (and rate) is zero for any finite value of $p_m^k(\mathbf{h})$, including the one in (2.30). For this reason, the power allocation in (2.30) is optimum regardless of the value $w_m^k(\mathbf{h})$. Interestingly, if $C_m^k(\cdot)$ corresponds to Shannon's capacity, the solution in (2.30) reduces to

$$p_m^k(\mathbf{h}, \boldsymbol{\lambda}) = \left[\frac{\rho_m}{\pi_m} - \frac{1}{h_m^k}\right]_0^{\hat{p}^k} \tag{2.31}$$

which corresponds to the celebrated *waterfilling* solution (with ρ_m/π_m being the waterfilling level) that maximizes the sum-capacity of a set of parallel orthogonal channels [Cov06].

As already mentioned, minimization w.r.t. $w_m^k(\mathbf{h})$ has to be handled more carefully. The first reason is that $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda})$ is linear in $w_m^k(\mathbf{h})$; hence the optimal scheduling cannot be found by differentiating w.r.t. $w_m^k(\mathbf{h})$. The second reason is that constraint (2.1), which was not dualized, couples the scheduling across users, so that the mere projection in (2.29) and (2.30) cannot be used here. And the third reason is that to find the optimum $w_m^{k*}(\mathbf{h}, \boldsymbol{\lambda})$, the optimum values $\{p_{m'}^{k*}(\mathbf{h}, \boldsymbol{\lambda})\}_{m'=1}^M$ need to be known. To solve for $w_m^{k*}(\mathbf{h}, \boldsymbol{\lambda})$, we first note that the terms in $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda})$ that depend on $w_m^k(\mathbf{h}, \boldsymbol{\lambda})$ are [cf. (2.27)]

$$-\sum_{m=1}^M \rho_m \mathbb{E}\left[\sum_{k=1}^K w_m^k(\mathbf{h}) C_m^k(\mathbf{h}, p_m^k(\mathbf{h}))\right] + \sum_{m=1}^M \pi_m \mathbb{E}\left[\sum_{k=1}^K w_m^k(\mathbf{h}) p_m^k(\mathbf{h})\right]. \tag{2.32}$$

Substituting for the optimum $p_m^{k*}(\mathbf{h}, \boldsymbol{\lambda})$, rearranging terms and defining the functional $f(m, k, \mathbf{h}, \boldsymbol{\lambda}) := \rho_m C_m^k(\mathbf{h}, p_m^{k*}(\mathbf{h}, p_m^{k*}(\mathbf{h}, \boldsymbol{\lambda}))) - \pi_m p_m^{k*}(\mathbf{h}, \boldsymbol{\lambda})$ (cf. Sec. 2.4.1), it is possible to rewrite (2.32) as

$$\mathbb{E}\left[\sum_{k=1}^K \left(-\sum_{m=1}^M w_m^k(\mathbf{h}) f(m, k, \mathbf{h}, \boldsymbol{\lambda})\right)\right]. \tag{2.33}$$

Based on the previous expression, it readily follows that finding the optimal $w_m^k(\mathbf{h}, \boldsymbol{\lambda})$ that minimizes (2.27) amounts to solving, for each $\mathbf{h}$ and k, the following problem:

$$\min_{\{w_m^k(\mathbf{h}, \boldsymbol{\lambda})\}_{m=1}^M} \; -\sum_{m=1}^M w_m^k(\mathbf{h}) f(m, k, \mathbf{h}, \boldsymbol{\lambda}) \tag{2.34a}$$

$$\text{s. to:} \quad \sum_{m=1}^M w_m^k(\mathbf{h}, \boldsymbol{\lambda}) \leq 1 \tag{2.34b}$$

$$w_m^k(\mathbf{h}, \boldsymbol{\lambda}) \geq 0 \qquad \forall m \tag{2.34c}$$

where (2.34b) and (2.34c) are the constraints that involve $w_m^k(\mathbf{h})$ and were not dualized.

Assuming that $f(m, k, \mathbf{h}, \boldsymbol{\lambda})$ is nonnegative for at least one user[4], the solution of the previous problem is straightforward and consists of setting $w_m^k(\mathbf{h}, \boldsymbol{\lambda}) = 1$ for the user m which maximizes $f(m, k, \mathbf{h}, \boldsymbol{\lambda})$, while setting $w_m^k(\mathbf{h}, \boldsymbol{\lambda}) = 0$ for all other users. This policy can be written in closed form using the indicator function as

$$w_m^{k*}(\mathbf{h}, \boldsymbol{\lambda}) = \mathbb{1}_{\{m = \arg\max_{m'}\{f(m', k, \mathbf{h}, \boldsymbol{\lambda}))\}\}}. \tag{2.35}$$

Substituting $\boldsymbol{\lambda} = \boldsymbol{\lambda}^*$ into (2.29), (2.30) and (2.35), yields the optimal solution in (2.7), (2.8) and (2.10), respectively.

[4] The convexity of $C_m^k(\cdot)$ can be used to rigorously show that $f(m, k, \mathbf{h}, \boldsymbol{\lambda})$ is always nonnegative.

Appendix B: Convergence and optimality of stochastic schemes

We begin by briefly reviewing some concepts from optimization theory and introducing notation. Then, we present a result on boundness of Lagrange multipliers (Lemma 1) to be used in subsequent proofs. Using Lemma 1, we prove Proposition 1 (Appendix B.1), and subsequently Proposition 3 (Appendix B.2). For brevity, several of the proofs presented here guarantee only convergence in the mean (hence, in probability), but more sophisticated tools can be used to establish convergence w.p.1 [Boy06], [Ber03, Prop. Sec. 8.2.2], [Rib10a].

The definition of a stochastic subgradient is reviewed first. Let f be a scalar, real-valued function of vector $\mathbf{x}$. Vector $\boldsymbol{\partial}(\mathbf{x}_0)$ is an unbiased stochastic subgradient of f at point $\mathbf{x}_0$ if and only if $f(\mathbf{x}) - f(\mathbf{x}_0) \geq \mathbb{E}^T[\boldsymbol{\partial}(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0)$ w.p.1. Getting back to our problem, we will use $D(\boldsymbol{\lambda})$ to denote the dual function of (3.6) [Ber99, Sec. 5.1.2]. Key for the subsequent proofs is the fact that the updates in (3.20)-(2.13) are finite-length averages of stochastic subgradients of $D(\boldsymbol{\lambda})$. To be more precise, let $\boldsymbol{\partial}[n]$ denote a $3M \times 1$ vector whose ith entry is

$$\partial_i[n] := \check{p}_i - \textstyle\sum_k p_i^{k*}(\mathbf{h}[n], \boldsymbol{\lambda}_L[n]) w_i^{k*}(\mathbf{h}[n], \boldsymbol{\lambda}_L[n]), \text{ if } 1 \leq i \leq M;$$

$$\partial_i[n] := \textstyle\sum_k r_{i-M}^{k*}(\mathbf{h}[n], \boldsymbol{\lambda}_L[n]) w_{i-M}^{k*}(\mathbf{h}[n], \boldsymbol{\lambda}_L[n]) - a_{i-M}^*[\boldsymbol{\lambda}_L[n]], \text{ if } M+1 \leq i \leq 2M;$$

$$\text{and } \partial_i[n] := a_{i-2M}^*[\boldsymbol{\lambda}_L[n]] - \check{a}_{i-2M}, \text{ if } 2M+1 \leq i \leq 3M.$$

It is easy to verify that for $L = 1$, $\boldsymbol{\partial}[n]$ comprises the correction terms in (3.20)-(2.13), so that (3.20)-(2.13) can be collectively written as the first-order vector recursion: $\boldsymbol{\lambda}_L[n+1] = [\boldsymbol{\lambda}_L[n] - \mu\boldsymbol{\partial}[n]]_0^\infty$. For the general case $L > 1$, define $\boldsymbol{\partial}_L[n] := W(\boldsymbol{\partial}[n], L)$ and write (3.20)-(2.13) as $\boldsymbol{\lambda}_L[n+1] = [\boldsymbol{\lambda}_L[n] - \mu\boldsymbol{\partial}_L[n]]_0^\infty$, where now the (3.20)-(2.13) correction terms are windowed averages of $\boldsymbol{\partial}[n]$.

Using results from optimization theory, it follows readily that $\boldsymbol{\partial}[n]$ is an unbiased stochastic gradient of $D(\boldsymbol{\lambda})$ at point $\boldsymbol{\lambda}_L[n]$ (cf. [Ber99, Ch. 6]). This implies by definition that $D(\boldsymbol{\lambda}) - D(\boldsymbol{\lambda}_L[n]) \geq \mathbb{E}^T[\boldsymbol{\partial}[n]](\boldsymbol{\lambda}_L[n] - \boldsymbol{\lambda})$ w.p.1, which is an important property to be used later on. Note however that $\boldsymbol{\partial}_L[n]$ for $L > 1$ may not be an *unbiased* stochastic subgradient of $D(\boldsymbol{\lambda})$. In fact, the subsequent proofs will reveal that $\boldsymbol{\partial}_L[n]$ is in general a *biased* estimate of the subgradient of $D(\boldsymbol{\lambda})$, and therefore results from convergence of ϵ-subgradients need to be used [Ber03, Sec. 4.3]. This will slightly complicate the convergence proofs. Finally, let $\tilde{\boldsymbol{\partial}}_L[n]$ denote the normalized multiplier error $\tilde{\boldsymbol{\partial}}_L[n] := (\boldsymbol{\lambda}_L[n+1] - \boldsymbol{\lambda}_L[n])/\mu$. In words, $\tilde{\boldsymbol{\partial}}_L[n] = \boldsymbol{\partial}_L[n]$ if $\boldsymbol{\lambda}_L[n] - \mu\boldsymbol{\partial}_L[n]$ is nonnegative so that the projection operator in $\boldsymbol{\lambda}_L[n+1] = [\boldsymbol{\lambda}_L[n] - \mu\boldsymbol{\partial}_L[n]]_0^\infty$ is not needed. Otherwise, the entries of $\boldsymbol{\partial}_L[n]$, which are large enough to render $\boldsymbol{\lambda}_L[n] - \mu\boldsymbol{\partial}_L[n]$ negative, are partially clipped in $\tilde{\boldsymbol{\partial}}_L[n]$. Based on this definition, we can write $\boldsymbol{\lambda}_L[n+1] = \boldsymbol{\lambda}_L[n] - \mu\tilde{\boldsymbol{\partial}}_L[n]$. This notation will be used at several points in the proofs to bypass the non-linearity of the projection operator in $\boldsymbol{\lambda}_L[n+1] = [\boldsymbol{\lambda}_L[n] - \mu\boldsymbol{\partial}_L[n]]_0^\infty$.

Now, we are ready to present a result on the boundness of the multipliers.

Lemma 1: *If the problem in* (3.6) *is strictly feasible and the constraint set in* (2.6b) *and* (2.6c) *is bounded, then there exists a finite constant* c *so that* $\mathbb{E}[\|\boldsymbol{\lambda}_L[n]\|] < c \ \forall n$.

This is a standard result for dual (sub)gradient algorithms, and its proof will only be sketched here. The proof follows the steps in, e.g., [Ned09, Lemma3], after writing $\boldsymbol{\partial}_L[n]$ as a sum of L terms and taking expectations. Intuitively speaking, this claim is reasonable because: a) feasibility implies that the optimum multipliers $\boldsymbol{\lambda}^*$ are finite; b) the subgradient iterations drive $\mathbb{E}[\boldsymbol{\lambda}_L[n]]$ closer to $\boldsymbol{\partial}^*$; and c) the maximum correction term is bounded. Since convergence in the mean implies convergence in probability, it is guaranteed that $\|\boldsymbol{\lambda}_L[n]\|$ is bounded in probability. As mentioned earlier, the ensuing appendixes will prove convergence in probability, but not convergence w.p.1. One of the key steps to show convergence w.p.1 is to prove that $\|\boldsymbol{\lambda}_L[n]\| < c \ \forall n$ w.p. 1. This can be done using a slight generalization of the supermartingale convergence theorem; see [Ber03,

Props. 8.2.10 and 8.2.11] and [Ber03, Prop. 8.2.11] for details, or [Rib10a] for a related argument.

Appendix B.1: Proof of Proposition 1

Proof of (i): Recall that the updates in (3.20)-(2.13) can be written in a single vector recursion as $\boldsymbol{\lambda}_L[n+1] = [\boldsymbol{\lambda}_L[n] - \mu\boldsymbol{\partial}_L[n]]_0^\infty$. Since the stochastic iteration is projected over the nonnegative orthant, it follows that $\boldsymbol{\lambda}_L[n+1] \geq \boldsymbol{\lambda}_L[n] - \mu\boldsymbol{\partial}_L[n]$. Applying this inequality recursively yields $\boldsymbol{\lambda}_L[n+1] \geq \boldsymbol{\lambda}_L[0] - \mu\sum_{l=1}^n \boldsymbol{\partial}_L[l]$. Dividing both sides by n and initializing with $\boldsymbol{\lambda}_L[0] \geq \mathbf{0}$, yields $(n^{-1})\boldsymbol{\lambda}_L[n+1] \geq -\mu(n^{-1})\sum_{l=1}^n \boldsymbol{\partial}_L[l]$. Since the multipliers are bounded (cf. Lemma 1), as $n \to \infty$ the left hand side of the last inequality goes to zero, which implies $0 \geq -\lim_{n\to\infty} n^{-1}\sum_{l=1}^n \boldsymbol{\partial}_L[l]$. Take now into account the following facts: (a) in (3.6) all the constraints are convex (in fact, linear); and, (b) The constraint violation averaged over the last L slots is represented by $-\boldsymbol{\partial}_L[l]$ (this follows from the definitions of $\boldsymbol{\partial}[l]$ and $\boldsymbol{\partial}_L[l]$). Based on these facts, $-\lim_{n\to\infty}(n^{-1})\sum_{l=1}^n \boldsymbol{\partial}_L[l]$ amounts to the time-averaged constraint violation; and therefore, the inequality $0 \geq -\lim_{n\to\infty}(n^{-1})\sum_{l=1}^n \boldsymbol{\partial}_L[l]$ implies that the stochastic solution is feasible, thus proving (i).

If it holds that after a time instant n_0 the multipliers of all active constraints are nonnegative, then the projection operator becomes transparent, which readily implies that $0 = \lim_{n\to\infty}(n - n_0)^{-1}\sum_{l=n_0}^n \boldsymbol{\partial}_L[l] = \lim_{n\to\infty} n^{-1}\sum_{l=1}^n \boldsymbol{\partial}_L[l]$. The latter together with the linearity of the constraints in (3.6) imply that the bounds in (2.14) and (2.14a) associated with the *active constraints* are satisfied with equality. For the problem in (3.6), it is easy to verify that the flow conservation and average power constraints are always active. However, the constraints in (2.6b) may or may not be active, depending on individual user rate requirements and on the system's operating conditions (average channel gain and power budgets).

Proof (ii): To show the bound in (2.15), we need first to prove the following lemma.

Lemma 2: *If the entries of $\boldsymbol{\partial}[n]$ are bounded, it holds that* $\lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^n \mathbb{E}\left[\boldsymbol{\lambda}_L^T[l]\mathbb{E}\left[\boldsymbol{\partial}[l]\right]\right] \leq \mathcal{O}(\mu)$. *Proof:* We first present the proof for $L = 1$. Based on that, we present the proof for $L = 2$, which can be easily generalized for $L > 2$.

The first step in proving the lemma for $L = 1$ consists of using the non-expansive property of the projection operator $[\cdot]_0^\infty$ in $\boldsymbol{\lambda}_L[n+1] = [\boldsymbol{\lambda}_L[n] - \mu\boldsymbol{\partial}_L[n]]_0^\infty$, to write

$$\|\boldsymbol{\lambda}_L[n+1]\|^2 \leq \|\boldsymbol{\lambda}_L[n]\|^2 + \mu^2\|\boldsymbol{\partial}_L[n]\|^2 - 2\mu\boldsymbol{\lambda}_L^T[n]\boldsymbol{\partial}_L[n]. \tag{2.36}$$

Supposing that the "path" (history) of the system up to instant n, namely $\boldsymbol{\lambda}_L[0], \ldots, \boldsymbol{\lambda}_L[n]$, is given, consider taking expectations on both sides of (2.36) over the future of the path, to obtain

$$\mathbb{E}\left[\|\boldsymbol{\lambda}_L[n+1]\|^2\right] \leq \|\boldsymbol{\lambda}_L[n]\|^2 + \mu^2\mathbb{E}\left[\|\boldsymbol{\partial}_L[n]\|^2\right] - 2\mu\boldsymbol{\lambda}_L^T[n]\mathbb{E}\left[\boldsymbol{\partial}_L[n]\right]. \tag{2.37}$$

Taking now expectations over any possible path $\boldsymbol{\lambda}_L[0], \ldots, \boldsymbol{\lambda}_L[n]$, yields

$$\mathbb{E}\left[\|\boldsymbol{\lambda}_L[n+1]\|^2\right] \leq \mathbb{E}\left[\|\boldsymbol{\lambda}_L[n]\|^2\right] + \mu^2\mathbb{E}\left[\|\boldsymbol{\partial}_L[n]\|^2\right] - 2\mu\mathbb{E}\left[\boldsymbol{\lambda}_L^T[n]\mathbb{E}\left[\boldsymbol{\partial}_L[n]\right]\right]. \tag{2.38}$$

Successively applying the last inequality and with $\Delta_L^{\max}$ upper bounding $\mathbb{E}[\|\boldsymbol{\partial}_L[n]\|^2]$, we arrive at

$$\mathbb{E}[\|\boldsymbol{\lambda}_L[n+1]\|^2] \leq \mathbb{E}[\|\boldsymbol{\lambda}_L[0]\|^2] + \mu^2\sum_{l=1}^n \Delta_L^{\max} - 2\mu\sum_{l=1}^n \mathbb{E}\left[\boldsymbol{\lambda}_L^T[l]\mathbb{E}\left[\boldsymbol{\partial}_L[l]\right]\right]. \tag{2.39}$$

Dividing by n and rearranging terms, yields

$$\frac{2\mu}{n}\sum_{l=1}^{n}\mathbb{E}\left[\boldsymbol{\lambda}_L^T[l]\mathbb{E}\left[\boldsymbol{\partial}_L[l]\right]\right] \leq \frac{\mu^2}{n}\sum_{l=1}^{n}\Delta_L^{\max} + \frac{1}{n}\left(\mathbb{E}[\|\boldsymbol{\lambda}_L[0]\|^2] - \mathbb{E}[\|\boldsymbol{\lambda}_L[n+1]\|^2]\right). \tag{2.40}$$

Since the multipliers are bounded (cf. Lemma 1), as $n \to \infty$ the last inequality reduces to

$$\lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n}\mathbb{E}\left[\boldsymbol{\lambda}_L^T[l]\mathbb{E}\left[\boldsymbol{\partial}_L[l]\right]\right] \leq \frac{\Delta_L^{\max}}{2}\mu. \tag{2.41}$$

For $L = 1$, it holds by definition that $\boldsymbol{\partial}_L[l] = \boldsymbol{\partial}[l]$. Substituting $\boldsymbol{\partial}_L[l] = \boldsymbol{\partial}[l]$ into (2.41), the bound in Lemma 2 follows readily.

The proof for $L = 2$ will rely on the definition of $\tilde{\boldsymbol{\partial}}_L$, and the bound in (2.41), which holds regardless of the value of L. Let us rewrite the left hand side of (2.41) as follows[5]

$$\lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n}\mathbb{E}\left[\boldsymbol{\lambda}_L^T[l]\mathbb{E}\left[\boldsymbol{\partial}_L[l]\right]\right] = \lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n}\mathbb{E}\left[\boldsymbol{\lambda}_L^T[l]\left(\frac{1}{2}\mathbb{E}\left[\boldsymbol{\partial}[l-1]\right] + \frac{1}{2}\mathbb{E}\left[\boldsymbol{\partial}[l]\right]\right)\right]$$

$$\begin{aligned}
&\overset{(a)}{=} \lim_{n\to\infty}\frac{1}{n}\sum_{l=2}^{n+1}\mathbb{E}\left[\frac{1}{2}\left(\boldsymbol{\lambda}_L[l-1] + \boldsymbol{\lambda}_L[l]\right)^T\mathbb{E}\left[\boldsymbol{\partial}[l-1]\right]\right] \\
&\overset{(b)}{=} \lim_{n\to\infty}\frac{2}{n+1}\sum_{l=2}^{n+1}\mathbb{E}\Big[\frac{1}{2}\left(\boldsymbol{\lambda}_L[l-1] + \boldsymbol{\lambda}_L[l-1] - \mu\tilde{\boldsymbol{\partial}}_L[l-1]\right)^T\mathbb{E}\left[\boldsymbol{\partial}[l-1]\right]\Big] \\
&\overset{(c)}{=} \lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n}\mathbb{E}\left[\boldsymbol{\lambda}_L^T[l-1]\mathbb{E}\left[\boldsymbol{\partial}[l-1]\right]\right] - \lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n}\mu\frac{1}{2}\mathbb{E}\left[\tilde{\boldsymbol{\partial}}_L^T[l-1]\mathbb{E}\left[\boldsymbol{\partial}[l-1]\right]\right]
\end{aligned} \tag{2.42}$$

where in (a) we have rearranged terms; in (b) we have used that $\boldsymbol{\lambda}_L[l] = \boldsymbol{\lambda}_L[l-1] - \mu\tilde{\boldsymbol{\partial}}_L[l-1]$ (cf. the definition of $\tilde{\boldsymbol{\partial}}_L[n]$); and in (c) we have shifted the summation index. Substituting (2.42) into (2.41), yields

$$\lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n}\mathbb{E}\left[\boldsymbol{\lambda}_L^T[l-1]\mathbb{E}\left[\boldsymbol{\partial}[l-1]\right]\right] - \lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n}\mu\frac{1}{2}\mathbb{E}\left[\tilde{\boldsymbol{\partial}}_L^T[l-1]\mathbb{E}\left[\boldsymbol{\partial}[l-1]\right]\right] \leq \frac{\Delta_L^{\max}}{2}\mu. \tag{2.43}$$

Since we have assumed that the multiplier updates are bounded, $|\mathbb{E}[\tilde{\boldsymbol{\partial}}_L^T[l-1]\mathbb{E}[\boldsymbol{\partial}[l-1]]]|$ is bounded by say $\tilde{\Delta}^{\max}$. Upon substituting this bound into (2.43) and rearranging terms, we have $\lim_{n\to\infty}(n^{-1})\sum_{l=1}^{n}\mathbb{E}\left[\boldsymbol{\lambda}_L^T[l-1]\mathbb{E}\left[\boldsymbol{\partial}[l-1]\right]\right] \leq (\Delta_L^{\max}/2 + \tilde{\Delta}^{\max}/2)\mu$, which completes the proof of Lemma 2. Note that although the lemma assumed that the entries of $\boldsymbol{\partial}[n]$ are bounded, the proof relied on the fact that second-order statistics of $\boldsymbol{\partial}[n]$ are bounded. Nevertheless, the assumption of the entries of $\boldsymbol{\partial}[n]$ being bounded does not entail a loss of generality from a practical perspective (arrival rates and the transmit powers and rates are always bounded in real systems), and therefore it was assumed through.

Having proved Lemma 2, we are ready to establish the bound in (2.15). Taking expectations

[5]To simplify the proof, we will not take into account that the window averaging operator renders the definition of $\boldsymbol{\partial}_L[l]$ different for $l = 1$. As already explained in footnote 2, such a difference is not relevant when one looks at the long-term average and takes $n \to \infty$.

on the left hand side of (2.15), we have

$$\begin{aligned}
&\sum_{m=1}^{M} U_m\left(\frac{1}{n}\sum_{l=1}^{n}\mathbb{E}[a_m^*(\boldsymbol{\lambda}_L[l])]\right)\\
&\overset{(a)}{=} \sum_{m=1}^{M} U_m\left(\frac{1}{n}\sum_{l=1}^{n}\mathbb{E}[\bar{a}_m^*(\boldsymbol{\lambda}_L[l])]\right)\\
&\overset{(b)}{\geq} \frac{1}{n}\sum_{l=1}^{n}\mathbb{E}\left[\sum_{m=1}^{M} U_m(\bar{a}_m^*(\boldsymbol{\lambda}_L[l]))\right]\\
&\overset{(c)}{=} \frac{1}{n}\sum_{l=1}^{n}\mathbb{E}\left[\left(\sum_{m=1}^{M} U_m(\bar{a}_m^*(\boldsymbol{\lambda}_L[l]))\right)+\boldsymbol{\lambda}_L^T[l]\mathbb{E}\left[\boldsymbol{\partial}[l]\right]\right]-\frac{1}{n}\sum_{l=1}^{n}\mathbb{E}\left[\boldsymbol{\lambda}_L^T[l]\mathbb{E}\left[\boldsymbol{\partial}[l]\right]\right]\\
&\overset{(d)}{=} \frac{1}{n}\sum_{l=1}^{n}\mathbb{E}\left[D\left(\boldsymbol{\lambda}_L[l]\right)\right]-\frac{1}{n}\sum_{l=1}^{n}\mathbb{E}\left[\boldsymbol{\lambda}_L^T[l]\mathbb{E}\left[\boldsymbol{\partial}[l]\right]\right]\\
&\overset{(e)}{\geq} D(\boldsymbol{\lambda}^*)-\frac{1}{n}\sum_{l=1}^{n}\mathbb{E}\left[\boldsymbol{\lambda}_L^T[l]\mathbb{E}\left[\boldsymbol{\partial}[l]\right]\right] \qquad (2.44)
\end{aligned}$$

where (a) is due to the definition of $a_m^*((\boldsymbol{\lambda}_L[l]))$; (b) holds because $U_m(\cdot)$ are concave; (c) follows by adding and subtracting the same term; (d) is due to the definition of the dual function [Ber99, Ch. 5]); and (e) relies on the fact that $D(\boldsymbol{\lambda}_L[l]) \geq D(\boldsymbol{\lambda}^*)$, which holds because the dual function is minimized at $\boldsymbol{\lambda}^*$.

Since the problem in (3.6) is convex and feasible, it has zero duality gap, and thus $D(\boldsymbol{\lambda}^*) = \mathsf{P}^*$. After substituting $D(\boldsymbol{\lambda}^*) = \mathsf{P}^*$ into (2.44) and letting $n \to \infty$, it holds that

$$\sum_{m} U_m\left(\lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n}\mathbb{E}[a_m^*(\boldsymbol{\lambda}_L[l])]\right) \geq \mathsf{P}^* - \lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n}\mathbb{E}\left[\boldsymbol{\lambda}_L^T[l]\mathbb{E}\left[\boldsymbol{\partial}[l]\right]\right]. \qquad (2.45)$$

Using the bound provided in Lemma 2, the statement in (ii) follows.

Appendix B.2: Proof of Proposition 3

The sketch of the proof proceeds in two steps. First, we show that the averaged iterates of the Lagrange multipliers converge to a neighborhood of the optimum ones. With $\bar{\boldsymbol{\lambda}}_L := \lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n}\mathbb{E}[\boldsymbol{\lambda}_L[l]]$, this will be accomplished by capitalizing on the fact that the dual function $D(\boldsymbol{\lambda})$ is continuous and $D\left(\bar{\boldsymbol{\lambda}}_L\right) - D(\boldsymbol{\lambda}^*) \leq \delta_D(\mu)$, where $\delta_D(\mu)$ is proportional to μ. This will be shown in Lemma 3 which is presented at the end of this appendix. Since the bound is proportional to μ, $\bar{\boldsymbol{\lambda}}_L$ belongs to the (level) set $\mathcal{B}_\lambda := \{\boldsymbol{\lambda} : D(\boldsymbol{\lambda}) - D(\boldsymbol{\lambda}^*) \leq \delta_D(\mu)\}$. Since $D(\boldsymbol{\lambda})$ is continuous, the size of $\mathcal{B}_\lambda$ shrinks[6] as $\mu \to 0$, and therefore $\|\bar{\boldsymbol{\lambda}}_L - \boldsymbol{\lambda}^*\| \to 0$ as $\mu \to 0$. The second step consists in using Proposition 2 to write $\rho_{m,L}[n] = \mu W(q_m[n], L) + \delta_m(n_0)$. Using the first step, it then follows that $|\mu \lim_{v\to\infty} v^{-1}\sum_{n=1}^{v}\mathbb{E}[W(q_m[n], L)] - \rho_{m,L}^*| \to 0$ as $\mu \to 0$. Swapping the order of $\mathbb{E}[\cdot]$ and $W(\cdot)$, we deduce that $|\mu \lim_{v\to\infty} v^{-1}\sum_{n=1}^{v}\mathbb{E}[q_m[n]] - \rho_{m,L}^*| \to 0$ as $\mu \to 0$, which is the claim of the proposition.

Lemma 3: *If the entries of $\boldsymbol{\partial}[n]$ are bounded, it holds that $D\left(\bar{\boldsymbol{\lambda}}_L\right) - D(\boldsymbol{\lambda}^*) \leq \mathcal{O}(\mu)$. Proof:* For brevity, we consider $L = 2$, but the proof can be easily extended for $L > 2$. Taking into account

[6] Expressions (or bounds) for $\delta_D(\mu)$ and the size of $\mathcal{B}_\lambda$ (thus for the value of $\|\bar{\boldsymbol{\lambda}}_L - \boldsymbol{\lambda}^*\|$) can be derived; e.g. for $L = 2$, $\delta_D(\mu)$ is expressed as in (2.53). Regarding the size of $\mathcal{B}_\lambda$, one can capitalize on the expression of $D(\boldsymbol{\lambda})$ associated with (3.6). Since a closed-form expression for $D(\boldsymbol{\lambda})$ may be difficult to obtain, bounds on $D(\boldsymbol{\lambda})$ (which in general are easier to find) can be used to bound the size of $\mathcal{B}_\lambda$. On the other hand, generic bounds on $\mathcal{B}_\lambda$ which do no exploit the specific form of (3.6), but only the fact that $D(\boldsymbol{\lambda})$ is a dual function can also be found along the lines of, e.g., [Ned09].

that $\boldsymbol{\lambda}_L[n+1] = [\boldsymbol{\lambda}_L[n] - \mu\boldsymbol{\partial}_L[n]]_0^\infty$, it follows that

$$\|\boldsymbol{\lambda}_L[n+1] - \boldsymbol{\lambda}^*\|^2 \quad \leq \quad \|\boldsymbol{\lambda}_L[n] - \boldsymbol{\lambda}^*\|^2 + \mu^2\|\boldsymbol{\partial}_L[n]\|^2 - 2\mu\boldsymbol{\partial}_L^T[n](\boldsymbol{\lambda}_L[n] - \boldsymbol{\lambda}^*) \tag{2.46}$$

where the inequality is due to the projection $[\cdot]_0^\infty$ operator. Considering that the path of the system up to instant n is given, taking expectations on both sides of (2.46) yields

$$\mathbb{E}\left[\|\boldsymbol{\lambda}_L[n+1] - \boldsymbol{\lambda}^*\|^2\right] \leq \|\boldsymbol{\lambda}_L[n] - \boldsymbol{\lambda}^*\|^2 + \mu^2\mathbb{E}\left[\|\boldsymbol{\partial}_L[n]\|^2\right] - 2\mu\mathbb{E}\left[\boldsymbol{\partial}_L^T[n]\right](\boldsymbol{\lambda}_L[n] - \boldsymbol{\lambda}^*) \tag{2.47}$$

For $L = 2$, we have that $\boldsymbol{\partial}_L[n] = (1/2)(\boldsymbol{\partial}[n] + \boldsymbol{\partial}[n-1])$, where as in Lemma 2, we ignore that the equality does not hold for $n = 1$ (cf. footnote 5). Based on this, the last term in (2.47) can be rewritten as

$$\begin{aligned}\mathbb{E}\left[\boldsymbol{\partial}_L^T[n]\right](\boldsymbol{\lambda}_L[n] - \boldsymbol{\lambda}^*) =& \frac{1}{2}\mathbb{E}\left[\boldsymbol{\partial}^T[n]\right](\boldsymbol{\lambda}_L[n] - \boldsymbol{\lambda}^*) \\ &+ \frac{1}{2}\mathbb{E}\left[\boldsymbol{\partial}^T[n-1]\right]\left(\boldsymbol{\lambda}_L[n-1] - \mu\tilde{\boldsymbol{\partial}}_L[n-1] - \boldsymbol{\lambda}^*\right).\end{aligned} \tag{2.48}$$

Using the fact that $\boldsymbol{\partial}[n]$ and $\boldsymbol{\partial}[n-1]$ are subgradients of the dual function at points $\boldsymbol{\lambda}_L[n]$ and $\boldsymbol{\lambda}_L[n-1]$, respectively, (2.48) can be bounded as

$$\begin{aligned}\mathbb{E}\left[\boldsymbol{\partial}_L^T[n]\right](\boldsymbol{\lambda}_L[n] - \boldsymbol{\lambda}^*) \leq& \frac{1}{2}\left(D(\boldsymbol{\lambda}[n]) - D(\boldsymbol{\lambda}^*)\right) \\ &+ \frac{1}{2}\left(D(\boldsymbol{\lambda}[n-1]) - D(\boldsymbol{\lambda}^*)\right) - \frac{\mu}{2}\mathbb{E}\left[\boldsymbol{\partial}^T[n-1]\right]\tilde{\boldsymbol{\partial}}[n-1].\end{aligned} \tag{2.49}$$

Substituting (2.49) into (2.47) yields

$$\begin{aligned}\mathbb{E}\left[\|\boldsymbol{\lambda}_L[n+1] - \boldsymbol{\lambda}^*\|^2\right] \leq \|\boldsymbol{\lambda}_L[n] - \boldsymbol{\lambda}^*\|^2 + \mu^2\mathbb{E}\left[\|\boldsymbol{\partial}_L[n]\|^2\right] \\ -2\mu\frac{1}{2}\left(D(\boldsymbol{\lambda}[n]) - D(\boldsymbol{\lambda}^*)\right) - 2\mu\frac{1}{2}\left(D(\boldsymbol{\lambda}[n-1]) - D(\boldsymbol{\lambda}^*)\right) + \mu^2\mathbb{E}\left[\boldsymbol{\partial}^T[n-1]\right]\tilde{\boldsymbol{\partial}}[n-1].\end{aligned}$$

Taking expectations in (2.9) over any possible path $\boldsymbol{\lambda}_L[0], \ldots, \boldsymbol{\lambda}_L[n]$, and with $\Delta_L^{\max}$ and $\tilde{\Delta}_L^{\max}$ denoting upper bounds on $\mathbb{E}[\|\boldsymbol{\partial}_L[n]\|^2]$ and $|\mathbb{E}[\tilde{\boldsymbol{\partial}}_L^T[n]\boldsymbol{\partial}_L[n]]|$, respectively, it follows that

$$\begin{aligned}\mathbb{E}\left[\|\boldsymbol{\lambda}_L[n+1] - \boldsymbol{\lambda}^*\|^2\right] \leq \mathbb{E}\left[\|\boldsymbol{\lambda}_L[n] - \boldsymbol{\lambda}^*\|^2\right] + \mu^2\Delta_L^{\max} \\ -\frac{\mu}{2}\mathbb{E}\left[D(\boldsymbol{\lambda}[n]) - D(\boldsymbol{\lambda}^*)\right] - \frac{\mu}{2}\mathbb{E}\left[D(\boldsymbol{\lambda}[n-1]) - D(\boldsymbol{\lambda}^*)\right] + \mu^2\tilde{\Delta}_L^{\max}.\end{aligned} \tag{2.50}$$

Applying successively (2.50), yields

$$\begin{aligned}\mathbb{E}[\|\boldsymbol{\lambda}_L[n+1] - \boldsymbol{\lambda}^*\|^2] \leq \mathbb{E}[\|\boldsymbol{\lambda}_L[0] - \boldsymbol{\lambda}^*\|^2] + \mu^2 \\ \sum_{l=1}^{n}(\Delta_L^{\max} + \tilde{\Delta}_L^{\max}) - 2\mu\sum_{l=1}^{n}\left(\mathbb{E}[D(\boldsymbol{\lambda}_L[l])] - D(\boldsymbol{\lambda}^*)\right).\end{aligned} \tag{2.51}$$

Dividing by n and rearranging terms, we arrive at

$$\begin{aligned}\frac{\mu^2}{n}\sum_{l=1}^{n}(\Delta^{\max} + \tilde{\Delta}^{\max}) \geq \frac{2\mu}{n}\sum_{l=1}^{n}\left(\mathbb{E}[D(\boldsymbol{\lambda}_L[l])] - D(\boldsymbol{\lambda}^*)\right) \\ +\frac{1}{n-1}\left(\mathbb{E}[\|\boldsymbol{\lambda}_L[n+1] - \boldsymbol{\lambda}^*\|^2] - \mathbb{E}[\|\boldsymbol{\lambda}_L[0] - \boldsymbol{\lambda}^*\|^2]\right).\end{aligned} \tag{2.52}$$

Since the multipliers are bounded (cf. Lemma 1), as $n \to \infty$ the last inequality yields

$$\begin{aligned}\frac{\mu(\Delta_L^{\max} + \tilde{\Delta}_L^{\max})}{2} &\geq \lim_{n\to\infty} \frac{1}{n} \sum_{l=1}^{n} (\mathbb{E}[D(\boldsymbol{\lambda}_L[l])]) - D(\boldsymbol{\lambda}^*) \\ &\geq D\left(\lim_{n\to\infty} \frac{1}{n} \sum_{l=1}^{n} \mathbb{E}[\boldsymbol{\lambda}_L[l]]\right) - D(\boldsymbol{\lambda}^*) \end{aligned} \tag{2.53}$$

where for the last inequality we have used that the dual function is convex. This concludes the proof.

Chapter 3

Resource Allocation for CRs under Probability-of-Interference Constraints

Efficient design of CRs calls for SUs implementing adaptive RA schemes that exploit knowledge of the channel state information (CSI), while at the same time limiting interference to PUs. This chapter introduces stochastic RA algorithms for both interweave and underlay CR paradigms. The algorithms are designed to maximize the weighted sum-rate of orthogonally transmitting SUs under average-power and probabilistic interference constraints. The latter are formulated either as short- or as long-term constraints, and guarantee that the probability of secondary transmissions interfering with primary receivers stays below a certain pre-specified level. The optimal schemes leverage CSI of the primary and secondary networks, as well as the Lagrange multipliers associated with the constraints.

3.1 Introduction

As CRs aim at controlling interference while leveraging favorable link conditions, knowledge of the CR-to-PU and CR-to-CR channels acquired during the sensing phase is instrumental. Based on this, CRs adapt available resources, namely power, rate, and scheduling coefficients, to the intended channels. The merits of exploiting statistical or instantaneous CSI for adaptive RA are well documented in wireless networking literature [Gol05, Ch. 9]. Nevertheless, CRs face the following additional *design challenges* (DC) [Mus09, Jaf07, Gha07, Kan09, Kan09, Urg09, Wan11, Bar11, DL11]:

- DC1) Extra constraints are needed to effect interference control;
- DC2) CR volatility may render statistical CSI outdated; and also,
- DC3) Instantaneous CSI of the PU network is difficult or impossible to acquire.

In order to address DC1, existing works limit CR-inflicted interference either through instantaneous (short-term) and average (long-term) transmit-power constraints [Kan09, Zha09, Zha10, Bar11]; or, by controlling the probability of interfering with PU transmissions, see, e.g., [Urg09, Che08b, Che08a, Wan11, Sol12, Dal11]. For this second case, most works have focused on short-term constraints, which are relatively easier to handle. Stochastic RA (RA) approaches [Mar09, Wan11], offer viable means to deal with DC2. As with general wireless networks, dual stochastic algorithms are particularly attractive because they are computationally simple, do not require knowledge of channel statistics, and exhibit robustness to channel variations; see

[Wan07, Rib10a] and references in [Mar09, Wan11]. Regarding DC3, most prior CR works consider noisy or quantized CSI [Mus09, Mar09, Sur10, He11]; a few consider *outdated* CSI for CRs [Sur10, Mus09, Che08a]; and very few incorporate mechanisms to *predict* the actual CSI [Che08b, Bar11].

The goal of the present chapter is to develop stochastic RA algorithms for both interweave and underlay paradigms that optimize sum-rate performance of a CR network, limit the *probability* of interfering with PUs (both short-term and *long-term* limits are investigated), and jointly account for *outdated* and noisy CSI. Probabilistic long-term interference constraints are adopted not only because they lead to improved performance, but also because uncertain information on the CR-to-PU channels renders short-term interference constraints infeasible (if the constraint has to hold with probability one) or grossly suboptimal (if the constraint holds probabilistically). Instantaneous CSI of the CR-to-CR links is assumed perfect, while that of CR-to-PU channels can be noisy and outdated. A simple first-order Markov model with additive white noise is used to capture such imperfections, but more complex models can be afforded too. Such models enable channel prediction and correction to track the CR-to-PU changing CSI, which is utilized by per-band orthogonal CR transmissions to adapt their power and rate loadings. The RA schemes are obtained as the solution of a weighted sum-average rate maximization subject to maximum "average power" and "probability of interference" constraints that come in two flavors: a short-term constraint ensuring that the probability of interference is kept below a pre-specified limit per time slot; and a novel long-term constraint guaranteeing the same for a fraction of time slots. Even though not all formulations are convex, it turns out that for all of them the duality gap is zero, meaning that the Langrangian relaxation is always optimal. Additionally, the operating conditions enable separation in the dual domain across users and frequency bands, which allows for optimal solvers with considerably reduced complexity. In all cases, the optimal RA scheme turns out to be a function of the *instantaneous* CSI of the CR-to-CR links, the (possibly outdated and noisy) CR-to-PU channels, and the optimum Lagrange multipliers, obtained via simple stochastic iterations that are robust to nonstationarities, and can even learn varying CSI on-the-fly – a highly desirable attribute for CR networks [Hay05, Mar09]. Extensions to scenarios with more than one CR network (each with several users) are of interest, but go beyond the scope of this thesis.

The rest of the chapter is organized as follows. Section II presents the CSI model, means to account for CSI imperfections, and pertinent operating conditions. A simplified adaptive RA optimization problem, devoid of interference constraints, is formulated and solved in Section III. Incorporation of various interference constraints and design of the corresponding algorithms are the subjects of Section IV. Section V outlines the low-complexity stochastic iterations needed to estimate the multipliers. Numerical examples and conclusions in Sections VI and VII wrap up this chapter.

3.2 Modeling

Consider a CR network of M SUs (indexed by m) transmitting opportunistically over K different frequency bands (indexed by k). For simplicity, suppose that: i) each band has identical bandwidth, and is licensed to a different PU; and ii) the CR network has a network controller (NC), which collects the CSI needed for channel-adaptive RA. Extensions to scenarios where those assumptions do not hold can be handled with a moderate increase in complexity.

3.2.1 Channel state information

Intuitively speaking, CSI in adaptive wireless systems entails channel-related information that must be: i) available to all users in the system; and, ii) relevant from an RA perspective. A key issue with CR systems is that CSI is heterogeneous, meaning that it is different for primary and

secondary networks. The reason is twofold. First, CSI availability for links involving PU and/or CR users is different [cf. i)]. Second, the impact CSI has on the design of RA is different [cf. ii)]. The CSI for CR-to-CR links will be assumed stationary and perfectly known; that is, at every instant, the instantaneous gain of SU links will be deterministically available. For notational purposes, the channel's instantaneous power gain between the mth secondary transmitter and its intended receiver over the kth frequency band at instant n is denoted by $h_{k,2}^m[n]$. Subscript "2" is used to emphasize that the channel pertains to *secondary* transceivers. If PU transmitters are located far away from SU receivers, $h_{k,2}^m[n]$ represents the squared magnitude of the instantaneous fading coefficient divided by the noise power in the kth band. If this is not the case, $h_{k,2}^m[n]$ represents the squared magnitude of the instantaneous fading coefficient divided by the sum of the *noise* power plus the instantaneous *interference* power caused by the kth primary transmitter.

Regarding the CSI corresponding to the PU network, it will not be always assumed perfectly known; e.g., because not all frequency bands are sensed at every time instant. As a result, knowledge of the primary CSI will be probabilistic and time variant. This assumption is well suited for scenarios where sensing the PU network state costs much more than sensing the state of the CR links; e.g., because PUs are too many, or they are possibly located far away from the CRs, or they are simply not willing to collaborate. The CSI model adopted by the NC for the PU network is different for interweave and underlay settings. Each of the cases is described in detail next.

Perfect and imperfect primary CSI in interweave networks

In the interweave setup, the NC *only needs to know* whether each frequency band is occupied or not. To capture this occupancy, let the Boolean variable a_k represent the *activity* of the PU network on the kth band, so that $a_k[n] = 1$ if at instant n the kth PU is active, and zero otherwise. Only the 2×1 belief vector $\mathbf{f}_{\mathbf{a}k}[n] := [\Pr\{a_k[n] = 0\}, \Pr\{a_k[n] = 1\}]^T$ is available, where the probability mass of $a_k[n]$ is based on the history of the system up to n. The belief can be estimated either beforehand or in real time. Next, an example of imperfect CSI in the PU network is considered along with means of estimating the corresponding belief vector.

Let $s_k[n]$ denote a Boolean variable which equals one if the kth band is sensed at instant n, and zero otherwise. Moreover, let $\tilde{a}_k[n]$ be the (perhaps noisy) measurement of $a_k[n]$ obtained at instant n, if $s_k[n] = 1$. Two main types of imperfect CSI are: i) outdated CSI (for the instants n when $s_k[n] = 0$); and ii) noisy CSI (due to errors in the sensing process that render $a_k[n] \neq \tilde{a}_k[n]$). To cope with outdated CSI, a model is needed to capture the dynamics of $a_k[n]$ across time which, for simplicity, are assumed here to follow a first-order Markov process [Bar11, Che08b]; see, e.g., [Zha11] for alternative models. Define the transition probability matrix $\mathbf{Q}$ with (i, j)th entry $Q_{ij} := \Pr\{a_k[n] = i \,|a_k[n-1] = j\}$, for $i, j = 0, 1$. In order to account for sensing errors, consider further the probabilities of miss detection and false alarm, namely $P_{MD} := \Pr\{\tilde{a}_k[n] = 0 \,|a_k[n] = 1\}$ and $P_{FA} := \Pr\{\tilde{a}_k[n] = 1 \,|a_k[n] = 0\}$; and use them to form the 2×1 vectors $\mathbf{q_1} := [1 - P_{FA},\ P_{MD}]^T$ and $\mathbf{q_0} := [P_{FA},\ 1 - P_{MD}]^T$.

Clearly, the CSI measurements are the observed states of a hidden Markov model (HMM), so that recursive Bayesian estimation can be implemented to obtain the instantaneous belief (posterior probability mass function of the unobserved states). In particular, the belief $\mathbf{f}_{\mathbf{a_k}}[n]$ is updated as follows:

- If $s_k[n] = 0$, then $\mathbf{f}_{\mathbf{a_k}}[n] = \mathbf{Q}\, \mathbf{f}_{\mathbf{a_k}}[n-1]$.
- If $s_k[n] = 1$ and $\tilde{a}_k[n] = 0$, then predict the belief vector as $\hat{\mathbf{f}}_{\mathbf{a_k}}[n] := \mathbf{Q}\, \mathbf{f}_{\mathbf{a_k}}[n-1]$; and using that $\tilde{a}_k[n] = 0$, correct $\hat{\mathbf{f}}_{\mathbf{a_k}}[n]$ via Bayes' rule to obtain ($[\cdot]_l$ stands for the lth entry of a vector)

$$[\mathbf{f}_{\mathbf{a_k}}[n]]_l = ([\mathbf{q_0}]_l [\hat{\mathbf{f}}_{\mathbf{a_k}}[n]]_l)/(\mathbf{q_0}^T \hat{\mathbf{f}}_{\mathbf{a_k}}[n]) \,. \tag{3.1}$$

- If $s_k[n] = 1$ and $\tilde{a}_k[n] = 1$, predict as before, and subsequently correct to find

$$[\mathbf{f}_{\mathbf{a_k}}[n]]_l = ([\mathbf{q_1}]_l[\hat{\mathbf{f}}_{\mathbf{a_k}}[n]]_l)/(\mathbf{q_1}^T\hat{\mathbf{f}}_{\mathbf{a_k}}[n]) \; . \tag{3.2}$$

Note that the described procedure resembles other recursive Bayesian models, such as the prediction-correction steps of a Kalman filter (only prediction if $s_k[n] = 0$, and prediction followed by correction when $s_k[n] = 1$). Different prediction-correction steps will be required if the model for the sensing error changes, the transition matrix $\mathbf{Q}$ is unknown or, if the dynamics of $a_k[n]$ are modeled differently. To be more specific about the latter, let τ_k denote the time passed between two changes of $a_k[n]$. Experimental studies, see [Zha11] and references therein, have shown that heavy-tailed distributions are proper alternatives to model τ_k (in contrast with Markov occupancy models, which give rise to exponentially distributed τ_k). Both *Pareto* and *lognormal* distributions are investigated in [Zha11]. Clearly, in those cases $a_k[n]$ is no longer Markovian and (3.1)-(3.2) are not optimal any more. However, the joint process $\{a_k[n], t_k[n]\}$, where $t_k[n]$ represents the time passed since the last time the value of $a_k[n]$ changed, can be modeled as Markovian, so that recursive Bayesian estimation can be employed again. These alternatives will be briefly explored through simulations in Section VI.

Perfect and imperfect primary CSI in underlay networks

In the underlay setup, the NC also needs to know the gains of the CR-to-PU channels. This implies that the primary CSI model in this case is different. Specifically, CSI here comprises information about the instantaneous squared fading coefficient between the mth CR and the kth PU divided by the noise power, which is denoted by $h^m_{k,1}$ (subscript "1" is used to emphasize that the link involves primary receivers). Note that $h^m_{k,2}$ accounts for the interference power, while $h^m_{k,1}$ does not. The reason is that while the interfering power generated by the PUs is a state variable, the one generated by the SU is a design variable. Clearly, if this CSI is perfect, then $h^m_{k,1}[n]$ is deterministically known at instant n. If imperfections are present, only the distribution of $h^m_{k,1}[n]$ (conditioned on all previous measurements) is available. The belief state then consists of the cumulative and the pdf denoted by $F_{h^m_{k,1}[n]}(h)$ and $f_{h^m_{k,1}[n]}(h)$, respectively. Depending on the operating conditions, the belief can be known beforehand or estimated over time. As in the interweave setup, the ensuing example highlights CSI imperfections in the underlay scenario, and the corresponding adaptive schemes to estimate the belief vector.

Define a Boolean variable $s^m_k[n]$ taking value 1 if $h^m_{k,1}$ is sensed at instant n, and 0 otherwise. Moreover, let $\tilde{h}^m_{k,1}[n]$ be the (possibly noisy) measurement of $h^m_{k,1}[n]$ obtained if $s^m_k[n] = 1$. Paralleling the previous example, two types of imperfections are possible: i) outdated CSI (for the instants n when $s^m_k[n] = 0$); and ii) noisy CSI (due to errors in the sensing process that cause $\tilde{h}^m_{k,1}[n] \neq h^m_{k,1}[n]$). The time evolution of $h^m_{k,1}[n]$ is assumed Markovian with $q^m_k(h_{new}, h_{old})$ denoting the probability of having $h^m_{k,1}[n+1] = h_{new}$, given that $h^m_{k,1}[n] = h_{old}$. Moreover, let $f^m_k(h, n)$ denote the pdf of $h^m_{k,1}[n] = h$. It then follows that $f^m_k(h, n+1) = \int_{\forall x} q^m_k(h, x) f^m_k(x, n) dx$. Next, in order to account for sensing errors, the following memoryless additive noise model is assumed: $\tilde{h}^m_{k,1}[n] = h^m_{k,1}[n] + v^m_k[n]$, where $v^m_k[n]$ stands for white noise with known pdf $f_{v^m_k}(v)$ independent of $h^m_{k,1}[n]$. This model has been used by the authors in a more advanced setup – interested readers are referred to [LR14a].

With these operating conditions, the observations follow again an HMM. Hence, the belief $f_{h^m_k[n]}(h)$ can be found using recursive Bayes estimates according to the following cases:

- If $s^m_k[n] = 0$, then $f_{h^m_k[n+1]}(h) = \int_{\forall x} q^m_k(h, x) f_{h^m_k[n]}(x) dx$.
- If $s^m_k[n] = 1$, then predict as $\hat{f}_{h^m_k[n+1]}(h) = \int q^m_k(h, x) f_{h^m_k[n]}(x) dx$, and use $\tilde{h}^m_k[n]$ to correct via Bayes' rule as

$$f_{h[n+1]}(h) = \frac{\hat{f}_{h_k^m[n+1]}(h) f_{v_k^m}(h-\tilde{h})}{\int_{\forall x} \hat{f}_{h_k^m[n+1]}(x) f_{v_k^m}(x-\tilde{h}) dx} \,. \tag{3.3}$$

Because in this case the number of unobserved HMM states is infinite (the channel is a continuous variable), the denominator in the update equation (3.3) is an integral. This is in contrast with (3.2), where the denominator was a finite sum, reflecting the fact that in the previous section the number of unobserved states is finite. From a practical perspective there are a few cases where those integrals can be found in closed form (e.g. Gaussian channels). For the remaining cases, an approximate technique (such as grid-based Bayesian estimators or particle filters) should be used.

Before moving to the proposed RA approach, it is worth reiterating the main points so far. The CSI model adopted by the NC is distinct for the primary and secondary networks. The secondary CSI consists of the CR-to-CR link gains, which account for primary interference; whereas the primary CSI is formed either by the PU activity vector alone (interweave setup), or, it is augmented by CR-to-PU channel gains (underlay setup), which do not account for secondary interference. Moreover, secondary CSI is assumed perfectly known, so that information about the instantaneous realization is deterministic; whereas primary CSI is allowed to be uncertain, so that information (belief state) about the instantaneous realization is probabilistic.

3.2.2 Resources at the secondary network

This subsection introduces the design variables to be adapted as a function of the *overall* CSI that is collectively denoted by $\mathbf{h}$. Define further a Boolean scheduling variable w_k^m taking the value 1, if the mth CR is scheduled to transmit over the kth band, and 0 otherwise. When $w_k^m = 1$, let p_k^m denote the instantaneous power transmitted over the kth band by the mth CR. Under bit error rate or capacity constraints, instantaneous rate and power variables are coupled. This rate-power coupling will be represented by the function $C_k^m(h_{k,2}^m, p_k^m)$. It will be assumed throughout that $C_k^m(h_{k,2}^m, \cdot)$ is given by Shannon's capacity formula $\log(1 + h_{k,2}^m p_k^m / \kappa_k^m)$, where κ_k^m represents the SNR-gap that depends on the coding scheme implemented [Gol05]. For systems that implement a relatively small number of adaptive modulation and coding (AMC) modes, the last formula can be replaced with a piecewise linear function combining the rates achieved by the modes (see, e.g., [Mar09], for details).

The secondary network operates in a block-by-block fashion, where the duration of each block corresponds to the coherence time of the fading channel. This way, per time slot n the NC uses the current CSI vector $\mathbf{h}$ to find w_k^m and p_k^m. Since $\mathbf{h}$ depends on n and $\{w_k^m, p_k^m\}$ depend on $\mathbf{h}$, $\{w_k^m, p_k^m\}$ will clearly vary across time. Henceforth, $\mathbf{h}$, $w_k^m(\mathbf{h})$, and $p_k^m(\mathbf{h})$ will be replaced by $\mathbf{h}[n]$, $w_k^m[n]$, and $p_k^m[n]$, whenever time dependence is to be stressed.

For this CR configuration, the goal is to develop adaptive RA algorithms leveraging the instantaneous secondary CSI and the generally uncertain primary CSI to determine which CR should transmit per band, and at what rate and power. An optimization problem will be formulated and solved in the ensuing section, first without interference constraints. Those will be incorporated in Section 3.4.

3.3 The Optimization Problem for Adaptive RA

To formulate the optimization problem associated with the novel RA approach, it is prudent to identify: i) the variables to be optimized, ii) the metric to be optimized, and iii) the constraints that must be satisfied. Section 3.2.2 identified $\{w_k^m, p_k^m\}$ as optimization variables. The metric to be optimized is the CRs' weighted sum-average rate given by $\bar{c} := \sum_{k,m} \mathbb{E}_{\mathbf{h}} \left[\beta^m w_k^m(\mathbf{h}) C_k^m(h_{k,2}^m, p_k^m(\mathbf{h})) \right]$, where $\mathbb{E}_{\mathbf{h}}$ stands for expectation over all CSI realizations, and $\beta^m > 0$ represents a user-dependent

priority coefficient. Note that only the rate of CR user-channel pairs for which $w_k^m(\mathbf{h}) = 1$ participate in forming $\bar{c}$. Other objective functions such as sum-utility rate could be used without changing the basic structure of the solution; see, e.g., [Wan07, Mar11a] for further details. Regarding the constraints, $\{p_k^m\}$ must be obviously nonnegative, while $\{w_k^m\}$ must belong to the set $\{0,1\}$. Moreover, since at most one CR transmits over each band k, it must hold that

$$\sum_k w_k^m(\mathbf{h}) \leq 1, \quad \forall k. \tag{3.4}$$

If the left-hand side (LHS) of (3.4) equals one, then one user accesses the channel (orthogonal access); otherwise, no user transmits either because all CR-to-CR channels are poor, or, because excessive interference is inflicted to the PU. The maximum average (long-term) power the mth CR can transmit is upper bounded; that is,

$$\mathbb{E}_{\mathbf{h}}\left[\sum_k w_k^m(\mathbf{h}) p_k^m(\mathbf{h})\right] \leq \check{p}^m, \quad \forall m. \tag{3.5}$$

Under these considerations, the optimal RA emerges as the solution of the following problem:

$$\bar{c}^\star := \max_{\{w_k^m(\mathbf{h}), p_k^m(\mathbf{h})\}} \sum_k \mathbb{E}_{\mathbf{h}}\left[\beta^m w_k^m(\mathbf{h}) C_k^m(h_{k,2}^m, p_k^m(\mathbf{h}))\right] \tag{3.6a}$$

$$\text{s. to}: (3.4),\ (3.5),\ w_k^m(\mathbf{h}) \in \{0,1\},\ \text{and}\ p_k^m(\mathbf{h}) \geq 0; \tag{3.6b}$$

where dependence of the optimization variables on $\mathbf{h}$ has been made explicit.

3.3.1 Optimal RA without interference constraints

Although the problem in (3.6) is non-convex, it can be trivially transformed (relaxed) into a convex one with identical KKT conditions[1]. In fact, the problem in (3.6) is a weighted sum-rate optimization of an uplink channel with orthogonal access. With π^m denoting the Lagrange multiplier associated with the constraint in (3.5), it has been shown that the solution of such a problem is (see, e.g., [Mar11c])

$$\varphi_k^m(p_k^m[n]) := \beta^m C_k^m(h_{k,2}^m[n], p_k^m[n]) - \pi^m[n] p_k^m[n], \tag{3.7}$$

$$p_k^{m\star}[n] := \left[\arg\max_{p_k^m[n]} \ \varphi_k^m(p_k^m[n])\right]_0^\infty \tag{3.8}$$

$$= \left[\frac{\beta^m}{\pi_m[n]} - \frac{\kappa_k^m}{h_{k,2}^m}\right]_0^\infty \tag{3.9}$$

$$w_k^{m\star}[n] := \mathbb{1}_{\{(m=\arg\max_l \varphi_k^l((p_k^{l\star}[n])))\wedge(\varphi_k^m(p_k^{m\star}[n])>0)\}}. \tag{3.10}$$

Key to understanding the solution of (3.6) is the definition of the functional $\varphi_k^m(\cdot)$ in (3.7). Intuitively, (3.7) can be interpreted as a user-quality indicator where the rate is a reward, the power a cost, and β^m and $\pi^m[n]$ their corresponding prices. Analytically, $\varphi_k^m(x)$ represents the contribution to the *Lagrangian* of (3.6) if the transmit-power is $p_k^m[n] = x$ and $w_k^m[n] = 1$.

Based on the definition of $\varphi_k^m(p_k^m[n])$, equation (3.8) reveals that $p_k^{m\star}[n]$ is found separately for each of the CR user-channel pairs. Similarly, (3.10) shows that finding the optimal scheduling variables $\{w_k^{m\star}[n]\}_{m=1}^M$ per channel k, requires no information from channels other than k. These

[1] There are two sources of non-convexity in (3.6). The first comes from $w_k^m \in \{0,1\}$, but can be relaxed to $w_k^m \in [0,1]$. As w_k^m only appears in linear terms, this relaxed solution coincides with the original one [Mar11c]. The second source corresponds to the monomials $w_k^m p_k^m$ and $w_k^m C_k^m$, for which one can introduce dummy implicit variables $\tilde{p}_k^m := w_k^m p_k^m$ in (3.6), and establish convexity using the properties of the perspective function. The resulting problem yields the same KKT conditions as those of (3.6); it is convex; and can be solved using a dual approach, see, e.g., [Mar11c], for details.

attractive properties hold thanks to the assumed orthogonal access in the secondary network and the definition of the objective in (3.6), both of which render the optimization problem in the dual domain separable across users and channels. Delving into the nuts-and-bolts of the optimal RA, consideration of a logarithmic rate-power function implies that (3.9) follows the well-known waterfilling solution [Gol05]; and (3.10) manifests that the user scheduling is opportunistic (as desired) and greedy (only the user with *highest* quality must be scheduled per band).

Finally, it is worth emphasizing that although traditionally $\pi^m[n]$ is set to a constant value π^{m*}, corresponding to the value that maximizes the dual function associated with (3.6) [Ber03], alternative (stochastic) methods can be used. Such an alternative is attractive especially for the CR setup considered here, and will be explored in Section 3.5.

3.4 Interference constraints

Different interference constraints are considered in this section along with ways the optimal RA approach must be modified in the constrained case. Attention is centered around constraints that limit the *probability of* CR transmitters to *interfere* with PU receivers. Other interference constraints (such as limiting the average interference power, or the rate loss for the primary network) could also be considered. Note that probabilistic constraints naturally account for CSI imperfections and, depending on their formulation, they can even exploit CSI variability.

When constraints on the probability of interference are included, there are two factors that significantly affect the design of optimum adaptive RA. The first is whether the interference constraints are formulated as instantaneous (short-term) or as average (long-term) constraints. The former require a certain probability of interference to hold for each and *every time instant*, while the latter allow PUs to be interfered at most over a maximum *fraction of time*. Clearly, instantaneous constraints are more restrictive than their average counterparts, which can exploit the so-called "cognitive diversity" of the primary CSI [Zha09, Zha10]. As a result, the total rate transmitted by the secondary users will be higher in the latter case. On the other hand, optimization problems under instantaneous interference constraints are easier to solve because such constraints are amenable to simplification. Differently, average interference constraints cannot be easily simplified, and a dual approach is often invoked to deal with them. The second factor is whether an interweave or an underlay setup is in operation. The definition of interference in each setup is different. In fact, it will be shown that underlay formulations will render the problem non-convex and thus, challenge the development of an efficient solver able to achieve optimal performance. Remarkably, for the formulations in this chapter, the optimization problem for the underlay setup exhibits zero duality gap, and the optimal solution can still be found with a moderate increase in terms of computational complexity.

Different formulations are considered for the probability of interference constraints because they will give rise to novel optimal RA schemes. But also because, upon comparing the different solutions, it will be possible to understand the differences among the considered alternatives, both theoretically and from a performance perspective. The first formulation considered is the one involving instantaneous interference constraints for both interweave and underlay setups. Subsequently, the interweave and underlay setups will be investigated separately under average interference constraints. In all formulations, the schemes will be designed assuming imperfect CSI and then specialized for the case of perfect CSI. To simplify derivations, schemes for the underlay setup will be developed assuming that the PU is always active. The minor modifications required when this assumption does not hold are discussed in the closing remark of Section 3.4.

3.4.1 Short-term interference constraints

To keep the interference to the primary network under control, a *maximum probability of interference*, call it $\breve{o}_k \in (0,1)$, is placed per band. Since this subsection focuses on short-term (instantaneous) interference constraints, such a limit is enforced $\forall n$.

Interweave networks

In this setup, interference occurs when $a_k[n] = 1$ (kth PU active over the kth band), and $\sum_m w_k^m[n] = 1$ (one CR transmits over the kth band). Then, the constraint on the probability can be formulated as $\Pr\{a_k[n] \sum_m w_k^m[n] = 1 \,|n\} \leq \breve{o}_k \;\forall n$. At time n, the only random quantity in the previous expression is $a_k[n]$. Hence, the constraint can be written as

$$\mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]\sum_m w_k^m[n]=1\}}\right] \leq \breve{o}_k \,. \tag{3.11}$$

Taking into account that $\sum_m w_k^m[n]$ is Boolean and deterministically known at time n, the constraint can be rewritten as $\mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right] \sum_m w_k^m[n] \leq \breve{o}_k$. Clearly, the expectation on the LHS corresponds to the second entry of the belief vector $[\mathbf{f}_{\mathbf{a_k}}[n]]_2$. Thus, $\sum_m w_k^m[n] = 1$ only if $[\mathbf{f}_{\mathbf{a_k}}[n]]_2 \leq \breve{o}_k$. This in turn implies that: i) there is no need to dualize the constraint, and therefore the expression for the link-quality indicator in (3.7) does not change; ii) the power allocation plays no role on the definition of the interference, and hence (3.8) still holds; and iii) to satisfy the interference constraint the optimal scheduling is now

$$w_k^{m*}[n] := \mathbb{1}_{\{[\mathbf{f}_{\mathbf{a_k}}[n]]_2 \leq \breve{o}_k\}} \cdot \mathbb{1}_{\{(\varphi_k^m[n]=\max_l \varphi_k^l[n]) \,\wedge\, (\varphi_k^m[n]>0)\}}. \tag{3.12}$$

In words, the "winner CR" can transmit only if the probability of the channel being occupied is less than $\breve{o}_k$. When the primary CSI is noisy and outdated, such a probability depends on the previous measurements and the accuracy of the sensor [cf. (3.1)-(3.2)]. On the other hand, if the primary CSI is perfect, $[\mathbf{f}_{\mathbf{a_k}}[n]]_2$ is either one or zero, and therefore transmissions can be allowed only if the channel is not occupied, i.e., it holds that $w_k^{m*}[n] := \mathbb{1}_{\{a_k[n]=0\}} \cdot \mathbb{1}_{\{(\varphi_k^m[n]=\max_l \varphi_k^l[n]) \,\wedge\, (\varphi_k^m[n]>0)\}}$.

Underlay networks

In this case, interference occurs when the received power at the PU due to CR transmissions exceeds a threshold Γ_k; i.e., if $w_k^m[n] > 0$ and $p_k^m[n] h_{k,1}^m[n] > \Gamma_k$, the constraint to be satisfied at every time n is $\Pr\{p_k^m[n] h_{k,1}^m[n] > \Gamma_k \,|n\} \leq \breve{o}_k$. Since at time n the only random quantity is now $h_{k,1}^m[n]$, the constraint can be rewritten as

$$E_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{p_k^m[n] h_{k,1}^m[n] > \Gamma_k\}}\right] \leq \breve{o}_k, \quad \forall k \tag{3.13}$$

or equivalently, $E_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{h_{k,1}^m[n] < \Gamma_k / p_k^m[n]\}}\right] \geq 1 - \breve{o}_k$. Using the belief for the primary CSI at time n, it follows that $F_{h_{k,1}^m[n]}(\Gamma_k / p_k^m[n]) \geq 1 - \breve{o}_k$. Upon defining $p_k^{m'}[n]$ as the root of $\breve{o}_k = F_{h_{k,1}^m[n]}(p_k^{m'}[n]/\Gamma_k)$, the last *inequality* amounts to the bound $p_k^m[n] \leq p_k^{m'}[n]$. In words, the interference constraint can be rewritten as a maximum (peak) power constraint.

This is very convenient, because while the original constraint in (3.13) is not convex, the maximum peak power constraint is convex. Since no multiplier is introduced to enforce the constraint, the Lagrangian remains the same, and thus $\varphi_k^m(p_k^m[n])$ is identical to that in (3.7). Similarly, the scheduling does not play a role in defining the interference so that the expression for $w_k^{m*}[n]$ in (3.10) holds true too. On the other hand, the expression for the optimum power in (3.8) needs to

be updated because it has to satisfy the constraint $p_k^m[n] \leq p_k^{m'}[n]$. Such a (box) constraint can be easily handled by a scalar projection, which readily yields

$$p_k^{m*}[n] := \left[\arg\max_{p_k^m[n]} \ \varphi_k^m(p_k^m[n]) \right]_0^{p_k^{m'}[n]}. \tag{3.14}$$

When the CSI is perfect, there is no uncertainty regarding $h_{k,1}^m[n]$; hence, the upper bound on the transmit-power is $p_k^{m'}[n] := h_{k,1}^m[n]/\Gamma_k$, and no interference is inflicted to the PU.

3.4.2 Long-term interference constraints in interweave systems

The previous subsection demonstrated that short-term interference constraints are easy to handle. In fact, for the interweave case the difficulty does not lie in how to satisfy the constraint, which is straightforward, but in estimating the probability of the PU being active. Here, the long-term probability of interfering with PUs is considered for the interweave setup. Since there is no easy way to enforce such a constraint, a dual relaxation will be used instead. It will be argued that regardless of CSI imperfections, the augmented optimization problem is convex and thus exhibits the following two properties: i) it can be tackled optimally using a dual approach, i.e., the duality gap is zero; and, ii) it is efficiently solvable.

Starting with the constraint formulation, recall that limiting the short-term probability of interference in an interweave setup consists in satisfying $\Pr\{\sum_m w_k^m[n]a_k[n] = 1 \ |n\} \leq \check{o}_k$, or equivalently, $E_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]\sum_m w_k^m[n]=1\}}\right] \leq \check{o}_k$ [cf. (3.11)]. In this section, the interest is in a long-term constraint so that all time instants are jointly considered. In this case, $\check{o}_k$ can be viewed as an upper bound on the fraction of time instants for which interference occurs. This implies that the solution needs to satisfy the following condition

$$\mathbb{E}_{\mathbf{h}}\left[\mathbb{1}_{\{a_k \sum_m w_k^m(\mathbf{h})=1\}}\right] \leq \check{o}_k, \quad \forall k. \tag{3.15}$$

Unlike (3.11), the expectation in (3.15) takes into account all CSI realizations. Note also that the LHS of (3.15) represents the joint probability of the PU being active *and* the NC scheduling one CR transmission. If one wants to limit the probability of one CR being active *provided that* the PU is active, then $\check{o}_k$ must be multiplied (re-scaled) by the stationary probability of the kth band being occupied by the corresponding PU.

When (3.15) is incorporated into (3.6), the augmented problem is still convex because: i) (3.15) can be rewritten as $\mathbb{E}_{\mathbf{h}}\left[\sum_m w_k^m(\mathbf{h})\mathbb{1}_{\{a_k=1\}}\right] \leq \check{o}_k$; and ii) the last inequality is convex (in fact linear) with respect to (w.r.t.) the only primary variable involved (i.e., w.r.t. w_k^m). As already mentioned, the approach to deal with the long-term interference constraint is to dualize it. To this end, let θ_k denote the Lagrange multiplier associated with the kth constraint in (3.15). The introduction of a new multiplier implies that the link-quality indicator needs to be redefined as

$$\varphi_k^m(p_k^m[n]) := \beta^m C_k^m(h_{k,2}^m[n], p_k^m[n]) - \pi^m[n]p_k^m[n] - \theta_k[n]\mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right]. \tag{3.16}$$

If the primary CSI is imperfect, then $\mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right] = [\mathbf{f}_{\mathbf{a_k}}[n]]_2$; when perfect, it is simply $a_k[n]$. The only difference between the definitions of the quality indicator in (3.7) and (3.16) is that on top of considering the trade-off between rate and power, (3.16) also penalizes CR transmissions that are likely to cause interference whose "price" is multiplied by the instantaneous (short-term) probability of interference. The structure of the indicator in (3.16) also shows the role of the secondary CSI in the RA (first term in the sum), the role of the primary CSI (third term in the sum), as well as the impact of CSI imperfections (specific expression for $\mathbb{E}_{a_k[n]}[\mathbb{1}_{\{a_k[n]=1\}}]$).

Upon substituting (3.16) into (3.8) and (3.10), the *expressions* for the optimal power in (3.8)

and the optimal scheduling in (3.10) still apply. However, this does not mean that *actual allocation* of resources is the same. While in the previous section transmissions never took place when $\mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right] > \check{o}_k$ [cf. (3.12)], the allocation in (3.16) allows for transmissions when the probability of interfering is high, provided that $\max_m\{\beta^m C_k^m(h_{k,2}^m[n], p_k^{m*}[n]) - \pi^m[n]p_k^{m*}[n]\}_{m=1}^M > \theta_k[n]\mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right]$. In other words, even if the scheduler knows that $a_k[n] = 1$, the secondary network can access the channel if the reward for the winner CR is high enough to exceed the cost of interfering represented by $\theta_k[n]$. Clearly, $\theta_k[n]$ is tuned to enforce that the percentage of interfering transmissions does not exceed the limit set by $\check{o}_k$ (a higher price for interfering means that secondary transmissions will be less frequent). Finally, since the new term in $\varphi_k^m(p_k^m[n])$ does not depend on $p_k^m[n]$, the equivalence between optimum power in (3.8) and the waterfilling interpretation is still valid [cf. (3.9)]. Hence, an important difference between the short-term and the long-term solutions for the interweave paradigm is the way in which scheduling decisions are made. Optimal scheduling for the short-term formulation does not take into account the benefit for the winner SU. Focus is placed first on the PU. Only if the interference caused to the PU is below a threshold, the winner SU can transmit [cf. (3.12) and (3.7)]. Differently, optimal scheduling for the long-term formulation is more flexible and weights both the benefit for the SU and the harm caused to the PU [cf. (3.10) and (3.16)].

3.4.3 Long-term interference constraints in underlay systems

As in Section 3.4.2, the approach to deal with a long-term constraint on the probability of interfering with PUs in the underlay setup, is to dualize it. Regardless of CSI imperfections, the interference constraints here render the optimization problem non-convex. However, the problem at hand has two attractive features: i) since the functions causing non-convexity are averaged across time, existing results can be adapted to show that the duality gap is zero; and, ii) the problem can still be separated in the dual domain, so that minimization of the Lagrangian can be efficiently performed. More details will be given soon.

To formulate the interference constraint, recall that limiting the short-term probability of interference in an underlay setup amounts to bounding $\Pr\{p_k^m[n]h_{k,1}^m[n] > \Gamma_k \,|n\} \leq \check{o}_k$, or equivalently, $\mathbb{E}_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{p_k^m[n]h_{k,1}^m[n]>\Gamma_k\}}\right] \leq \check{o}_k$. For such a long-term bound, all channel realizations (time instants) must be accounted for, along with the CR causing interference. This can be accomplished by writing the constraint as

$$\mathbb{E}_{\mathbf{h}}\left[\sum_m w_k^m(\mathbf{h})\mathbb{1}_{\{p_k^m(\mathbf{h})h_{k,1}^m(\mathbf{h})>\Gamma_k\}}\right] \leq \check{o}_k, \quad \forall k. \tag{3.17}$$

Similar to (3.15), averaging over all $\mathbf{h}$ in (3.17) clearly implies that the constraint need not be satisfied for every CSI realization $\mathbf{h}$, but only on the average.

When (3.17) is incorporated into (3.6), the augmented problem is non-convex and thus challenging to solve. Remarkably, since the functions responsible for the non-convexity are averaged across time, existing results can be leveraged to show that the duality gap is zero. A rigorous proof can be obtained after adapting the results in either [Rib10b] or [Raj11, App. A] for the problem at hand[2]. The fact of having zero-duality gap implies that dual methods can be used to relax the constraints without loss of optimality. However, the (unconstrained) Lagrangian is still non-convex and thus challenging to minimize. Next, the optimal RA for this scenario is developed and shown how it allows for an efficient minimization of the Lagrangian. Let ϑ_k denote the Lagrange multiplier associated with the kth constraint in (3.17). As in the interweave case, introduction of

[2]The basic idea is that non-convexity comes from a constraint of the form $\mathbb{E}_{\mathbf{x}}[g(\mathbf{y},\mathbf{x})]$, where $g(\mathbf{y},\mathbf{x})$ is a non-convex function w.r.t. $\mathbf{y}$, and $\mathbf{x}$ is a random process with infinite support. Here $\mathbf{y}$ is the power; $\mathbf{x}$ is the CSI; and $g(\mathbf{y},\mathbf{x})$ is $\mathbb{E}_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{p_k^m[n]h_{k,1}^m[n]>\Gamma_k\}}\right]$. The interested readers are referred to [Rib10b] and [Raj11, App. A].

a new multiplier modifies the Lagrangian structure, and thus the link-quality indicator has to be modified accordingly as

$$\varphi_k^m(p_k^m[n]) := \beta^m C_k^m(h_{k,2}^m[n], p_k^m[n]) - \pi^m[n] p_k^m[n] - \vartheta_k[n] \mathbb{E}_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{p_k^m[n] h_{k,1}^m[n] > \Gamma_k\}}\right]. \quad (3.18)$$

As its counterpart in (3.16), the quality indicator in (3.18) considers both secondary and primary CSI and trades off rate reward with power and the cost of interfering. Indeed, the only difference between (3.16) and (3.18) is the expression for the instantaneous probability of interfering. Here, $\mathbb{E}_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{p_k^m[n] h_{k,1}^m[n] > \Gamma_k\}}\right]$ corresponds to $1 - F_{h_{k,1}^m[n]}(\Gamma_k / p_k^m[n])$ when CSI is imperfect, and to $\mathbb{1}_{\{p_k^m[n] h_{k,1}^m[n] > \Gamma_k\}}$ when CSI is perfect. As in Section 3.4.2, upon replacing (3.7) with (3.18), the expressions for the optimal power in (3.8) and the optimal scheduling in (3.10) remain the same. However, the equivalence between (3.8) and (3.9) no longer holds. This is because the third term in (3.18) depends on the transmit power and the optimal power in (3.9) is found by optimizing only the two first terms. In fact, the power optimization when (3.18) is substituted into (3.8) is challenging because the third term renders $\varphi_k^m(\cdot)$ non-concave. However, since optimizing $\varphi_k^m(\cdot)$ involves a single (scalar) variable, efficient methods to solve the optimization can be employed. Once $\{p_k^{m*}[n]\}_{m=1}^M$ are obtained, finding $\{w_k^{m*}[n]\}_{m=1}^M$ just requires the evaluation of closed-form expressions [cf. (3.10)]. In other words, because in the dual domain the problem can be separated across users and channels, optimizing the Lagrangian does not require optimizing a non-convex problem over a $2MK$-dimensional space; but instead, MK closed forms and MK *one-dimensional* non-convex problems must be solved. Recall that the factors enabling separability in the dual domain were the orthogonal access adopted by SUs within the CR network, and the definition of the metric to be optimized (summation across users) under the long-term constraints.

Indeed, when CSI is perfect, power optimization is straightforward and proceeds as follows. Let $p_k^{m''}[n] := h_{1,k}^m[n]/\Gamma_k$ be the maximum transmit-power by which interference is avoided; and let $\hat{p}_k^{m*}[n]$ denote the optimal power in (3.9), which ignores the interference constraint. Then, it holds that

$$p_k^{m*}[n] := \begin{cases} \hat{p}_k^{m*}[n] & \text{if } (\hat{p}_k^{m*}[n] < p_k^{m''}[n]) \ \vee \ (\varphi_k^m(\hat{p}_k^{m*}[n]) > \varphi_k^m(p_k^{m'}[n])) \\ p_k^{m''}[n] & \text{otherwise}. \end{cases} \quad (3.19)$$

In words, if the cost of interfering is too high, transmit-power is constrained not to exceed $p_k^{m''}[n]$. However, if the cost of interfering is low enough (or the reward of the CR transmission is high enough), $p_k^{m*}[n]$ is allowed to exceed the upper bound.

When the primary CSI is imperfect, evaluating $F_{h_k^m[n]}$ dominates the complexity of power optimization. Unless $f_{h_k^m[n]}$ (which is the derivative of $F_{h_k^m[n]}$) is monotonic, the optimization is non-convex. However, if the number of stationary points of $f_{h_k^m[n]}$ is small (which holds true for most practical distributions), the number of local optima of $\varphi_k^m(\cdot)$ will be small too. In this case, all of them can be found, and the global optimum can be subsequently selected.

Remark 1 *The schemes for the underlay setup in Sections 3.4.1.2 and 3.4.3 have been developed under the assumption that PUs are always active, meaning that* $a_k[n] = 1$. *If this is not the case, interference only occurs if* $p_k^m[n] h_{k,1}^m[n] > \Gamma_k$ *and* $a_k[n] = 1$. *Assuming that* $h_{k,1}^m[n]$ *and* $a_k[n]$ *are independent, the only modification required is to replace the instantaneous probability of interference* $\mathbb{E}_{h_{k,1}^m[n]}[\mathbb{1}_{\{p_k^m[n] h_{k,1}^m[n] > \Gamma_k\}}]$ *with* $\mathbb{E}_{h_{k,1}^m[n]}[\mathbb{1}_{\{p_k^m[n] h_{k,1}^m[n] > \Gamma_k\}}] \mathbb{E}_{a_k[n]}[\mathbb{1}_{\{a_k[n]=1\}}]$.

3.5 Estimating the optimum Lagrange multipliers

Different methods can be used to estimate $\pi^m[n]$, $\theta_k[n]$, and $\vartheta_k[n]$. Since the duality gap is

zero, one approach is to set $\pi^m[n] = \pi^{m*}$, $\theta_k[n] = \theta_k^*$ and $\vartheta_k[n] = \vartheta_k^*$, where $\{\pi^{m*}, \theta_k^*, \vartheta_k^*\}$ are the values which optimize the dual function associated with (3.6). Clearly, the RA resulting after substituting those values into (3.7)-(3.19) would be the optimal solution for (3.6) [Ber03]. The main limitations of this approach are that: i) $\{\pi^{m*}, \theta_k^*, \vartheta_k^*\}$ need to be found through numerical search[3] which, at every step, requires averaging over all possible states of $\mathbf{h}$ (including channel imperfections); and ii) every time channel statistics or the number of users change, $\{\pi^{m*}, \theta_k^*, \vartheta_k^*\}$ must be recomputed. Recently, alternative approaches that rely on stochastic approximation iterations have been proposed to obtain the multipliers [Mar09, Wan11]. These approaches do not aim at the optimal $\{\pi^{m*}, \theta_k^*, \vartheta_k^*\}$, but estimates that are updated at every time instant, and remain sufficiently close to $\{\pi^{m*}, \theta_k^*, \vartheta_k^*\}$. The main advantages of these approaches, especially for CR settings, are: i) their computational complexity is very low; and, ii) they can cope with non-stationary channels. The latter is very convenient when the PU transmitters are close to the SU receivers. The price paid is that the resulting RA schemes are slightly suboptimal. Specifically, with μ_π, μ_θ and μ_ϑ denoting sufficiently small, constant stepsizes, the following iterations yield the desired multipliers $\forall n$

$$\pi^m[n+1] = \left[\pi^m[n] - \mu_\pi(\check{p}^m - \textstyle\sum_k w_k^{m*}[n] p_k^{m*}[n])\right]_0^\infty \tag{3.20}$$

$$\theta_k[n+1] = \left[\theta_k[n] - \mu_\theta(\check{o}_k - \mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right] \textstyle\sum_m w_k^{m*}[n])\right]_0^\infty \tag{3.21}$$

$$\vartheta_k[n+1] = \left[\vartheta_k[n] - \mu_\vartheta(\check{o}_k - \mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right] \mathbb{E}_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{p_k^m[n] h_{k,1}^m[n] > \Gamma_k\}}\right] \textstyle\sum_m w_k^{m*}[n]\right]_0^\infty. \tag{3.22}$$

Recall that the expression for the instantaneous probability of interference in (3.21) and (3.22) is different for the cases of perfect and imperfect CSI. From an optimization point of view, the updates in (3.20)-(3.22) form an *unbiased* stochastic subgradient of the dual function of (3.6); see [Ber03]. Using also that the updates in (3.20)-(3.22) are *bounded*, it can be shown that the sample average of the stochastic RA: i) is feasible; and, ii) incurs minimal performance loss relative to the optimal solution of (3.6). Rigorously stated, define $\mu := \max\{\mu_\pi, \mu_\theta, \mu_\vartheta\}$; $\bar{p}^m[n] := \frac{1}{n}\sum_{l=1}^n \sum_k w_k^{m*}[l] p_k^{m*}[l]$; $\bar{c}[n] := \frac{1}{n}\sum_{l=1}^n \sum_{k,m} \beta^m w_k^{m*}[l] C_k^m$ $(h_{k,2}^m[l], p_k^{m*}[l])$; and $\bar{o}_k[n] := \frac{1}{n}\sum_{l=1}^n \sum_m w_k^{m*}[l] \mathbb{1}_{\{a_k[l]=1\}}$ (interweave) or $\bar{o}_k[n] := \frac{1}{n}\sum_{l=1}^n \sum_m w_k^{m*}[l] \mathbb{1}_{\{a_k[n]=1\}}\, \mathbb{1}_{\{p_k^{m*}[l] h_{k,1}^m[l] > \Gamma_k\}}$ (underlay). It then holds with probability one that[4] as $n \to \infty$: i) $\bar{p}^m[n] = \check{p}^m$ and $\bar{o}_k[n] = \check{o}_k$, and ii) $\bar{c}[n] \geq \bar{c}^* - \delta(\mu)$, where $\delta(\mu) \to 0$ as $\mu \to 0$.

3.6 Simulated tests

The default simulation parameters are as follows: $M = 5$, $K = 10$, $\beta^m = 1$, $\check{p}^m = 2$, $\kappa_k^m = 1$, $\check{o}_k = 4\%$, and $\Gamma_k = 0.5$. Amplitudes of the secondary links are Rayleigh (so that $h_{k,2}^m[n]$ are exponential) distributed, and the average SNR for all users and bands is $\mathbb{E}_{\mathbf{h}}[h_{k,2}^m] = 9$. The primary CSI model is $h_{k,1}^m[n] = |H_{k,1}^m[n]|^2$, where $H_{k,1}^m[n]$ is low-pass equivalent, complex Gaussian distributed (CGD) with zero mean and unit variance. Real and imaginary parts are independent, so that the amplitude is Rayleigh (and likewise $h_{k,1}^m[n]$ is exponential) distributed. The time correlation model is $H_{k,1}^m[n] = \sqrt{\rho} H_{k,1}^m[n-1] + \sqrt{1-\rho} z_k^m[n]$, with $\rho = 0.95$ and $z_k^m[n]$ white, CGD with zero mean and unit variance. Measurement noise $v_k^m[n]$ is CGD, with zero mean and variance 0.01. The NC senses $H_{k,1}^m[n]$ every $N_h = 6$ slots. The PU activity model is simulated

[3] A classical dual subgradient with diminishing stepsize [Ber03, Ch. 6] would work; see, e.g., [Mar11c] for a related case.

[4] A proof of this result provided can be derived following the lines of [Rib10a, Mar11a].

with the following parameters: $Q_{00} = 0.95$, $Q_{01} = 0.10$, $Q_{10} = 0.05$, and $Q_{11} = 0.90$; $P_{FA} = 3\%$ and $P_{MD} = 2\%$; and the NC senses $a_k[n]$ every $N_a = 3$ slots. Since the optimality and feasibility of the developed schemes has been established theoretically, the simulation parameters and test cases have been chosen to illustrate relevant properties of the developed schemes.

Test Case 1: optimality and feasibility. Table 3.1 lists the average weighted sum-rate, power, and interference probability for an *interweave* CR network implementing nine different RA schemes. The first three solve (3.6) under a short-term interference constraint (STIC): S1) is a genie-aided scheme in which the true is known; S2) is the optimal one developed in this work that accounts for CSI imperfections; and S3) is a scheme adopting error-free CSI with $a_k[n + n_a] = \tilde{a}_k[n]$, and $h^m_{k,1}[n + n_h] = \tilde{h}^m_{k,1}[n]$, for $n_a = 0, 1,..., N_a - 1$, $n_h = 0,..., N_h - 1$. The following three S4), S5) and S6) are the counterparts of S1), S2) and S3) under a long-term interference constraint (LTIC). For further comparison, three more are considered: S7) a scheme with no instantaneous information of the primary CSI, since it relies only on statistical CSI [Sol12, Bar11]; S8) a scheme that solves (3.6) ignoring the interference constraints [Mar12b]; and S9) a scheme that accounts for CSI imperfections, and solves (3.6) guaranteeing that the *average* interfering power at the PUs is less than Γ_k [Mus09, Zha09].

Table 3.1: Interweave CR with $N_a = 3$, $N_h = 6$, $P_{FA} = 1\%$, $P_{MD} = 2\%$, $\check{o}_k = 4\%$, $\mathrm{Var}\{v^m_k[n]\} = 0.01$, $\Gamma_k = 0.5$. Meaning of codes used in row "Comments": C1=STIC enforced, long-term $\bar{o}_k$ shown for illustrative purposes; C2=STIC often violated; C3=LTIC violated.

	S1	S2	S3	S4	S5	S6	S7	S8	S9
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.5
$\bar{c}$	16.3	6.9	16.8	17.6	16.9	17.8	5.3	23.1	17.6
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	0.05%	3.6%	4.0%	4.0%	7.3%	4.0%	39.6%	39.9%
$(1/K)\sum_k \bar{o}_k$ (estim.)	0.0%	0.05%	0.0%	4.0%	4.0%	4.0%	4.0%	−−	−−
Comments	C1	C1	C1,C2			C3		C2, C3	C2, C3

The results corroborate the analytical claims and illustrate the advantages of the developed algorithms. The novel schemes satisfy the constraints, while those ignoring CSI errors violate them; and outperform the suboptimal schemes, especially the one based on statistical knowledge of the CR-to-PU channels. It is worth noticing how S2 (long-term constraint) yields a higher maximum than S1 (short-term). Indeed, S1 over-satisfies the long-term interference constraint, while S2 satisfies the constraint tightly. Finally, the results confirm that the probability of interference estimated by the novel algorithms using the stochastic updates of the belief state corresponds to the actual one. Since our results guarantee that the long-term constraints are satisfied as $n \to \infty$, small discrepancies may occur when the number of simulated time instants is not high enough.

Table 3.2: Underlay CR with $N_a = 3$, $N_h = 6$, $P_{FA} = 1\%$, $P_{MD} = 2\%$, $\check{o}_k = 4\%$, $\mathrm{Var}\{v^m_k[n]\} = 0.01$, $\Gamma_k = 0.5$. Meaning of codes used in row "Comments": see Table 3.1.

	S1	S2	S3	S4	S5	S6	S7	S7'	S8	S9
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.2	1.0	0.5
$\bar{c}$	21.4	20.3	21.5	22.8	21.5	22.8	13.1	11.2	23.1	17.6
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	3.2%	14.4%	4.0%	2.7%	17.0%	4.0%	3.9%	37.5%	14.1%
$(1/K)\sum_k \bar{o}_k$ (estim.)	0.0%	3.4%	0.0%	4.0%	4.0%	4.0%	4.0%	4.0%	−−	−−
Comments	C1	C1	C1,C2			C3		C1	C2, C3	C2, C3

Table 3.3: Interweave and underlay CR with $N_a = 5$, $N_h = 10$, $P_{FA} = 5\%$, $P_{MD} = 3\%$, $\check{o}_k = 2\%$, $\mathrm{Var}\{v_k^m[n]\} = 0.02$, $\Gamma_k = 0.25$. Meaning of codes used in row "Comments": see Table 3.1. The results for the interweave setup are shown in the first (top) half of the table and the ones for the underlay setup are shown in the second (bottom) half.

	S1	S2	S3	S4	S5	S6	S7	S7'	S8	S9
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	--	1.0	0.5
$\bar{c}$	16.7	3.2	16.5	17.2	10.7	16.9	4.3	--	23.2	17.6
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	0.05%	7.0%	2.0%	2.0%	8.8%	2.0%	--	39.6%	39.2%
$(1/K)\sum_k \bar{o}_k$ (estim.)	0.0%	0.05%	0.0%	2.0%	2.0%	2.0%	2.0%	--	--	--
Comments	C1	C1	C1,C2			C3		--	C2, C3	C2, C3
Setup	Interweave									
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	0.3	0.1	1.0	0.5
$\bar{c}$	19.3	11.4	19.2	22.1	15.7	22.1	6.4	5.9	23.1	17.7
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	1.9%	21.8%	2.0%	2.1%	22.0%	2.1%	2.0%	38.0%	26.4%
$(1/K)\sum_k \bar{o}_k$ (estim.)	0.0%	2.0%	0.0%	2.0%	2.0%	2.0%	2.0%	2.0%	--	--
Comments	C1	C1	C1,C2			C3		C1	C2, C3	C2, C3
Setup	Underlay									

Table 3.4: Results for different simulation setups. The CR paradigm and the parameters which are different from those in the default test case are described in row "Setup".

	S1	S2	S3	S4	S5	S6	S7	S7'	S8	S9
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	--	1.0	0.5
$\bar{c}$	8.8	4.1	8.8	9.4	9.0	9.4	3.7	--	11.4	7.7
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	0.05%	3.8%	4.0%	4.0%	7.2%	4.0%	--	40.0%	36.8%
Setup	Interweave: $\mathbb{E}_{\mathbf{h}}[h_{k,2}^m] = 2$									
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.2	1.0	0.5
$\bar{c}$	10.3	9.6	10.3	11.2	10.5	11.2	5.4	3.9	11.5	7.7
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	1.7%	14.2%	4.0%	3.9%	15.7%	4.0%	3.6%	38.0%	13.5%
Setup	Underlay: $\mathbb{E}_{\mathbf{h}}[h_{k,2}^m] = 2$									
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	--	1.0	0.3
$\bar{c}$	10.2	3.6	10.3	11.0	10.4	11.0	2.8	--	14.7	8.8
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	0.05%	3.8%	4.0%	4.0%	7.3%	4.0%	--	39.7%	40.0%
Setup	Interweave: $M = 5$, $K = 5$									
$(1/M)\sum_m \bar{p}_m$	0.8	0.8	0.8	0.8	0.8	0.8	0.8	--	0.8	0.5
$\bar{c}$	15.2	6.1	15.3	16.3	15.6	16.4	5.0	--	21.1	17.4
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	0.05%	3.8%	4.0%	4.0%	7.3%	4.0%	--	39.8%	40.1%
Setup	Interweave: $\check{p}^4 = \check{p}^5 = 0.5$									

Table 3.5: Interweave CR with different models for the activity of the PUs. MM represents the Markov model considered in this work. PM represents the Pareto model considered in [Zha11]. The stationary distribution of a_k is the same in both cases. The simulation setup is the same than that in Table 3.1. Meaning of codes used in row "Comments": see Table 3.1.

	MM: S4	MM: S5	MM: S6	PM: S4	PM: S5	PM: S6
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0
$\bar{c}$	22.8	21.5	22.8	22.8	21.3	22.8
$(1/K)\sum_k \bar{o}_k$ (actual)	4.0%	3.6%	17.0%	4.0%	3.4%	17.1%
$(1/K)\sum_k \bar{o}_k$ (estimated)	4.0%	4.0%	4.0%	4.0%	4.0%	4.0%
Comments			C3			C3

Table 3.2 is the counterpart of Table 3.1 for an *underlay* system. The additional scheme S7' (which is the counterpart of S7 for the case of a STIC) is also tested. Such a scheme is not appropriate for an interweave setup, but it has been considered for underlay CR networks [Sol12]. The results confirm the previous findings. The main observation is that the underlay schemes achieve higher sum-rate than the interweave ones. This is reasonable because CRs in underlay operation have more opportunities to transmit (secondary transmissions with sufficiently low transmit-power do not cause interference even if the PU is active). Results in Tables 3.3-3.5 summarize further numerical tests assessing performance of the novel schemes over a wide range of parameter values, including a non-Markov model for the PUs activity [Zha11] in Table 3.5. These not only confirm the previous conclusions, but also show that when a more demanding setup is simulated (pronounced CSI imperfections and/or strict interference constraints) then: i) the impact of CSI imperfections on $\bar{c}$ is larger (cf. S4, S5, and S6); ii) the interference constraints are more difficult to be satisfied; iii) the performance gain of the LTIC schemes relative to the STIC ones is more pronounced; and iv) the performance gain of our underlay schemes (S2, S5) relative to their interweave counterparts is larger too.

Test Case 2: dynamic behavior of the stochastic schemes. The dynamic behavior of the stochastic iterates is analyzed in this simulation, focusing on S2 in Table 3.1. Figure 3.1 comprises four subplots, each depicting the evolution over time of a different subset of variables. Subplot (a) corresponds to the average power consumption $\bar{p}_m[n]$, and subplot (b) to the long-term probability of interference $\bar{o}_k[n]$ (cf. Section 3.5). Dashed lines mark the performance when the optimal multipliers are known, while solid lines correspond to the proposed stochastic RA algorithms. Subplots (c) and (d) depict the instantaneous value of the Lagrange multipliers $\pi^m[n]$ and $\theta_k[n]$, respectively (in this case dashed lines correspond to the optimum multiplier values).

Subplots (a) and (b) show that the considered constraints are satisfied (with equality), and the stochastic RA converges in a few hundred iterations. The Lagrange multipliers plotted in (c) and (d) suggest that after an initial phase during which the multipliers approach the optimal value, they never converge but hover around the optimum value. This is reasonable because (c) and (d) are instantaneous estimates while (a) and (b) are running averages. Finally, it is worth noting that in order to attain comparable rate of convergence, the stepsizes used to update $\pi^m[n]$ and $\theta_k[n]$ are considerably different ($\mu_\pi = 10^{-2}$ versus $\mu_\theta = 10^{-1}$).

3.7 Concluding summary

Stochastic RA algorithms were developed for wireless cognitive radios communicating over fading links in interweave and underlay settings, and account for imperfections present in the sensing and CSI acquisition phase. Short-term and long-term interference constraints were enforced to account for such imperfections. Although not all formulated problems were convex, enticingly they all turned out to have zero-duality gap, and could thus be solved with manageable complexity. Stochastic algorithms were introduced to: i) estimate and predict the probability of interference; and, ii) estimate the optimum multipliers for the average power and interference constraints.

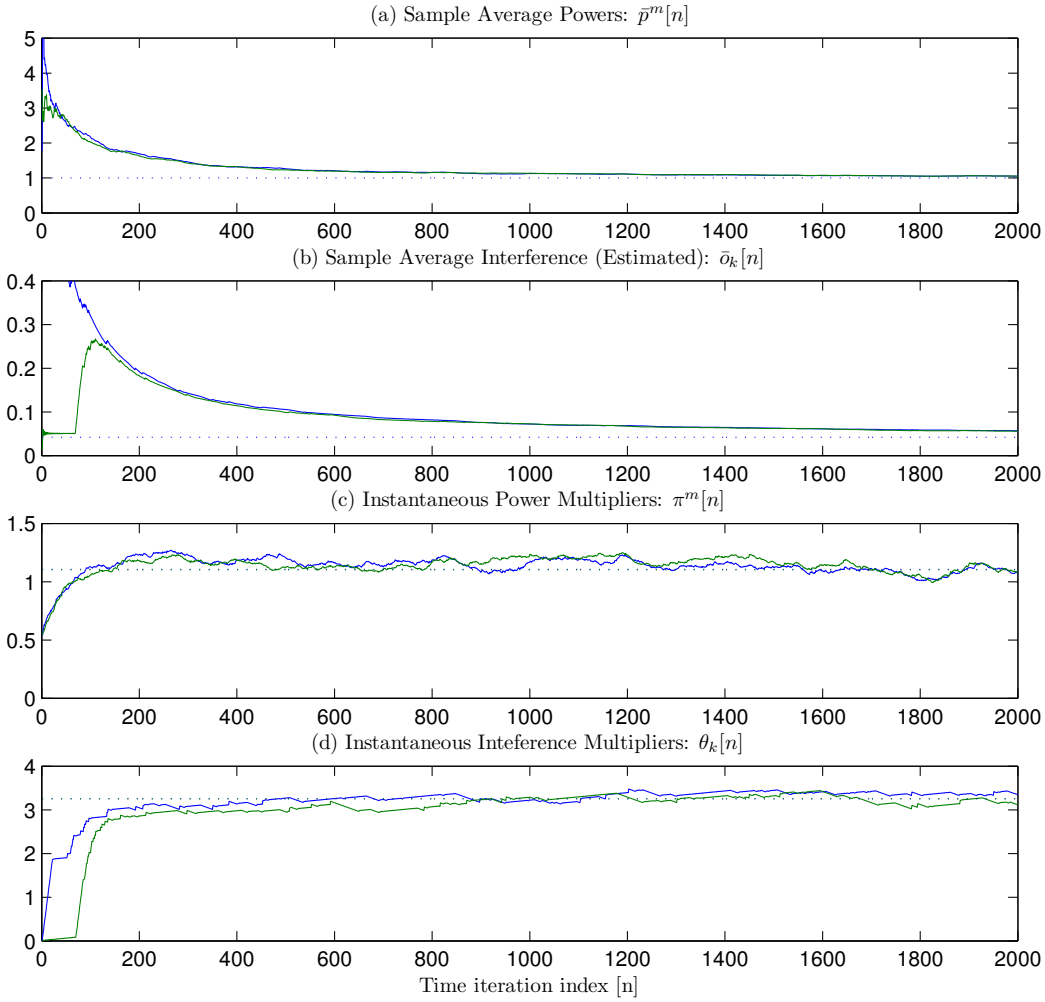

Figure 3.1: Trajectories of different primal and dual variables for scheme S2 in Table 3.1: (a) sample average power; (b) sample average interference; (c) (instantaneous) power multiplier; (d) (instantaneous) interference multiplier. To help visualization, only the multipliers of users $m = 1, 2$ and channels $k = 5, 6$ are plotted. Dashed lines correspond to the optimal (constant) values.

Chapter 4

Jointly Optimal Sensing and Resource Allocation for Interweave CRs

Successful deployment of CRs requires efficient sensing of the spectrum and dynamic adaptation of the available resources according to the sensed (imperfect) information. While the previous chapter focused on optimal RA for a CR network, in this chapter the focus is on the *joint* design of the sensing and resource allocation tasks. The investigated cognitive radio consists of multiple SUs that access orthogonally a set of frequency bands originally devoted to PUs. The schemes are designed to optimize the tradeoff between the sensing cost and the SUs throughput (weighted sum rate), and limit their power consumption as well as the probability of interfering with the PUs. The joint design is addressed using nonlinear optimization and dynamic programming, which is able to leverage the time correlation in the activity of the primary network. A two-step strategy is implemented: it first finds the optimal RA for any sensing scheme and then uses that solution as input to solve for the optimal sensing policy. The two-step strategy is optimal, gives rise to intuitive optimal policies, and entails a computational complexity much lower than that required to solve the original formulation. Also, a stochastic solution is developed, which is adaptive to non-stationary fading conditions.

4.1 Introduction

Effective operation of CRs requires the implementation of two critical tasks: i) sensing the spectrum seeking transmit opportunities and ii) dynamically adapting the available resources according to the sensed information [Hay05]. To carry out the *sensing task* two important challenges are: C1) the presence of errors in the measurements that lead to errors in the channel occupancy detection and thus render harmless SU transmissions impossible; and C2) the inability to sense the whole of the time-frequency lattice due to scarcity of resources (time, energy, or sensing devices). Two additional challenges that arise to carry out the *RA task* are: C3) the need of the RA algorithms to deal with channel imperfections such as noise or quantization; and C4) the selection of metrics that properly quantify the reward for the SUs and the harm for the PUs in case of interference.

4.1.1 Related Work

Many alternatives have been proposed in the CR literature to deal with these challenges. Different forms of imperfect CSI, such as quantized or noisy CSI, have been used to deal with C1 [Mus09]. However, in the context of CR, fewer works have considered the fact that the CSI may be not only noisy but also outdated, or have incorporated those imperfections into the design

Table 4.1: Summary of most significant notation in Chapter 4.

Symbol	Meaning
n	Time slot index
M , m	Number of users / User index
K , k	Number of channels / Channel index
$h_k^m[n]$	Fading gain for user m on channel k
$a_k[n]$	Presence of PU in channel k
$\mathbf{P}_k$	Transition probability matrix of $a_k[n]$
$s_k[n]$, $z_k[n]$	Sensing decision / Sensing output
P_k^{FA}, P_k^{MD}	False alarm / missed detection probabilities
$b_k[n]$, $B_k[n]$	Pre-decision / post decision belief on $a_k[n]$
$w_k^m[n]$	Scheduling variable (1 if user m occupies channel k)
$p_k^m[n]$	Nominal power of user m in channel k
$\breve{p}^m[n]$	Maximum average power consumed by user m
$\breve{o}_k[n]$	Maximum prob. of interference on channel k
π^m , θ_k	Lagrange multipliers associated with (4.4), (4.5)
γ	Discount factor $0 < \gamma < 1$
ξ_k	Cost of sensing channel k
$L_k(\mathbf{h}_k[n])$	Instantaneous reward indicator (IRI)
$Q_k(b_k[n], s_k[n])$	Q-function of the POMDP associated with channel k

of RA algorithms [Che08b]. The inherent tradeoff between sensing cost and throughput gains in C2 has been discussed in [Lia08] and designs that account for such a tradeoff based on convex optimization [Wan11] and SDP [Che08b] for specific system setups have been proposed. Regarding C3, many works consider that the CSI is imperfect, but only a few exploit the statistical model of these imperfections (especially for the time correlation) to mitigate them [Che08b, Mar12a]. Finally, different alternatives have been considered to deal with C4 and limit the harm that the SUs cause to the PUs [Gon11]. The most widely used is to set limits on the peak (instantaneous) and average interfering power. Some works also have imposed limits on the rate loss that PUs experience [Mar08b, Mar12c], while others look at limiting the instantaneous or average probability of interfering the PU (bounds on the short-term or long-term outage probability) [Urg09, Mar12a].

Regardless of the challenges addressed and the formulation chosen, the sensing and RA policies have been traditionally designed separately. Each of the tasks has been investigated thoroughly and relevant results are available in the literature. However, a globally optimum solution requires a joint design capable of leveraging the interactions between those two tasks. Clearly, more accurate sensing enables more efficient RA, but at the expense of higher time and/or energy consumption [Lia08]. Early works dealing with joint design of sensing and RA are [Zha07b] and [Che08b]. In such works, imperfections in the sensors and time correlation of the state information of the primary network (SIPN) are considered, and the sensing design is modeled as a POMDP [Kae98], [Wie12, Ch. 12], [Ber95].The design of the RA in these works amounts simply to user scheduling (i.e., selecting the user transmitting on each channel). Under mild conditions, the authors establish that a separation principle holds in the design of the optimal access and sensing policies. Additional works addressing the joint design of sensing and RA, and considering more complex operating conditions, were published more recently [Wan11, Kim10]. For a single SU operating multiple

fading channels, [Wan11] relies on convex optimization to optimally design both the RA and the indices of the channels to be sensed at every time instant. Assuming that the number of channels that can be sensed at every instant is fixed and that the PU activity is independent across time, the author establishes that the channels to sense are those that can potentially yield a higher reward for the secondary user. Joint optimal design is also pursued in [Kim10], although for a very different setup. Specifically, [Kim10] postulates that at each slot, the CR must compute the fraction of time devoted to sense the channel and the fraction devoted to transmit in the bands which are found to be unoccupied. Clearly, a tradeoff between sensing accuracy and transmission rate emerges. The design is formulated as an optimal stopping problem, and solved by means of Lagrange relaxation of SDP [Cas97]. However, none of these two works takes into account the temporal correlation of the SIPN.

4.1.2 Objective and Contributions

In this chapter we design the sensing and the RA policies *jointly* while accounting for the challenges C1-C4. The specific operating conditions considered are described next. We analyze an *interweave* CR with multiple SUs and PUs. SUs are able to adapt their transmit power and rate, and access orthogonally[cf. Sec. 3.3] a set of frequency bands originally devoted to PU transmissions. *Orthogonally* here means that if a SU is transmitting, no other SU can be active in the same band. The schemes are designed to maximize the sum-average rate of the SUs while adhering to constraints that *limit* the maximum "average power" that SUs transmit and the average "probability of interfering" the PUs. It is assumed that the CSI of the SU links is instantaneous and free of errors, while the CSI of the PUs activity is outdated and noisy. A simple first-order Markov model is used to characterize such imperfections. Sensing a channel band entails a given cost, and at each instant the system has to decide which channels (if any) are sensed.

The main novelty (and contribution) of this work is the combined use of SDP and dual nonlinear optimization techniques to design the jointly optimal sensing and RA schemes. The requirement for SDP techniques comes because the activity of PUs is assumed to be correlated across time, so that sensing a channel has an impact not only for the current instant, but also for future time instants [Zha07b]. To solve the joint design, a two-step strategy is implemented. In the first step, we obtain an analytical expression for the performance achieved by the *optimal* RA as a function of the state of the system after the sensing task (this expression is valid for *any* fixed sensing scheme). This (sub-) problem was solved in Chapter 3, as well as [Mar11b, Mar12a]. In the second step, the *analytical expression* obtained in the first step is used as input to obtain the optimal sensing policy. This two-step strategy has a double motivation. First, while the joint design is non convex and involves the formulation of a SDP problem (also referred to as dynamic program (DP), for brevity) techniques, the problem to be solved in the first step (optimal RA for a fixed sensing scheme) can be recast as a convex one. Second, when the *expression* for the optimal RA performance is substituted back into the original joint design, the resulting sensing optimization problem (which does need to be solved using SDP) has a more favorable structure. More specifically, while the original design problem was a constrained DP, the problem to be solved in the second step is an unconstrained DP which can be solved separately for each of the channels. These facts will make our problem computationally affordable without entailing a loss of optimality [cf. Sec. 4.3].

The rest of the chapter is organized as follows. Sec. II describes the system setup and introduces notation. The optimization problem that gives rise to the optimal sensing and RA schemes is formulated in Sec. 4.3. The solution for the optimal RA given the sensing scheme is presented in Sec. 4.4. The optimization of the sensing scheme is addressed in Sec. 4.5, formulating the problem in the SDP framework and developing its solution. Numerical simulations validating the theoretical claims and providing insights on our optimal schemes are presented in Sec. 4.7. Sec. 4.8 summarizes the main properties of our jointly optimal RA and sensing policies and concludes

the chapter.

4.2 System setup and state information

The section begins by briefly describing the system setup and the main operation steps (tasks that the system runs at every time slot). Then, the model for the CSI, which will play a critical role in the problem formulation, is explained in detail. The resources that SUs will adapt as a function of the CSI are described in the last part of the section.

We consider a CR scenario with several PUs and SUs. The frequency band of interest (portion of spectrum that is licensed to PUs, or the subset of this shared with the SUs) is divided into K frequency-flat orthogonal subchannels (indexed by k). Each of the M SUs (indexed by m) opportunistically accesses any number of these channels during a time slot (indexed by n). Opportunistic here means that the user accessing each channel will vary with time as a function of the current CSI, with the objective of optimally utilizing the available channel resources. For simplicity, we assume that there exists a NC which acts as a central scheduler and will also perform the task of sensing the medium for primary presence. The scheduling information will be forwarded to the mobile stations through a parallel feedback channel. The results hold for one-hop (either cellular or any-to-any) setups.

Next, we briefly describe the operation of the system. A more detailed description will be given in Sec. 4.3, which will rely on the notation and problem formulation introduced in the following sections. Before starting, it is important to clarify that we focus on systems where the SIPN is more difficult to acquire than the state information of the secondary network (SISN). As a result, we will assume that SISN is error-free and acquired at every slot n, while SIPN is not. With these considerations in mind, the CR operates as follows. At every slot n the following tasks are run sequentially: T1) the NC acquires the SISN; T2) the NC relies on the output of T1 (and on previous measurements) to decide which channels to sense (if any), then the output of the sensing is used to update the SIPN; and T3) the NC uses the outputs of T1 and T2 to find the optimal RA for instant n. Overheads associated with acquisition of the SISN and notification of the optimal RA to the SUs are considered negligible. Such an assumption facilitates the analysis, and it is reasonable for scenarios where the SUs are deployed in a relatively small area which allows for low-cost signaling transmissions.

4.2.1 State information and sensing scheme

Let us begin by introducing the model for the SISN. Let $\tilde{h}_k^m[n]$ be the square magnitude of the fading coefficient of the channel between the mth SU and its intended receiver on frequency k during slot n. With $\sigma_k^m[n]$ denoting the corresponding noise-plus-interference power, $h_k^m[n] := \tilde{h}_k^m[n]/\sigma_k^m[n]$ is defined as the noise-normalized power gain for the mth SU on frequency k. The stochastic process $h_k^m[n]$ will be assumed to be independent and identically distributed (i.i.d.) across time. The values of $h_k^m[n]$ for all m and k constitute the SISN at slot n. The SISN is assumed perfect, so that the values of $h_k^m[n]$ at every time slot n are known with no errors. This assumption may be unrealistic, but it is made to focus on the challenges due to SIPN imperfections, which are always more severe. Nonetheless, comments on how to modify the RA when this assumption does not hold true will be provided in Sec. 4.4.1.

The SIPN accounts for the channel occupancy. It will be assumed that the primary system contains one user per channel. This assumption keeps the modeling simple and it is accurate for some practical scenarios, e.g. a primary system of mobile telephony where a single narrow-band channel is assigned to a single user during the course of a call. Since the CR under study follows the interweave paradigm (i.e., if an SU accesses the channel when the PU is active, interference takes place regardless of the transmit power), it suffices to know whether a given channel is

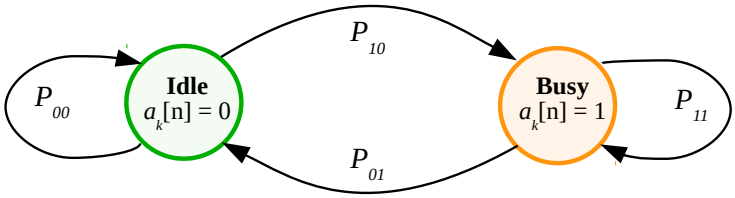

Figure 4.1: Two-state Gilbert-Elliot channel model and transition probabilities.

occupied or not [Gol09]. This way, when a PU is not active, opportunities for SUs to transmit in the corresponding channel arise. The primary system is not assumed to collaborate with the secondary system. Hence, from the point of view of the SUs, the behavior of PUs is a stochastic process independent of $h_k^m[n]$. With these considerations in mind, the presence of the primary user in channel k at time n is represented by the binary state variable $a_k[n]$ (0/idle, 1/busy). Each primary user's behavior will be modeled as a simple, discrete-time, Gilbert-Elliot channel model, so that $a_k[n]$ is assumed to remain constant during the whole time slot, and then change according to a two-state, time invariant Markov chain.

The Markovian property will be useful to keep the DP modeling simple and will also be exploited to recursively keep track of the SIPN. Nonetheless, more refined models can be considered without paying a big computational price [Mar12a, Zha11]. With $P_k^{xy} := \Pr(a_k[n] = x | a_k[n-1] = y)$, the dynamics for the Gilbert-Elliot model are fully described by the 2×2 Markov transition matrix $\mathbf{P}_k := [P_k^{00}, P_k^{01}; P_k^{10}, P_k^{11}]$. Sec. 4.8 discusses the implications of relaxing some of these assumptions.

While knowledge of $h_k^m[n]$ at instant n was assumed to be perfect (deterministic), knowledge of $a_k[n]$ at instant n is assumed to be imperfect (probabilistic). Two important sources of imperfections are: i) errors in the sensing process and ii) outdated information (because the channels are not always sensed). To model the sensing task, let $s_k[n]$ denote a binary design variable which is 1 if the kth channel is sensed at time n, and 0 otherwise. Moreover, let $z_k[n]$ denote the output of the sensor if indeed $s_k[n] = 1$; i.e., if the kth channel has been sensed. We will assume that the output of the sensor is binary and may contain errors. To account for asymmetric errors, the probabilities of false alarm $P_k^{FA} = \Pr(z_k[n] = 1 | a_k[n] = 0)$ and missed detection $P_k^{MD} = \Pr(z_k[n] = 0 | a_k[n] = 1)$ are considered.

Clearly, the specific values of P_k^{FA} and P_k^{MD} will depend on the detection technique the sensors implement [Yuc09] and the sensor parameters (operating point). For simplicity, P_k^{FA} and P_k^{MD} are considered known and time invariant[1]. As already mentioned, the sensing imperfections render the knowledge of $a_k[n]$ at instant n probabilistic; in other words, $a_k[n]$ is a partially observable state variable. The knowledge about the value of $a_k[n]$ at instant n will be referred to as the instantaneous belief. For a given instant n, two different types of belief are considered: the *pre-decision* belief $b_k[n]$ and the *post-decision* belief $B_k[n]$. Intuitively, $b_k[n]$ contains the information about $a_k[n]$ before the sensing decision has been made (i.e., at the beginning of task T2), while $B_k[n]$ contains the information about $a_k[n]$ once $s_k[n]$ and $z_k[n]$ (if $s_k[n] = 1$) are known (i.e., at the end of task T2). Mathematically, if $\mathcal{H}_n$ represents the history of all sensing decisions and measurements up to and including instant n, i.e., $\mathcal{H}_n := \{s_k[0], z_k[0], \ldots, s_k[n], z_k[n]\}$; then $b_k[n] := \Pr(a_k[n] = 1 | \mathcal{H}_{n-1})$ and $B_k[n] := \Pr(a_k[n] = 1 | \mathcal{H}_n)$. Provided that the Markov matrix

[1] This is reasonable if: i) the primary-secondary fading conditions are stationary and ii) complete information about their statistics (but not about their instantaneous values) is available. Under such conditions, the optimal operating point of the sensor is constant during the CR operation [Gha05]. Nonetheless, our results can be adapted to handle time-variant $P_k^{FA}[n]$ and $P_k^{MD}[n]$. Specifically, results in section 4.4 are not affected, and results in Section 4.5 can be adapted by accounting for the distribution of $P_k^{FA}[n]$ and $P_k^{MD}[n]$.

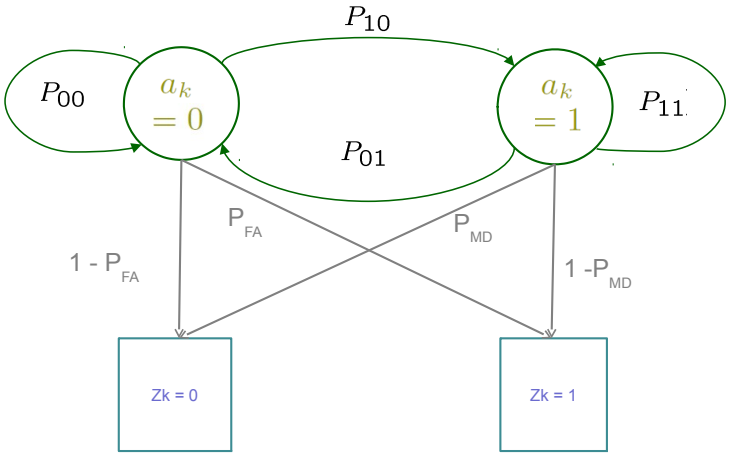

Figure 4.2: Observation model for the PU state with asymmetric errors.

$\mathbf{P}_k$ is known, the expression to get the pre-decision belief at time slot n is

$$b_k[n] = \mathcal{P}_k(B_k[n-1]) := P_k^{10}(1 - B_k[n-1]) + P_k^{11}B_k[n-1] \tag{4.1}$$

Differently, the expression to get $B_k[n]$ depends on the sensing decision $s_k[n]$. If $s_k[n] = 0$, no additional information is available, so that the post-decision belief will equal the pre-decision one. If $s_k[n] = 1$, the belief is updated according to the sensing outcome. Mathematically,

$$B_k[n] = \mathcal{U}(b_k[n], s_k[n], z_k[n]) := \begin{cases} b_k[n] & s_k[n] = 0 \\ \frac{P^{MD}b_k[n]}{P^{MD}b_k[n]+(1-P^{FA})(1-b_k[n])} & s_k[n] = 1, z_k[n] = 0 \\ \frac{(1-P^{MD})b_k[n]}{(1-P^{MD})b_k[n]+P^{FA}(1-b_k[n])} & s_k[n] = 1, z_k[n] = 1 \end{cases} \tag{4.2}$$

Note that (4.1) and (4.2) correspond to the prediction and update steps of a Bayesian recursive estimator, respectively. If no information about $a_k[0]$ is available, $b_k[0]$ can be initialized as the corresponding stationary probability. Nevertheless, the effects of an incorrect estimation of the PU presence vanish as new prediction and update steps are performed.

In a nutshell, the actual state of the primary and secondary networks is given by the random processes $a_k[n]$ and $h_k^m[n]$, which are assumed to be mutually independent. The operating conditions of our CR are such that at instant n, the value of $h_k^m[n]$ is perfectly known, while the SIPN is formed by $b_k[n]$ and $\mathbf{b}_k^S[n]$, which are a *probabilistic* description (estimator) of $a_k[n]$. The system will perform the sensing and RA tasks based on the available SISN and SIPN. In particular, the sensing decision will be made based on $h_k^m[n]$ and $\mathbf{b}_k[n]$, while the RA will be implemented based on $h_k^m[n]$ and $\mathbf{b}_k^S[n]$.

4.2.2 Resources at the secondary network

We consider a secondary network where users implement adaptive modulation and power control, and share the available channels orthogonally. To describe the channel access scheme (scheduling) rigorously, let $w_k^m[n]$ be a Boolean variable so that $w_k^m[n] = 1$ if SU m accesses channel k and zero otherwise. Moreover, let $p_k^m[n]$ be a nonnegative variable denoting the *nominal* power

assigned for SU m to transmit in channel k, and let $C_k^m[n]$ be its corresponding rate. We say that the $p_k^m[n]$ is a nominal power in the sense that power is consumed only if the user is actually accessing the channel. Otherwise the power is zero, so that the actual (effective) power user m loads in channel k can be written as $w_k^m[n]p_k^m[n]$.

The transmission bit rate is obtained through Shannon's capacity formula [Li01]: $C_k^m[n] := C_k^m(h_k^m[n], p_k^m[n]) := \log_2(1 + h_k^m[n]p_k^m[n]/\Gamma)$ where Γ is a signal-to-noise ratio (SNR) gap that accounts for the difference between the theoretical capacity and the actual rate achieved by the modulation and coding scheme the SU implements. This is a bijective, nondecreasing, concave function with $p_k^m[n]$ and it establishes a relationship between power and rate in the sense that controlling $p_k^m[n]$ implies also controlling $C_k^m[n]$.

The fact of the access being orthogonal implies that, at any time instant, at most one SU can access the channel. Mathematically,

$$\sum_m w_k^m[n] \leq 1 \quad \forall k, n. \tag{4.3}$$

Note that (4.3) allows for the event of all $w_k^m[n]$ being zero for a given channel k. That would happen if, for example, the CR infers that, at instant n, it is very likely that channel k is occupied by a PU.

4.3 Problem statement

The approach in this chapter is to design the sensing and RA schemes as the solution of a judiciously formulated optimization problem. Consequently, it is critical to identify: i) the design (optimization) variables, ii) the state variables, iii) the constraints that design and state variables must obey, and iv) the objective of the optimization problem.

The first two steps were accomplished in Sec. 4.2, stating that the design variables are $s_k[n]$, $w_k^m[n]$ and $p_k^m[n]$ (recall that there is no need to optimize over $C_k^m[n]$); and that the state variables are $h_k^m[n]$ (SISN), and $\mathbf{b}_k[n]$ and $\mathbf{b}_k^S[n]$ (SIPN).

Moving to step iii), the constraints that the variables need to satisfy can be grouped into two classes. The first class is formed by constraints that account for the system setup. This class includes constraint (4.3) as well as the following constraints that were implicitly introduced in the previous section: $s_k[n] \in \{0, 1\}$, $w_k^m[n] \in \{0, 1\}$ and $p_k^m[n] \geq 0$. The second class is formed by constraints that account for QoS. In particular, we consider the following two constraints. The first one is a limit on the maximum average (long-term) power an SU can transmit. By enforcing an average consumption constraint, opportunistic strategies are favored because energy can be saved during deep fadings (or when the channel is known to be occupied) and used during transmission opportunities. Transmission opportunities are time slots where the channel is certainly known to be idle and the fading conditions are favorable. Mathematically, with $\check{p}^m$ denoting such maximum value, the average power constraint is written as:

$$\mathbb{E}\left[\lim_{N\to\infty}(1-\gamma)\sum_{n=0}^{N-1}\gamma^n \textstyle\sum_k w_k^m[n]p_k^m[n]\right] \leq \check{p}^m, \quad \forall m, \tag{4.4}$$

where $0 < \gamma < 1$ is a discount factor such that more emphasis is placed in near future instants. This exponentially decaying window also facilitates accommodating potential non-stationarities. The factor $(1-\gamma)$ ensures that the averaging operator is normalized; i.e., that $\lim_{N\to\infty}\sum_{n=0}^{N-1}(1-\gamma)\gamma^n = 1$. As explained in more detail in Sec. 4.5, using an exponentially decaying average is also useful from a mathematical perspective (convergence and existence of the solution is guaranteed).

While the previous constraint guarantees QoS for the SUs, we also need to guarantee a level of QoS for the PUs. As explained in the introduction, there are different strategies to limit

the interference that SUs cause to PUs; e.g., by imposing limits on the interfering power at the PUs, or on the rate loss that such interference generates [Mar12a]. In this work, we will guarantee that the *long-term* probability of a PU being interfered by SUs is below a certain prespecified threshold $\check{o}_k$. Mathematically, we require $\Pr\{\sum_m w_k^m = 1 | a_k = 1\} \leq \check{o}_k$ for each band $k = 1, \ldots, K$. Using the definition of conditional probability, the constraint can be rewritten as $\Pr\{\sum_m w_k^m = 1, a_k = 1\}/\Pr\{a_k = 1\} \leq \check{o}_k$ and, capitalizing on the fact that both a_k and $\sum_m w_k^m$ are Boolean variables:

$$\mathbb{E}\left[\lim_{N\to\infty} \sum_{n=0}^{N-1} (1-\gamma)\gamma^n a_k[n] \sum_m w_k^m[n]\right] / A_k \leq \check{o}_k, \quad \forall k, \tag{4.5}$$

where A_k, which is assumed known, denotes the stationary probability of the kth band being occupied by the corresponding primary user. Writing the constraint in this form reveals its underlying convexity. + Before moving to the next step, two clarifications are in order. The first one is on the practicality of (4.5). Constraints that allow for a certain level of interference are reasonable because error-free sensing is unrealistic. Indeed, our model assumes that even if channel k is sensed as idle, there is a probability P_k^{MD} of being occupied. Moreover, when the interference limit is formulated as a long-term constraint (as it is in our case), there is an additional motivation for the constraint. The system is able to exploit the so-called interference diversity [Zha10]. Such diversity allows SUs to take advantage of very favorable channel realizations even if they are likely to interfere PUs. To balance the outcome, SUs will be conservative when channel realizations are not that favorable and may remain silent even if it is likely that the PU is not present. The second clarification is that we implicitly assumed that SU transmissions are possible even if the PU is present. The reason is twofold. First, the fact that a SU transmitter is interfering a PU receiver, does not necessarily imply that the reciprocal is true. Second, since the NC does not have any control over the power that primary transmitters use, the interfering power at the secondary receiver could be incorporated into $h_k^m[n]$ as an additional source of noise.

The fourth (and last) step to formulate the optimization problem is to design the metric (objective) to be maximized. Different utility (reward) and cost functions can be used to such purpose. As mentioned in the introduction, focus is put in optimizing the tradeoff between the weighted sum rate of the SUs and the cost associated with sensing. Specifically, we consider that every time that channel k is sensed, the system has to pay a price $\xi_k > 0$. We assume that such a price is fixed and known beforehand, but time-varying prices can be accommodated into our formulation too (see Sec. 4.6.3 for additional details). This way, the sensing cost at time n is $U_S[n] \coloneqq \sum_k \xi_k s_k[n]$. Similarly, we define the utility for the SUs at time n as

$$U_{SU}[n] \coloneqq \sum_k \left(\sum_m \beta^m w_k^m[n] C_k^m(h_k^m[n], p_k^m[n]) \right)$$

where $\beta^m > 0$ is a user-priority coefficient. Based on these definitions, the utility for our CR at time n is $U_{SU}[n] - U_S[n]$. Finally, we aim to maximize the long-term utility of the system denoted by $\bar{U}_T$ and defined as

$$\bar{U}_T \coloneqq \mathbb{E}\left[\lim_{N\to\infty} \sum_{n=0}^{N-1} (1-\gamma)\gamma^n U_{SU}[n] - U_S[n]\right]. \tag{4.6}$$

With these notational conventions, the optimal $s_k^\star[n]$, $w_k^{m\star}[n]$ and $p_k^{m\star}[n]$ will be obtained

as the solution of the following constrained optimization problem.

$$\mathsf{P}^*_{DP} := \max_{\{s_k[n], w_k^m[n], p_k^m[n]\}} \quad \bar{U}_T \tag{4.7a}$$

$$\text{s. to:} \quad (4.3),\ w_k^m[n] \in \{0,1\},\ p_k^m[n] \geq 0, \tag{4.7b}$$

$$(4.4),\ (4.5) \tag{4.7c}$$

$$s_k[n] \in \{0,1\}. \tag{4.7d}$$

The constraints in (4.7) have been gathered into three groups: (4.7b) refers to constraints that affect the RA (i.e., $w_k^m[n]$ and $p_k^m[n]$) and need to hold at every time instant; (4.7c) refers to constraints that affect the RA and need to hold in the long term; and (4.7d) affects the design variables involved in the sensing task ($s_k[n]$).

The main difficulty in solving (4.7) is that the solution for all time instants has to be found jointly. The key reason is that sensing decisions at instant n have an impact not only at that instant, but at future instants as well. As a result, a separate per-slot optimization approach is not optimal in the long term. More specifically, (4.7) belongs to the class of sequential decision problems that needs to be solved by means of SDP. This techniques usually give rise to algorithms of high complexity, so that a careful analysis must be performed to keep the problem tractable. In this work, the algorithmic strategy exploits some details of the problem structure that will help to reduce the complexity of the solution.

Common practice to render SDP problems tractable is: i) analyzing the problem within a well-studied framework; and ii) looking for approximation strategies that allow to reduce the computational load in exchange for a small loss of optimality. The problem at hand will be analyzed within the framework of POMDPs. MDPs are a class of DPs where state transitions and average rewards only depend on the current state-action pair. POMDPs can be viewed as a generalization of MDPs where the system state is not known perfectly; only an observation (affected by errors, missing data, or ambiguity) is available. By using a belief variable [cf. Sec. 4.2.1] a POMDP can be recast as an MDP [Bra03].

In addition, a two-step strategy will considerably reduce the computational burden to solve (4.7) without sacrificing optimality. To explain this strategy, let us revisit (and further clarify) the operation of the system. In Sec. 4.2 it was explained that, at each slot n, our CR had to implement three main tasks: T1) acquisition of the SISN, T2) sensing and update of the SIPN, and T3) resource allocation and transmission. In what follows, task T2 is split into 3 subtasks, so that the CR runs five sequential steps for each channel k:

- T1) $h_k^m[n]$ is acquired;
- T2.1) $b_k[n]$ is computed using $\mathbf{P}_k$ and $B_k[n-1]$ via (4.1);
- T2.2) $h_k^m[n]$ and $b_k[n]$ are used to find $s_k^*[n]$;
- T2.3) $b_k[n]$ and $z_k[n]$ (for the channels where $s_k^*[n] = 1$) are used to get $B_k[n]$ via (4.2);
- T3) $h_k^m[n]$ and $B_k[n]$ are used to find the optimal value of $w_k^{m*}[n]$ and $p_k^{m*}[n]$, and the SUs transmit accordingly.

The two-step strategy to solve (4.7) proceeds as follows. In step I, we obtain the expression for the optimal $w_k^m[n]$ and $p_k^m[n]$ for any sensing scheme. The resultant optimization problem is simpler than the original one in (4.7) because the dimensionality of the optimization space is smaller and the terms in (4.7) that depend only on $s_k[n]$ can be dropped. More importantly, if the sensing is not optimized, a per-slot optimization can be rendered optimal (see discussion on the per-slot separability of the Lagrangian in Sec. IV.A). In step II, we substitute the output of

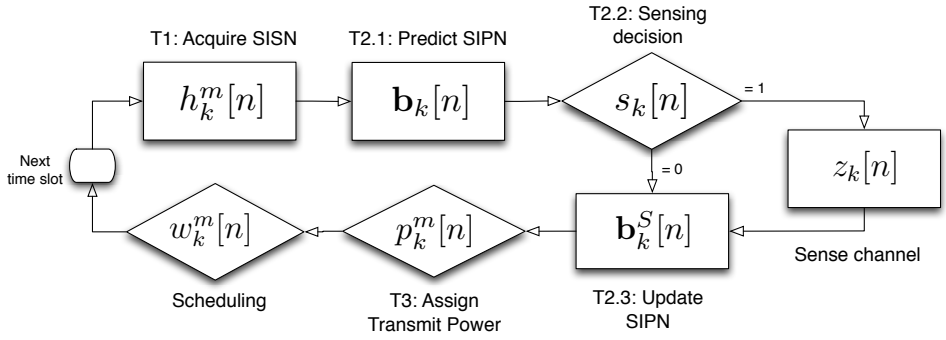

Figure 4.3: Sequential operation of the CR system.

step I into (4.7) and solve for the optimal $s_k[n]$. Note that this does not entail a loss of optimality because the solution of step I is a *function* of the sensing scheme, and the latter is optimized in step II. Mathematically, for a generic function $f(x,y)$, the approach amounts to find $(x^*,y^*) = \arg\min_{x,y} f(x,y)$ as follows: i) $x^*(y) = \arg\min_x f(x,y)$ and ii) $y^* = \arg\min_y f(x^*(y),y)$. The last (trivial) step is to find x^* as $x^* = x^*(y^*)$. Here, the RA variables correspond to x and the sensing variables correspond to y. Clearly, the output of step I will be used in T2.2 (to find the optimal sensing) and in T3 (to find the optimal RA once the optimal sensing is known). The output of step II will be used in T2.2.[2] The optimization in step I (RA) is addressed next, while the one in step II (sensing) is addressed in Sec. 4.5.

4.4 Optimal RA for the secondary network

According to the previous explanation, the objective of this section is to design the optimal RA (scheduling and powers) for a fixed sensing policy. Solving this problem is convenient because: i) it corresponds to one of the tasks our CR has to implement; ii) it is a much simpler problem than the original problem in (4.7), indeed the problem in this section has a smaller dimensionality and, more importantly, can be recast as a convex optimization problem; and iii) it will serve as an input for the design of the optimal sensing, simplifying the task of finding the global solution of (4.7).

Because in this section the sensing policy is considered given (fixed), $s_k[n]$ is not a design variable, and all the terms that depend only on $s_k[n]$ can be ignored. Specifically, the constraint in (4.7d) and the contribution of the sensing cost $U_S[n]$ to the total utility in (4.7a) can be dropped. As a result, we define the new objective to optimize as

$$\bar{U}_{SU} := \sum\nolimits_{k,m} \mathbb{E}\Big[\lim_{N\to\infty} \sum_{n=0}^{N-1} (1-\gamma)\gamma^n \beta^m w_k^m[n] C_k^m(h_k^m[n], p_k^m[n])\Big]$$

and aim at solving the following problem [cf. (4.7)]

$$\mathsf{P}^*_{RA} := \max_{\{w_k^m[n], p_k^m[n]\}} \quad \bar{U}_{SU} \tag{4.8a}$$

$$\text{s. to}: \ \text{(4.7b)}, \ \text{(4.7c)}. \tag{4.8b}$$

A slightly modified version of this problem was solved in Chapter 3 and published in [Mar12a]. For this reason, we organize the remaining of this section into two parts. The first one adapts the results from Chapter 3, presenting the optimal RA. The second part is devoted to introduce new variables that will serve as input for the design of the optimal sensing in Sec. 4.5.

[2] Note that the steps to obtain the optimal solution and those to implement the optimal solution during the CR operation do not follow the same order.

4.4.1 Solving for the RA

It can be shown that the problem in (4.8) can be reformulated as a convex one; see [Mar12a] as well as [Mar11c] for details. Specifically, if the constraint $w_k^m[n] \in \{0,1\}$ is relaxed to yield $w_k^m[n] \in [0,1]$ and an auxiliary (dummy) variable $x_k^m[n] := p_k^m[n]w_k^m[n]$ is introduced, then the problem is convex in $x_k^m[n]$ and $w_k^m[n]$. Moreover, with probability one (w.p.1) the solution to the relaxed problem is feasible (hence, optimal) for the original problem [Mar12a, Mar11c]. The approach to solve the reformulated version of (4.8) is to dualize the long-term constraints in (4.7c). For such a purpose, let π^m and θ_k be the Lagrange multipliers associated with the (now convex) constraints (4.4) and (4.5), respectively, and define the following auxiliary variables:

$$\tilde{p}_k^m[n] := \left[(\dot{C}_k^m)^{-1}\left(h_k^m[n], \pi^m/\beta^m\right)\right]_+, \tag{4.9}$$

$$L_{SU,k}^m[n] := \beta^m C_k^m(h_k^m[n], \tilde{p}_k^m[n]) - \pi^m \tilde{p}_k^m[n], \text{ and} \tag{4.10}$$

$$L_k^m[n] := L_{SU,k}^m[n] - \theta_k B_k[n]. \tag{4.11}$$

Using (4.9)–(4.11), it can be shown that the optimal $w_k^{m*}[n]$ and $x_k^{m*}[n]$ that solve the convex version of (4.8) are

$$w_k^{m*}[n] := \mathbb{1}_{\{(L_k^m[n]=\max_q L_k^q[n]) \wedge (L_k^m[n]>0)\}}. \tag{4.12}$$

and $x_k^{m*}[n] := \tilde{p}_k^m[n]w_k^{m*}[n]$. The latter readily implies that the optimal power allocation can be written as $p_k^{m*}[n] := \tilde{p}_k^m[n]$.

The auxiliary variables $L_{SU,k}^m[n]$ and $L_k^m[n]$ are useful to express the optimal RA, but also to gain insights on how the optimal RA operates. Both variables can be viewed as instantaneous reward indicators (IRIs) representing the reward that can be obtained if $w_k^m[n]$ is set to one. The indicator $L_{SU,k}^m[n]$ considers only SISN and represents the best achievable tradeoff between the rate and power transmitted by the SU. Note that the risk of interfering the PU is not considered in $L_{SU,k}^m$, but it is considered in $L_k^m[n]$ through the addition of the interference-aware term $-\theta_k B_k[n]$. According to (4.12), only the SU with highest IRI can access the channel; moreover, if all users obtain a negative IRI, then none of them should access the channel (meaning that the utility for the SUs does not compensate for the risk of interfering the PU). The expressions in (4.9)–(4.12) also reveal the favorable structure of the optimal RA. The only parameters coupling users, channels and time slots are the multipliers (prices) π^m and θ_k. Once they are known, the Lagrangian of the convex version of (4.8) is separable, and the optimal RA at time n can be found using only information at time n.

Before moving to the next section, two comments are in . The first one is related to the computation of the Lagrange multipliers. Several methods to set the value of the dual variables π^m and θ_k are available. Since, after relaxation, the problem has zero duality gap, there exists a *constant* (stationary) optimal value for each multiplier, denoted as π^{m*} and θ_k^*, such that substituting $\pi^m = \pi^{m*}$ and $\theta_k = \theta_k^*$ into (4.9) and (4.11) yields the optimal solution to the RA problem. Optimal Lagrange multipliers are rarely available in closed form and they have to be found through numerical search, for example, by using a dual subgradient method aimed at maximizing the dual function associated with (4.8) [Ber99]. The second comment addresses the assumption of perfect SISN. Key to dealing with this issue is to acquire the instantaneous belief of the SISN, which is denoted as $H_k^m[n]$ and basically corresponds to the instantaneous distribution of the actual $h_k^m[n]$ at time n. Once $H_k^m[n]$ is available, the presence of imperfections has to be incorporated into the SU rate-power function $C(h_k^m[n], p_k^m[n])$, now $C(H_k^m[n], p_k^m[n])$, which is substituted into (4.9) and (4.10). Several alternatives to design $C(H_k^m[n], p_k^m[n])$ arise. Under a robust (worst-case) approach, the updated rate-power function would be $C(H_k^m[n], p_k^m[n]) = \min_{h_k^m[n] \in H_k^m[n]} C(h_k^m[n], p_k^m[n])$. Under a ergodic approach, it would be $C(H_k^m[n], p_k^m[n]) = \mathbb{E}_{h_k^m[n] \in H_k^m[n]}[C(h_k^m[n], p_k^m[n])]$. In any case, the basic structure (separability) of the RA remains the same.

4.4.2 RA as input for the design of the optimal sensing

The optimal solution in (4.9)-(4.10) will serve as input for the algorithms that design the optimal sensing scheme. For this reason, we introduce some auxiliary notation that will simplify the mathematical derivations in the next section. Observe that the optimal value of (4.8) can be written as

$$\mathsf{P}^{*}_{RA} = \mathbb{E}\Big[\lim_{N\to\infty}\sum_{n=0}^{N-1}(1-\gamma)\gamma^{n}\textstyle\sum_{k}\mathcal{R}_{k}(\mathbf{h}_{k}[n],B_{k}[n])\Big]$$

with

$$\mathcal{R}_{k}(\mathbf{h}_{k}[n],B_{k}[n]) := \Big[L_{SU,k}(\mathbf{h}_{k}[n]) - \theta_{k}B_{k}[n]\Big]_{+} \tag{4.13}$$

and

$$L_{SU,k}(\mathbf{h}_{k}[n]) := \max_{m}\big\{L^{m}_{SU,k}(\mathbf{h}_{k}[n])\big\}. \tag{4.14}$$

where $\mathbf{h}_{k}[n]$ is a vector containing $h^{m}_{k}[n]\ \forall m$. From the point of view of the SIPN, $\mathcal{R}_{k}(\mathbf{h}_{k}[n]B_{k}[n])$ can be viewed as the (stochastic) contribution to the Lagrangian of (4.8) at instant n when $p^{m}_{k}[n] = p^{m*}_{k}[n]$ and $w^{m}_{k}[n] = w^{m*}_{k}[n]$ for all k and m, or as an instantaneous *expected* reward indicator, where the expectation is carried over the uncertainty on $a_{k}[n]$ (recall that $B_{k}[n]$ accounts for the corresponding probabilities). Clearly, (4.13) encapsulates (via $B_{k}[n]$) the way in which the sensing affects the optimal RA and P^{*}_{RA}. Equally important, while the value of $B_{k}[n]$ is only available after making the sensing decision, the value of $L_{k}[n]$ is available before making such a decision. In other words, sensing decisions do not have an impact on $\max_{m,p}\big\{\beta^{m}C^{m}_{k}(h^{m}_{k}[n],p) - \pi^{m}p\big\}$, but only on $B_{k}[n]$. These properties will be exploited in the next section.

4.5 Optimal Sensing

The aim of this section is to leverage the results of Secs. 4.3 and 4.4 to design the optimal sensing scheme. Recall that $s_{k}[n]$ has an impact not only on $B_{k}[n]$, but also on future beliefs. This in turn implies that future sensing decisions are affected by the current decision, so that the sensing decisions across time form a string of events that has to be optimized jointly. The section is organized as follows. First, the optimal RA policy obtained in Sec. 4.4 is substituted into the original optimization problem presented in Sec. 4.3. It is shown that the design of the optimal sensing amounts to solving a set of separate unconstrained DPs (Sec. 4.5.1). Then, the solution to each of the formulated DPs (Sec. 4.6) is obtained. It turns out that the optimal sensing depends on the current channel realizations $\mathbf{h}_{k}[n]$ (SISN) and the pre-decision belief $b_{k}[n]$ (SIPN). Also, the optimization leverages a *value function* that quantifies the future reward for time slots $n' > n$ (more details will be given soon). Intuitively, a channel should be sensed if there is uncertainty on the actual channel occupancy (SIPN) and the potential reward for the secondary network is high enough (SISN). The expression for the optimal sensing provided at the end of this section will corroborate this intuition.

4.5.1 Formulating the optimal sensing problem

The aim of this section is to formulate the optimal decision problem as a standard unconstrained DP. Since the expressions for the optimal RA hold for any sensing scheme $s_{k}[n]$, the aim here is to obtain $s^{*}_{k}[n]$. Key to accomplish this are two facts. The first one is that, instead of solving (4.7), it suffices to solve for all n

$$\mathsf{P}^{*}_{DP} := \max_{\{s_{k}[n]\in\{0,1\}\}}\sum_{n=0}^{\infty}\gamma^{n}\sum_{k}\mathbb{E}\big[\mathcal{R}_{k}(\mathbf{h}_{k}[n],B_{k}[n]) - \xi_{k}s_{k}[n]\big], \tag{4.15}$$

which can be optimized separately per channel. The second fact is that the instantaneous reward $\mathcal{R}_k[n]$ depends on the sensing only through the belief. As a result, the value of $s_k[n]$ impacts the term in (4.15) corresponding to instant n via $B_k[n]$ and $\xi_k s_k[n]$, but the terms in (4.15) corresponding to instants $n' > n$ only via $b_k[n']$. The problem at hand is indeed a POMDP: current actions that improve the information about the current state have also an impact on the information about the state in future instants and thus, on future rewards.

The three main differences between (4.15) and the original formulation in (4.7) are that now: i) the only optimization variables are $\{s_k[n]\}$; ii) because the optimal RA fulfills the constraints in (4.7b) and (4.7c), the only constraints that need to be enforced are (4.7d), which simply require $s_k[n] \in \{0,1\}$; and iii) as a result of the Lagrangian relaxation of the DP, the objective has been augmented with the terms accounting for the dualized constraints.

To be rigorous and solve this sequential optimization, we leverage techniques described in [Bus10, Ch. 2]. Let us first identify the generic elements of a POMDP in (4.15). The state space is the Cartesian product of the supports of $\mathbf{h}_k[n]$ and $b_k[n]$ (replacing the partially observable state $a_k[n]$). The transition functions that describe the dynamics of the system over time are the functions (4.1) – (4.2). The action space is the set of values that $s_k[n]$ can take; there is no need to include $p_k^m[n]$ and $w_k^m[n]$ in the POMDP state space because i) their values do not impact the future states, and ii) their optimal expression, as a function of $s_k[n]$, was already found.

The POMDP in (4.15) can be split into k POMDPs without complicating constraints, each having $\mathcal{R}_k(B_k[n], \mathbf{h}[n]) - \xi_k s_k[n]$ as a reward function. Since the latter depends on $B_k[n]$, which in turn depends on $s_k[n]$, to design $s_k[n]$ we will need to compute the *a priori* expectation of the reward function conditioned on $s_k[n]$. Let us define for brevity

$$\begin{aligned}\bar{\mathcal{R}}_k(b_k[n], \mathbf{h}, s) &:= \mathbb{E}_{\mathbf{z}}\left[\mathcal{R}_k(\mathbf{h}, B_k[n])\middle| b_k[n], s\right] - \xi_k s_k[n] \\ &= \int_0^1 f_B(B|b_k, s_k)\Big[L_k(\mathbf{h}_k) - \theta_k B_k[n]\Big]_+ dB - \xi_k s_k[n].\end{aligned} \tag{4.16}$$

which accounts for the sensor-outcome uncertainty prior to sensing. Key to computing this function will be the expressions to update the belief presented in Sec. 4.2.1. Specifically, expressions in (4.1)–(4.2) describe how the future beliefs depend on the sensing decision, the current belief, and a random variable (the outcome of the sensing if the channel is indeed sensed).

Suppose momentarily that the sensing is designed as $s_k^*[n] = \arg\max_{s\in\{0,1\}} \bar{\mathcal{R}}_k(b_k[n], \mathbf{h}_k[n], s)$. Despite being computationally simple, this approach (typically referred to as *myopic policy*) ignores the impact that current sensing decisions have in future time instants.

Rather than a sequence of sensing decisions, we seek a *policy*, i.e. a mapping from the state space to the action space of the POMDP. Once the *optimal* policy is obtained, obtaining the optimal sequence of decisions is straightforward, since it boils down to evaluating at each time slot the policy for the current state.

Since (4.15) is an infinite horizon problem with $\gamma < 1$, the optimal policy is stationary and its existence is guaranteed [Pow07]. Stationarity implies that the expression for $s_k^*[n]$ is time-invariant, i.e., it depends on $(b_k[n], h_k[n])$ but not on n. Since in our problem the state information is formed by the SISN and the SIPN, $V_k(\cdot)$ should be written as $V_k(B_k[n], \mathbf{h}_k[n])$.

To account for future time instants we leverage the concept of value function [Bus10]. Conceptually, a value function quantifies the expected sum reward on channel k for all future instants. To be specific, for any sensing policy Π, there exists a value function (more specifically, a Q-function[3]) $Q_k^{\Pi}(b_k, \mathbf{h}_k, s_k)$ representing the discounted, expected reward resulting from performing sensing action s_k and following policy Π sequentially afterwards. The policy optimizing (4.15)

[3]To express the cost-to-go or value function, we prefer the Q-function form (value function corresponding to a state-action pair) over the V-function form (value function corresponding to a state [Bus10, Sec. 2.2.1]) because it facilitates the mathematical analysis and the design of an online algorithm.

is then denoted as Π^*, while the associated Q-function is denoted as $Q_k^{\Pi*}(b_k, \mathbf{h}_k, s_k)$ or simply $Q_k^*(b_k, \mathbf{h}_k, s_k)$. The existence of this stationary policy is guaranteed by the discounted formulation in (4.15). Once $Q_k^*(b_k, \mathbf{h}_k, s_k)$ is available, Π^* is

$$s_k^*[n] = \arg \max_{s\in\{0,1\}} \{Q_k^*(b_k[n], \mathbf{h}_k[n], s)\}, \tag{4.17}$$

for every k, n. The computation of the optimal sensing decision $s_k^*[n]$ at time n boils down to evaluating the Q-function in (4.17) for $s_k[n] = 0$ and $s_k[n] = 1$.

So far, the joint design has been reduced to computing $Q_k^*(b_k, \mathbf{h}_k, s_k)$. This is accomplished in the ensuing section.

4.6 Computing the optimal Q-function

Since the joint design is separable across channels and the method for computing Q_k^* is the same for every k, subindex k will be dropped for clarity. Moreover, since the optimal policy is stationary, time index n will be dropped too. Every variable refers to instant n except *prime* variables, which refer to $n + 1$.

4.6.1 Offline method

The first step to compute $Q^*(b, \mathbf{h}, s)$ is to write the Bellman equation for the optimal policy [Bus10, Eq. (2.18)]

$$Q^*(b, \mathbf{h}, s) = \bar{\mathcal{R}}(b, \mathbf{h}, s) + \mathbb{E}_{b',\mathbf{h}'}\Big[\gamma \max_{s'\in\{0,1\}} Q^*(b', \mathbf{h}', s')\Big|b, s\Big]. \tag{4.18}$$

Note that the expectation over b' implies averaging over the distribution of the sensor outcome $z_k[n]$. The first summand is the expected short-term reward conditioned to $s_k[n] = s$, while the second one is the expected long-term sum reward over all future time instants, conditioned to $s_k[n] = s$ and that every future decision is optimal. The Bellman equation expresses the condition that the Q-function must satisfy in order to be optimal (and stationary) and provides a way to compute it iteratively.

Two methods to compute the optimal Q-function will be developed. The first one is an off-line method, based on the value iteration algorithm, that relies on (4.18) to iteratively compute Q^*. More specifically, the model-based Q-iteration algorithm [Bus10, Sec. 2.3.1] is employed. With ℓ denoting an iteration index, this amounts to computing for every $(b, \mathbf{h}, s)$

$$Q_{\ell+1}(b, \mathbf{h}, s) = \bar{\mathcal{R}}(b, \mathbf{h}, s) + \mathbb{E}_{b',\mathbf{h}}\Big[\gamma \max_{s'\in\{0,1\}} Q_\ell(b', \mathbf{h}', s')\Big|b, s\Big]. \tag{4.19}$$

To reduce the cost of computing $Q_\ell(b, \mathbf{h}, s)$, we will define a function with a smaller dimensionality, such that $Q_\ell(b, \mathbf{h}, s)$ can be obtained from it. Specifically, let define

$$\mathcal{Q}_\ell(b, s) := \mathbb{E}_{b',\mathbf{h}'}[\max_{s'\in\{0,1\}} Q_\ell(b', \mathbf{h}', s')|b, s]$$

and rewrite (4.18) as

$$Q_\ell(b, \mathbf{h}, s) = \bar{\mathcal{R}}(b, \mathbf{h}, s) + \gamma \mathcal{Q}_\ell(b, s). \tag{4.20}$$

Substituting (4.20) into (4.17) yields

$$s^* = \arg \max_{s\in\{0,1\}} \left\{\bar{\mathcal{R}}(b, \mathbf{h}, s) + \gamma \mathcal{Q}^*(b, s)\right\}. \tag{4.21}$$

Substituting (4.20) into both sides of (4.19) and simplifying yields

$$\mathcal{Q}_{\ell+1}(b,s) = \mathbb{E}_{b',\mathbf{h}'}\Big[\max_{s'\in\{0,1\}} \big\{\bar{\mathcal{R}}(b',\mathbf{h}',s') + \gamma \mathcal{Q}_\ell(b',s')\big\}\big|b,s\Big]. \tag{4.22}$$

Note that $\mathcal{Q}_\ell$ now depends only on two *scalar* variables. Provided that the distribution of B conditioned on (b,s) is known, and that $b_k[n+1]$ depends deterministically on $B_k[n]$ [cf. (4.1)], the iterate in (4.22) can be rewritten as

$$\mathcal{Q}_{\ell+1}(b,s) = \mathbb{E}_{\mathbf{h}'}\Big[\int_{B=0}^{1} f_B(B|b,s) \max_{s'\in\{0,1\}} \big\{\bar{\mathcal{R}}(\mathcal{P}(B),\mathbf{h}',s') + \gamma \mathcal{Q}_\ell(\mathcal{P}(B),s')\big\}\, dB\Big]. \tag{4.23}$$

The Q-iteration algorithm initializes the Q-function at an arbitrary Q_0 and at each iteration ℓ updates the Q-function indirectly by using (4.23) to update $\mathcal{Q}_\ell$. For simplicity we choose $\mathcal{Q}_0 = 0$ which corresponds to $Q_0(b,\mathbf{h},s) = \bar{\mathcal{R}}(b,\mathbf{h},s)$. Convergence of $\mathcal{Q}_\ell$ to $\mathcal{Q}^*$ when $\ell \to \infty$ can be shown based on the fact that the mapping defined in (4.19) is contractive with factor $\gamma_Q < 1$ in the infinity norm [Bus10, Ch. 2]. Summarizing, an iterative off-line method has been proposed to compute the Q-function. The computational cost has been lowered by splitting Q into two terms [cf. (4.20)]. The first one can be computed directly, while the second one still has to be computed iteratively, but has much smaller dimensionality. To implement this method, the following information was assumed to be known: (i) the Markov matrix for the primary occupancy $\mathbf{P}_k$; (ii) the distribution of z conditioned on the SIPN; and (iii) the distribution of $\mathbf{h}$.

4.6.2 Online method

The second method to compute the Q-function will further reduce the computational complexity, and bypass the need to know the distribution of $\mathbf{h}$. Stochastic approximation is leveraged to design an online algorithm. With $\hat{\mathcal{Q}}_n(b,s)$ representing the online approximation of $\mathcal{Q}^*(b,s)$, the stochastic update is

$$\begin{aligned}\hat{\mathcal{Q}}_n(b,s) =& (1-\alpha)\hat{\mathcal{Q}}_{n-1}(b,s) \\ &+ \alpha \int_0^1 f_B(B|b,s) \max_{s'\in\{0,1\}} \big\{\bar{\mathcal{R}}(\mathcal{P}(B),\mathbf{h}[n],s') + \gamma \hat{\mathcal{Q}}_{n-1}(\mathcal{P}(B),s')\big\}\, dB,\end{aligned} \tag{4.24}$$

where $\alpha \in (0,1]$ is a learning rate. The proposed rule is a variant of the Q-learning algorithm. More specifically, it is a model-free value iteration [Bus10, Sec. 2.3.2]. The proposed update is model-free for $\mathbf{h}_k$, but it is still model-based for b_k. The main advantages of this mixed algorithm are that it does not need to know the distribution of $\mathbf{h}_k[n]$, makes the system robust to channel gain non-stationarities, and avoids the need for exploration that affects Q-learning in more general situations. The price to pay is a small loss of optimality which is typically present in stochastic and online algorithms.

The presented Q-iteration (off-line) and Q-learning (online) algorithms will be evaluated using numerical simulations in the next section.

Let us summarize the most relevant properties (several of them were already pointed out) of the optimal sensing policy in (4.17): i) it can be found separately for each of the channels; ii) since it amounts to a decision problem, we only have to evaluate the long-term aggregate reward if $s_k[n] = 1$ (the channel is sensed at time n) and that if $s_k[n] = 0$ (the channel is not sensed at time n), and make the decision which gives rise to a higher reward; iii) the reward takes into account not only the sensing cost but also the utility and QoS for the SUs (joint design); iv) the sensing at instant n is found as a function of both the instantaneous and the future reward (the problem is a DP); vi) the instantaneous reward depends on both the current SISN and the current SIPN, while the future reward depends on the current SIPN and not on the current SISN; and vii)

to quantify the future reward, the Q-function is required. The input of this function is the SIPN. Additional insights on the optimal sensing will be given in Sec. 4.8.

To conclude this section, some comments regarding the computational complexity associated with the optimal sensing policy are in order. It is assumed here that the optimal Lagrange multipliers are known. Since there are K channels and computing $\bar{\mathcal{R}}(\cdot)$ entails a maximization over M terms [cf. (4.12)],the per-slot complexity associated with the RA task is $O(KM)$ [cf. (4.12)]. Regarding the sensing policy, its complexity depends on whether the form $Q(\dot{)}$ or $\mathcal{Q}(\dot{)}$ is adopted. If $Q^*(\cdot)$ is available, computing s^* is immediate; note, however, that the form $\mathcal{Q}(\dot{)}$ is preferable rather than $Q(\dot{)}$. If $\mathcal{Q}^*(\dot{)}$ is available the cost of computing s^* is similar to that of computing $\bar{\mathcal{R}}(\cdot)$, i.e. comparable to the RA complexity. Regarding the computational cost of calculating $\mathcal{Q}^*(\dot{)}$, we will first analyze the complexity of computing it offline, and then the complexity of the online iterations. Note that the functional estimation involved requires an approximation: we adopt here the simplest approach based on uniform sampling and nearest-neighbor interpolation. Since the offline iteration requires numerical integration over all possible $\mathbf{h}$, the complexity of a single iteration of (4.23) is $O(N_b^2 N_h KM)$ where N_b is the number of samples of the interval $[0,1]$ as the support of b, and N_h is the number of samples of $\mathbf{h}$ used to estimate the expectation. This results into an offline phase of $O(N_b^2 N_h KM)$ iterations, and $O(KM)$ complexity per slot. The online iteration does not require any off-line computation, but runs an iteration of (4.24) at every time slot, incurring $O(N_b^2 KM)$ cost per slot. Considering that an offline Q-iteration algorithm usually takes a large number of iterations to converge to an optimal solution, the moderate increase in the per-slot complexity of the online algorithm is worthwhile for most practical scenarios.

4.6.3 Sensing cost

To account for the cost of sensing a given channel, the additive and constant cost ξ_k was introduced. So far, we considered that the value of ξ_k was pre-specified. However, the value of ξ_k can be tuned to represent physical properties of the CR. Let us provide three illustrative examples. Example 1: If the NC invests power P_k^{NC} in sensing channel k, then ξ_k can be set to $\xi_k = \pi^{NC} P_k^{NC}$, where π^{NC} stands for the Lagrange multiplier associated with a long-term power constraint on the NC (see, e.g. [LR14a] for a related work that uses this criterion). Example 2: Consider a setup for which the long-term rate of sensing is limited; mathematically, this can be accomplished by imposing that $\mathbb{E}[\lim_{N\to\infty} \sum_{n=0}^{N-1} (1-\gamma)\gamma^n s_k[n]] \leq \eta$, where η represents the maximum sensing rate (say 10%). Let ρ_k be the Lagrange multiplier associated with such a constraint, in this scenario ξ_k should be set to $\xi_k = \rho_k$. Example 3: Suppose now that to sense the channel, the NC needs a fraction ψ_k of total duration of the slot (that, otherwise, would have been used for SU transmissions). In this scenario, $\xi_k[n] = \psi_k L_k[n]$ (time-varying opportunity cost). Linear combinations and stochastic versions of any of those costs are possible too. Similarly, if a collaborative sensing scheme is assumed, aggregation of costs across users can also be considered.

4.7 Numerical results

Numerical experiments to corroborate the theoretical findings and gain insights on the optimal policies are implemented in this section. Since an *RA scheme* similar to the one presented in this chapter was analyzed in Chapter 3 and also in [Mar12a], the focus is on analyzing the properties of the optimal *sensing scheme*.

The experiments are grouped into two test cases. In the first one, we compare the performance of our algorithms with that of other existing (suboptimal) alternatives. Moreover, we analyze the behavior of the sensing schemes and assess the impact of variation of different parameters (correlation of the PUs activity, sensing cost, sensor quality, and average SNR). In the second test case, we provide a graphical representation of the sensing functions in the form of two-dimensional

decision maps. Such representation will help us to understand the behavior of the optimal schemes.

The parameters for the default test case are listed in Table 4.2. We consider $K = 4$ channels, each of them with different values for the sensor quality, the sensing cost and the QoS requirements. In most cases, the value of $\check{o}_k$ has been chosen to be larger than the value of P_k^{MD} (so that the cognitive diversity can be effectively exploited), while the values of the remaining parameters have been tailored such that the test case yields illustrative results. The secondary link gains are Rayleigh distributed, their average SNR is 5dB, and the frequency selectivity is such that the gains are uncorrelated across channels. We consider $M = 4$ SUs, and their average power limits are set to $[\check{p}_1, \check{p}_2, \check{p}_3, \check{p}_4] = [20, 16, 18, 10]$. The discount factor γ is set to 0.95, so that the autocorrelation function of $a_k[n]$ and the implemented window have comparable length. The multipliers associated with the system parameters have been calculated by gradient descent, using a Monte Carlo approach to average over the channel processes. Unless otherwise stated, the remaining parameters are set to one.

Table 4.2: Parameters of the system under test.

k	P_k^{FA}	P_k^{MD}	$\mathbf{P}_k$	ξ_k	$\check{o}_k$
1	0.09	0.08	[0.95, 0.05; 0.02, 0.98]	1.00	0.30
2	0.09	0.08	[0.95, 0.05; 0.02, 0.98]	1.80	0.05
3	0.05	0.03	[0.95, 0.05; 0.02, 0.98]	1.00	0.10
4	0.05	0.03	[0.95, 0.05; 0.02, 0.98]	1.80	0.10

Test case 1: Optimality and performance analysis. The objective here is twofold. First, we want to numerically demonstrate that our schemes are indeed optimal. Second, we are also interested in assessing the loss of optimality incurred by suboptimal schemes with low computational burden. Specifically, the optimal sensing scheme is compared with the three suboptimal alternatives described next. A) A myopic policy, which is implemented by setting $\mathcal{Q}_k(b, s) = 0 \ \forall (b, s)$, equivalent to the greedy sensing and RA technique proposed in [Wan11], which is optimal for a CR scenario whose PU activity is not time-correlated, but is suboptimal for the test cases at hand. B) A policy that replaces the infinite horizon value function with a "horizon-1" value function. (i.e., it makes the sensing decision at time n considering the expected reward for instants n and $n+1$). C) A rule-of-thumb sensing scheme implementing the simple (separable) decision function: $s_k[n] = \mathbb{1}_{\{L_k[n] \in [\xi_k, \theta_k - \xi_k]\}} \mathbb{1}_{\{B_k[n] \in [\mathbf{b}_k^S(A_k,0), \mathbf{b}_k^S(A_k,1)]\}}$. This simple policy senses the channel only if two conditions hold: a) the IRI $L_k[n]$ is neither too low nor too high (so that the scheduling decision is uncertain), and b) the instantaneous belief $B_k[n]$ is close to the long-term belief A_k (so that the information given by the observation will modify the value of $B_k[n]$ substantially).

To evaluate the effect of the variation of different parameters, we run 4 experiments. In each experiment, one of the parameters is swept and the rest remain constant (with the values listed in Table 4.2): a) average PU state transition time (by modifying $\mathbf{P}_k$); b) sensing cost; c) probability of sensing error (by modifying P_k^{FA} and P_k^{MD}) and d) average SNR. Results are plotted in Fig. 4.4. The slight lack of monotonicity observed in the curves is due to the fact that simulations have been run using a Monte-Carlo approach. As expected, the optimal sensing scheme achieves the best performance for all test cases. Moreover, Figs. 4.4a and 4.4b reveal that the "horizon-1" value function approximation constitutes a good approximation to the optimal value function in two cases: i) when the expected PU transition time is short (low time correlation) and ii) when the sensing cost is relatively small. The performance of the myopic policy is shown to be far from the optimal. This finding is in disagreement with the results obtained for simpler models in the opportunistic spectrum access literature [Zha07b] where it was suggested that the myopic policy could be a good approximation to solve the associated POMDP efficiently. This is probably due to the CR models considered being substantially different (the RA schemes in this chapter are

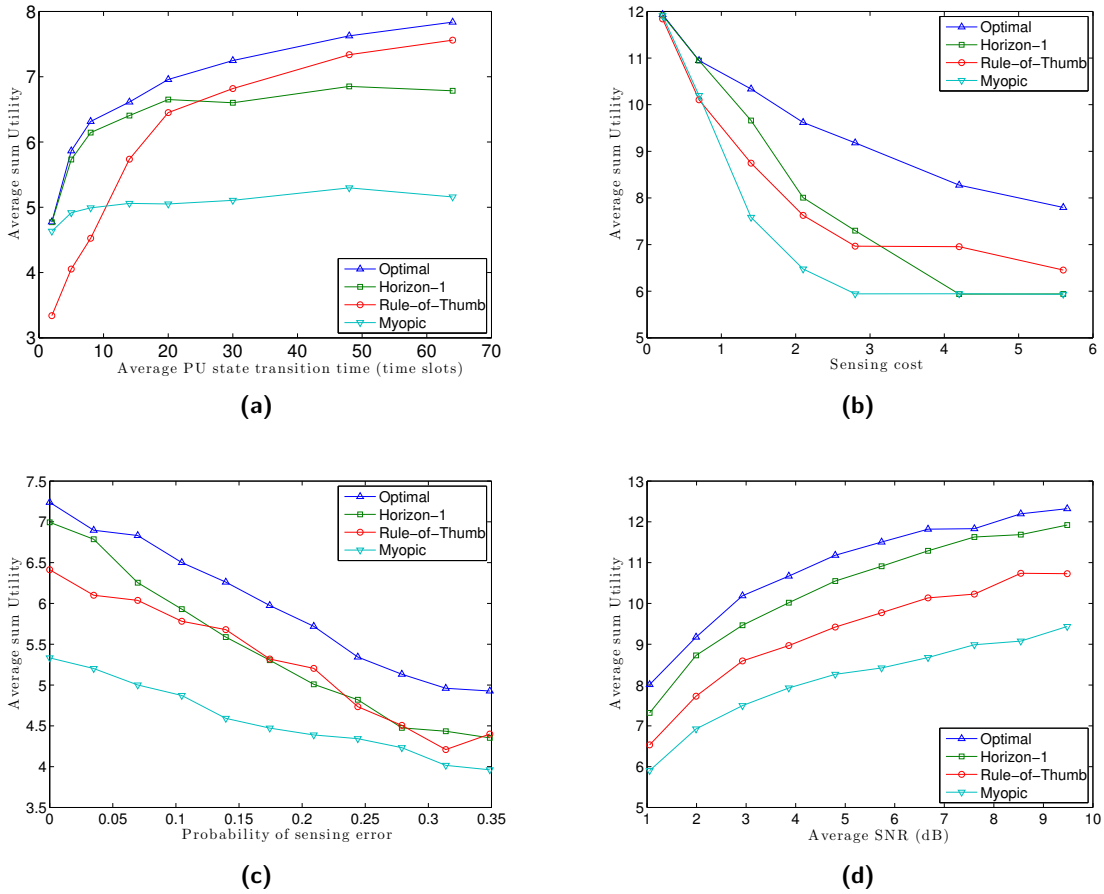

Figure 4.4: Performance of the optimal scheme vs. some suboptimal schemes for variations in (a) expected primary transition time; (b) sensing cost; (c) probability of error; (d) average SNR.

more complex and the interference constraints are formulated differently). In fact, the only cases where the myopic policy seems to approximate the optimal performance are: i) if $\xi_k \to 0$, this is unsurprising because in that case the optimal policy is to sense at every time instant; and ii) if the PUs activity is not correlated across time (which was the assumption in [Wan11]).

Fig. 4.4c suggests that the benefits of implementing the optimal sensing policies are stronger when sensors are inaccurate. In other words, the proposed schemes can help to soften the negative impact of deploying low-quality (low-cost) sensing devices. Finally, results in Fig. 4.4d also suggest that changes in the average SNR between SU and NC, have similar effects on the performance of all analyzed schemes.

Test case 2: Sensing decision maps. To gain insights on the behavior of the optimal sensing schemes, Fig. 4.5 plots the sensing decisions as a function of $B_k[n]$ and $L_{SU,k}[n]$. Simulations are run using the parameters for the default test case (see Table 4.2) and each subplot corresponds to a different channel k. Since the domain of the sensing decision function is two dimensional, the function itself can be visually displayed as a decision map. Two main regions are identified, one corresponding to the pairs $(B_k[n], L_{SU,k}[n])$ which give rise to $s_k[n] = 1$, and one corresponding to the pairs giving rise to $s_k[n] = 0$. Moreover, the region where $s_k[n] = 0$ is split into two subregions, the first one corresponding to $\sum_m w_k^m[n] = 1$ (i.e., when there is a user accessing the channel) and the second one when $\sum_m w_k^m[n] = 0$ (i.e., when the system decides that no user will access the channel). Note that for the region where $s_k[n] = 1$, the access decision basically depends on the outcome of the sensing process $z_k[n]$ (if fact, it can be rigorously shown that

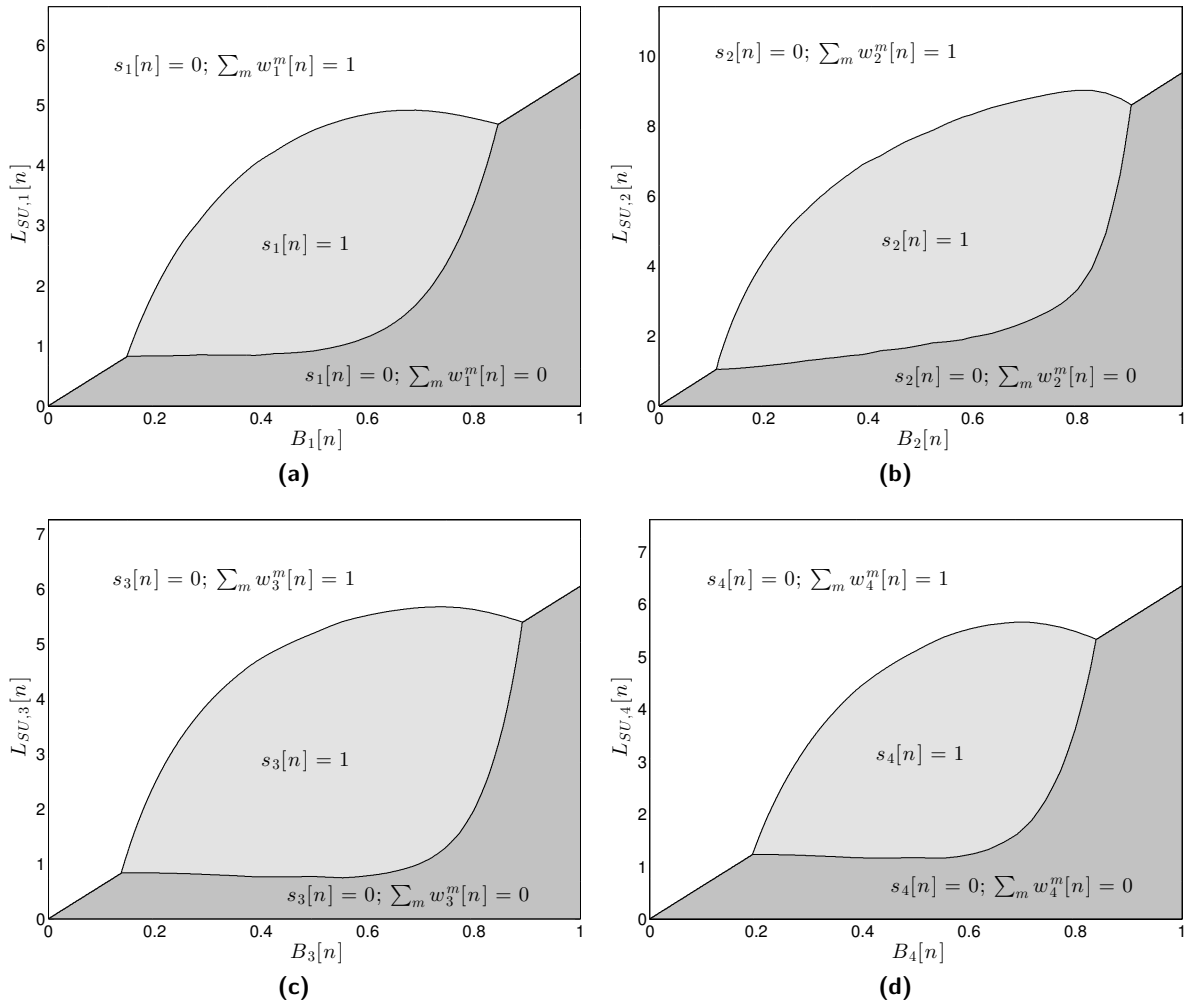

Figure 4.5: Decision maps (regions) for the four channels in the default test case (see Table 4.2). The light gray area in the center corresponds to the sensing decision.

$\sum_m w_k^m[n] = 1$ if and only if $z_k[n] = 1$).

Upon comparing the different subplots, it is clear that the size and shape of the $s_k[n] = 1$ region depend on $\mathbf{P}_k$, P_k^{FA}, P_k^{MD}, ξ_k, and $\check{o}_k$. For example, the simulations reveal that channels with stricter interference constraint need to be more frequently sensed and thus the sensing region is larger: Fig. 4.5a vs. Fig. 4.5b. They also reveal that if the sensing cost ξ_k increases, then the sensing region becomes smaller: Fig. 4.5c vs. Fig. 4.5d. This is an expected result: more expensive sensing implies more conservative sensing decisions.

Test case 3: Stochastic update of the Q-function. This experiment analyzes the performance of the online stochastic policy (4.24) for a non-stationary scenario. In particular, we consider a single-channel system and set the simulation time to 50000 slots. During the first 25000 slots, the channel conditions are those of $k = 1$ for the previous experiment. At $n = 25000$, the channel conditions switch to those of $k = 2$. Fig. 4.6 depicts the achieved utility, averaged using a rectangular window of length 9000.

The stochastic policy performs better than the myopic one and close to the optimal off-line policy corresponding to each of the two cases considered. Utility gaps and speed of converge depend heavily on the simulated scenario and, especially, on the correlation of the primary user behavior (stronger correlation brings slower convergence).

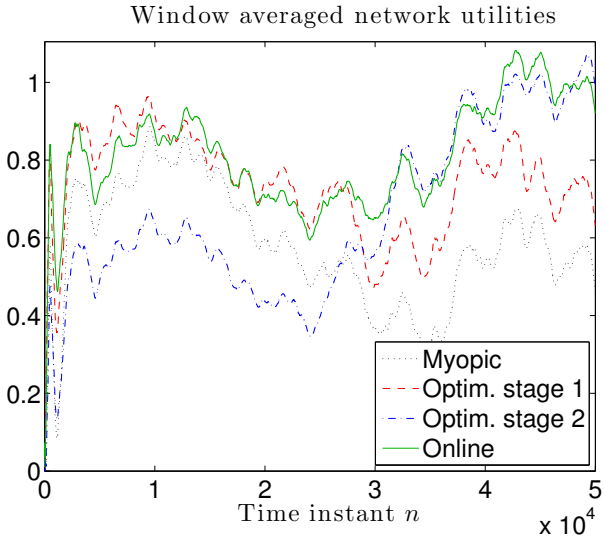

Figure 4.6: Performance comparison of the optimal and stochastic iterates.

4.8 Summarizing conclusions and alternative CR models

The aim of this chapter was to design jointly optimal RA and sensing schemes for an interweave CR with multiple primary and secondary users. Since, due to time correlation in the SIPN, sensing decisions were coupled across time, the problem fell into the class of DP (more precisely, POMDP). To address the complexity typically associated with DP, both the objective and the QoS constraints were formulated as long-term (discounted) time averages (which are less restrictive than their short-term counterparts and, hence, give rise to a better global objective), and the constraints were handled using duality and dual decomposition. Additionally, a discounted, infinite time-horizon formulation was chosen for the DP (giving rise to stationary value functions). A two-step approach that solved first for the optimal RA for *any* sensing scheme and, then, solved for the optimal sensing, was implemented. As a result, the DP finally solved had a state space much smaller than that of the original formulation. In particular, the optimization was separable across channels, partially separable across SUs, and separable across time. Both the value functions and the multipliers accounted for the effect of sensing and RA across time. Provided that the stationary value function and multipliers were available (they are found off-line via numerical search during the initialization phase), the online implementation of the optimal schemes entailed very low computational complexity. Numerical results showed that both optimal and near-optimal policies corresponding to the formulated DP perform significantly better than myopic policies, conversely to the model studied in [Zha07b].

These results have been extended in two additional works [LR13, LR14a]. The first one extends the interweave CR model studied in this chapter, but instead of considering only binary (0/1) sensing decisions, the CR also has to decide how many samples are taken from the received PU signal. As a result, a finer control over the sensing cost (which is proportional to the number of samples taken) and the potential improvement in the belief is implemented.

The second work that capitalizes on these results addresses an *underlay* CR. As the interference SUs cause to the PUs will depend on the fading gains of the CR-to-PU channels, a POMDP approach (similar to that used in this chapter) is proposed to keep track of them. The nature of such gains requires that the sensing task is carried out by the SUs (instead of the NC), increasing the complexity of the sensing decision. To overcome such complexity, a greedy algorithm which is scalable with the number of SUs is proposed and evaluated to reach a near-optimal sensing policy.

Chapter 5

Two-timescale dispatch of power distribution networks

Smart distribution grids should efficiently integrate stochastic renewable resources while effecting voltage regulation. The design of energy management schemes is challenging, one of the reasons being that energy management is a multistage problem where decisions are not all made at the same timescale and must account for the variability during real-time operation. The joint dispatch of slow- and fast-timescale controls in a distribution grid is considered here. The substation voltage, the energy exchanged with a main grid, and the generation schedules for small diesel generators have to be decided on a slow timescale; whereas optimal photovoltaic inverter set points are found on a more frequent basis. While inverter and looser voltage regulation limits are imposed at all times, tighter bus voltage constraints are enforced on the average or in probability, thus enabling more efficient renewable integration. Upon reformulating the two-stage grid dispatch as a stochastic convex-concave problem, two distribution-free schemes are put forth. An average dispatch algorithm converges provably to the optimal two-stage decisions via a sequence of convex quadratic programs. Its non-convex probabilistic alternative entails solving two slightly different convex problems and is numerically shown to converge. Numerical tests on a real-world distribution feeder verify that both novel schemes yield lower costs over competing alternatives.

5.1 Introduction

With increasing renewable generation, energy management of power distribution grids is becoming a computationally challenging task. Solar energy from photovoltaic (PV) units can change significantly over one-minute intervals. The power inverters found in PV units can be commanded to curtail active power generation or adjust their power factor within seconds [Liu08], [Car08]. At a slower timescale, distribution grid operators exchange energy with the main grid hourly or on a 10-minute basis, and may experience cost penalties upon deviating from energy market schedules [Var11]. Moreover, voltage regulation equipment and small diesel generators potentially installed in microgrids respond at the same slower timescale. As a result, comprehensive designs to optimize such diverse tasks call for *multistage* smart grid dispatch solutions.

Spurred by demand-response programs and the use of PV inverters to accomplish various grid tasks [Tur11], *single-stage* dispatch schemes for distribution grids have been an active area of research. Power inverters can be controlled using localized rules for voltage regulation, see e.g., [Zha13a, Kek15c, Bol15, Zhuar]. Assuming two-way communication between buses and the utility operator, dispatching a distribution system can be posed as an optimal power flow (OPF) problem. Centralized schemes use nonlinear program solvers [Pau11]; or rely on convex relaxations of the full ac model of balanced [Far13b, Gan15], or unbalanced grids [Dal14]. Distributed solvers with reduced computational complexity have been devised in [Dal13, Zha15, Penar].

Table 5.1: Summary of the main symbols in Chapter 5

Symbol	Meaning
t	Fast-timescale time slot index
t	Number of slots in a slow-timescale period
n, N	Line/bus index, total number of lines
$\mathbf{v}$	Bus voltage squared magnitudes
v_0^a	Voltage squared magnitude at substation
$p_{0,t}$	active power injected at substation
p_0^a	active power bought in advance
$p_{0,t}^\delta$	active power bought in actual time
$\mathbf{p}, \mathbf{q}$	active/reactive net power injections
$\mathbf{p}_t^l, \mathbf{q}_t^l$	active/reactive demand load injections
$\mathbf{p}^d$	active injections from diesel generators
$\mathbf{p}_t^r, \mathbf{q}_t^r$	active/reactive injections from PV inverters
$\overline{\mathbf{p}}_t^r$	Available active power at inverters
β	Power block (advance) market price
γ_b, γ_s	Actual-time buy/sell prices
$C^t(\cdot)$	Cost of energy traded in real time
$C_{\mathrm{D}}(\cdot)$	Diesel generation cost
φ_0	Parameter limiting power factor at the substation
φ_n^d	Parameter limiting power factor at diesel gen. n
φ_n	Parameter limiting power factor at inverter n
$\overline{s}_n$	Maximum apparent power at inverter n
$\overline{S}_n$	Maximum apparent power flowing at line n
$\mathcal{V}_A$	Narrower region for voltages
$\mathcal{V}_B$	Broader region for voltages
$\boldsymbol{\xi}$	collects all random variables $\{\mathbf{p}^l, \mathbf{q}^l, \overline{\mathbf{p}}^r, \gamma_s, \gamma_b\}$

Nevertheless, the efficient and secure operation of distribution grids involves decisions at different timescales. A dynamic programming approach for a two-stage dispatch is suggested in [Far12]: The taps of voltage regulators are set on a slow timescale and remain fixed for consecutive shorter time slots over which elastic loads are dispatched; yet the flexibility of loads is assumed known *a priori*. Alternatively, centrally computed OPF decisions can be communicated to buses at a slow timescale, while on a faster timescale, PV power electronics are adjusted to optimally track variations in renewable generation and demand [Dor14, Dalar]. Relying on approximate grid models, the latter schemes yield a fully localized real-time implementation. However, they presume smooth system transitions and dispatch slow-responding units for a single deterministic fast-timescale scenario.

Multistage dispatching under uncertainty is routinely used in transmission systems and microgrids [Con10]. Robust approaches find optimal slow-timescale decisions for the worst-case fast-timescale outcome; see [Zha13b] and references therein. To avoid the conservativeness of robust schemes, probabilistic approaches postulate a probability density function (pdf) for demand, wind

generation, and system contingencies to find day-ahead grid schedules [Bou05], [Bie14]. The risk-limiting dispatch framework adjusts multistage decisions as the variance of the random variables involved decreases while approaching actual time [Var11]. Decisions can be efficiently calculated only for convenient pdfs for a network-constrained risk-limiting dispatch under additional transmission congestion assumptions [Zha14a]. As a third alternative, stochastic sample approximation approaches yield optimal slow-timescale decisions using samples drawn from the postulated pdf; see e.g., [Bou08], [Zha14b].

Returning to distribution grids, PV inverters could be overloaded sporadically in time and across buses to accommodate solar fluctuations and prevent overvoltages [Mou13]. The spatiotemporal overloading of power system components (such as inverters, bus voltages, line flows) could thus constitute an additional means for integrating renewables in smart grids. Nonetheless, ensuring that overloading occurs sparingly couples decisions across time. The single-stage scheme of [Wanar] finds optimal PV set points while limiting time averages of overloaded quantities. The latter approach has been also adopted in [LR15] for dispatching a transmission system under a day-ahead/real-time market setup under load shedding.

Jointly dispatching slow- and fast-timescale distribution grid resources under *average* or *probabilistic* constraints over fast-timescale decisions is considered here. Our contributions are three-fold. First, using an approximate grid model, the expected cost over a slow control period is minimized while inverter and looser bus voltage constraints are satisfied at all fast-timescale slots. Further, tighter voltage limits are enforced either on the average or in probability across successive fast-timescale slots (see Sec. 5.2). Two-stage grid dispatch is formulated as a convex-concave optimization in Sec. 5.3. Second, adopting a stochastic saddle-point approximation scheme from [Nem09], the provably convergent algorithm of Sec. 5.4 finds the optimal slow-timescale decisions in the case of average constraints. Third, for the case of non-convex *probabilistic* constraints, an algorithm solving two similar convex problems for each fast-timescale period is put forth in Sec. 5.5. Albeit the related expected recourse function enjoys zero-duality gap [Rib10b], the overall two-stage dispatch is not convex-concave; and the algorithm's performance is only numerically validated. Both algorithms require only samples rather than pdfs of loads and solar generation, and involve solving simple convex quadratic programs. Numerical tests in Sec. 5.6 on a 56-bus feeder corroborate the validity of our findings [1].

5.2 Problem Formulation

Consider a distribution grid whose energy needs are procured by distributed renewable generation, distributed conventional (small diesel) generators, and the main grid. The distribution grid operator aims at serving load at the minimum cost while respecting voltage regulation and network constraints. Energy is exchanged with the main grid at whole-sale electricity prices through the feeder bus. To effectively integrate stochastic renewable generation, the focus here is on short-term grid dispatch. To that end, the distribution grid is operated at two timescales: a slower timescale corresponds to 5- or 10-min real-time energy market intervals, while the inverters found in PVs are controlled at a faster timescale of say 10-sec intervals. One period of the slower timescale is comprised by T faster time slots indexed by $t = 1, \ldots, T$.

The grid is operated as a radial network with $N+1$ buses rooted at the substation bus indexed by $n = 0$. The distribution line feeding bus n is also indexed by n for $n = 1, \ldots, N$. Let $p_{n,t}$ and $q_{n,t}$ denote respectively the net active and reactive power injections at bus n and slot t; the N-dimensional vectors $\mathbf{p}_t$ and $\mathbf{q}_t$ collect the net injections at all buses except for the substation. Diesel generators are dispatched at the slower timescale to generate $\mathbf{p}^d$ throughout the subsequent

[1] Regarding *notation*, the convention of lower-(upper-)case boldface letters denoting column vectors (matrices) holds in this chapter with the exception of the power flow vectors, which are uppercase. Calligraphic letters are used to denote sets.

T slots at unit power factor. During slot t, PVs can contribute solar generation up to $\overline{\mathbf{p}}_t^r$ that is modeled as a random process. Smart inverters perform active power curtailment and reactive power compensation by following the set points $\mathbf{p}_t^r$ and $\mathbf{q}_t^r$ commanded by the utility operator. Load demands $\mathbf{p}_t^l$ and $\mathbf{q}_t^l$ are also modeled as random processes. To simplify the exposition, $(\mathbf{p}_t^l, \mathbf{q}_t^l)$ are assumed inelastic and known at the beginning of slot t; although elastic loads can be incorporated without any essential differences. The operator buys a power block p_0^a from the main grid at the slow timescale, which can be adjusted to $p_{0,t} := p_0^a + p_{0,t}^\delta$ in actual time.

Voltage regulation is effected by controlling (re)active power injections at slot t. Let $v_{n,t}$ denote the squared voltage magnitude at bus n and slot t, and $\mathbf{v}_t$ the vector collecting $\{v_{n,t}\}_{n=1}^N$. The substation voltage v_0^a is controlled at the slower timescale [Far12], while voltage magnitudes at all buses must adhere to voltage regulation standards, e.g., ANSI C84.1 and EN50160 in [ans11], [en504]. These standards differentiate between a narrower voltage regulation range denoted here by $\mathcal{V}_A$ in which voltages should lie most of the time; and a wider range $\mathcal{V}_B$ (with $\mathcal{V}_A \subset \mathcal{V}_B$) whom voltages should not exceed at any time. One of the goals of this work is to leverage this flexibility to design dispatch schemes that: i) guarantee that voltages lie in $\mathcal{V}_B$ at all times, while ii) they belong to $\mathcal{V}_A$ in a stochastic fashion. To this end, two alternative schemes are presented, the difference between them being how constraint ii) is formulated. The first scheme guarantees that the *average* voltage lies in $\mathcal{V}_A$, whereas the second one maintains the *probability* of under-/over-voltage at a specified low value.

5.2.1 Grid modeling

To capture voltage and network limitations, the distribution grid is captured by the approximate linear distribution flow (LDF) model, which is briefly reviewed next [Bar89]. Let $\mathbf{r}$ and $\mathbf{x}$ be accordingly the vectors of line resistances and reactances across lines. Define also the branch-bus incidence matrix $\tilde{\mathbf{A}} \in \mathbb{R}^{N\times(N+1)}$ whose (i,j)-th entry is

$$\tilde{A}_{ij} = \begin{cases} +1 & , \text{ if } j-1 \text{ is the source bus of line } i \\ -1 & , \text{ if } j-1 \text{ is the destination bus of line } i \\ 0 & , \text{ otherwise.} \end{cases} \tag{5.1}$$

Partition $\tilde{\mathbf{A}}$ into its first column and the *reduced* branch-bus incidence matrix $\mathbf{A}$ as $\tilde{\mathbf{A}} = [\mathbf{a}_0 \ \mathbf{A}]$. Ignoring line losses, the LDF model asserts that the vectors of active and reactive line power flows at time t can be approximated by

$$\mathbf{P}_t = \mathbf{F}^\top \mathbf{p}_t \text{ and } \mathbf{Q}_t = \mathbf{F}^\top \mathbf{q}_t \tag{5.2}$$

where $\mathbf{F} := \mathbf{A}^{-1}$. Moreover, the squared voltage magnitudes can be expressed as [Bar89], [Far13a], [Kek15b]

$$\mathbf{v}_t = 2\mathbf{R}\mathbf{p}_t + 2\mathbf{X}\mathbf{q}_t + v_0^d \mathbf{1} \tag{5.3}$$

where $\mathbf{R} := \mathbf{F}\mathrm{dg}(\mathbf{r})\mathbf{F}^\top$ and $\mathbf{X} := \mathbf{F}\mathrm{dg}(\mathbf{x})\mathbf{F}^\top$. Let us define the voltage regulation regions

$$\mathcal{V}_A := \{\mathbf{v} : \ \underline{v}_A \mathbf{1} \le \mathbf{v} \le \overline{v}_A \mathbf{1}\} \tag{5.4a}$$

$$\mathcal{V}_B := \{\mathbf{v} : \ \underline{v}_B \mathbf{1} \le \mathbf{v} \le \overline{v}_B \mathbf{1}\} \tag{5.4b}$$

with $\overline{v}_B \ge \overline{v}_A$ and $\underline{v}_B \le \underline{v}_A$. Compliance with $\mathcal{V}_A$ can be imposed either on the average as $\mathbb{E}_t[\mathbf{v}_t] \in \mathcal{V}_A$, or in probability as $\Pr\{\mathbf{v}_t \in \mathcal{V}_A\} \ge 1-\alpha$ for some small α. Either way, safe grid operation requires that $\mathbf{v}_t \in \mathcal{V}_B$ at all times t. Within the optimization horizon, the random processes involved (demand and renewable generation) can be assumed ergodic, i.e., their time averages converge to their ensemble averages. For this reason, voltage constraints pertaining to

$\mathcal{V}_A$ will be referred to as *ergodic*.

According to (5.2), if $\mathbf{f}_n$ is the n-th column of $\mathbf{F}$, the squared power flow on line n can be written as $P_{n,t}^2 = \mathbf{p}_t^\top \mathbf{f}_n \mathbf{f}_n^\top \mathbf{p}_t$ and $Q_{n,t}^2 = \mathbf{q}_t^\top \mathbf{f}_n \mathbf{f}_n^\top \mathbf{q}_t$. Imposing the upper limit $\overline{S}_n$ on the apparent flow on line n is thus expressed as the convex quadratic constraint

$$\mathbf{p}_t^\top \mathbf{f}_n \mathbf{f}_n^\top \mathbf{p}_t + \mathbf{q}_t^\top \mathbf{f}_n \mathbf{f}_n^\top \mathbf{q}_t \leq \overline{S}_n^2. \tag{5.5}$$

Assuming voltage magnitudes to be close to unity, active power losses can be approximated as $\sum_{n=1}^N r_n(P_{n,t}^2+Q_{n,t}^2)$ [Sul14], which from (5.2)–(5.3), can be equivalently expressed as $\mathbf{P}_t^\top \mathrm{dg}(\mathbf{r})\mathbf{P}_t + \mathbf{Q}_t^\top \mathrm{dg}(\mathbf{r})\mathbf{Q}_t = \mathbf{p}_t^\top \mathbf{R}\mathbf{p}_t + \mathbf{q}_t^\top \mathbf{R}\mathbf{q}_t$. Thus, the active power injection at the substation is approximately

$$p_{0,t} = -\mathbf{1}^\top \mathbf{p}_t + \mathbf{p}_t^\top \mathbf{R}\mathbf{p}_t + \mathbf{q}_t^\top \mathbf{R}\mathbf{q}_t \tag{5.6}$$

Regarding smart inverters, the tuple $(p_{n,t}^r, q_{n,t}^r)$, which denotes the power injection from the inverter located on bus n at slot t, should belong to the feasible set

$$\Omega_{n,t} := \big\{(p_{n,t}^r, q_{n,t}^r) : 0 \leq p_{n,t}^r \leq \overline{p}_{n,t}^r, \tag{5.7a}$$

$$|q_{n,t}^r| \leq \phi_n p_{n,t}^r, \tag{5.7b}$$

$$(p_{n,t}^r)^2 + (q_{n,t}^r)^2 \leq \overline{s}_n^2\big\} \tag{5.7c}$$

that is random and time-variant due to the variability of $\overline{p}_{n,t}^r$. Constraint (5.7a) limits the active power generation according to the available solar power; constraint (5.7b) enforces the lower limit $\cos(\arctan(\phi_n))$ on the power factor (lagging or leading); and (5.7c) limits the inverter apparent power.

5.2.2 Operation costs

If PVs owners are compensated at price $\boldsymbol{\pi}$ for the active power surplus they inject into the distribution grid, the related utility cost at slot t is $C_{\mathrm{PV}}(\mathbf{p}_t^r) := \boldsymbol{\pi}^\top[\mathbf{p}_t^r - \mathbf{p}_t^l]_+$ with $[\cdot]_+ := \max\{0, \cdot\}$ applied entrywise on vector $\mathbf{p}_t^r - \mathbf{p}_t^l$. The diesel generation cost is represented by $C_{\mathrm{D}}(\mathbf{p}^d)$. Regarding energy transactions with the main grid, the power block p_0^a bought in advance is charged at a fixed and known price β. Deviating from p_0^a by $p_{0,t}^\delta$ at slot t is charged at

$$C^t(p_{0,t}^\delta) := \gamma_b[p_{0,t}^\delta]_+ - \gamma_s[-p_{0,t}^\delta]_+ \tag{5.8}$$

for known prices (γ_b, γ_s). To avoid arbitrage, it is assumed that $0 < \gamma_s < \beta < \gamma_b$; see e.g., [Var11], [Zha14a]. Then, the deviation charge can also be expressed as $C^t(p_{0,t}^\delta) = \max\{\gamma_b p_{0,t}^\delta, \gamma_s p_{0,t}^\delta\}$, which is certainly convex [Zha13b].

5.2.3 Optimal grid dispatch

Depending on the way compliance with voltage regulation region $\mathcal{V}_A$ is enforced, two grid dispatch formulations are developed next. Commencing with the *average dispatch*, the optimal grid operation is posed as

$$\mathsf{P}_a^* := \min \; C_{\mathrm{D}}(\mathbf{p}^d) + \beta p_0^a + \mathbb{E}_t\left[C^t(p_{0,t}^\delta) + C_{\mathrm{PV}}(\mathbf{p}_t^r)\right] \tag{5.9a}$$

$$\text{s.to:} \;\; \mathbf{p}_t = \mathbf{p}_t^r - \mathbf{p}_t^l + \mathbf{p}^d \tag{5.9b}$$

$$\mathbf{q}_t = \mathbf{q}_t^r - \mathbf{q}_t^l \tag{5.9c}$$

$$p_{0,t} = p_0^a + p_{0,t}^\delta \tag{5.9d}$$

$$p_{0,t} \geq -\mathbf{1}^\top \mathbf{p}_t + \mathbf{p}_t^\top \mathbf{R}\mathbf{p}_t + \mathbf{q}_t^\top \mathbf{R}\mathbf{q}_t \tag{5.9e}$$

$$\mathbf{p}_t^\top \mathbf{f}_n \mathbf{f}_n^\top \mathbf{p}_t + \mathbf{q}_t^\top \mathbf{f}_n \mathbf{f}_n^\top \mathbf{q}_t \leq \overline{S}_n, \ \forall n \in \mathcal{N} \tag{5.9f}$$

$$\underline{\mathbf{p}}^d \leq \mathbf{p}^d \leq \overline{\mathbf{p}}^d \tag{5.9g}$$

$$(p_{n,t}^r, q_{n,t}^r) \in \Omega_{n,t}, \ \forall n \in \mathcal{N} \tag{5.9h}$$

$$\underline{v}_0 \leq v_0^a \leq \overline{v}_0 \tag{5.9i}$$

$$\mathbf{v}_t = 2\mathbf{R}\mathbf{p}_t + 2\mathbf{X}\mathbf{q}_t + v_0^a \mathbf{1} \tag{5.9j}$$

$$\mathbf{v}_t \in \mathcal{V}_B \tag{5.9k}$$

$$\mathbb{E}_t[\mathbf{v}_t] \in \mathcal{V}_A \tag{5.9l}$$

$$\text{over} \quad v_0^a, p_0^a, \mathbf{p}^d, \{\mathbf{p}_t, \mathbf{q}_t, \mathbf{v}_t, \mathbf{p}_t^r, \mathbf{q}_t^r, p_{0,t}, p_{0,t}^\delta\}_{t=1}^T.$$

The slow-timescale variables $\{v_0^a, p_0^a, \mathbf{p}^d\}$ are set in advance, and remain fixed throughout the T subsequent control slots over which the fast-timescale variables $\{\mathbf{p}_t, \mathbf{q}_t, \mathbf{v}_t, \mathbf{p}_t^r, \mathbf{q}_t^r, p_{0,t}, p_{0,t}^\delta\}_{t=1}^T$ are implemented. The latter variables depend on the randomness of slot t as well as slow-timescale decisions.

Alternatively to (5.9), optimal grid operation can be posed as a *probabilistic dispatch* that is identical to (5.9) with the exception that (5.9l) is replaced by the probabilistic constraint

$$\Pr\{\mathbf{v}_t \notin \mathcal{V}_A\} \leq \alpha \tag{5.10}$$

for some small parameter $\alpha > 0$, say $\alpha = 0.05$. The optimal cost for the probabilistic dispatch will be denoted by P_p^*.

The objective function in (5.9a) involves the cost of energy dispatched at the slow timescale plus the average fast-timescale energy management cost. Nodal (re)active power balance is ensured via (5.9b)–(5.9c). Constraint (5.9e) accounts for the active power losses upon relaxing the quadratic equation in (5.6) to a convex inequality without loss of optimality. Constraint (5.9f) limits the apparent power flow at each line n based on (5.5). Constraints (5.9i)–(5.9l) are voltage regulation constraints: In detail, (5.9j) relates squared voltage magnitudes to power injections [cf. (5.3)]; (5.9i) constraints the substation bus voltage; and (5.9k) constraints voltages in $\mathcal{V}_B$. While (5.9l) maintains the average voltage magnitudes in $\mathcal{V}_A$, its alternative in (5.10) limits the probability of voltage magnitudes being outside $\mathcal{V}_A$.

5.3 Problem Analysis

To facilitate algorithmic developments, the problem in (5.9) is expressed in a more compact form next. Collect the slow-timescale variables in vector $\mathbf{z}^\top := [v_0^a, p_0^a, \mathbf{p}^d]$; the fast-timescale variables at slot t in $\mathbf{y}_t^\top := [\mathbf{p}_t, \mathbf{q}_t, \mathbf{v}_t, \mathbf{p}_t^r, \mathbf{q}_t^r, p_{0,t}, p_{0,t}^\delta]$; and the random variables involved at slot t in $\boldsymbol{\xi}_t^\top := [\overline{\mathbf{p}}_t^r, \mathbf{p}_t^l, \mathbf{q}_t^l]$.

The constraints in (5.9) can be then classified into four groups:

1. Constraints involving fast-timescale variables only, such as (5.9c), (5.9f), (5.9h), and (5.9k), that will be abstracted as $\mathbf{y}_t \in \mathcal{Y}_t$.

2. Constraints (5.9g) and (5.9i) that involve slow-timescale variables only, and they will be denoted as $\mathbf{z} \in \mathcal{Z}$.

3. The *linear* constraints (5.9b), (5.9d), and (5.9j), coupling slow- and fast-timescale variables as well as random variables. These constraints are collectively expressed as $\mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t$ for appropriate matrices $\mathbf{K}$, $\mathbf{B}$, and $\mathbf{H}$.

4. The ergodic constraints (5.9l) and (5.10) depend on the voltage sequence $\{\mathbf{v}_t\}_{t=1}^T$, hence coupling decisions across time. A substantial difference between (5.9l) and (5.10) is that the

latter is a non-convex constraint.

Based on this grouping, the two dispatch problems can be compactly rewritten as

$$\mathsf{P}^*_{(a,p)} := \min_{\mathbf{z},\{\mathbf{y}_t\}_{t=1}^T} \quad f(\mathbf{z}) + \mathbb{E}_t\left[g_t(\mathbf{y}_t)\right] \tag{5.11a}$$

$$\text{s.to:} \quad \mathbf{z} \in \mathcal{Z} \tag{5.11b}$$

$$\mathbf{y}_t \in \mathcal{Y}_t \qquad \forall t \tag{5.11c}$$

$$\mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t \qquad \forall t \tag{5.11d}$$

$$\mathbb{E}_t\left[\mathbf{h}(\mathbf{y}_t)\right] \le \mathbf{0} \tag{5.11e}$$

where $f(\mathbf{z}) := C_{\mathrm{D}}(\mathbf{p}^d) + \beta p_0^a$ and $g_t(\mathbf{y}_t) := C^t(p_{0,t}^\delta) + C_{\mathrm{PV}}(\mathbf{p}_t^r)$. For the average dispatch, the optimal cost in (5.11) is P_a^* and the function in (5.11e) is $\mathbf{h}(\mathbf{y}_t) = [\mathbf{v}_t - \overline{v}_A\mathbf{1}, \underline{v}_A\mathbf{1} - \mathbf{v}_t]$. For the probabilistic dispatch, the optimal cost is P_p^* and the function in (5.11e) is $\mathbf{h}(\mathbf{y}_t) = \mathbb{1}\{\mathbf{v}_t \notin \mathcal{V}_A\} - \alpha$.

The optimal values for the slow-timescale variables $\mathbf{z}$ must be decided in advance. Once the optimal $\mathbf{z}$ is found, it remains fixed over the slow-timescale interval. The fast-timescale decisions $\mathbf{y}_t(\mathbf{z})$ for slot t depend on $\mathbf{z}$, while the subscript t indicates their dependence on the realization $\boldsymbol{\xi}_t$. Both the average and the probabilistic dispatch are stochastic programming problems with recourse [Var11]. Their costs can be decomposed as $\mathsf{P}^*_{(a,p)} = \min_{\mathbf{z}\in\mathcal{Z}} f(\mathbf{z}) + G_{(a,p)}(\mathbf{z})$, where the so termed *expected recourse function* is defined as

$$G_{(a,p)}(\mathbf{z}) := \min_{\{\mathbf{y}_t\in\mathcal{Y}_t\}} \quad \mathbb{E}_t\left[g_t(\mathbf{y}_t)\right] \tag{5.12a}$$

$$\text{s.to:} \quad \mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t \qquad \forall t \tag{5.12b}$$

$$\mathbb{E}_t\left[\mathbf{h}(\mathbf{y}_t)\right] \le \mathbf{0}. \tag{5.12c}$$

Since problem (5.12) depends on $\mathbf{z}$, its minimizer can be written as $\{\mathbf{y}_t^*(\mathbf{z})\}_{t=1}^T$ and the recourse function as $G_{(a,p)}(\mathbf{z}) = \mathbb{E}_t[g_t(\mathbf{y}_t^*(\mathbf{z}))]$. The ensuing two sections solve the average and the probabilistic dispatches.

5.4 Average Dispatch Algorithm

This section tackles problem (5.11) with the ergodic constraint in (5.11e), for which $\mathbf{h}(\mathbf{y}_t) = [\underline{v}_A\mathbf{1} - \mathbf{v}_t, \mathbf{v}_t - \overline{v}_A\mathbf{1}]$. Although convex, problem (5.11) is challenging due to the coupling across $\{\mathbf{y}_t\}_{t=1}^T$ and between fast- and slow-timescale variables. Dual decomposition is adopted to resolve the coupling across $\{\mathbf{y}_t\}_{t=1}^T$. The partial Lagrangian function for (5.12) is

$$L_a(\{\mathbf{y}_t\}, \boldsymbol{\nu}) := \mathbb{E}_t\left[g_t(\mathbf{y}_t) + \boldsymbol{\nu}^\top\mathbf{h}(\mathbf{y}_t)\right] \tag{5.13}$$

with the entries of $\boldsymbol{\nu}$ being the multipliers associated with the upper and lower per-bus constraints in (5.11e). The corresponding dual function is

$$D_a(\boldsymbol{\nu}; \mathbf{z}) := \min_{\{\mathbf{y}_t\in\mathcal{Y}_t\}} \quad L_a(\{\mathbf{y}_t\}, \boldsymbol{\nu}) \tag{5.14}$$

$$\text{s.to: } \mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t \quad \forall t.$$

Observe that after dualizing, the minimization in (5.14) is separable over the realizations $\{\boldsymbol{\xi}_t\}$. Precisely, the optimal fast-timescale variable for fixed $(\boldsymbol{\nu}, \mathbf{z})$ and for a specific realization $\boldsymbol{\xi}_t$ can be found by solving:

$$\mathbf{y}_t^*(\boldsymbol{\nu}, \mathbf{z}) \in \arg\min_{\mathbf{y}_t\in\mathcal{Y}_t} \; g_t(\mathbf{y}_t) + \boldsymbol{\nu}^\top\mathbf{h}(\mathbf{y}_t) \tag{5.15a}$$

$$\text{s.to: } \mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t. \tag{5.15b}$$

For future reference, let us also define $\boldsymbol{\lambda}_t^*(\boldsymbol{\nu}, \mathbf{z})$ as the optimal Lagrange multiplier associated with (5.15b). If $\boldsymbol{\nu}$ is partitioned as $\boldsymbol{\nu}^\top = [\underline{\boldsymbol{\nu}}^\top, \overline{\boldsymbol{\nu}}^\top]$ with $\underline{\boldsymbol{\nu}}$ corresponding to constraint $\mathbb{E}_t[\mathbf{v}_t] \geq \underline{v}_A\mathbf{1}$ and $\overline{\boldsymbol{\nu}}$ to $\mathbb{E}_t[\mathbf{v}_t] \leq \overline{v}_A\mathbf{1}$, then (5.15) simplifies to

$$\begin{aligned} \mathbf{y}_t^*(\boldsymbol{\nu}, \mathbf{z}) \in \operatorname{argmin}\ & C^t(p_{0,t}^\delta) + C_{\mathrm{PV}}(\mathbf{p}_t^r) + (\overline{\boldsymbol{\nu}} - \underline{\boldsymbol{\nu}})^\top \mathbf{v}_t \\ \text{s.to: } & (5.9b) - (5.9f), (5.9h), (5.9j), (5.9k) \\ \text{over } & \{\mathbf{p}_t, \mathbf{q}_t, \mathbf{v}_t, \mathbf{p}_t^r, \mathbf{q}_t^r, p_{0,t}, p_{0,t}^\delta\} \end{aligned} \tag{5.16}$$

and can be solved as a convex quadratic program. Given the optimal pair $(\boldsymbol{\nu}^*, \mathbf{z}^*)$, the optimal fast-timescale variables $\mathbf{y}_t$ can be thus found for any $\boldsymbol{\xi}_t$.

Back to finding the optimal primal and dual slow-timescale variables, note that the dual problem associated with (5.14) is

$$\boldsymbol{\nu}^* := \arg\max_{\boldsymbol{\nu} \geq \mathbf{0}}\ D_a(\boldsymbol{\nu}; \mathbf{z}). \tag{5.17}$$

Duality theory asserts that (5.17) is a convex problem. Moreover, assuming a strictly feasible point exists for (5.12), strong duality implies that $G_a(\mathbf{z}) = D_a(\boldsymbol{\nu}^*, \mathbf{z})$. Due to the latter, the original problem in (5.11) can be transformed to:

$$\min_{\mathbf{z} \in \mathcal{Z}} f(\mathbf{z}) + G_a(\mathbf{z}) = \min_{\mathbf{z} \in \mathcal{Z}} \{f(\mathbf{z}) + \max_{\boldsymbol{\nu} \geq \mathbf{0}} D_a(\boldsymbol{\nu}; \mathbf{z})\} \tag{5.18a}$$

$$= \min_{\mathbf{z} \in \mathcal{Z}} \max_{\boldsymbol{\nu} \geq \mathbf{0}} \tilde{f}_a(\boldsymbol{\nu}, \mathbf{z}) \tag{5.18b}$$

where the auxiliary function $\tilde{f}_a$ is defined as:

$$\tilde{f}_a(\boldsymbol{\nu}, \mathbf{z}) := f(\mathbf{z}) + D_a(\boldsymbol{\nu}; \mathbf{z}). \tag{5.19}$$

Being a dual function, $D_a(\boldsymbol{\nu}; \mathbf{z})$ is a concave function of $\boldsymbol{\nu}$. At the same time, $D_a(\boldsymbol{\nu}; \mathbf{z})$ is a perturbation function with respect to $\mathbf{z}$; and hence, it is a convex function of $\mathbf{z}$ [Boy04]. Recall that $f(\mathbf{z})$ is a convex function of $\mathbf{z}$ too. Therefore, the auxiliary function $\tilde{f}_a(\boldsymbol{\nu}, \mathbf{z})$ is convex in $\mathbf{z}$ and concave in $\boldsymbol{\nu}$. Because of the randomness of $\{\boldsymbol{\xi}_t\}$, function $D_a(\boldsymbol{\nu}; \mathbf{z})$ in (5.14) is stochastic. Consequently, problem (5.18b) is a stochastic convex-concave saddle point problem [Boy04], [Nem09].

To solve (5.18b), we rely on the stochastic saddle-point approximation method of [Nem09]. The method involves the subgradient of $\tilde{f}_a$ with respect to $\mathbf{z}$, and its supergradient with respect to $\boldsymbol{\nu}$. Upon viewing $D_a(\boldsymbol{\nu}, \mathbf{z})$ in (5.14) as a perturbation function of $\mathbf{z}$, the subgradient of $\tilde{f}_a$ with respect to $\mathbf{z}$ is [Boy04]

$$\partial_{\mathbf{z}} \tilde{f}_a = \partial_{\mathbf{z}} f(\mathbf{z}) + \mathbf{K}^\top \mathbb{E}_t[\boldsymbol{\lambda}_t^*(\boldsymbol{\nu}, \mathbf{z})]. \tag{5.20}$$

By definition of the dual function, the supergradient of $\tilde{f}_a$ with respect to $\boldsymbol{\nu}$ is

$$\partial_{\boldsymbol{\nu}} \tilde{f}_a = \mathbb{E}_t[\mathbf{h}(\mathbf{y}_t^*(\boldsymbol{\nu}, \mathbf{z}))]. \tag{5.21}$$

The stochastic saddle point approximation method of [Nem09] involves primal-dual subgradient iterates with the expectations in (5.20)–(5.21) being replaced by their instantaneous estimates based on a single realization $\boldsymbol{\xi}_k$. Precisely, the method involves the iterates over k:

$$\boldsymbol{\nu}^{k+1} := [\boldsymbol{\nu}^k + \mathrm{dg}(\boldsymbol{\mu}_k)\mathbf{h}(\mathbf{y}_k^*(\boldsymbol{\nu}^k, \mathbf{z}^k))]_+ \tag{5.22a}$$

$$\mathbf{z}^{k+1} := [\mathbf{z}^k - \mathrm{dg}(\boldsymbol{\epsilon}_k)(\partial_z f(\mathbf{z}^k) + \mathbf{K}^\top \boldsymbol{\lambda}_k^*(\boldsymbol{\nu}^k, \mathbf{z}^k))]_{\mathcal{Z}} \tag{5.22b}$$

Algorithm 1 Average Dispatch Algorithm (ADA)

1: Initialize $(\mathbf{z}^0, \boldsymbol{\nu}^0)$.
2: **repeat** for $k = 0, 1, \ldots$
3: Draw sample $\boldsymbol{\xi}_k$.
4: Find $(\mathbf{y}_k^*(\boldsymbol{\nu}^k, \mathbf{z}^k), \boldsymbol{\lambda}_k^*(\boldsymbol{\nu}^k, \mathbf{z}^k))$ by solving (5.15).
5: Update $(\mathbf{z}^{k+1}, \boldsymbol{\nu}^{k+1})$ via (5.22).
6: Compute sliding averages $(\tilde{\mathbf{z}}^k, \tilde{\boldsymbol{\nu}}^k)$ through (5.23).
7: **until** convergence of $(\tilde{\mathbf{z}}^k, \tilde{\boldsymbol{\nu}}^k)$.
8: **Output** $\mathbf{z}^* = \tilde{\mathbf{z}}^k$ and $\boldsymbol{\nu}^* = \tilde{\boldsymbol{\nu}}^k$.

where the operator $[\cdot]_{\mathcal{Z}}$ projects its argument onto $\mathcal{Z}$; and vectors $\boldsymbol{\mu}_k = \boldsymbol{\mu}_0/\sqrt{k}$ and $\boldsymbol{\epsilon}_k = \boldsymbol{\epsilon}_0/\sqrt{k}$ collect respectively the primal and dual step sizes for positive $\boldsymbol{\mu}_0$ and $\boldsymbol{\epsilon}_0$. At every iteration k, the method draws a realization $\boldsymbol{\xi}_k$ and solves (5.15) for the tuple $(\boldsymbol{\xi}_k, \boldsymbol{\nu}^k, \mathbf{z}^k)$ to acquire $(\mathbf{y}_k^*(\boldsymbol{\nu}^k, \mathbf{z}^k), \boldsymbol{\lambda}_k^*(\boldsymbol{\nu}^k, \mathbf{z}^k))$ and perform the primal-dual updates in (5.22). The method finally outputs the sliding averages of the updates as:

$$\tilde{\mathbf{z}}^k := \big(\textstyle\sum_{i=\lceil k/2\rceil}^{k} \mathbf{z}^i/\sqrt{i}\big)/\big(\textstyle\sum_{i=\lceil k/2\rceil}^{k} 1/\sqrt{i}\big) \tag{5.23a}$$

$$\tilde{\boldsymbol{\nu}}^k := \big(\textstyle\sum_{i=\lceil k/2\rceil}^{k} \boldsymbol{\nu}^i/\sqrt{i}\big)/\big(\textstyle\sum_{i=\lceil k/2\rceil}^{k} 1/\sqrt{i}\big). \tag{5.23b}$$

The proposed scheme converges to the value $\tilde{f}_a(\boldsymbol{\nu}^*, \mathbf{z}^*)$ obtained at a saddle point $(\boldsymbol{\nu}^*, \mathbf{z}^*)$ asymptotically in the number of iterations k [Nem09, Sec. 3.1].

Upon convergence of the iterates in (5.23), the slow-timescale variables $\mathbf{z}^*$ have been derived together with the optimal Lagrange multiplier $\boldsymbol{\nu}^*$ related to constraint (5.12c). The grid operator can implement $\mathbf{z}^*$, and the fast-timescale decisions $\mathbf{y}_t^*$ for a realization $\boldsymbol{\xi}_t$ can be found by solving (5.16). The average dispatch algorithm (ADA) is summarized as Alg. 1.

5.5 Probabilistic Dispatch Algorithm

The probabilistic version of problem (5.11) is considered next. Here, the ergodic constraint (5.11e) reads $h(\mathbf{y}_t) = \mathbb{1}\{\mathbf{v}_t \notin \mathcal{V}_A\} - \alpha$. Despite the non-convexity of the probabilistic constraint, (5.12) can still be solved optimally. However, optimality for (5.11) cannot be guaranteed. A heuristic solution is detailed next by adapting the solution of Sec. 5.4.

To that end, dual decomposition is used here as well. If ν is the scalar Lagrange multiplier associated with constraint (5.11e), the partial Lagrangian function for (5.12) is now $L_p(\{\mathbf{y}_t\}, \nu) := \mathbb{E}_t\left[g_t(\mathbf{y}_t) + \nu(\mathbb{1}\{\mathbf{v}_t \notin \mathcal{V}_A\} - \alpha)\right]$. The corresponding dual function, fast-timescale problem, and dual problem are defined analogously to (5.14), (5.15), and (5.17). The indicator function renders $L_p(\{\mathbf{y}_t\}, \nu)$ non-convex. Surprisingly enough though, under the practical assumption that $\{\boldsymbol{\xi}_t\}$ follows a continuous pdf, problem (5.12) enjoys zero duality gap; see [Rib10b, Th. 1].

The additional challenge here is the non-convexity of the Lagrangian minimization:

$$\mathbf{y}_t^*(\nu, \mathbf{z}) \in \arg\min_{\mathbf{y}_t \in \mathcal{Y}_t} \; g_t(\mathbf{y}_t) + \nu \mathbb{1}\{\mathbf{v}_t \notin \mathcal{V}_A\} \tag{5.24}$$
$$\text{s.to:} \;\; \mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t.$$

Because the indicator function takes only the values $\{0, 1\}$ however, the solution to (5.24) can be found by solving a pair of slightly different convex problems. The first problem is

$$\mathbf{y}_{t,A}^*(\mathbf{z}) \in \arg\min_{\mathbf{y}_t \in \mathcal{Y}_t} \; g_t(\mathbf{y}_t) \tag{5.25a}$$
$$\text{s.to:} \;\; \mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t \tag{5.25b}$$

Algorithm 2 Probabilistic Dispatch Algorithm (PDA)

Initialize $(\mathbf{z}^0, \nu^0)$.
repeat for $k = 0, 1, \ldots$
  Draw sample $\boldsymbol{\xi}_k$.
  Find $(\mathbf{y}^*_{k,B}(\nu^k, \mathbf{z}^k), \boldsymbol{\lambda}^*_{k,B}(\nu^k, \mathbf{z}^k))$ by solving (5.26).
  Set $\mathbf{y}^*_t(\nu, \mathbf{z}) := \mathbf{y}^*_{t,B}(\mathbf{z})$ and $\boldsymbol{\lambda}^*_t(\nu, \mathbf{z}) := \boldsymbol{\lambda}^*_{t,B}(\mathbf{z})$.
  if $\mathbf{v}^*_{k,B}(\mathbf{z}) \notin \mathcal{V}_A$, **then** find $\mathbf{y}^*_{k,A}(\nu^k, \mathbf{z}^k)$ and $\boldsymbol{\lambda}^*_{k,A}(\nu^k, \mathbf{z}^k)$ by solving (5.25).
    if $g_t(\mathbf{y}^*_{t,A}(\mathbf{z})) \leq g_t(\mathbf{y}^*_{t,B}(\mathbf{z})) + \nu$, **then** set $\mathbf{y}^*_t(\nu, \mathbf{z}) := \mathbf{y}^*_{t,A}(\mathbf{z})$ and $\boldsymbol{\lambda}^*_t(\nu, \mathbf{z}) := \boldsymbol{\lambda}^*_{t,A}(\mathbf{z})$.
    end if
  end if
  Update $(\mathbf{z}^{k+1}, \nu^{k+1})$ via (5.27).
  Compute sliding averages $(\tilde{\mathbf{z}}^k, \tilde{\nu}^k)$ through (5.23).
until convergence of $(\tilde{\mathbf{z}}^k, \tilde{\nu}^k)$.
Output $\mathbf{z}^* = \tilde{\mathbf{z}}^k$ and $\nu^* = \tilde{\nu}^k$.

$$\mathbf{v}_t \in \mathcal{V}_A \tag{5.25c}$$

whereas the second problem ignores constraint $\mathbf{v}_t \in \mathcal{V}_A$ as

$$\mathbf{y}^*_{t,B}(\mathbf{z}) \in \arg\min_{\mathbf{y}_t \in \mathcal{Y}_t} \quad g_t(\mathbf{y}_t) \tag{5.26a}$$

$$\text{s.to:} \quad \mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t. \tag{5.26b}$$

From the point of view of (5.24), if the voltages in $\mathbf{y}^*_{t,B}(\mathbf{z})$ do not belong to $\mathcal{V}_A$, the solution to the second problem will incur an additional cost quantified by ν. Observe that neither problem (5.25) nor (5.26) depend on ν, while their complexity is similar to the one problem (5.15). Suppose that (5.25) and (5.26) have been solved and let $\boldsymbol{\lambda}^*_{t,A}(\mathbf{z})$ and $\boldsymbol{\lambda}^*_{t,B}(\mathbf{z})$ denote the optimal multipliers associated with (5.25b) and (5.26b), respectively. Then, problem (5.24) can be neatly tackled by identifying two cases:

(c1) If $g_t(\mathbf{y}^*_{t,A}(\mathbf{z})) > g_t(\mathbf{y}^*_{t,B}(\mathbf{z})) + \nu$, then $\mathbf{y}^*_{t,B}(\mathbf{z})$ is a minimizer of (5.24) as well and voltages are allowed to lie outside $\mathcal{V}_A$. In this case, set $\mathbf{y}^*_t(\nu, \mathbf{z}) := \mathbf{y}^*_{t,B}(\mathbf{z})$ and $\boldsymbol{\lambda}^*_t(\nu, \mathbf{z}) := \boldsymbol{\lambda}^*_{t,B}(\mathbf{z})$. This case includes instances where problem (5.25) is infeasible for which $g_t(\mathbf{y}^*_{t,A}(\mathbf{z})) = \infty$.

(c2) If $g_t(\mathbf{y}^*_{t,A}(\mathbf{z})) \leq g_t(\mathbf{y}^*_{t,B}(\mathbf{z})) + \nu$, then $\mathbf{y}^*_{t,A}(\mathbf{z})$ minimizes (5.24) too and voltages lie within $\mathcal{V}_A$. In this case, set $\mathbf{y}^*_t(\nu, \mathbf{z}) := \mathbf{y}^*_{t,A}(\mathbf{z})$ and $\boldsymbol{\lambda}^*_t(\nu, \mathbf{z}) := \boldsymbol{\lambda}^*_{t,A}(\mathbf{z})$.

Case (c2) covers also instances where $\mathbf{v}^*_{t,B}(\mathbf{z})$ happens to lie in $\mathcal{V}_A$. In these particular instances, $\mathbf{y}^*_{t,B}(\mathbf{z})$ serves as a minimizer of (5.25) too. Then, it follows that $g_t(\mathbf{y}^*_{t,A}(\mathbf{z})) = g_t(\mathbf{y}^*_{t,B}(\mathbf{z})) \leq g_t(\mathbf{y}^*_{t,B}(\mathbf{z})) + \nu$ for $\nu \geq 0$. This implies that one can solve (5.26) first and, if $\mathbf{v}^*_{t,B}(\mathbf{z}) \in \mathcal{V}_A$, there is no need to solve problem (5.25).

To find the optimal slow-timescale variables under the probabilistic dispatch, the stochastic primal-dual iterations of Sec. 5.4 are adapted here as

$$\nu^{k+1} := [\nu^k + \mu_k(\mathbb{1}\{\mathbf{v}^*_k(\nu^k, \mathbf{z}^k) \notin \mathcal{V}_A\} - \alpha)]_+ \tag{5.27a}$$

$$\mathbf{z}^{k+1} := [\mathbf{z}^k - \mathrm{dg}(\boldsymbol{\epsilon}_k)(\partial_z f(\mathbf{z}^k) + \mathbf{K}^\top \boldsymbol{\lambda}^*_k(\nu^k, \mathbf{z}^k)]_{\mathcal{Z}}. \tag{5.27b}$$

The probabilistic dispatch algorithm (PDA) is tabulated as Alg. 2. Because function $G_p(\mathbf{z})$ is not necessarily convex, the iterates in (5.27) are not guaranteed to converge to a minimizer of (5.11). The practical performance of PDA in finding $\mathbf{z}^*$ is numerically validated in Sec. 5.6.

5.6 Numerical Tests

The proposed grid dispatches were tested on a 56-bus Southern California Edison (SCE) distribution feeder [Gan15]. 5-MW PVs were added on buses 44 and 50; both with 6-MVA inverters enabling power factors as low as 0.83 (leading or lagging) at full solar generation. The prices for the energy exchange with the main grid were $\beta = 37$ \$/MWh; $\gamma_b = 45$ \$/MWh, and $\gamma_s = 19$ \$/MWh. Diesel generators with capacity $\overline{p}_n^d = 0.5$ MW were sited on buses 10, 18, 21, 30, 36, 43, 51, and 55. The cost of diesel generation was $C_{\mathrm{D}}(\mathbf{p}^d) = \sum_{n=1}^{N}(30p_n^d + 15(p_n^d)^2)$ \$/h with $\mathbf{p}^d$ expressed in MW. Apparent power flows were limited to 7 MVA. The voltage operation limits were set to $\underline{v}_A = 0.98^2$, $\overline{v}_A = 1.02^2$, $\underline{v}_B = 0.97^2$, and $\overline{v}_B = 1.03^2$, expressed per unit (pu) with respect to a voltage base of 12 kV. (Re)active nodal loads were Gaussian distributed with the nominal load of the SCE benchmark as mean value, and standard deviation of 0.2 times the nominal load. The solar energy generated at each PVs was drawn uniformly between 0.5 and 1 times the actual power PVs rating.

ADA was run with step sizes proportional to $1/\sqrt{k}$ with initial values $\epsilon_0^{v_0} = 4{\cdot}10^{-5}$, $\epsilon_0^{p_0} = 4{\cdot}10^{-1}$, $\epsilon_0^{p_d} = 6 \cdot 10^{-3}$, and $\mu_0 = 225$, to account for different dynamic ranges. The iterates for primal and dual variables as well as their corresponding sliding averages are depicted in Fig. 5.2. Primal and dual slow-timescale variables hover in a small range whose width diminishes with time. Their sliding averages converge asymptotically. The algorithm reaches a practically meaningful solution within 5,000 iterations. Buses 44 and 50 are prone to overvoltages since they host PVs generation, and buses 2 and 15 are prone to under-voltages; thus yielding non-zero dual variables for the average upper and lower voltage constraints, respectively.

PDA was tested using the same simulation setup for $\alpha = 0.05$ and $\mu_0 = 1$. Figure 5.3 shows the convergence of primal and dual variables, and the probability of voltages deviating from $\mathcal{V}_A$. Granted that the probabilistic constraint in (5.10) applies collectively to all buses, the under-/over-voltage probabilities on a per-bus basis is depicted in Fig. 5.4. The occurrences of overvoltage seem to be shared primarily among buses 40–56 which are neighboring to the PVs buses 40 and 55. On the contrary, buses 10–16 being electrically far from both the substation and PVs, experience under-voltage with a small probability.

The effect of the average versus the probabilistic constraint on voltage magnitudes was evaluated next. After slow-timescale variables $\mathbf{z}$ had converged, fast-timescale variables $\mathbf{y}_t$ were calculated for 6,000 instances of $\boldsymbol{\xi}_t$ using both ADA and PDA. The histograms of the voltage magnitudes on two representative buses are presented in Fig. 5.5. Under PDA, the average voltage on bus 15 is slightly higher than the average voltage obtained by ADA. In exchange, the instantaneous value of the voltage on bus 15 stays within $\mathcal{V}_A$ with higher probability. A similar behavior is observed for the overvoltage instances on PVs bus 40.

ADA and PDA were finally compared to three alternative schemes. The first two, henceforth called *approximate average* and *approximate probabilistic* dispatches, obtained $\mathbf{z}$ by setting loads and solar generation to their expected values, while variables $\boldsymbol{\nu}$ were calculated via dual stochastic subgradient, and $\{\mathbf{y}_t\}_{t=1}^T$ were found by solving either (5.15) or (5.24), depending on whether the setting is average or probabilistic. The third *deterministic* dispatch found $\mathbf{z}$ as the approximate schemes do, and $\{\mathbf{y}_t\}_{t=1}^T$ by enforcing $\mathbf{v}_t \in \mathcal{V}_A$ at all times. Note that the three proposed alternatives provide *feasible* solutions satisfying voltage regulation constraints. The five dispatches were tested under five scenarios: Scenario 1 is the setup described earlier. Scenario 2 involved the tighter voltage limits $\underline{v}_A = 0.99^2$ and $\overline{v}_A = 1.01^2$. Scenarios 3, 4, and 5 were generated by scaling the mean value and the standard deviation for loads of scenario 1 by 0.5, 1.5, and 2, respectively. Figure 5.6 shows the expected operation costs for all five scenarios. ADA (PDA) yielded the lowest cost under all scenarios in the average (probabilistic) setting as expected. In all test cases, ADA yielded a slightly lower objective than PDA for $\alpha = 0.05$. The loss of optimality entailed by the approximate average and probabilistic schemes is due to the suboptimal choice of $\mathbf{z}$. The

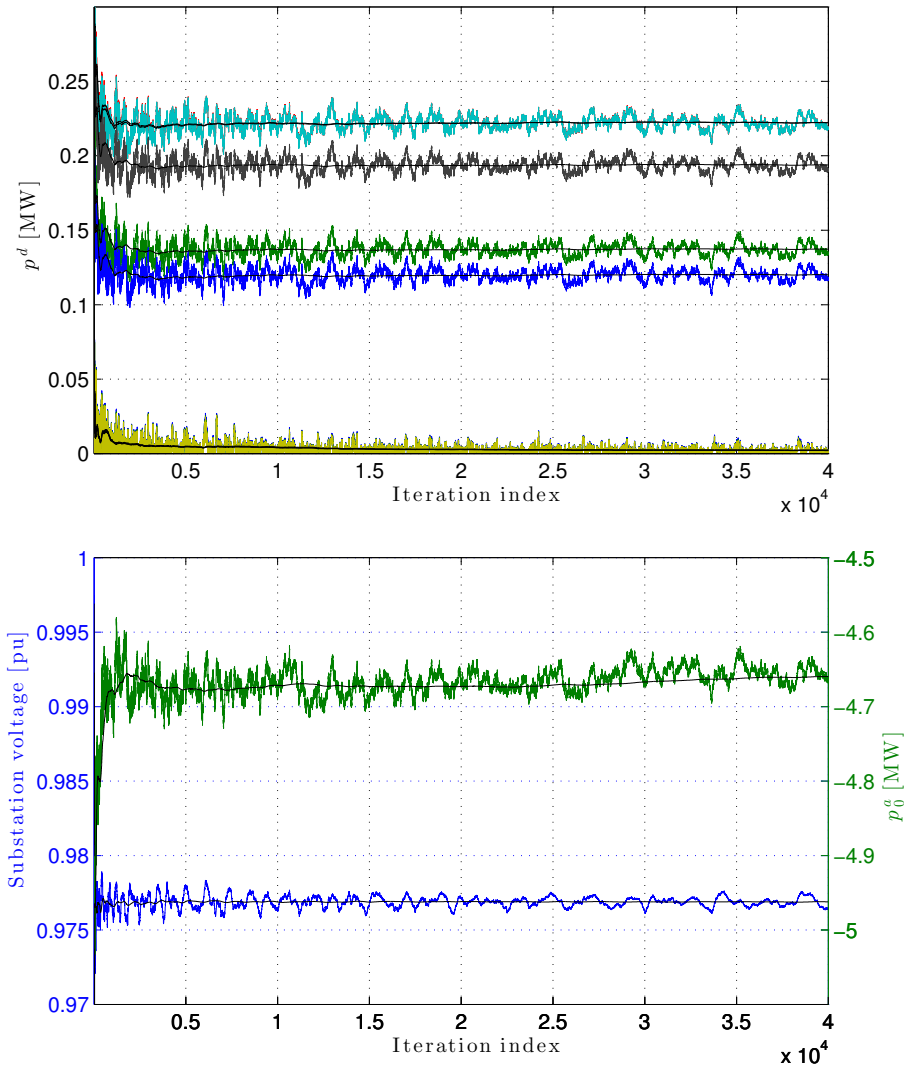

Figure 5.1: Convergence of primal variables for ADA: (top) diesel generation; (bottom) substation voltage v_0 (left y-axis) and energy exchange p_0^a (right y-axis). Sliding averages of optimization variables are depicted too.

deterministic scheme entailed an additional loss of optimality by preventing the occasional violation of $\mathcal{V}_A$.

5.7 Summarizing conclusions and alternative grid models

By nature of renewable generation, electromechanical component limits, and the manner markets operate, energy management of smart distribution grids involves decisions at slower and faster timescales. Since slow-timescale controls remain fixed over multiple PVs operation slots, decisions are coupled across time in a stochastic manner. To accommodate solar energy fluctuations, voltages have been allowed to be sporadically overloaded; hence introducing coupling of fast-timescale variables on the average or in probability. Average voltage constraints have resulted in a stochastic convex-concave problem, whereas non-convex probabilistic constraints were tackled using dual decomposition and convex optimization. Efficient algorithms for finding both slow and fast controls using only random samples have been put forth. The two developed novel solvers converge in terms of the primal and dual variables, and have attained lower operational costs compared to

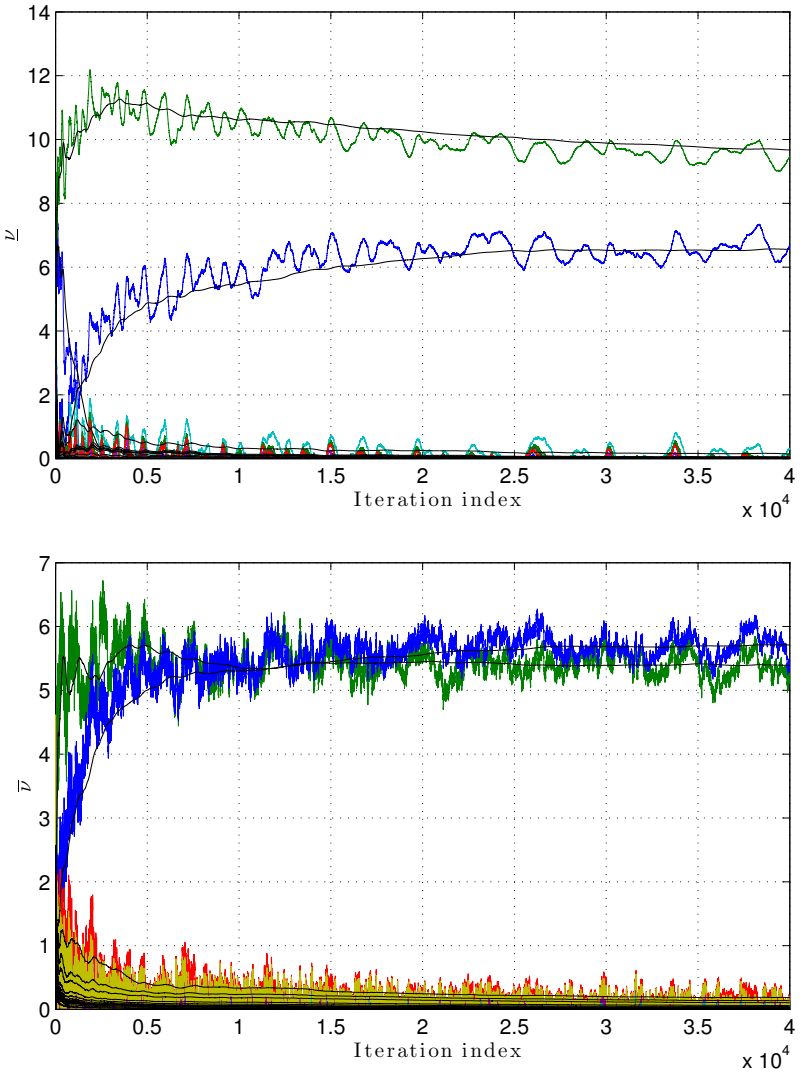

Figure 5.2: Convergence of dual variables for ADA: (top) dual variables associated with average lower voltage limits for all buses; and (bottom) dual variables associated with average upper voltage limits for all buses. Sliding averages of optimization variables are depicted too.

deterministic alternatives.

Results in this chapter have been adapted to a microgrid model in the transmission level [LR15] where limits to the expected load not served (ELNS) are imposed in a two-stage dispatch problem. Similarly to this chapter, the dispatch is reformulated as a stochastic saddle point problem and solved using a stochastic approximation primal-dual method. Leveraging the proposed formulation and using sensitivity analysis, the marginal cost of reducing the ELNS imposed limits is identified as the price of reliability.

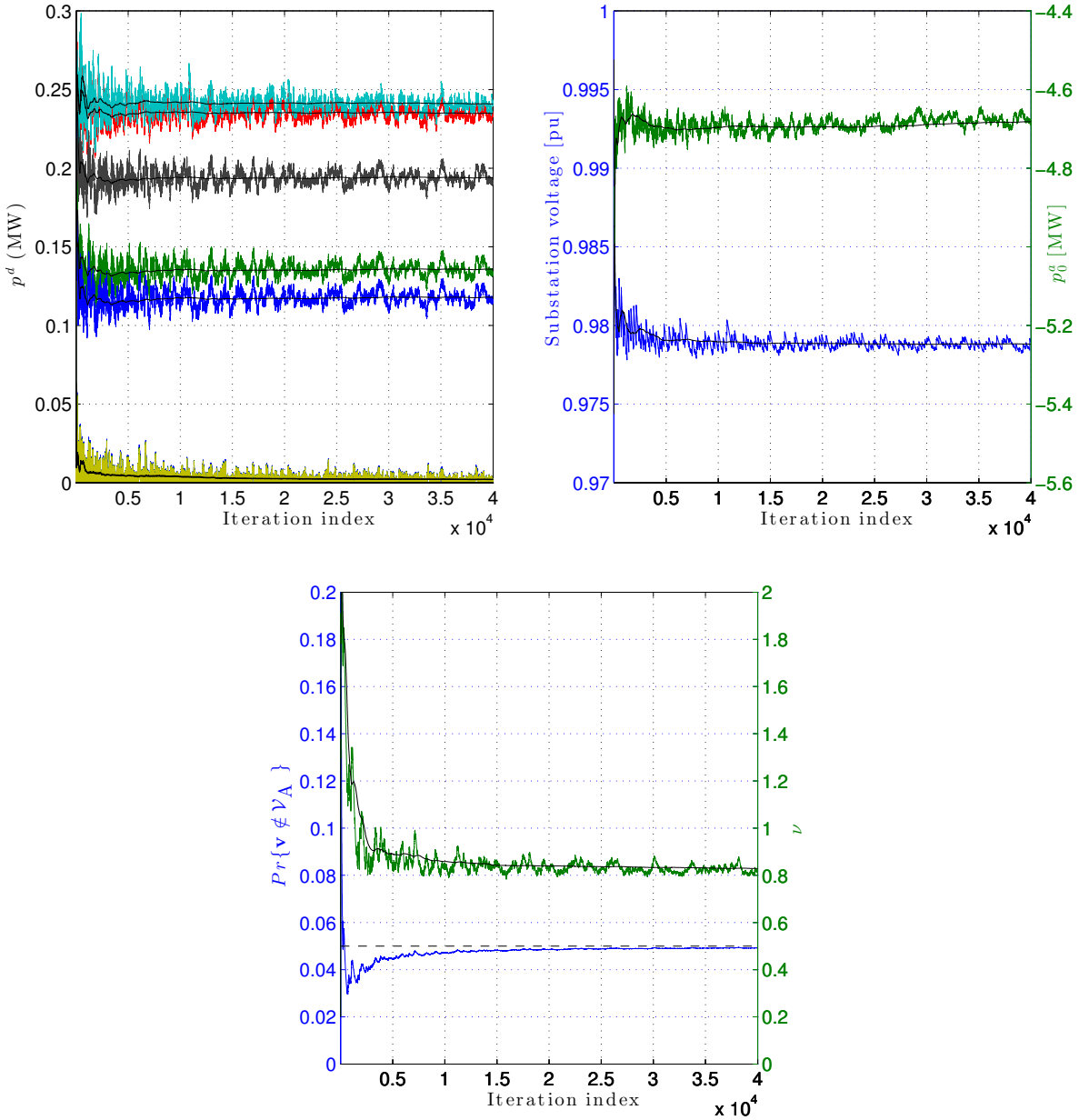

Figure 5.3: Convergence for PDA: (left) diesel generation; (middle) substation voltage (left y-axis) and energy exchange p_0^a (right y-axis); and (right) dual variable related to probabilistic constraint (left y-axis) and under-/over-voltage probability (right y-axis). Sliding averages of optimization variables are shown too.

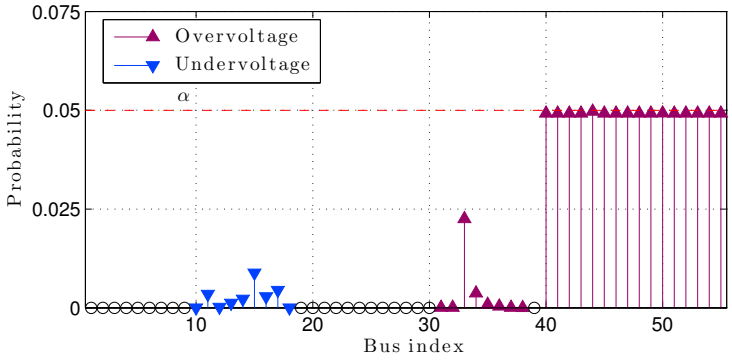

Figure 5.4: Per-bus probability of under-/over-voltages.

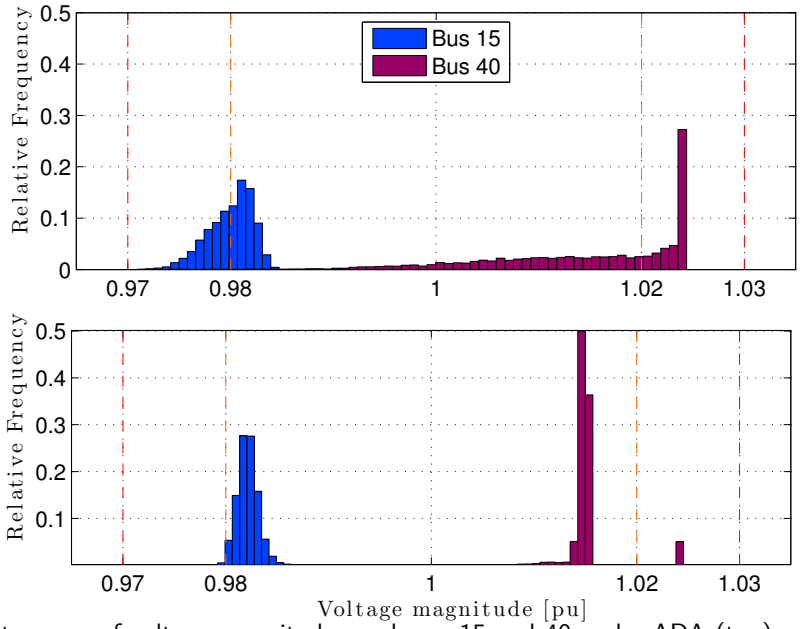

Figure 5.5: Histograms of voltage magnitudes on buses 15 and 40 under ADA (top) and PDA (bottom). Dashed lines show regulation limits $\mathcal{V}_A$ and $\mathcal{V}_B$.

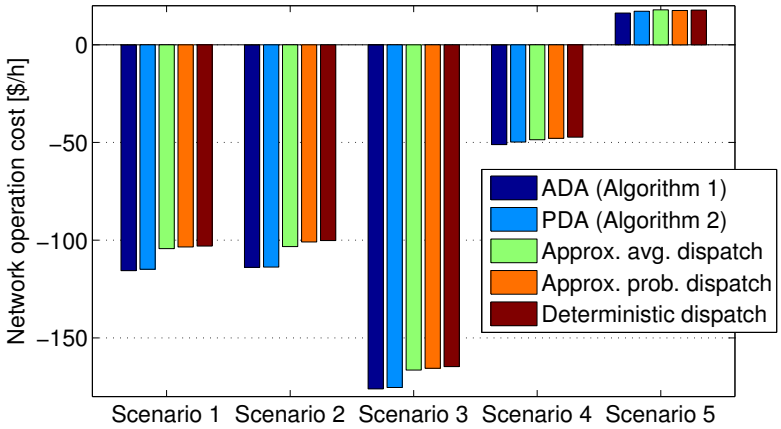

Figure 5.6: Performance for ADA, PDA, approximate average, approximate probabilistic, and deterministic scheme.

Chapter 6

Conclusions

The goal of this thesis was to blend nonlinear constrained optimization and stochastic signal processing to design (near optimal) schemes to optimize and monitor dynamic networks.

Cross-layer resource allocation (RA) algorithms were designed in chapter 2 to allocate resources (flows, channel access, power and rates) in a cellular system, where users transmit over a set of orthogonal channels. Both uplink and downlink setups were addressed. The developed resource allocation strategy depends on the instantaneous fading, the queue lengths, and user-specific weights. The Lagrange multipliers were estimated online using stochastic approximation iterations. A correspondence was established, for each user, between the transmit queue length and the estimated Lagrange multiplier associated with the average flow conservation constraint. This relationship still held for windowed average queue lengths and stochastic multipliers, allowing to analyze stability and average delay performance.

Numerical experiments in chapter 2 confirmed that, under mild conditions (stationarity of the fading gain process), there is a proportionality constant between the delay and the stepsize for any given setup. This allowed to propose a mechanism to effect delay priorities among users by tuning the stepsize of individual users. Experiments also pointed out that variations in the multiplier averaging window length do not necessarily have a monotonic impact on the system performance. It seems though, that long averaging windows may help to mitigate the negative effects of large stepsizes (used to effect lower delays).

The presence of average power consumption constraints was common ground for the cross-layer and cognitive radio (CR) designs undertaken in chapters 2, 3, and 4. Such constraints were addressed by means of dual decomposition and stochastic dual gradient schemes. A similar strategy was followed to deal with probability of interference constraints in chapters 3 and 4.

Moreover, in chapter 3 the probabilistic interference constraint was tailored to account for imperfections present in the sensing and channel state information (CSI) acquisition phase. Two versions of the interference constraint were investigated. The first one was a short-term constraint that took into account CSI imperfections to limit the *instantaneous* probability of interference. The second type of interference constraint was long-term and capitalized on the diversity of the interfering channel to upper-bound the fraction of time during which interference occurred. It was shown that the optimal schemes maximize a functional which accounts for the quality of the secondary links (in terms of rate), the transmission power, and the probability of interfering with primary users. Several of those terms were multiplied by Lagrange multipliers whose value depended on the history of the system and the requirements of the primary and secondary networks. Stochastic algorithms were introduced to estimate and predict the probability of interference, and to estimate the optimum value of the multipliers.

The result of zero-duality gap between the presented resource allocation optimizations and their Lagrangian duals allowed to solve the non convex problems that showed up in chapter 3 with manageable complexity. Beyond this, the availability of a multiplier that objectively represents the penalty for interfering a primary user (PU) motivated the joint optimization of RA and sensing schemes for a CR in chapter 4. Note also that the same zero-duality gap property appears again in chapter 5 for a probability constraint appearing in an electric power network.

The aim of Chapter 4 was to design jointly optimal RA and sensing schemes for an interweave CR with multiple primary and secondary users. Given the time correlation in the state information of the primary network (SIPN), sensing decisions were coupled across time, so that the problem fell into the class of stochastic dynamic programming (SDP). Moreover, since the sensing decisions impacted the observability of the PU activity, the problem was formulated as a partially observable Markov decision process (POMDP). To address the associated complexity, both the objective and the quality of service (QoS) constraints were formulated as long-term (discounted) time averages (which are less restrictive than their short-term counterparts and, hence, give rise to a better global objective). Additionally, a discounted, infinite time-horizon formulation was chosen for the POMDP, giving rise to stationary value functions. The same dual decomposition methods that were leveraged in Chapter 3 facilitated the optimization.

A two-step approach that solved first for the optimal RA for *any* sensing scheme and, then, solved for the optimal sensing, was implemented. As a result, the SDP solved in last instance had a state space much smaller than that of the original formulation. In particular, the optimization was separable across channels, partially separable across secondary users (SUs), and separable across time. A formulation based on Q-functions (instead of the commonly employed V-functions) was chosen because it would facilitate the implementation of an adaptive scheme to estimate the Q-functions online in non-stationary environments. It turned out that the optimal sensing at time n depended on: the state information of the secondary network (SISN) and SIPN at n; the stationary Lagrange *multipliers* associated with the dualized constraints; and the stationary Q-function associated with the long-term objective. Note also that both the Q-functions and the multipliers accounted for the effect of sensing and RA corresponding to future time instants. The expressions for the optimal RA and the optimal sensing policies were intuitive and relatively simple.

For each time instant, the CR had to run five consecutive steps: 1) acquisition of the SISN; 2) prediction of SIPN to form the *pre-decision* belief; 3) implementation of the optimal sensing decision based on 1-2; 4) update of the SIPN (using the output of the sensing) to form the *post-decision* belief; 5) implementation of the optimal RA. The expression for the RA was taken into account in 3) and the one for the sensing policy was considered in 5); testament to the fact that this was a joint design. Provided that the stationary value function and multipliers were available (they could be found off-line via numerical search during the initialization phase), the online implementation of the optimal schemes entailed very low computational complexity. Numerical results showed that both optimal and near-optimal policies corresponding to the formulated dynamic program (DP) perform significantly better than myopic policies, conversely to the model studied in [Zha07b].

Energy management of smart distribution grids involving decisions at slower and faster timescales was investigated in Chapter 5. The two-timescale system model is pertinent due to the joint presence of electromechanical components and fast-adapting inverters attached to renewable energy sources such as PV. Since slow-timescale controls remained fixed over multiple PV operation slots, decisions were coupled across time in a stochastic manner. To accommodate solar energy fluctuations, voltages were allowed to be sporadically overloaded; hence introducing coupling of fast-timescale variables on the average or in probability. Average voltage constraints resulted in a stochastic convex-concave problem, whereas non-convex probabilistic constraints were tackled using dual decomposition and convex optimization. Efficient algorithms for finding both slow and fast controls using only random samples were put forth. The two novel solvers proposed proved to be convergent in terms of the primal and dual variables. Since probabilistic constraints were

enforced grid-wise, voltage limits on individual buses could be slightly oversatisfied; nevertheless, they attained significantly lower operational costs when compared to deterministic alternatives.

6.1 Future work directions

Several directions to extend the results presented in this thesis to more diverse aspects of resource allocation in dynamic wireless networks and smart grids are presented below.

Possible lines of work to extend the CR research in chapters 3 and 4 are:

- Considering more complex models for the CSI such as imperfect SISN and non-Markovian models for the PU activity. The latter calls for efficient ways to represent and update the belief, e.g., by applying recursive Bayesian estimation to the augmented models outlined at the end of Sec. 3.2.1.

- Incorporating additional sources of correlation (correlation across time for the SISN and correlation across channels for the SIPN), motivating the study of more challenging formulations of the POMDP.

- Addressing the jointly optimal design for underlay CR networks, by means of formulations that limit the average interfering power or the PU average rate loss due to the interference. In such cases, information about the channel gains between the SUs and PUs would be required.

- Developing distributed schemes that address the problem of cooperative sensing as well as the problem of distributed RA is another interesting extension. The achieved design of stochastic sensing and RA allocation policies would facilitate that such distributed algorithms cope with noise and delay in the network state information (NSI) exchanged by the nodes. For some of these extensions, designs based on suboptimal but low-complexity solutions are an alternative worth exploring.

- Studying low-complexity stochastic estimations is also of interest from the algorithmic perspective. Alternating iterations that attain joint convergence of the Lagrange multipliers and value functions, together with alternatives such as reinforcement learning [Wie12] techniques or online parametric functional estimation [Pow07, Ch. 8] can be explored to this end.

Regarding the research in smart grid optimization, interesting avenues to extend the results in chapter 5 are:

- Considering additional sources of variability such as real-time electricity prices in the energy transactions with the main grid.

- Enforcing probabilistic constraints on a per-bus basis, which is expected to entail lower costs than doing so grid-wise. Also, the strategy of allowing an occasional overloading is not necessarily restricted to the case of voltages: other components such as the inverters in photovoltaic (PV) systems can also be subject to similar ergodic constraints. In all these cases, the Lagrangian minimization carried out in each iteration becomes a combinatorial problem that deserves a more cautious study.

- Formulation of decentralized implementations, as it can robustify the power network to the outage of central controllers, and may also reduce the traffic of user-sensitive information between the network buses and controlling entities, helping to maintain user security and privacy.

- Including voltage regulating devices in the models, which would require implementing advanced decision support schemes. Their optimal design would probably require SDP techniques related to those implemented in Chapter 4.
- Incorporation of demand-response capable appliances and networked energy storage, and the design of their optimal energy scheduling policies.
- Joint management of power networks and other energy facilities such as district-heating networks. In the same line, the joint optimization of power generation and activity schedules in industry or hospitals is of interest too.

Last but not least, it is also worth exploring the application of the tools studied in this thesis in other practical cyber-physical systems, such as intelligent transport networks.

References

Bibliography

[Ami05] Amin S.M. and Wollenberg B.F., Toward a smart grid: power delivery for the 21st century, *IEEE power and energy magazine*, volume 3, 5, pp. 34–41, 2005.

[ans11] C84.1-1995 Electric Power Systems and Equipment Voltage Ratings (60 Herz), 2011.

[Bar89] Baran M. and Wu F., Network reconfiguration in distribution systems for loss reduction and load balancing, *IEEE Trans. Power Delivery*, volume 4, 2, pp. 1401–1407, 1989.

[Bar11] Barbarossa S., Carfagna A., Sardellitti S., Omilipo M., and Pescosolido L., Optimal radio access in femtocell networks based on markov modeling of interferers' activity, in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3212–3215, IEEE, 2011.

[Ber95] Bertsekas D.P., *Dynamic programming and optimal control*, volume 1, Athena Scientific Belmont, MA, 1995.

[Ber99] Bertsekas D.P., *Nonlinear programming*, Athena scientific Belmont, 1999.

[Ber03] Bertsekas D., Nedic A., and Ozdaglar A.E., *Convex Analysis and Optimization*, Athena Scientific, 2003.

[Ber11] Bertsimas D., Brown D.B., and Caramanis C., Theory and applications of robust optimization, *SIAM review*, volume 53, 3, pp. 464–501, 2011.

[Bie14] Bienstock D., Chertkov M., and Harnett S., Chance-Constrained Optimal Power Flow: Risk-Aware Network Control under Uncertainty, *SIAM Rev.*, volume 56, 3, pp. 461–495, 2014.

[Bir85] Birge J.R., Decomposition and partitioning methods for multistage stochastic linear programs, *Operations Research*, volume 33, 5, pp. 989–1007, 1985.

[Bol15] Bolognani S., Carli R., Cavraro G., and Zampieri S., Distributed Reactive Power Feedback Control for Voltage Regulation and Loss Minimization, *IEEE Trans. Automat. Contr.*, volume 60, 4, pp. 966–981, 2015.

[Bou05] Bouffard F., Galiana F.D., and Conejo A.J., Market-clearing with Stochastic Security–Part I: Formulation, *IEEE Trans. Power Syst.*, volume 20, 4, pp. 1818–1826, 2005.

[Bou08] Bouffard F. and Galiana F.D., Stochastic Security for Operations Planning With Significant Wind Power Generation, *IEEE Trans. Power Syst.*, volume 23, 2, pp. 306–316, 2008.

[Boy04] Boyd S. and Vandenberghe L., *Convex Optimization*, Cambridge University Press, New York, NY, 2004.

[Boy06] Boyd S. and Mutapcic A., Subgradient methods, *Lecture notes of EE364b, Stanford University, Winter Quarter*, volume 2007, 2006.

[Bra03] Braziunas D., Pomdp solution methods, *University of Toronto, Tech. Rep*, 2003.

[Bus10] Busoniu L., Babuska R., De Schutter B., and Ernst D., *Reinforcement learning and dynamic programming using function approximators*, volume 39, CRC press, 2010.

[Car08] Carvalho P.M.S., Correia P.F., and Ferreira L.A., Distributed Reactive Power Generation Control for Voltage Rise Mitigation in Distribution Networks, *IEEE Trans. Power Syst.*, volume 23, 2, pp. 766–772, 2008.

[Cas97] Castanon D.A., Approximate dynamic programming for sensor management, in *Decision and Control, 1997., Proceedings of the 36th IEEE Conference on*, volume 2, pp. 1202–1207, IEEE, 1997.

[Che06] Chen L., Low S.H., Chiang M., and Doyle J.C., Cross-layer congestion control, routing and scheduling design in ad hoc wireless networks, 2006.

[Che08a] Chen Y., Yu G., Zhang Z., Chen H.h., and Qiu P., On cognitive radio networks with opportunistic power control strategies in fading channels, *IEEE Transactions on Wireless Communications*, volume 7, 7, pp. 2752–2761, 2008.

[Che08b] Chen Y., Zhao Q., and Swami A., Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors, *IEEE Transactions on Information Theory*, volume 54, 5, pp. 2053–2071, 2008.

[Chi07] Chiang M., Low S.H., Calderbank A.R., and Doyle J.C., Layering as optimization decomposition: A mathematical theory of network architectures, *Proceedings of the IEEE*, volume 95, 1, pp. 255–312, 2007.

[Con10] Conejo A.J., Carrión M., and Morales J.M., *Decision making under uncertainty in electricity markets*, Springer, 2010.

[Cov06] Cover T.M. and Thomas J.A., Elements of information theory 2nd edition, 2006.

[Dal11] Dall'Anese E., Kim S.J., Giannakis G.B., and Pupolin S., Power control for cognitive radio networks under channel uncertainty, *IEEE Transactions on Wireless Communications*, volume 10, 10, pp. 3541–3551, 2011.

[Dal13] Dall'Anese E., Zhu H., and Giannakis G.B., Distributed Optimal Power Flow for Smart Microgrids, *IEEE Trans. Smart Grid*, volume 4, 3, pp. 1464–1475, 2013.

[Dal14] Dall'Anese E., Dhople S.V., and Giannakis G.B., Optimal dispatch of photovoltaic inverters in residential distribution systems, *IEEE Trans. Sustain. Energy*, volume 5, 2, pp. 487–497, 2014.

[Dalar] Dall'Anese E. and Simonetto A., Optimal Power Flow Pursuit, *IEEE Trans. Smart Grid*, 2016 (to appear).

[DD12] De Domenico A., Strinati E.C., and Di Benedetto M.G., A survey on MAC strategies for cognitive radio networks, *IEEE Communications Surveys & Tutorials*, volume 14, 1, pp. 21–44, 2012.

[DL11] Di Lorenzo P. and Barbarossa S., A bio-inspired swarming algorithm for decentralized access in cognitive radio, *IEEE Transactions on Signal Processing*, volume 59, 12, pp. 6160–6174, 2011.

[Dor14] Dorfler F., Simpson-Porco J.W., and Bullo F., Breaking the hierarchy: Distributed control & economic optimality in microgrids, 2014, URL `http://arxiv.org/pdf/1401.1767v1.pdf`, (under review).

[en504] EN 50160: Voltage Characteristics of Public Distribution Systems, 2004.

[Ery06] Eryilmaz A. and Srikant R., Joint congestion control, routing, and MAC for stability and fairness in wireless networks, *IEEE Journal on Selected Areas in Communications*, volume 24, 8, pp. 1514–1524, 2006.

[Fan12] Fang X., Misra S., Xue G., and Yang D., Smart grid—The new and improved power grid: A survey, *IEEE communications surveys & tutorials*, volume 14, 4, pp. 944–980, 2012.

[Far10] Farhangi H., The path of the smart grid, *IEEE power and energy magazine*, volume 8, 1, pp. 18–28, 2010.

[Far12] Farivar M., Neal R., Clarke C., and Low S., Optimal inverter VAR control in distribution systems with high PV penetration, in *Proc. IEEE Power & Energy Society General Meeting*, San Diego, CA, 2012.

[Far13a] Farivar M., Chen L., and Low S., Equilibrium and Dynamics of Local Voltage Control in Distribution Systems, in *Proc. IEEE Conf. on Decision and Control*, pp. 4329–4334, Florence, Italy, 2013.

[Far13b] Farivar M. and Low S., Branch Flow Model: Relaxations and Convexification — Part I, *IEEE Trans. Power Syst.*, volume 28, 3, pp. 2554–2564, 2013.

[Gan15] Gan L., Li N., Topcu U., and Low S.H., Exact Convex Relaxation of Optimal Power Flow in Radial Networks, *IEEE Trans. Automat. Contr.*, volume 60, 1, pp. 72–87, 2015.

[Gao09] Gao N. and Wang X., Stochastic resource allocation over fading multiple access and broadcast channels, in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pp. 1–5, IEEE, 2009.

[Gat10] Gatsis N., Ribeiro A., and Giannakis G.B., A class of convergent algorithms for resource allocation in wireless fading networks, *IEEE Transactions on Wireless Communications*, volume 9, 5, pp. 1808–1823, 2010.

[Gat12] Gatsis N., *Resource Management in Wireless Networks and the Smart Power Grid*, Ph.D. thesis, UNIVERSITY OF MINNESOTA, 2012.

[Gat14] Gatsis N. and Marques A.G., A stochastic approximation approach to load shedding in power networks, in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6464–6468, IEEE, 2014.

[Geo06] Georgiadis L., Neely M.J., and Tassiulas L., *Resource allocation and cross-layer control in wireless networks*, Now Publishers Inc, 2006.

[Gha05] Ghasemi A. and Sousa E.S., Collaborative spectrum sensing for opportunistic access in fading environments, in *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005.*, pp. 131–136, IEEE, 2005.

[Gha07] Ghasemi A. and Sousa E.S., Fundamental limits of spectrum-sharing in fading environments, *IEEE Transactions on Wireless Communications*, volume 6, 2, pp. 649–658, 2007.

[Gia06] Giannoulis A., Tsoukatos K.P., and Tassiulas L., Lightweight cross-layer control algorithms for fairness and energy efficiency in CDMA ad-hoc networks, in *2006 4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, pp. 1–8, IEEE, 2006.

[Gia13] Giannakis G.B., Kekatos V., Gatsis N., Kim S.J., Zhu H., and Wollenberg B., Monitoring and Optimization for Power Grids: A Signal Processing Perspective, *IEEE Signal Processing Mag.*, volume 30, 5, pp. 107–128, 2013.

[Gol05] Goldsmith A., *Wireless communications*, Cambridge university press, 2005.

[Gol09] Goldsmith A., Jafar S.A., Maric I., and Srinivasa S., Breaking spectrum gridlock with cognitive radios: An information theoretic perspective, *Proceedings of the IEEE*, volume 97, 5, pp. 894–914, 2009.

[Gon11] Gong X., Vorobyov S.A., and Tellambura C., Joint bandwidth and power allocation in cognitive radio networks under fading channels, in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2976–2979, IEEE, 2011.

[Hay05] Haykin S., Cognitive radio: brain-empowered wireless communications, *IEEE journal on selected areas in communications*, volume 23, 2, pp. 201–220, 2005.

[He11] He Y. and Dey S., Power allocation in spectrum sharing cognitive radio networks with quantized channel information, *IEEE Transactions on Communications*, volume 59, 6, pp. 1644–1656, 2011.

[Hua11] Huang L. and Neely M.J., Delay reduction via Lagrange multipliers in stochastic network optimization, *IEEE Transactions on Automatic Control*, volume 56, 4, pp. 842–857, 2011.

[Jaf07] Jafar S.A. and Srinivasa S., Capacity limits of cognitive radio with distributed and dynamic spectral activity, *IEEE Journal on selected Areas in Communications*, volume 25, 3, pp. 529–537, 2007.

[Jia05] Jiang Z., Ge Y., and Li Y., Max-utility wireless resource management for best-effort traffic, *IEEE Transactions on Wireless Communications*, volume 4, 1, pp. 100–111, 2005.

[Joh06] Johansson B., Soldati P., and Johansson M., Mathematical decomposition techniques for distributed cross-layer optimization of data networks, *IEEE Journal on Selected Areas in Communications*, volume 24, 8, pp. 1535–1547, 2006.

[Kae98] Kaelbling L.P., Littman M.L., and Cassandra A.R., Planning and acting in partially observable stochastic domains, *Artificial intelligence*, volume 101, 1, pp. 99–134, 1998.

[Kal94] KALL P. and WALLACE S.W., *Stochastic Programming*, John Wiley and Sons, Chichester, 1994.

[Kan09] KANG X., LIANG Y.C., NALLANATHAN A., GARG H.K., and ZHANG R., Optimal power allocation for fading channels in cognitive radio networks: Ergodic capacity and outage capacity, *IEEE Transactions on Wireless Communications*, volume 8, 2, pp. 940–950, 2009.

[Kek14] KEKATOS V., GIANNAKIS G.B., and BALDICK R., Grid topology identification using electricity prices, in *2014 IEEE PES General Meeting| Conference & Exposition*, pp. 1–5, IEEE, 2014.

[Kek15a] KEKATOS V., WANG G., CONEJO A.J., and GIANNAKIS G.B., Stochastic Reactive Power Management in Microgrids With Renewables, *IEEE Trans. Power Syst.*, volume 30, 6, pp. 3386–3395, 2015.

[Kek15b] KEKATOS V., ZHANG L., GIANNAKIS G.B., and BALDICK R., Accelerated Localized Voltage Regulation in Single-Phase Distribution Grids, in *Proc. IEEE Intl. Conf. on Smart Grid Commun.*, Miami, FL, 2015.

[Kek15c] KEKATOS V., ZHANG L., GIANNAKIS G.B., and BALDICK R., Voltage Regulation Algorithms for Multiphase Power Distribution Grids, *IEEE Trans. Power Syst.*, 2015, URL `http://arxiv.org/abs/1508.06594`, (submitted).

[Kel98] KELLY F.P., MAULLOO A.K., and TAN D.K., Rate control for communication networks: shadow prices, proportional fairness and stability, *Journal of the Operational Research society*, volume 49, 3, pp. 237–252, 1998.

[Kim10] KIM S.J. and GIANNAKIS G.B., Sequential and cooperative sensing for multi-channel cognitive radios, *IEEE Transactions on Signal Processing*, volume 58, 8, pp. 4239–4253, 2010.

[Kun10] KUNDUR D., FENG X., LIU S., ZOURNTOS T., and BUTLER-PURRY K.L., Towards a framework for cyber attack impact analysis of the electric smart grid, in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, pp. 244–249, IEEE, 2010.

[Kus03] KUSHNER H. and YIN G.G., *Stochastic approximation and recursive algorithms and applications*, volume 35, Springer Science & Business Media, 2003.

[Lee06] LEE J.W., MAZUMDAR R.R., and SHROFF N.B., Opportunistic power scheduling for dynamic multi-server wireless systems, *IEEE Transactions on Wireless Communications*, volume 5, 6, pp. 1506–1515, 2006.

[Li01] LI L. and GOLDSMITH A.J., Capacity and optimal resource allocation for fading broadcast channels. I. Ergodic capacity, *IEEE Transactions on Information Theory*, volume 47, 3, pp. 1083–1102, 2001.

[Lia08] LIANG Y.C., ZENG Y., PEH E.C., and HOANG A.T., Sensing-throughput tradeoff for cognitive radio networks, *IEEE transactions on Wireless Communications*, volume 7, 4, pp. 1326–1337, 2008.

[Lin05] LIN Y.H. and CRUZ R.L., Opportunistic link scheduling, power control, and routing for multi-hop wireless networks over time varying channels, in *Proc. Allerton Conf. Commun., Control, and Computing*, pp. 976–985, Monticello, IL, 2005.

[Lin06a] Lin X. and Shroff N.B., The impact of imperfect scheduling on cross-layer congestion control in wireless networks, *IEEE/ACM transactions on networking*, volume 14, 2, pp. 302–315, 2006.

[Lin06b] Lin X., Shroff N.B., and Srikant R., A tutorial on cross-layer optimization in wireless networks, *IEEE Journal on Selected areas in Communications*, volume 24, 8, pp. 1452–1463, 2006.

[Liu06] Liu P., Berry R.A., and Honig M.L., A fluid analysis of a utility-based wireless scheduling policy, *IEEE Transactions on Information Theory*, volume 52, 7, pp. 2872–2889, 2006.

[Liu08] Liu E. and Bebic J., Distribution System Voltage Performance Analysis for High-Penetration Photovoltaics, Technical report, National Renewable Energy Laboratory, 2008, URL `https://www1.eere.energy.gov/solar/pdfs/42298.pdf`.

[Lov08] Love D.J., Heath R.W., Lau V.K., Gesbert D., Rao B.D., and Andrews M., An overview of limited feedback in wireless communication systems, *IEEE Journal on Selected Areas in Communications*, volume 26, 8, pp. 1341–1365, 2008.

[Low99] Low S.H. and Lapsley D.E., Optimization flow control—I: basic algorithm and convergence, *IEEE/ACM Transactions on Networking (TON)*, volume 7, 6, pp. 861–874, 1999.

[Low03] Low S.H., A duality model of TCP and queue management algorithms, *IEEE/ACM Transactions On Networking*, volume 11, 4, pp. 525–536, 2003.

[LR10] Lopez-Ramos L.M., Marques A.G., Ramos J., and Caamano A.J., Cross-layer resource allocation for downlink access using instantaneous fading and queue length information, in *2010 IEEE Globecom Workshops*, pp. 1212–1216, IEEE, 2010.

[LR13] Lopez-Ramos L.M., Marques A.G., and Ramos J., Soft-decision sequential sensing for optimization of interweave Cognitive Radio networks, in *2013 IEEE 14th Workshop on Signal Proc. Advances in Wireless Comms. (SPAWC)*, pp. 235–239, IEEE, 2013.

[LR14a] Lopez-Ramos L.M., Marques A.G., and Ramos J., Joint sensing and resource allocation for underlay cognitive radios, in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7283–7287, IEEE, 2014.

[LR14b] Lopez-Ramos L.M., Marques A.G., and Ramos J., Jointly Optimal Sensing and Resource Allocation for Multiuser Interweave Cognitive Radios, *IEEE Transactions on Wireless Communications*, volume 13, 11, pp. 5954–5967, 2014.

[LR15] Lopez-Ramos L.M., Kekatos V., Marques A.G., and Giannakis G.B., Microgrid Dispatch and Price of Reliability using Stochastic Approximation, in *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1131–1135, 2015.

[LR16] Lopez-Ramos L.M., Kekatos V., Marques A.G., and Giannakis G.B., Two-timescale stochastic dispatch for power distribution networks, *IEEE Transactions on Smart Grid (submitted)*, 2016.

[Mar08a] Marques A.G., Giannakis G.B., Digham F.F., and Ramos F.J., Power-efficient wireless OFDMA using limited-rate feedback, *IEEE Transactions on Wireless Communications*, volume 7, 2, pp. 685–696, 2008.

[Mar08b] Marques A.G., Wang X., and Giannakis G.B., Optimal stochastic dual resource allocation for cognitive radios based on quantized CSI, in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2801–2804, IEEE, 2008.

[Mar09] Marques A.G., Wang X., and Giannakis G.B., Dynamic resource management for cognitive radios using limited-rate feedback, *IEEE Transactions on signal processing*, volume 57, 9, pp. 3651–3666, 2009.

[Mar10] Marques A.G., Giannakis G.B., and Ramos J., Stochastic cross-layer resource allocation for wireless networks using orthogonal access: Optimality and delay analysis, in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3154–3157, IEEE, 2010.

[Mar11a] Marques A.G., Gatsis N., and Giannakis G.B., Optimal cross-layer design of wireless fading multi-hop networks, *Cross Layer Designs in WLAN Systems*, 2011.

[Mar11b] Marques A.G., Giannakis G.B., Lopez-Ramos L.M., and Ramos J., Stochastic resource allocation for cognitive radio networks based on imperfect state information, in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3196–3199, IEEE, 2011.

[Mar11c] Marques A.G., Giannakis G.B., and Ramos J., Optimizing orthogonal multiple access based on quantized channel state information, *IEEE Transactions on Signal Processing*, volume 59, 10, pp. 5023–5038, 2011.

[Mar11d] Marques A.G., Lopez-Ramos L.M., Giannakis G.B., and Ramos J., Adaptive underlay cognitive radios with imperfect CSI and probabilistic interference constraints, in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2011 4th IEEE International Workshop on*, pp. 161–164, IEEE, 2011.

[Mar12a] Marques A.G., Lopez-Ramos L.M., Giannakis G.B., and Ramos J., Resource allocation for interweave and underlay CRs under probability-of-interference constraints, *IEEE Journal on Selected Areas in Communications*, volume 30, 10, pp. 1922–1933, 2012.

[Mar12b] Marques A.G., Lopez-Ramos L.M., Giannakis G.B., Ramos J., and Caamaño A.J., Optimal cross-layer resource allocation in cellular networks using channel- and queue-state information, *IEEE Trans. on Vehicular Tech.*, volume 61, 6, pp. 2789–2807, 2012.

[Mar12c] Marques A.G., Lopez-Ramos L.M., and Ramos J., Cognitive radios with ergodic capacity guarantees for primary users, in *Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), 2012 7th International ICST Conf. on*, pp. 147–152, IEEE, 2012.

[Mar13] Marques A.G., Figuera C., Rey-Moreno C., and Simo-Reigadas J., Asymptotically optimal cross-layer schemes for relay networks with short-term and long-term constraints, *IEEE Transactions on Wireless Communications*, volume 12, 1, pp. 333–345, 2013.

[Mar14] Marques A.G., Dall'Anese E., and Giannakis G.B., Cross-layer optimization and receiver localization for cognitive networks using interference tweets, *IEEE Journal on Selected Areas in Communications*, volume 32, 3, pp. 641–653, 2014.

[Mou13] Moursi M.S.E., Xiao W., and Kirtley J.L., Fault ride through capability for grid interfacing large scale PV power plants, *IET Gener. Transm. Dis.*, volume 7, 9, pp. 1027–1036, 2013.

[Mus09] Musavian L. and Aïssa S., Fundamental capacity limits of cognitive radio in fading environments with imperfect channel information, *IEEE Transactions on Communications*, volume 57, 11, pp. 3472–3480, 2009.

[Ned09] Nedic A. and Ozdaglar A., Approximate primal solutions and rate analysis for dual subgradient methods, *SIAM Journal on Optimization*, volume 19, 4, pp. 1757–1780, 2009.

[Nee05] Neely M.J., Modiano E., and Rohrs C.E., Dynamic power allocation and routing for time-varying wireless networks, *IEEE Journal on Selected Areas in Communications*, volume 23, 1, pp. 89–103, 2005.

[Nem09] Nemirovski A., Juditsky A., Lan G., and Shapiro A., Robust stochastic approximation approach to stochastic programming, *SIAM J. Optim.*, volume 19, 4, pp. 1574–1609, 2009.

[Pal06] Palomar D.P. and Chiang M., A tutorial on decomposition methods for network utility maximization, *IEEE Journal on Selected Areas in Communications*, volume 24, 8, pp. 1439–1451, 2006.

[Pau11] Paudyal S., Canizares C.A., and Bhattacharya K., Optimal operation of distribution feeders in smart grids, *IEEE Trans. Ind. Applicat.*, volume 10, 58, pp. 4495—4503, 2011.

[Penar] Peng Q. and Low S.H., Distributed Optimal Power Flow Algorithm for Radial Networks, I: Balanced Single Phase Case, *IEEE Trans. Smart Grid*, 2016 (to appear).

[Pfl12] Pflug G.C., *Optimization of stochastic models: the interface between simulation and optimization*, volume 373, Springer Science & Business Media, 2012.

[Pow07] Powell W.B., *Approximate Dynamic Programming: Solving the curses of dimensionality*, volume 703, John Wiley & Sons, 2007.

[Raj11] Rajawat K., Gatsis N., and Giannakis G.B., Cross-layer designs in coded wireless fading networks with multicast, *IEEE/ACM Transactions on Networking*, volume 19, 5, pp. 1276–1289, 2011.

[Rib10a] Ribeiro A., Ergodic stochastic optimization algorithms for wireless communication and networking, *IEEE Transactions on Signal Processing*, volume 58, 12, pp. 6369–6386, 2010.

[Rib10b] Ribeiro A. and Giannakis G.B., Separation principles in wireless networking, *IEEE Transactions on Information Theory*, volume 56, 9, pp. 4488–4505, 2010.

[Rob51] Robbins H. and Monro S., A stochastic approximation method, *The annals of mathematical statistics*, pp. 400–407, 1951.

[Rus97] Ruszczyński A., Decomposition methods in stochastic programming, *Mathematical programming*, volume 79, 1-3, pp. 333–353, 1997.

[Sha08] Shakkottai S., Shakkottai S.G., and Srikant R., *Network optimization and control*, Now Publishers Inc, 2008.

[Sol06] Soldati P., Johansson B., and Johansson M., Proportionally fair allocation of end-to-end bandwidth in STDMA wireless networks, in *Proceedings of the 7th ACM international symposium on Mobile ad hoc networking and computing*, pp. 286–297, ACM, 2006.

[Sol12] Solares J.R.A., Rezki Z., and Alouini M.S., Optimal power allocation of a sensor node under different rate constraints, in *2012 IEEE International Conference on Communications (ICC)*, pp. 2334–2338, IEEE, 2012.

[Sto05] Stolyar A.L., Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm, *Queueing Systems*, volume 50, 4, pp. 401–457, 2005.

[Sul14] Sulc P., Backhaus S., and Chertkov M., Optimal Distributed Control of Reactive Power Via the Alternating Direction Method of Multipliers, *IEEE Trans. Energy Conversion*, volume 29, 4, pp. 968–977, 2014.

[Sur10] Suraweera H.A., Smith P.J., and Shafi M., Capacity limits and performance analysis of cognitive radio with imperfect channel knowledge, *IEEE Transactions on Vehicular Technology*, volume 59, 4, pp. 1811–1822, 2010.

[Tas92] Tassiulas L. and Ephremides A., Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks, *IEEE transactions on automatic control*, volume 37, 12, pp. 1936–1948, 1992.

[Tur11] Turitsyn K., Sulc P., Backhaus S., and Chertkov M., Options for Control of Reactive Power by Distributed Photovoltaic Generators, *Proc. IEEE*, volume 99, 6, pp. 1063–1073, 2011.

[Urg09] Urgaonkar R. and Neely M.J., Opportunistic scheduling with reliability guarantees in cognitive radio networks, *IEEE Transactions on Mobile Computing*, volume 8, 6, pp. 766–777, 2009.

[Var11] Varaiya P.P., Wu F.F., and Bialek J.W., Smart Operation of Smart Grid: Risk-limiting dispatch, *Proc. IEEE*, volume 99, 1, pp. 40–57, 2011.

[Wan07] Wang X., Giannakis G.B., and Marques A.G., A unified approach to QoS-guaranteed scheduling for channel-adaptive wireless networks, *Proceedings of the IEEE*, volume 95, 12, pp. 2410–2431, 2007.

[Wan11] Wang X., Joint sensing-channel selection and power control for cognitive radios, *IEEE Transactions on Wireless Communications*, volume 10, 3, pp. 958–967, 2011.

[Wanar] Wang G., Kekatos V., Conejo A.J., and Giannakis G.B., Ergodic Energy Management Leveraging Resource Variability in Distribution Grids, *IEEE Trans. Power Syst.*, 2016 (to appear).

[Wie12] Wiering M. and van Otterlo M., *Reinforcement Learning: State-of-the-art*, Springer, 2012.

[Won99] Wong C.Y., Cheng R.S., Lataief K.B., and Murch R.D., Multiuser OFDM with adaptive subcarrier, bit, and power allocation, *IEEE Journal on selected areas in communications*, volume 17, 10, pp. 1747–1758, 1999.

[Woo12] Wood A.J. and Wollenberg B.F., *Power generation, operation, and control*, John Wiley & Sons, 2012.

[Yuc09] Yucek T. and Arslan H., A survey of spectrum sensing algorithms for cognitive radio applications, *IEEE communications surveys & tutorials*, volume 11, 1, pp. 116–130, 2009.

[Zha07a] Zhao Q. and Sadler B.M., A survey of dynamic spectrum access, *IEEE signal processing magazine*, volume 24, 3, pp. 79–89, 2007.

[Zha07b] Zhao Q., Tong L., Swami A., and Chen Y., Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework, *IEEE Journal on selected areas in communications*, volume 25, 3, pp. 589–600, 2007.

[Zha09] Zhang R., On peak versus average interference power constraints for protecting primary users in cognitive radio networks, *IEEE Transactions on Wireless Communications*, volume 8, 4, pp. 2112–2120, 2009.

[Zha10] Zhang R., Liang Y.C., and Cui S., Dynamic resource allocation in cognitive radio networks, *IEEE Signal Processing Magazine*, volume 27, 3, pp. 102–114, 2010.

[Zha11] Zhang X. and Su H., Opportunistic spectrum sharing schemes for CDMA-based uplink MAC in cognitive radio networks, *IEEE Journal on Selected Areas in Communications*, volume 29, 4, pp. 716–730, 2011.

[Zha13a] Zhang B., Dominguez-Garcia A., and Tse D., A local control approach to voltage regulation in distribution networks, in *Proc. North American Power Symposium*, Manhattan, KS, 2013.

[Zha13b] Zhang Y., Gatsis N., and Giannakis G.B., Robust Energy Management for Microgrids With High-Penetration Renewables, *IEEE Trans. Sustain. Energy*, volume 4, 4, pp. 944–953, 2013.

[Zha14a] Zhang B., Rajagopal R., and Tse D., Network Risk Limiting Dispatch: Optimal Control and Price of Uncertainty, *IEEE Trans. Automat. Contr.*, volume 59, 9, pp. 2442–2456, 2014.

[Zha14b] Zhang Y. and Giannakis G.B., Efficient Decentralized Economic Dispatch for Microgrids with Wind Power Integration, in *In Proc. IEEE Green Tech. Conf.*, Corpus Christi, TX, 2014.

[Zha15] Zhang B., Lam A.Y.S., Domínguez-García A.D., and Tse D., An Optimal and Distributed Method for Voltage Regulation in Power Distribution Systems, *IEEE Trans. Power Syst.*, volume 30, 4, pp. 1714–1726, 2015.

[Zhu12] Zhu H. and Giannakis G.B., Sparse overcomplete representations for efficient identification of power line outages, *IEEE Transactions on Power Systems*, volume 27, 4, pp. 2215–2224, 2012.

[Zhuar] Zhu H. and Liu H.J., Fast Local Voltage Control Under Limited Reactive Power: Optimality and Stability Analysis, *IEEE Trans. Power Syst.*, 2016, (to appear).

Acronyms

ADA	average dispatch algorithm
AMC	adaptive modulation and coding
AP	access point
BER	bit error rate
CR	cognitive radio
CSI	channel state information
DER	distributed energy resource
DP	dynamic program
HMM	hidden Markov model
KKT	Karush-Kuhn-Tucker
LDF	linear distribution flow
LHS	left-hand side
LTIC	long-term interference constraint
MDP	Markov decision process
NC	network controller
NSI	network state information
OPF	optimal power flow
PDA	probabilistic dispatch algorithm
pdf	probability density function
POMDP	partially observable Markov decision process
PU	primary user
PV	photovoltaic
QoS	quality of service
QSI	queue state information
RA	resource allocation
RES	renewable energy source
SA	stochastic approximation
SAA	sample average approximation
SCE	Southern California Edison
SDP	stochastic dynamic programming
SIPN	state information of the primary network
SISN	state information of the secondary network
SNR	signal-to-noise ratio
STIC	short-term interference constraint
SU	secondary user