QoS-aware Resource Provisioning in Virtualized Wireless Networks

Vikas Jumba

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

April 2016

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Engineering.

© 2016 Vikas Jumba
Dedicated to my parents.
Abstract

Wireless virtualization is emerging as a promising paradigm to tackle the issues of spectrum-crisis and network ossification via enabling common shared substrate of wireless resources among service providers, commonly referred as slices. Due to random nature of wireless channels and limited resources, virtualized wireless network (VWN) requires an efficient resource provisioning policy to operate. The objective of this thesis is to study and propose quality-of-service-aware (QoS-aware) resource provisioning policies applicable to VWNs.

The first part of the thesis focuses on the design of resource provisioning policies to satisfy instantaneous requirements (e.g., minimum reserved rates and resources) of slices and minimize the VWN outage due to infeasibility. At first, an optimal algorithm for resource provisioning is developed to simultaneously satisfy the rate and resource (power and sub-carriers) based requirements of two groups of slices. Afterwards, to deal with the issue of infeasibility in VWNs due to limited wireless resources, an admission control policy is proposed. In this context, an optimal admission control algorithm is developed to dynamically adjust the requirements of slices according to channel state information (CSI) and priorities of slices. Finally, to further improve the feasibility region, a joint power, sub-carrier and antenna allocation algorithm is developed for VWN with massive multiple input multiple output (MIMO) setup, where a base station (BS), equipped with large number of antennas, serves users belonging to different slices.

The second part of the thesis focuses on more realistic design of VWNs. Specifically, the issues of random, bursty traffic arrival in users’ queue, energy-efficiency and uncertain CSI at the BS, often experienced in practice, are addressed through resource provisioning policies. At first, to improve end-user service experience, cross-layer resource provisioning policies are presented. In this context, a dynamic resource provisioning policy, adaptable to both CSI and queue state information (QSI) of VWNs, is proposed to maintain the stable queue state of VWN. Afterwards, to minimize energy consumption, a resource provisioning policy is proposed to satisfy the maximum average packet transmission delay in VWNs to offer reliable end-users’ experience. Finally, the issue of uncertain CSI at the BS due to estimation errors is addressed through an energy-efficient robust resource provisioning policy. The total energy consumption of VWN is easily controllable via the cost factors of slices in the proposed policy. Simulations are performed to deeply analyze the effects of system parameters on VWN’s performance.
Sommaire

La virtualisation des réseaux sans fil est considérée comme un modèle prometteur pour régler la crise spectrale et l’ossification des réseaux. Notamment, celle-ci se fait en partageant, entre les fournisseurs de service, des ressources sans fil communes, généralement référé en tant que tranches. De plus, à cause de la nature aléatoire des canaux radio et des ressources limitées, un réseau sans fil virtualisé (VWN) demande une provision efficace des ressources pour fonctionner. L’objectif de cette thèse est donc d’étudier et de proposer un algorithme de provision de ressources en tenant compte de la qualité de service (Qos-aware) applicable aux VWNs.

La première partie de cette thèse se concentre sur la conception des politiques de provisionnement des ressources pour satisfaire aux exigences instantanées (ex., débit et ressource bases réservé) des tranches et minimiser l’interruption de service des VWN dû à l’infaisabilité. D’abord, un algorithme optimal de provisionnement de ressource est développé pour satisfaire simultanément au débit et aux ressources (puissance et sous porteuses) requis par deux groupes de tranches. Ensuite, pour faire face à la question de l’infaisabilité des VWNs causées par les ressources sans fil limitées, une politique de contrôle d’admission est proposée. Dans ce contexte, un algorithme de contrôle d’admission optimal est développé pour ajuster dynamiquement les exigences des tranches, selon les informations d’état de canal (CSI) et les priorités des tranches. Enfin, pour améliorer encore la région de faisabilité, un algorithme d’allocation conjointe de puissances, de sous porteuses et d’antennes est développé pour les VWNs avec une configuration de multiples antennes à grande échelle, où la station de base (BS), équipée avec un large nombre d’antennes, sert des utilisateurs appartenant à différentes tranches.

La deuxième partie de cette thèse propose une conception réaliste des VWNs. Plus précisément, les questions de l’arrivée aléatoire de trafic en rafales dans la file d’attente, de l’efficacité énergétique des utilisateurs, de l’incertitude de la CSI à la BS, souvent éprouvées en pratique, sont abordées à travers des politiques de provisionnement des ressources. Premièrement, pour améliorer l’expérience de l’utilisateur final, les politiques de provisionnement des ressources inter-couches sont présentés. Dans ce contexte, une politique de provisionnement dynamique des ressources, adaptable à la fois aux informations d’état de canal (CSI) et de file d’attente (QSI) des VWNs, est proposée pour maintenir la stabilité des files d’attente du VWNs. Ensuite, pour minimiser la consommation d’énergie, une politique de provisionnement des ressources est proposée pour satisfaire au maximum des retards moyens de transmission de paquets, afin d’offrir une expérience fiable aux utilisateurs finaux des VWNs. Enfin, la question de la CSI incertaine au BS en raison d’erreurs
d’estimation est abordée à travers une politique de provisionnement des ressources qui est à la fois robuste et efficace énergétiquement. La consommation totale d’énergie du VWN est facilement contrôlable via les facteurs de coût de tranches dans la politique proposée. Des simulations sont effectuées pour analyser en profondeur les effets des paramètres du système sur la performance des VWNs.
Acknowledgments

I would like to express a sincere thanks to my supervisor, Prof. Tho Le-Ngoc, for offering me an opportunity to pursue research during my MEng. study at McGill University. Without his invaluable support and guidance, I could hardly imagine the completion of this thesis. His deep knowledge, vast research experience and encouragements have always motivated me for a meaningful and concrete research.

I also wish to express my special gratitude to Dr. Saeedeh Parsaefard and Dr. Mahsa Derakhshani for their invaluable research coaching and collaboration. I am privileged to learn from their research experience and technical knowledge during this research journey. Thank you for inculcating me with the spirit of research and adding values in my academic career. I would also like to acknowledge the financial support received from Natural Sciences and Engineering Research Council of Canada (NSERC) and Grad Excellence Award in Electrical and Computer Engineering (McGill University).

I am also thankful to my colleagues and friends at McGill university, especially members of the Broadband Communication Research Laboratory: Yahya, Xiaowei, Atoosa, Rouzhu, Sanjeewa and Gowdemy, for their beneficial discussions in lab meetings. I am grateful to Rajesh, Alfred, Daniel and Dr Quang Dung Ho for proofreading my thesis. Specially, I thank Ahmed for helping me with the French translation of the thesis abstract.

My love and gratitude go to my parents and siblings for your unconditional love and trust in me. Your guidance, support and moral values always help me throughout my life. I hope this accomplishment makes all of you proud.
# Contents

1 Introduction 1

1.1 Current Trends and Future Perspectives of Wireless Networks . . . . . . . . . . 1
1.2 Wireless Virtualization: Background and Motivations . . . . . . . . . . . . . . . 2
1.3 Technical Challenges in Virtualized Wireless Networks . . . . . . . . . . . . . . 3
1.4 Literature Review . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
1.5 Thesis Objectives, Contributions and Organization . . . . . . . . . . . . . . . . 8

2 Joint Resource Provisioning and Admission Control in VWN 10

2.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10
2.2 Network Model and Problem Formulation . . . . . . . . . . . . . . . . . . . . . 12
2.3 Joint Power and Sub-Carrier Allocation . . . . . . . . . . . . . . . . . . . . . . 14
2.4 Admission Control . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
2.5 Numerical Results and Discussions . . . . . . . . . . . . . . . . . . . . . . . . . 18
   2.5.1 Performance of Slice Provisioning Algorithm . . . . . . . . . . . . . . . 19
   2.5.2 Performance of Admission Control Algorithm . . . . . . . . . . . . . . 21
2.6 Concluding Remarks . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 26

3 Resource Provisioning with Massive MIMO in VWN 27

3.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27
3.2 Network Model and Problem Formulation . . . . . . . . . . . . . . . . . . . . . 29
3.3 Proposed Algorithm . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32
3.4 Numerical Results and Discussions . . . . . . . . . . . . . . . . . . . . . . . . . 35
3.5 Concluding Remarks . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 38
Contents

4 Resource Provisioning with Stable Queue Control for VWN
4.1 Introduction .......................................................... 39
4.2 Network Model and Dynamic Resource Provisioning Problem ...... 41
4.3 Dynamic Resource Provisioning Algorithm ................................ 43
4.4 Numerical Results and Discussions .................................. 47
4.5 Concluding Remarks .................................................. 50

5 Delay-Aware and Energy-Efficient Resource Allocation in VWN
5.1 Introduction .......................................................... 51
5.2 Network Model and Problem Formulation ............................. 53
5.3 Proposed Algorithm .................................................. 54
5.4 Numerical Results and Discussions .................................. 57
5.5 Concluding Remarks .................................................. 61

6 Energy-Efficient Robust Resource Provisioning in VWN
6.1 Introduction .......................................................... 63
6.2 Network Model and Problem Formulation ............................. 65
6.3 Robust EE Resource Provisioning Algorithm .......................... 67
  6.3.1 Inner Optimization Problem .................................... 68
  6.3.2 Robust Slice Provisioning Algorithm ............................ 69
6.4 Numerical Results and Discussions .................................. 72
6.5 Concluding Remarks .................................................. 75

7 Conclusion
7.1 Summary ............................................................ 76
7.2 Suggested Future Works .............................................. 77

References 79
## List of Figures

2.1 Illustration of VWN with one BS and two groups of slices. .......................... 12  
2.2 Rate versus Number of Users in slices. .................................................. 21  
2.3 Pr(outage) versus $P_{\text{max}}$. .............................................................. 22  
2.4 $\varepsilon_{gr}$ versus $R_{gs}^{\text{rev}}$ for Scenario 1. ................................... 23  
2.5 $\varepsilon_{gr}$ versus $R_{gs}^{\text{rev}}$ and $\alpha_g$ for Scenario 2. ....................... 24  
3.1 Total rate vs. $M_{g}^{\text{max}}$. ................................................................. 35  
3.2 Total rate vs. $R_{gs}^{\text{rev}}$. ................................................................. 36  
3.3 Total rate vs. $\tau/T_c$. ..................................................................... 37  
3.4 Total Rate vs. $c^P_g$ and $c^M_g$ ......................................................... 38  
4.1 Locations of users of three slices. ......................................................... 48  
4.2 Total Rate versus $V$ and $P_{\text{max}}$. ................................................. 48  
4.3 Total Rate versus $V$ and $l$. ............................................................ 49  
4.4 Total Rate versus $V$ and $R_{gs}^{\text{rev}}$. ............................................. 49  
4.5 Queue Lengths versus Slot iterations. .................................................. 50  
5.1 Total transmit power versus delay $T_{n_g}$ for $G = 2$ slices, $N_1 = N_2 = 2$ users, $K = 64$ sub-carriers with $R_{g1}^{\text{rev}} = R_{g2}^{\text{rev}} = 1.0$ bps/Hz and $\mu = 5$ packets/slot ........................................... 59  
5.2 Total transmit power versus $\mu$ for $G = 2$ slices, $N_1 = N_2 = 2$ users, $K = 64$ sub-carriers with $T_{n_g} = 0.5$ and $R_{g1}^{\text{rev}} = R_{g2}^{\text{rev}} = 1.0$ bps/Hz ............................................ 60  
5.3 Convergence of Lagrange Multipliers ................................................. 62  
6.1 EE factor versus number of users $N$. .................................................. 73  
6.2 EE versus number of subcarriers $K$. .................................................. 73  
6.3 EE factor versus $\varepsilon$ and $C^E_g$ for cell-center users scenario ............ 74
6.4 EE factor versus $\varepsilon$ and $C^E_g$ for cell-edge users scenario . . . . . . . . . . . . . . . . . . 74
List of Tables

2.1 Different Sets to Evaluate Performance of Algorithm 1 ........................................... 19
2.2 Parameters of Different Sets for Admission Control Algorithm. ............................ 22
2.3 $\varepsilon_{gr}, \varepsilon_{gk}$ and $\varepsilon_{gp}$ with $\alpha_{g1} = \alpha_{g2} = 1$ for Scenario 3 ....................... 24
2.4 $\varepsilon_{gr}, \varepsilon_{gk}$ and $\varepsilon_{gp}$ for Different Sets in Scenario 3 ................................. 25
4.1 Dynamic Resource Provisioning Algorithm ............................................................... 45
5.1 Different Sets based on $R_{g}^{sv}$ and $L_{ng}$ ................................................................. 58
## List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWGN</td>
<td>Additive White Gaussian Noise</td>
</tr>
<tr>
<td>bps</td>
<td>bits per second</td>
</tr>
<tr>
<td>BS</td>
<td>Base station</td>
</tr>
<tr>
<td>cdf</td>
<td>Cumulative Distribution Function</td>
</tr>
<tr>
<td>CDI</td>
<td>Channel Distribution Information</td>
</tr>
<tr>
<td>CSI</td>
<td>Channel State Information</td>
</tr>
<tr>
<td>dB</td>
<td>Decibel</td>
</tr>
<tr>
<td>EE</td>
<td>Energy Efficiency</td>
</tr>
<tr>
<td>EPA</td>
<td>Equal Power Allocation</td>
</tr>
<tr>
<td>Hz</td>
<td>Hertz</td>
</tr>
<tr>
<td>i.i.d.</td>
<td>independent and identically distributed</td>
</tr>
<tr>
<td>Imperf</td>
<td>Imperfect</td>
</tr>
<tr>
<td>KKT</td>
<td>Karush–Kuhn–Tucker</td>
</tr>
<tr>
<td>LTE</td>
<td>Long Term Evolution</td>
</tr>
<tr>
<td>MAC</td>
<td>Media Access Control</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple Input and Multiple Output</td>
</tr>
<tr>
<td>MMSE</td>
<td>Minimum Mean Square Error</td>
</tr>
<tr>
<td>MRC</td>
<td>Maximum Ratio Combining</td>
</tr>
<tr>
<td>OFDMA</td>
<td>Orthogonal Frequency Division Multiple Access</td>
</tr>
<tr>
<td>pdf</td>
<td>probability density function</td>
</tr>
<tr>
<td>Perf</td>
<td>Perfect</td>
</tr>
<tr>
<td>PHY</td>
<td>Physical</td>
</tr>
<tr>
<td>Pr(Outage)</td>
<td>Outage Probability</td>
</tr>
<tr>
<td>PRB</td>
<td>Physical Resource Block</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality-of-Service</td>
</tr>
<tr>
<td>QSI</td>
<td>Queue State Information</td>
</tr>
<tr>
<td>RAN</td>
<td>Radio Access Network</td>
</tr>
<tr>
<td>RV</td>
<td>Random Variable</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>UMTS</td>
<td>Universal Mobile Telecommunications Service</td>
</tr>
<tr>
<td>VWN</td>
<td>Virtualized Wireless Network</td>
</tr>
<tr>
<td>w.r.t.</td>
<td>with respect to</td>
</tr>
<tr>
<td>WiMAX</td>
<td>Wireless Interoperability for Microwave Access</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless Local Area Network</td>
</tr>
</tbody>
</table>
List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_g$</td>
<td>Admission control priority factor of slice $g$</td>
</tr>
<tr>
<td>$\alpha_{ng}'$</td>
<td>i.i.d. packet arrival process for user $n_g$ in time slot $t$</td>
</tr>
<tr>
<td>$\mu_{ng}$</td>
<td>Average packet arrival rate at the queue of user $n_g$</td>
</tr>
<tr>
<td>$B$</td>
<td>Total bandwidth available in VWN</td>
</tr>
<tr>
<td>$B_k$</td>
<td>Equal bandwidth of each sub-carrier</td>
</tr>
<tr>
<td>$B_c$</td>
<td>Coherence bandwidth</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Path loss exponent</td>
</tr>
<tr>
<td>$c_g^M, c_g^P$</td>
<td>Pricing factors for the number of allocated antennas and the transmit power of slice $g$, respectively</td>
</tr>
<tr>
<td>$c_g^E$</td>
<td>Energy-cost coefficient of slice $g$</td>
</tr>
<tr>
<td>$D_{ng}$</td>
<td>Distance of user $n_g$ from BS</td>
</tr>
<tr>
<td>$d_{ng}$</td>
<td>Large scale fading coefficient of user $n_g$</td>
</tr>
<tr>
<td>$\delta_\nu$</td>
<td>Step size for Lagrange multiplier $\nu$</td>
</tr>
<tr>
<td>$\delta_{ng,k}(\cdot)$</td>
<td>Function of a contamination error, interference and noise</td>
</tr>
<tr>
<td>$\varepsilon_g$</td>
<td>Admission control variable for slice $g$</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>Admission control vector of all slices</td>
</tr>
<tr>
<td>$\varepsilon_{gk}$</td>
<td>Admission control parameter for minimum reserved sub-carriers of slice $g$</td>
</tr>
<tr>
<td>$\varepsilon_{gp}$</td>
<td>Admission control parameter for minimum reserved power of slice $g$</td>
</tr>
<tr>
<td>$\varepsilon_{gr}$</td>
<td>Admission control parameter for minimum reserved rate of slice $g$</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>A constant with small value</td>
</tr>
<tr>
<td>$\epsilon_{ng}$</td>
<td>Uncertainty error bound in CSI estimation</td>
</tr>
<tr>
<td>$\mathcal{E}_{ng}$</td>
<td>Uncertainty error region</td>
</tr>
<tr>
<td>$\mathbf{f}_{ng,k}^{\text{Imperf}}$</td>
<td>Precoding vector of user $n_g$ on sub-carrier $k$ for imperfect CSI</td>
</tr>
<tr>
<td>$\mathbf{f}_{ng,k}^{\text{Perf}}$</td>
<td>Precoding vector of user $n_g$ on sub-carrier $k$ for perfect CSI</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>$\gamma_{ng,k}$</td>
<td>SNR of user $ng$ on sub-carrier $k$ in any time slot</td>
</tr>
<tr>
<td>$\gamma_{t,ng,k}$</td>
<td>SNR of user $ng$ on sub-carrier $k$ in time slot $t$</td>
</tr>
<tr>
<td>$\gamma_{\text{Perf},ng,k}$</td>
<td>SINR of user $ng$ on sub-carrier $k$ for perfect CSI case without any pilot contamination error</td>
</tr>
<tr>
<td>$\gamma_{\text{Imperf},ng,k}$</td>
<td>SINR of user $ng$ on sub-carrier $k$ for imperfect CSI case due to pilot contamination error</td>
</tr>
<tr>
<td>$g_{ng,k,m_{ng,k}}$</td>
<td>Small scale fading coefficient of user $ng$ on sub-carrier $k$ and antenna $m_{ng,k}$</td>
</tr>
<tr>
<td>$G_1, G_1$</td>
<td>Set and number of rate-based slices</td>
</tr>
<tr>
<td>$G_2, G_2$</td>
<td>Set and number of resource-based slices</td>
</tr>
<tr>
<td>$G, G$</td>
<td>Set and number of all slices</td>
</tr>
<tr>
<td>$h_{ng,k}$</td>
<td>Channel gain of user $ng$ on sub-carrier $k$ in any time slot</td>
</tr>
<tr>
<td>$h_{ng}$</td>
<td>Channel gain vector for user $ng$ of size $1 \times K$</td>
</tr>
<tr>
<td>$h_g$</td>
<td>Channel gain vector for slice $g$ of size $1 \times ng$</td>
</tr>
<tr>
<td>$\bar{h}_{ng,k}$</td>
<td>Channel gain vector for VWN of size $1 \times G$</td>
</tr>
<tr>
<td>$h_{t,ng,k}$</td>
<td>Channel gain of user $ng$ on sub-carrier $k$ in time slot $t$</td>
</tr>
<tr>
<td>$\tilde{h}<em>{ng,k}, \tilde{h}</em>{ng}$</td>
<td>Uncertain channel gain of user $ng$ on sub-carrier $k$ and corresponding uncertain channel vector for user $ng$ of size $1 \times K$, respectively.</td>
</tr>
<tr>
<td>$\bar{h}<em>{ng,k}, \bar{h}</em>{ng}$</td>
<td>Estimated channel gain of user $ng$ on sub-carrier $k$ and corresponding vector for user $ng$ of size $1 \times K$, respectively, at the BS in the context of uncertain CSI modeling</td>
</tr>
<tr>
<td>$\hat{h}<em>{ng,k}, \hat{h}</em>{ng}$</td>
<td>Channel estimation error of user $ng$ on sub-carrier $k$ and corresponding vector for user $ng$ of size $1 \times K$, respectively, in the worst-case uncertain CSI model</td>
</tr>
<tr>
<td>$h_{ng,k,m_{ng,k}}$</td>
<td>Channel gain of user $ng$ on sub-carrier $k$ for antenna $m_{ng,k}$</td>
</tr>
<tr>
<td>$h_{ng,k}$</td>
<td>Channel gain vector for user $ng$ on sub-carrier $k$ of size $1 \times m_{ng,k}$</td>
</tr>
<tr>
<td>$h^\text{Imperf}_{ng,k}$</td>
<td>Estimated channel vector including pilot contamination error of user $ng$ on sub-carrier $k$</td>
</tr>
<tr>
<td>$I_{ng,k}$</td>
<td>Interference from other users to $ng$</td>
</tr>
<tr>
<td>$K, K$</td>
<td>Set and number of OFDMA sub-carriers in VWN</td>
</tr>
<tr>
<td>$K^\text{rsv}_g$</td>
<td>Minimum reserved number of sub-carrier for slice $g$</td>
</tr>
<tr>
<td>$l$</td>
<td>The size of a transmission packet in bits per Hz per slot</td>
</tr>
<tr>
<td>$L_{ng}$</td>
<td>Average packet size at the queue of user $ng$</td>
</tr>
<tr>
<td>$\lambda, \nu_k, \nu_g, \psi_g, \phi_g, \zeta_{ng}, \rho_k$</td>
<td>Lagrangian multipliers</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>$\rho, \phi, \zeta$</td>
<td>Lagrangian multiplier vectors</td>
</tr>
<tr>
<td>$m_{n_g,k}$</td>
<td>Allocated antennas to user $n_g$ on sub-carrier $k$</td>
</tr>
<tr>
<td>$M_{n_g}$</td>
<td>Antenna allocation vector for user $n_g$ of size $1 \times k$</td>
</tr>
<tr>
<td>$M_g$</td>
<td>Antenna allocation vector for slice $g$ of size $1 \times n_g$</td>
</tr>
<tr>
<td>$M$</td>
<td>Total number of antennas available at the BS</td>
</tr>
<tr>
<td>$M_{g \text{min}}, M_{g \text{max}}$</td>
<td>Minimum and maximum numbers of reserved antennas of slice $g$, respectively</td>
</tr>
<tr>
<td>$N, N$</td>
<td>Set and number of all users in VWN</td>
</tr>
<tr>
<td>$N_g, N_g$</td>
<td>Set and number of users in slice $g$</td>
</tr>
<tr>
<td>$p_{g \text{rsv}}$</td>
<td>Minimum reserved power for slice $g$</td>
</tr>
<tr>
<td>$p_{n_g,k}$</td>
<td>Allocated power to user $n_g$ on sub-carrier $k$ in any time slot</td>
</tr>
<tr>
<td>$P_{n_g}$</td>
<td>Power allocation vector for user $n_g$ of size $1 \times K$</td>
</tr>
<tr>
<td>$P_g$</td>
<td>Power allocation vector for slice $g$ of size $1 \times n_g$</td>
</tr>
<tr>
<td>$P$</td>
<td>Power allocation vector for VWN of size $1 \times G$</td>
</tr>
<tr>
<td>$p_{n_g,k}^t$</td>
<td>Allocated power to user $n_g$ on sub-carrier $k$ in time slot $t$</td>
</tr>
<tr>
<td>$P_{\text{max}}$</td>
<td>Maximum down-link transmit power limit of BS</td>
</tr>
<tr>
<td>$P_{n_g \text{max}}$</td>
<td>Maximum up-link transmit power limit of user $n_g$</td>
</tr>
<tr>
<td>$Q_{t,n_g}$</td>
<td>Queue length of user $n_g$ in time slot $t$</td>
</tr>
<tr>
<td>$R_{g \text{rsv}}$</td>
<td>Minimum reserved rate for slice $g$</td>
</tr>
<tr>
<td>$\bar{R}_{g \text{rsv}}$</td>
<td>Average minimum reserved rate for slice $g$ over a transmission frame</td>
</tr>
<tr>
<td>$R_{n_g}$</td>
<td>Rate of user $n_g$ on all sub-carriers in any time slot</td>
</tr>
<tr>
<td>$R_{t,n_g}$</td>
<td>Rate of user $n_g$ on all sub-carriers in time slot $t$</td>
</tr>
<tr>
<td>$\bar{R}_{n_g}$</td>
<td>Average rate of user $n_g$ over a transmission frame</td>
</tr>
<tr>
<td>$\hat{R}_{n_g}$</td>
<td>Rate of user $n_g$ under uncertain CSI</td>
</tr>
<tr>
<td>$R_{n_g,k}$</td>
<td>Rate of user $n_g$ on sub-carrier $k$</td>
</tr>
<tr>
<td>$R_{n_g,k}^\text{Perf}$</td>
<td>Rate of user $n_g$ on sub-carrier $k$ for perfect CSI case</td>
</tr>
<tr>
<td>$R_{n_g,k}^\text{Imperf}$</td>
<td>Rate of user $n_g$ on sub-carrier $k$ for imperfect CSI case</td>
</tr>
<tr>
<td>$R_{\text{opt}}$</td>
<td>Optimum rate</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Noise power of all users over all sub-carriers</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Duration of up-link pilot</td>
</tr>
<tr>
<td>$T_{\text{max}}$</td>
<td>Maximum value of time slot index</td>
</tr>
<tr>
<td>$T_{n_g}$</td>
<td>Average traffic delay of user $n_g$</td>
</tr>
<tr>
<td>$T_i$</td>
<td>Duration of each time slot</td>
</tr>
</tbody>
</table>
### List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_c$</td>
<td>Coherence interval time</td>
</tr>
<tr>
<td>$U_g$</td>
<td>Utility of slice $g$</td>
</tr>
<tr>
<td>$t$</td>
<td>time-slot index</td>
</tr>
<tr>
<td>$\mathcal{T}$</td>
<td>Set of all time slots</td>
</tr>
<tr>
<td>$V$</td>
<td>Trade-off parameter in Lyapunov drift-plus-penalty algorithm</td>
</tr>
<tr>
<td>$w_{ng,k}$</td>
<td>Sub-carrier assignment indicator variable, specifies if sub-carrier $k$ is assigned to user $n_g$ or not in any time slot</td>
</tr>
<tr>
<td>$w_{ng}$</td>
<td>Vector of sub-carrier assignment indicator variables for user $n_g$ of size $1 \times K$</td>
</tr>
<tr>
<td>$w_g$</td>
<td>Vector of sub-carrier assignment indicator variables for slice $g$ of size $1 \times n_g$</td>
</tr>
<tr>
<td>$w$</td>
<td>Vector of sub-carrier assignment indicator variables for VWN of size $1 \times G$</td>
</tr>
<tr>
<td>$w^t_{ng,k}$</td>
<td>Sub-carrier $k$ assignment indicator to user $n_g$ in time slot $t$</td>
</tr>
<tr>
<td>$x_{ng,k}$</td>
<td>Transmit symbol of user $n_g$ on sub-carrier $k$</td>
</tr>
<tr>
<td>$\mathcal{X}$</td>
<td>Exponential random variable</td>
</tr>
<tr>
<td>$y_{ng,k}$</td>
<td>Up-link received sample at BS with perfect CSI</td>
</tr>
<tr>
<td>$y_{ng,k}^{\text{Imperf}}$</td>
<td>Up-link received sample at BS with imperfect CSI</td>
</tr>
<tr>
<td>$z_{ng,k}$</td>
<td>Vector of AWGN</td>
</tr>
<tr>
<td>$Z^t_g$</td>
<td>Virtual queue parameter for slice $g$ in time slot $t$</td>
</tr>
</tbody>
</table>
Mathematical Operators

- \{a, b\} \quad \text{Set of continuous values from } a \text{ to } b
- [a, b] \quad \text{Set of discrete values from } a \text{ to } b
- \mathcal{A} \cup \mathcal{B} \quad \text{Union of sets } \mathcal{A} \text{ and } \mathcal{B}
- \{a \cdots b\} \quad \text{Ellipsis to denote intermediate values from } a \text{ to } b
- a \in \mathcal{A} \quad a \text{ belongs to } \mathcal{A}
- \sum_{a \in \mathcal{A}} \quad \text{sum over all entities in set } \mathcal{A}
- f(x) \quad \text{Function of } x
- \mathcal{L}(\cdot) \quad \text{Lagrange function}
- \mathcal{D}(\cdot) \quad \text{Dual function}
- \frac{\partial f}{\partial x} \quad \text{Partial derivative of } f \text{ with respect to } x
- a < b \quad a \text{ is less than } b
- a \leq b \quad a \text{ is less than or equal to } b
- a > b \quad a \text{ is greater than } b
- a \geq b \quad a \text{ is greater than or equal to } b
- \ln(\cdot) \quad \text{Natural logarithm}
- \log_a(\cdot) \quad \text{Logarithm to the base } a
- \max[a, b] \quad \text{Maximum of the values of } a \text{ and } b
- \min[a, b] \quad \text{Minimum of the values of } a \text{ and } b
- [x]^b_a \quad \min\{b, \max\{x, a\}\}
- [x]^+ \quad \max\{x, 0\}
- a \ll b \quad a \text{ is very smaller than } b
- a \gg b \quad b \text{ is very smaller than } a
- \mathbb{C}^{a \times b} \quad \text{A space of all } a \times b \text{ matrices with complex entries}
- \mathbb{E}_x\{\cdot\} \quad \text{Expectation operator over variable } x
Mathematical Operators

\[ \sqrt{a} \] Square root of \( a \)

\[ \lfloor a \rfloor \] The largest integer less than or equal to the value of \( a \)

\[ \| x \| \] Euclidean norm of matrix or vector \( x \)

\[ a \xrightarrow{a.s.} \] Almost sure convergence

\[ \Pr \{ a \} \] Probability of event \( a \)
Chapter 1

Introduction

1.1 Current Trends and Future Perspectives of Wireless Networks

The recent proliferation of Internet services and applications in wireless devices has triggered a rapid evolution in existing wireless networks to satisfy the QoS requirements of multimedia applications. As an outcome of this evolution, diverse heterogeneous networks such as universal mobile telecommunications service (UMTS), long term evolution (LTE), and wireless interoperability for microwave access (WiMAX) have been developed with specific operating strategies to focus on different services and QoS requirements. For instance, wireless local area networks (WLANs) provide high speed video services while third-generation (3G) networks offer high quality voice services and LTE tries to increase the flexibility of wireless networks by introducing more simplified architecture. Since these network models are tightly coupled with hardware devices to follow specific operating strategies that are suitable to different service requirements, it is hard to interconnect or converge them in order to combine advantages in an efficient way [1].

As a result of tight coupling in the current model of wireless networks, service providers and the infrastructure providers overlap each other. This overlap acts as barrier for third party service providers to fully exploit the underutilized wireless resources available from infrastructure providers to satisfy their QoS requirements. Consequently, due to this limitation of the wireless model, most of the valuable wireless resources such as spectrum are wasted or remain underutilized.

To withstand the rapid growth of wireless subscribers, new emerging wireless standards and services are being proposed to achieve and improve the QoS of wireless users. However, as the existing wireless networks are inflexible due to their deep rooting in hardware, commonly re-
ferred to as ossification of network, it is highly expensive to deploy innovative wireless standards and protocols for testing purposes.

Finally, the range of QoS requirements of end-users has broadened due to shift from voice oriented wireless devices to application or multimedia oriented devices. For instance, the network requirements of video applications differ from voice or data applications. The network cost is growing to meet these requirements due to infrastructure and operational costs. However, as the wireless market model generates revenue based on services rather than infrastructure, the cost of revenue generation from network infrastructure is stagnant [2].

1.2 Wireless Virtualization: Background and Motivations

To resolve the aforementioned issues in existing wireless networks due to their obstinate design, some fundamental modifications in the current model are necessary. Inspired by the success of wired network virtualization, the idea of wireless virtualization is proposed to break the coupling between infrastructure and services to increase the efficiency of resource utilization in existing wireless networks [2,3].

Virtualization in generic terms is defined as the abstraction and sharing of resources among different service providers in a manner that is transparent to the users. Network virtualization revolutionized wired networks in recent decades by enabling the convergence of diverse networks in a single physical network. The concept of network virtualization extends the idea of machine virtualization, first proposed by IBM in the 1960’s in the form of a time sharing computing system [4, 5]. The proposed solution enabled the sharing of physical hardware among multiple virtual machines via an extra software layer known as virtual machine monitor. Motivated by the success of virtualization technology, many solutions such as Oracle VirtualBox [6] and Xen hypervisor [7] implement the basic idea of virtualization in modern computing platforms to offer hardware efficient services.

Witnessing the success of machine virtualization, researchers adopted the idea of network virtualization to break the ossification of wired networks. As a result, depending on the Internet layers, various virtualization projects are proposed. For instance, VNET [8] virtualizes network data-link layer by implementing a VLAN [9] spread over wide area using layer-2 tunneling. The VNET servers, running on the physical machines, adopt the concept of a packet filter to route the packets to an appropriate destination. In order to improve the end-to-end QoS experience in the network layer, AGAVE [10] proposes the concept of network planes that enable parallel Internets
Several network virtualization projects provide research testbeds for new technologies. For instance, VINI [11], GENI [12] and 4WARD [13] allow researchers to create an isolated, realistic network environment to evaluate new standards and protocols. Similar to wired network virtualization, the idea of VWNs can be beneficial for wireless networks in tackling the existing challenges of wireless network.

Like network virtualization, wireless virtualization enables flexibility and modularity in the existing wireless networks via sharing a physical substrate of wireless resources such as power, BS, spectrum, sub-carriers and antennas among various slices. As a consequence of sharing a physical substrate, wireless virtualization provides opportunities for efficient resource utilization and solves the problem of spectrum crisis. A well-implemented VWN provides isolation among slices such that any change in one slice does not affect the operations of other slices, enables customization in QoS requirements of slices, and efficiently utilizes wireless resources and thus decouple the functionality of service providers from the infrastructure providers.

Furthermore, VWN can provide separate slices to act as testbeds for innovations in wireless networks, thereby allowing testing of new standards without the need of a highly expensive infrastructure setup. The idea of sharing common wireless resources by means of VWNs, allows network operators to expand or shrink the range of their networks without any significant investments, thereby, opening new business opportunities for network operators. From a security perspective, VWNs can offer extra security to special network domains such as those for military and banks, by enabling these networks in slices separate from other domains. Moreover, wireless virtualization offers a platform to support coexistence of different wireless standards on a single physical substrate to support the convergence of heterogeneous networks and reduce the operational as well as setup costs of wireless networks. Wireless virtualization is also a suitable candidate for future wireless networks to enable custom services and abstraction in the existing model of wireless networks [2, 14]. However, to fully harness the benefits of wireless virtualization, it is very important to design efficient resource provisioning policies which enables abstraction and isolation in VWNs.

1.3 Technical Challenges in Virtualized Wireless Networks

Wireless virtualization provides a paradigm shift from an infrastructure oriented wireless network model to a service oriented model. Since wireless resources such as spectrum, power and band-
width are limited, resource provisioning policy is essential to satisfy the diverse QoS requirements of slices in VWNs. Moreover, due to the inherent stochastic characteristics of wireless channels, highly dynamic slice environment, user mobility and varying number of active users, a resource provisioning policy is essential to maintain the quality of VWN operation and prevent any performance degradation. Typically, a resource provisioning policy aims to provide the mathematical formulation to translate the QoS requirements of slices into the fractional amount of physical resources required to meet those requirements. A well-designed resource provisioning policy is essential to providing isolation between multiple slices and efficient resource utilization in VWNs. However, the feature of coexistence of slices with different QoS requirements raises a new set of technical challenges for resource provisioning problems of VWNs.

First, it is challenging to deliver an optimized QoS for each slice in a VWN, when unique, isolated slices share the resources of a common physical network and works in isolation to each other. It is well known that each slice in a VWN can have different QoS requirements based on the end-users’ applications. For instance, some slices need a minimum reserved rate, referred to as rate-based QoS, while others may simply require resource-based QoS, where a minimum fraction of physical resources such as power, sub-carriers and antennas are reserved. Due to random behavior of slices and limited wireless resources, it is hard to satisfy the diverse QoS requirements of slices, which results in service outage caused by infeasibility, degrading the performance of the VWN. Moreover, in practice, wireless networks have limited queue backlogs and traffic arrival is random in nature. Therefore, to offer continuous and reliable services, a cross-layer optimization setup including both physical (PHY) and media access control (MAC) layer parameters is desired for VWNs’ resource provisioning policies. However, since cross-layer problems contains optimization variables from multiple dimensions and are non-convex in nature, the optimize solution is expensive to compute.

Second, it is hard to maintain performance of VWNs while operating in an energy-efficient mode. The growing interest in high data rate multimedia applications has increased the energy consumption of wireless networks. In fact, BSs across the globe produce over a hundred million tons of carbon dioxide every year [15]. Such a high energy consumption in VWNs is not only a threat to the environment but can also add a significant amount to VWNs’ operational cost for slice owners, thereby nullifying one of the benefits of low operational cost of VWNs. Moreover, due to rising concern of green communications, future networks are expected to follow energy-efficient design. Since VWNs are a strong candidate to become part of future networks, it is essential that resource provisioning for VWNs adopts energy-efficient algorithms. However, it is
challenging to support diverse QoS requirements of slices while maximizing the energy efficiency of VWNs.

Third, in order to optimize the resource utilization in VWNs, resource provisioning policies require some information exchange of CSI at the BSs. However, since information exchange cannot be perfect due to the limited available capacity, it is challenging to design robust resource provisioning policies in VWNs. Often in practice, due to the random nature of wireless channel or users’ mobility, the CSI estimation at the BS is erroneous or uncertain. This uncertainty degrades the performance of VWNs and hence results in inefficient resource utilization.

In this thesis, we will show how the above challenges for VWNs can be addressed with an appropriate resource provisioning policies for VWNs.

1.4 Literature Review

There are plenty of resource provisioning schemes proposed in literature for traditional wireless networks in comparison to the virtualized wireless network; therefore, in the following, we will first review the related works of this thesis in the traditional wireless networks. Due to several advantages of orthogonal frequency division multiple access (OFDMA) in wireless networks, a considerable number of existing works investigate the resource allocations in OFDMA-based wireless networks. For instance, authors in [16] propose a power and sub-carrier allocation algorithm for multi-cell underlay cognitive radio networks while considering the signal-to-noise ratio (SINR) constraints at primary users. A sub-carrier, power and rate allocation algorithm for uplink OFDMA transmission is proposed in [17] which considers per-user fairness. Another work [18] proposes a near optimal joint power and sub-carrier allocations algorithm with low complexity for uplink OFDMA systems.

Another group of works utilize the spatial gains of massive MIMO in traditional wireless networks to enhance the QoS experiences of end-users. In order to maximize the spectral efficiency in multi-cell wireless network, [19] uses a massive MIMO setup and proposes a joint optimal resource allocation scheme for training duration, training signal power and data signal power for up-link transmission. Since antenna sub-systems as well as signal processing algorithms at massive MIMO BSs are heavy sources of energy consumption, many works propose energy-efficient solutions for massive MIMO setups. For instance, the authors of [20] propose an optimal solution for power allocation, number of activated antennas and the pilot assignment for multicellular massive MIMO setup. The resource provisioning solution maximizes the overall system energy
efficiency measured in bits/Joule under the effect of pilot contamination during the channel estimation. Another research work [21] applies fractional programming to reach an optimal solution for power, sub-carrier and antenna allocations, maximizing the energy-efficiency of systems with massive MIMO setups for perfect as well as imperfect CSI scenarios.

Some existing works in traditional wireless networks focus on providing reliable and continuous services in practice, where networks have limited queue backlog and traffic arrival is bursty. The authors in [22] propose a power control policy to maximize the average rate of femto-cells under the constraint of stable queues of macro-cell users by applying a Lyapunov drift function. Also based on Lyapunov drift function theory, [23] provides an optimized solution for power allocation and routing for wireless channels with time varying traffic that stabilizes the system with guaranteed delay. Some resource provisioning works in traditional wireless networks focus on the issue of uncertainty in CSI at the BS due to users’ mobility, or any other estimation error. For instance, [24] proposes a robust power allocation scheme under the constraint of predefined outage probability. The authors of [25] study a robust resource allocation for a MIMO based cognitive radio system to protect the spectral resources of primary users from CSI uncertainties due to secondary users under the global interference temperature constraints. Furthermore, the authors of [26, 27] develop resource provisioning schemes based on the worst-case robust optimization theory to deal with uncertain CSI at the BS, where the error in CSI is assumed to be bounded within a specific region.

In spite of the fact that few existing works in traditional wireless networks address the challenges of energy efficiency, bursty traffic arrival and uncertainty in CSI estimation. Nevertheless, these works are not directly applicable to a VWN setup due to diverse QoS requirements of slices and the presence of isolated instances of virtual networks using the shared resources of the physical network. To the best of our knowledge, none of the existing works address these issues in a VWN setup. However, existing works related to VWNs target specific QoS requirements of slices and propose resource provisioning schemes.

A group of existing works in wireless virtualization enable radio access network (RAN) sharing in wireless networks via virtualizing the BS. NVS [28] virtualizes WiMAX BS and proposes slice scheduling to satisfy both bandwidth-based and resource-based reservations. To enable customized solutions within each slice, NVS includes a generic framework for flow scheduler which is customized according to slices’ requirements via simple programming interfaces. However, NVS focuses on satisfying the accumulative requirements of slices while instantaneous (each time slot-based) requirements of slices are ignored, which are also essential in VWNs. More-
over, NVS implementation requires a modification in the existing code of the MAC scheduler. To deal with this issue, CellSlice [29] extends the NVS solution via remotely controlling the traffic and thus the scheduling decisions of BS from gateways without any modification in the existing scheduler in the BS. VNTS [30]—another WiMAX based virtualization solution—implements a virtualized substrate for down-link transmission and allocates fractions of BS resources to the slices.

Due to growing number of studies about tentative roles and potential benefits of wireless virtualization in future networks [3,31,32], a group of works study virtualization of LTE, the latest evolution of 3GPP, to analyze its performance and feasibility under various resource provisioning schemes. For instance, authors in [33] propose an additional layer of hypervisor, similar to node virtualization, on top of physical layer of LTE network. The hypervisor schedules the LTE air interface resources, i.e., physical resource blocks (PRBs), among virtual eNodeBs based on the received information such as channel conditions, QoS requirements and load from virtual eNodeBs. In [34], a resource provisioning problem is solved using game theory wherein virtual eNodeBs make bids for power resources to the hypervisor acting as an auctioneer, who allocates power based on the rate requirements of users.

Some of the works in VWN apply game theory in resource provisioning schemes to maximize the benefits to slices and network owners. A multi-objective resource allocation scheme for OFDMA-based VWNs is presented in [35], where the profits of mobile network operator and multiple slices are maximized. The formulated resource allocation problem is solved via equilibrium theory in which a solution is obtained at an equilibrium point. Authors of [36] apply the framework of a bankruptcy game in VWNs, where the PRBs among virtual mobile networks are allocated using the Shapley value method [37]. Another work [38] formulates the resource allocation problem based on game theory in cloud-based VWNs, where mobile users, acting as players, bid for transmit powers and number of sub-channels to maximize their spectrum efficiency. The optimized solution is guaranteed via convergence of game setup to the Nash equilibrium. Furthermore, the authors of [39] model the interaction between slices, network operator, and users as an auction game where the network operator is responsible for spectrum management on a higher level, and each slice focuses on QoS management for its own users.
1 Introduction

1.5 Thesis Objectives, Contributions and Organization

Clearly, wireless virtualization is a revolutionary idea to circumvent the existing challenges of wireless networks via enabling flexibility and customization in its operation to support new requirements of wireless users. To fully realize the potential benefits of wireless virtualization, it is essential to overcome the aforementioned challenges of VWNs. Taking these challenges into account, and the limits of existing works in VWNs, the objective of this thesis is to study and develop efficient resource provisioning schemes to support different QoS requirements of slices in VWNs. Specifically, this thesis focuses on minimum reserved rates and resources (power, sub-carriers and antennas) in addition to the queue stability and delay-based QoS parameters of VWN. Furthermore, we extend the study of QoS-aware resource provisioning in VWNs to more realistic scenarios, where the resource provisioning policies are energy-efficient and robust against the uncertainty in CSI estimation at the BS.

The rest of the thesis is organized as follows. Chapter 2-6 contains the key contributions of this thesis, followed by Chapter 7, which concludes the thesis with potential future works.

As noted before, VWNs are desired to perform well while supporting the diverse QoS requirements of slices. Chapter 2 focuses on this issue via proposing a joint resource provisioning and admission control policy suitable to VWNs. In this chapter, we present the resource provisioning formulation for OFDMA-based VWNs to satisfy the instantaneous requirements of two groups of slices with resource-based (power and sub-carriers) and rate-based requirements. The formulated problem is simplified via variable transformations and relaxation techniques to develop an optimized iterative algorithm for power and sub-carrier allocation. Afterwards, to enhance VWNs’ performance, an admission control algorithm is presented that minimizes the VWNs’ infeasibility and reduces the service outage experience of slices in VWNs. The proposed admission control algorithm dynamically adjusts the requirements of slices according to CSI and priorities of slices.

Even though admission control minimizes the outage experience in VWNs, it has a risk of degrading the QoS experience of slices and may not be suitable under all situations. To deal with this issue, in Chapter 3, we utilize the multiplexing gains of massive MIMO technology in the design of a resource provisioning policy for VWNs. In particular, to maximize a sum utility of massive MIMO-aided VWN, a joint power, sub-carrier, and antenna allocation problems is presented for both perfect and imperfect channel knowledge cases. Simulation results reveal the benefits of applying a large number of antennas in this setup.

Moving further in delivering the optimized QoS for each slice, Chapter 4 extends the resource
provisioning policies to cross-layer resource provisioning schemes that are adaptive to both PHY layer and MAC layer parameters, i.e., CSI and QSI, respectively. As noted before, due to the limited queue size and random traffic arrival in users’ queues, cross-layer resource provisioning policies are essential in realistic VWNs to offer continuous and reliable services to VWNs’ users. Focusing on this aspect, Chapter 4 presents a dynamic resource provisioning policy for stable queue control in VWNs. Specifically, the total average rate of a VWN over a transmission frame is maximized while guaranteeing a minimum average required rate of each slice and a stable-queue constraint of the VWN. Finally, an iterative algorithm is developed for joint power and sub-carrier allocations via applying the elegant theory of Lyapunov quadratic approximation and drift-plus-penalty algorithm.

Still considering the cross-layer resource provisioning policies for VWNs to improve the QoS experience of slices in practice, Chapter 5 focuses on delay sensitive design of VWNs for future networks with energy-efficient and delay-aware resource provisioning design. Specifically, in Chapter 5, we propose a delay-aware resource provisioning policy for energy-efficient VWNs. The set objective is to minimize the total average transmit power while satisfying the minimum required average rate of each slice and maximum average packet transmission delay for each user. To obtain a simple and efficient algorithm for resource provisioning, the cross-layer dependency of constraints is reduced to only physical layer dependency via an equivalent constraint approach. Simulation results show the performance of the VWN under different system parameters, e.g., average packet transmission delay and minimum average slice rate on total consumed power in the VWN.

Continuing with the energy-efficient design of VWNs’ resource provisioning policies, the issue of performance degradation due to uncertainty in CSI estimation at the BSs of VWNs is addressed in Chapter 6. Specifically, an energy-efficient and robust resource provisioning policy is proposed to deal with uncertainty in CSI estimation in VWNs. In this context, uncertain CSI is modeled as the sum of its true estimated value and an error assumed to be bounded in a specific uncertainty region. With this modeling of uncertain CSI, the overall energy efficiency (EE) of the VWN is maximized in terms of a slice utility function. Eventually, an iterative algorithm is developed to study the effects of system parameters through simulations.

Finally, Chapter 7 provides concluding remarks and potential future works.
Chapter 2

Joint Resource Provisioning and Admission Control in VWN

This chapter studies joint resource provisioning and admission control in VWNs, where one base station of an OFDMA-based wireless network is virtualized into two types of slices with resource-based and rate-based reservations. Aiming to maximize the total rate of a VWN, the resource provisioning optimization problems are formulated by guaranteeing a minimum requirement for each slice. Via constraint relaxation and variable transformations, an iterative algorithm is developed for power and sub-carrier allocation. Due to the channel variations, VWNs suffer from non-zero outage probability, i.e., slice requirements cannot always be met. To prevent this issue, we present an admission control algorithm in which slice requirements are dynamically adjusted based on CSI. The simulation results demonstrate the effectiveness of our proposed algorithms.

2.1 Introduction

To increase the spectrum efficiency, wireless network virtualization has recently emerged as an enabler for resource management and service customization among slices belonging to different service providers on a shared physical network [2]. In a VWN, different slices can share physical network resources (e.g., BS) and wireless resources (e.g., sub-carriers and power) with guaranteed isolated flows and customized services. Generally, each slice comprises a set of users and has its own QoS requirement. Due to the diverse QoS requirements of slices, wireless resource

\[ \text{Parts of this chapter have been presented in [40].} \]
limitations, and wireless channel variations, resource provisioning among slices is challenging, yet essential in order to satisfy the QoS requirements of slices and improve the overall VWN performance [3, 28, 39, 41, 42].

Recently, the design of resource provisioning for VWNs has received growing attention. For instance, in [33], a resource allocation algorithm is proposed among virtual operators in LTE-based VWNs. Another resource management scheme is studied in [28], by introducing two types of slices, including rate-based slices and resource-based slices, where the minimum rate and minimum network resources are preserved for slices, respectively. Furthermore, in [39], interaction between slices, network operator, and users is modeled as an auction game where the network operator is responsible for spectrum management on a higher level, while each slice focuses on QoS management for its own users.

In a VWN, due to the stochastic nature of wireless channels and limited available resources, there always exists a non-zero probability that QoS requirements of each slice cannot be satisfied. In this chapter, we address this issue by introducing a joint resource provisioning and admission control policy for an OFDMA-based VWN. Taking into account the minimum required rate and resources of rate-based and resource-based slices, respectively, a joint power and sub-carrier optimization problem is formulated with an objective to maximize the total network rate.

The formulated resource provisioning problem involves binary variables and suffers from high computational complexity. To develop an efficient joint power and sub-carrier allocation algorithm, relaxation and variable transformation are applied to convert the problem into a convex optimization problem. Then, we apply a two-level iterative algorithm to solve the Lagrangian dual function of the convexified problem. In the first loop, the Lagrange multipliers are updated based on gradient descent approach, and in the second loop, the optimal power and sub-carrier are derived via KKT conditions.

In addition, we investigate an admission control problem. In the proposed admission control policy, the minimum QoS requirements are adjusted based on CSI of users in each slice, and priority of each slice. Via simulation results, we demonstrate the effectiveness of our proposed algorithms compared to other sub-optimal solutions, e.g., the equal transmit power for all sub-carriers. We also show how admission control policy can adjust the minimum QoS of different slices based on CSI and slice priorities.

The rest of this chapter is organized as follows. In Section 2.2, the system model and problem formulations are introduced, followed by Section 2.3, where iterative algorithms are presented. Section 2.4 introduces the admission control policy and its solution. In Section 2.5, the simulation
results are presented and Section 2.6 provides some concluding remarks.

2.2 Network Model and Problem Formulation

We consider the down-link of a VWN where the coverage of specific area is provided by one central BS, serving two specific sets of slices: 1) Rate-based group, $\mathcal{G}_1 = \{1, \ldots, G_1\}$, where the minimum reserved rate $R_{rsv}^g$ is provided for slice $g \in \mathcal{G}_1$; 2) Resource-based group, $\mathcal{G}_2 = \{1, \ldots, G_2\}$, where the minimum required resources, including $K_{rsv}^g$ sub-carriers and $P_{rsv}^g$ power, are provided for each slice $g \in \mathcal{G}_2$. Each slice $g \in \mathcal{G}$ where $\mathcal{G} = \mathcal{G}_1 \cup \mathcal{G}_2$ has a set of users denoted by $\mathcal{N}_g = \{1, \ldots, N_g\}$ and total number of users on the network is equal to $N = \sum_{g \in \mathcal{G}} N_g$. A typical example of such a network is illustrated in Fig. 2.1.

The total bandwidth of $B$ Hz is shared between users of all slices through OFDMA scheme, which is divided into a set of $\mathcal{K} = \{1, \ldots, K\}$ sub-carriers. The bandwidth of each sub-carrier, $B_k = B/K$, is assumed to be small compared to the coherent bandwidth $B_c$ of the wireless channel. Therefore, the channel gain $h_{n_g,k}$ of user $n_g$ on sub-carrier $k$ can be considered flat. Without loss of generality, we assume that the noise power of all users over all sub-carriers is
equal to $\sigma$.

In order to avoid inter-user interference within a cell, we consider orthogonal sub-carrier assignment, i.e., each sub-carrier is assigned exclusively to one user. Mathematically, $C_1 : \sum_{g \in G} \sum_{n_g \in N_g} w_{n_g,k} \leq 1, \forall k \in K$ where $w_{n_g,k} \in \{0, 1\}$ indicates that sub-carrier allocation for user $n_g$ on sub-carrier $k$. More specifically, $w_{n_g,k} = 1$ indicates that sub-carrier $k$ is assigned to user $n_g$, and otherwise $w_{n_g,k} = 0$.

Let $p_{n_g,k}$ be the allocated transmit power on sub-carrier $k$ for user $n_g$. Due to the transmit power limitation of BS, the allocated powers of all users over all sub-carriers are subject to $C_2 : \sum_{g \in G} \sum_{n_g \in N_g} \sum_{k \in K} w_{n_g,k} p_{n_g,k} \leq P_{\text{max}}$, where $P_{\text{max}}$ is the maximum transmit power of BS.

Consequently, the rate of user $n_g \in N_g$ is equal to $R_{n_g}(P, w) = \sum_{k \in K} w_{n_g,k} \log_2 \left(1 + \frac{p_{n_g,k} h_{n_g,k}}{\sigma^2}\right)$ for all $n_g \in N_g$, where $P = [P_1, \ldots, P_G]$ is the allocated power vector of all users, $P_g = [P_{n_1}, \ldots, P_{n_G}]$, and $P_{n_g} = [p_{n_g,1}, \ldots, p_{n_g,K}]$. Furthermore, $w = [w_1, \ldots, w_G]$ is the sub-carrier assignment vector for all users where $w_g = [w_{n_1}, \ldots, w_{n_G}]$, and $w_{n_g} = [w_{n_1}, \ldots, w_{n_g,K}]$.

The minimum required resources of each slice $g \in G_2$ can be represented as

$$C_3 : \sum_{n_g \in N_g} \sum_{k \in K} w_{n_g,k} \geq \frac{K_{rsv}^g}{\gamma_2} \quad \text{and} \quad \sum_{n_g \in N_g} \sum_{k \in K} w_{n_g,k} p_{n_g,k} \geq P_{\text{rsv}}^g$$

Also, the minimum required rate of each slice $g \in G_1$ can be represented as

$$C_4 : \sum_{n_g \in N_g} \sum_{k \in K} w_{n_g,k} \log_2 \left(1 + \frac{p_{n_g,k} h_{n_g,k}}{\sigma^2}\right) \geq R_{rsv}^g.$$

Aiming to provide the minimum slice requirements, i.e., minimum rate and sub-carriers, and to maximize overall network rate, the optimization problem of VWN can be presented as

$$\max_{w, P} \sum_{g \in G} \sum_{n_g \in N_g} R_{n_g}(P, w),$$

subject to $C_1 - C_4$.

The optimization problem (2.1) suffers from two major issues: 1) High computational complexity due to discrete variables, i.e., sub-carriers assignment; 2) Non-zero probability to have the empty feasibility set due to $C_4$ which depends on the CSI of different users. In the following sections, we first propose an algorithm to solve (2.1). Also, for dealing with the latter issue, we propose
Joint Resource Provisioning and Admission Control in VWN

2.3 Joint Power and Sub-Carrier Allocation

Since the efficient approach to find an optimal solution is very appealing from implementation perspective, in this section, we propose an approach to transform (2.1) into a convex optimization problem. First, we relax $w_{ng,k}$ as a continuous variable in interval $[0, 1]$. In the new definition, $w_{ng,k}$ indicates the portion of time that sub-carrier $k$ is assigned to user $ng$ for a specific transmission frame. Consequently, C1 is changed to

$$\tilde{C}_1 : w_{ng,k} \in [0, 1], \text{ and } \sum_{g \in \mathcal{G}} \sum_{ng \in \mathcal{N}_g} w_{ng,k} \leq 1, \forall k \in \mathcal{K}.$$ 

Considering $x_{ng,k} = w_{ng,k}p_{ng,k}$ and the alternative rate function

$$\tilde{R}_{ng}(x, w) = \sum_{k \in \mathcal{K}} w_{ng,k} \log_2 (1 + \frac{x_{ng,k}h_{ng,k}}{\sigma w_{ng,k}})$$

for all $ng \in \mathcal{N}_g$ and $g \in \mathcal{G}$, then

$$\tilde{C}_4 : \sum_{ng \in \mathcal{N}_g} \tilde{R}_{ng} \geq R_{rs}^{\text{rv}}, \forall g \in \mathcal{G}_1$$

Therefore, (2.1) is transformed into

$$\max_{x, w} \sum_{g \in \mathcal{G}} \sum_{ng \in \mathcal{N}_g} \tilde{R}_{ng}, \quad \text{(2.3)}$$

subject to: $\tilde{C}_1$, $\tilde{C}_2$, $\tilde{C}_3$ and $\tilde{C}_4$.

Since $\tilde{R}_{ng}(x, w)$ belongs to the class of convex functions represented as $f(x, y) = x \log_2(1+y/x)$ for $x, y \geq 0$, (2.3) is a convex optimization problem [43]. Since the duality gap is zero for convex optimization problems [44], we solve the dual problem of (2.3) to develop subcarrier and power allocation algorithm. Let $\lambda, \nu_k, \nu_g, \psi_g$, and $\phi_g$ represent Lagrange multipliers for constraints $\tilde{C}_1$, $\tilde{C}_2$, $\tilde{C}_3$ and $\tilde{C}_4$, which should satisfy $\lambda(\sum_{g \in \mathcal{G}} \sum_{ng \in \mathcal{N}_g} \sum_{k \in \mathcal{K}} x_{ng,k} - P_{\text{max}}) = 0$, $\nu_g(\mathcal{K}_g^{\text{rs}} - \sum_{ng \in \mathcal{N}_g} \sum_{k \in \mathcal{K}} w_{ng,k}) = 0$ for all $g \in \mathcal{G}_2$, $\psi_g(P_{\text{rs}} - \sum_{ng \in \mathcal{N}_g} \sum_{k \in \mathcal{K}} p_{ng,k}) = 0$ for all $g \in \mathcal{G}_2$, and $\phi_g(\tilde{R}_g^{\text{rs}} - \sum_{ng \in \mathcal{N}_g} \tilde{R}_{ng}) = 0$ for all $g \in \mathcal{G}_1$. Consequently, the Lagrange function for
(2.3) is

\[
\mathcal{L}(w, x, \lambda, \nu_k, \nu_g, \psi_g, \phi_g) = - \sum_{g \in G} \sum_{n_g \in N_g} \tilde{R}_{n_g} + \lambda \left( \sum_{g \in G} \sum_{n_g \in N_g} \sum_{k \in K} x_{n_g, k} - P_{\text{max}} \right) \\
+ \sum_{k \in K} \nu_k \left( \sum_{g \in G} \sum_{n_g \in N_g} w_{n_g, k} - 1 \right) + \sum_{g \in G_2} \nu_g \left( K_{g}^{\text{rsv}} - \sum_{n_g \in N_g} \sum_{k \in K} w_{n_g, k} \right) \\
+ \sum_{g \in G_2} \psi_g \left( P_{g}^{\text{rsv}} - \sum_{n_g \in N_g} \sum_{k \in K} x_{n_g, k} \right) + \sum_{g \in G_1} \phi_g \left( R_{g}^{\text{rsv}} - \sum_{n_g \in N_g} \tilde{R}_{n_g} \right).
\]

(2.4)

Our proposed iterative algorithm to solve (2.3) is summarized in Alg. 1. In the first step of Alg. 1, power vector \(P_{n_g}\), sub-carrier vector \(w\), and step sizes \(\delta_\mathcal{V}\) for all \(\mathcal{V} \in \{\phi_g, \lambda, \nu_g, \psi_g\}\) are initialized, where \(\mathcal{V} \in \{\phi_g, \lambda, \nu_g, \psi_g\}\) represents the set of dual variables of (6.8). Afterwards, the main body of Alg. 1 consists of two loops: 1) Outer loop: the values of dual variables \(\mathcal{V}\) are updated according to gradient descent method with small step sizes for each iteration. 2) Inner loop: the values of dual variables from outer loop is used to calculate the values of primal variables (i.e., \(P\) and \(w\)). For example, for \(x_{n_g, k}\), we have

\[
\frac{\partial \mathcal{L}}{\partial x_{n_g, k}} = \lambda - \frac{(1 + \phi_g)}{\ln(2)} \left( \frac{w_{n_g, k} h_{n_g, k}}{\sigma w_{n_g, k} + x_{n_g, k} h_{n_g, k}} \right) - \psi_g = 0,
\]

then

\[
P_{n_g, k} = \begin{cases} 
\left[ (1 + \phi_g) - \frac{\sigma}{h_{n_g, k}} \right] P_{\text{max}} & \text{if } g \in G_1, \\
\left[ \frac{1}{\ln(2) \lambda} - \frac{\sigma}{h_{n_g, k}} \right] P_{\text{max}} & \text{if } g \in G_2,
\end{cases}
\]

(2.5)

where \([x]_a^b = \min\{b, \max\{x, a\}\}\). Applying KKT conditions, we obtain the following necessary condition for \(w_{n_g, k}, \forall k \in K, n_g \in N_g, [45, 46]\),

\[
\begin{cases} 
w_{n_g, k} = 0 & \frac{\partial \mathcal{L}}{\partial w_{n_g, k}} < 0 \\
0 < w_{n_g, k} < 1 & \frac{\partial \mathcal{L}}{\partial w_{n_g, k}} = 0 \\
w_{n_g, k} = 1 & \frac{\partial \mathcal{L}}{\partial w_{n_g, k}} > 0
\end{cases}
\]

where
\[
\frac{\partial L}{\partial w_{n_g,k}} = \begin{cases} 
\nu_k - (1 + \phi_g) \left( \log_2 (1 + \gamma_{n_g,k}) - \frac{\gamma_{n_g,k}}{(1+\gamma_{n_g,k}) \ln(2)} \right), & \forall g \in \mathcal{G}_1, \\
\nu_k - \log_2 (1 + \gamma_{n_g,k}) + \frac{\gamma_{n_g,k}}{(1+\gamma_{n_g,k}) \ln(2)} - \nu_g, & \forall g \in \mathcal{G}_2.
\end{cases}
\]

Therefore, from KKT conditions, we obtain
\[
\rho_{n_g,k} = \begin{cases} 
(1 + \phi_g) \left( \log_2 (1 + \gamma_{n_g,k}) - \frac{\gamma_{n_g,k}}{(1+\gamma_{n_g,k}) \ln(2)} \right), & \forall g \in \mathcal{G}_1, \\
\nu_g + \log_2 (1 + \gamma_{n_g,k}) - \frac{\gamma_{n_g,k}}{(1+\gamma_{n_g,k}) \ln(2)}, & \forall g \in \mathcal{G}_2,
\end{cases}
\] (2.6)

where \( \gamma_{n_g,k} = \frac{x_{n_g,k} h_{n_g,k}}{\sigma w_{n_g,k}} \). From C1, we select user \( n_g \) for which \( \rho_{n_g,k} \) is maximum, and allocate the sub-carrier \( k \) to that user [43]. This process is repeated until Alg. I converges to stable values of primal and dual variables.

Algorithm 1: Slice Provisioning Algorithm

**Initialization:** Set \( w^*(l = 0) = 1, P^*_n (l = 0) = P_{\text{max}} / K, \forall n_g \in \mathcal{N}_g, g \in \mathcal{G}, l_{\text{max}}, i_{\text{max}} \) and \( 0 < \epsilon_m \ll 1 \) for \( m = \{1, \ldots, 4\} \).

**Step 1:** Outer loop iteration \( l \):
\[
\lambda(l + 1) = \left[ \lambda(l) + \delta \lambda \frac{\partial \lambda}{\partial \lambda} \right]^+, \\
\nu_g(l + 1) = \left[ \nu_g(l) + \delta \nu_g \frac{\partial \nu_g}{\partial \nu_g} \right]^+, & \forall g \in \mathcal{G}_1, \\
\psi_g(l + 1) = \left[ \psi_g(l) + \delta \psi_g \frac{\partial \psi_g}{\partial \psi_g} \right]^+, & \forall g \in \mathcal{G}_2, \\
\phi_g(l + 1) = \left[ \phi_g(l) + \delta \phi_g \frac{\partial \phi_g}{\partial \phi_g} \right]^+, & \forall g \in \mathcal{G}_1.
\]

**Step 2:** Inner loop iteration \( i \):
- Update \( P^*(i + 1) \) according to (2.5).
- Calculate \( \rho_{n_g,k} \) according to (2.6).
- select \( n_g, k = \max[\rho_{n_g,k}] \) and set \( w_{n_g,k} = 1 \), otherwise \( w_{n_g,k} = 0, \forall k \in \mathcal{K} \).
- When \( \| P(i + 1) - P(i) \| \leq \epsilon_1 \), or \( i > i_{\text{max}} \), Stop.
- Otherwise goto **Step 2**.
- When \( \| \phi_g(l + 1) - \phi_g(l) \| \leq \epsilon_2 \) and \( \| \lambda(l + 1) - \lambda(l) \| \leq \epsilon_3 \) and \( \| \psi_g(l + 1) - \psi_g(l) \| \leq \epsilon_4 \) or \( l > l_{\text{max}} \), Stop. Otherwise goto **Step 1**.

**2.4 Admission Control**

Due to user mobility, path-loss, shadowing, random nature of wireless channel, and transmit power limitations at the BS, there always exists a non-zero outage probability when C4 is infe-
2 Joint Resource Provisioning and Admission Control in VWN

Sensible for network. In other words, even with allocating all the power and sub-carries to one slice, $R_{gs}^{sv}$ cannot hold, and consequently, $R_{gs}^{sv}$ and $K_{gs}^{sv}$ cannot be satisfied. Therefore, the admission control policy is required to dynamically adapt the values of $R_{gs}^{sv}$, $K_{gs}^{sv}$ and $P_{gs}^{sv}$ according to CSI of users. We introduce $\varepsilon = [\varepsilon_1, \ldots, \varepsilon_g, \ldots, \varepsilon_G]$, as admission control variables, and $\alpha_g$ as a priority factor for slice $g \in G$. More specifically,

$$\varepsilon_g = \begin{cases} 
\varepsilon_gr, & \forall g \in G_1, \\
\varepsilon_gk + \varepsilon_gp, & \forall g \in G_2,
\end{cases} \quad (2.7)$$

Furthermore, $\varepsilon_gr$, $\varepsilon_gk$ and $\varepsilon_gp$ represent admission control variables for $R_{gs}^{sv}$, $K_{gs}^{sv}$ and $P_{gs}^{sv}$, respectively. Consequently, we transform C3 and C4 into

$$C5 : \sum_{n_g \in N_g} \sum_{k \in K} w_{n_g,k} \geq K_{gs}^{sv} - \varepsilon_gk, \quad \sum_{n_g \in N_g} \sum_{k \in K} w_{n_g,k}p_{n_g,k} \geq P_{gs}^{sv} - \varepsilon_gp, \quad \forall g \in G_2$$

and

$$C6 : \sum_{n_g \in N_g} R_{n_g} \geq R_{gs}^{sv} - \varepsilon_gr, \quad \forall g \in G_1,$$

respectively. Via these transformations, the minimum required rate and resources for slices can be adjusted using the following admission control optimization problem.

$$\max_{x, w, \varepsilon} \sum_{g \in G} \left( \sum_{n_g \in N_g} R_{n_g} - \alpha_g \varepsilon_g \right), \quad (2.8)$$

subject to C1, C2, C5 and C6.

In (2.8), the new set of variables, $\varepsilon$, is utilized to transfer the non-feasible optimization problem in (2.1) into the feasible one in (2.8). However, we deploy the term $\alpha_g \varepsilon_g$ in the objective function as a price of admission control policy. Therefore, the VWN tries to maintain the feasibility of resource provisioning problem with the minimum values of admission control variable which also depend on the priority of each slice, i.e., $\alpha_g$. Therefore, high-priority slices have higher VWN price than low-priority slices. This issue will be further studied in Section V. To solve (2.8) which involve the discrete variables similar to (2.1), we deploy the approach of previous section.
to transform the binary variables into the continuous one as follows,

$$\max_{x,w,\varepsilon} \sum_{g \in G} \left( \sum_{n_g \in N_g} \tilde{R}_{n_g} - \alpha_g \varepsilon_g \right),$$

subject to \(\tilde{C}1, C2, C5\) and

\[ \tilde{C}6 : \sum_{n_g \in N_g} \tilde{R}_{n_g} \geq R^\text{env} - \varepsilon_{gr}, \forall g \in G_1.\]

Now, we can apply Alg. 1 with minor modifications to solve (2.9). Due to the space limitation, we just highlight the main differences. In this case, similar to Alg. 1, the outer loop updates the Lagrange multipliers associated to \(\tilde{C}1, C2, C5\) and \(\tilde{C}6\). In the inner loop, in addition to power vector \(P\) and sub-carrier vector \(w\), the admission control vector is also updated as

$$\varepsilon_g(i + 1) = \left[ \varepsilon_g(i) - \delta \varepsilon_g \frac{\partial L}{\partial \varepsilon_g} \right]^+, \forall g \in G.$$

(2.10)

where \(\frac{\partial L}{\partial \varepsilon_g} = \alpha_g - \phi_g, \forall g \in G_1\) and \(\frac{\partial L}{\partial \varepsilon_g} = \alpha_g - \nu_g - \psi_g, \forall g \in G_2\).

2.5 Numerical Results and Discussions

In this section, we evaluate the performance of our proposed problems via MATLAB simulations results. Consider a BS virtualized into two slices where \(K = 64, B_k = 200 \text{ KHz}, \sigma = -173 \text{ dBm}\). Furthermore, \(h_{n_g,k} = X D_{n_g}^{-\beta}\) where \(\beta = 3\) is the path loss exponent, \(X\) is exponential random variable with mean one, and \(D_{n_g}\) is the distance of user \(n_g\) from BS. For comparison, we use equal power allocation (EPA) algorithm where the total available signal power is shared equally among optimal selected sub-carriers [47] since its performance approaches that of the global optimum at the high SINR scenario. All the results are presented in terms of total average rate over 100 channel realizations. For both Alg. 1 and EPA, when C4 and C5 do not hold, the total rate is set to zero. In the following, we present simulation results to study the effects of different system parameters on the proposed algorithms.
2.5.1 Performance of Slice Provisioning Algorithm

To evaluate the performance of Alg. 1, we set $P_{\text{max}} = 15\, \text{dB}$ and $D_{n_g} \in [0.3, 0.9]$ for both slices. Furthermore, we consider five sets assuming different values for C3 and C4, given in Table 2.1. Set 1 has relaxed constraints, i.e., $R^\text{RSV}_g$, $K^\text{RSV}_g$ and $P^\text{RSV}_g$ in C3 and C4 have small values. Set 2 has normal constraints, i.e., $R^\text{RSV}_g$, $K^\text{RSV}_g$ and $P^\text{RSV}_g$ are neither very small nor large. In contrast, Set 3 has the most restricted constraints for both slices $g_1$ and $g_2$, i.e., $R^\text{RSV}_g$, $K^\text{RSV}_g$ and $P^\text{RSV}_g$ have large values. In other words, the QoS requirements of both $g_1$ and $g_2$ are high. Set 4 is restricted for $g_1$ but relaxed for $g_2$, and vice versa for Set 5. Note that in Table 2.1, $P^\text{RSV}_g$ shows the percentage of $P_{\text{max}}$. Note that the values of these constraints can control the feasibility region of (2.1) and we will show that how these values can reduce or increase the total rate of VWN.

<table>
<thead>
<tr>
<th>Set</th>
<th>$R^\text{RSV}_g$ (bps/Hz)</th>
<th>$K^\text{RSV}_g$</th>
<th>$P^\text{RSV}<em>g$ (% of $P</em>{\text{max}}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>1.0</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>1.5</td>
<td>32</td>
<td>40</td>
</tr>
<tr>
<td>4</td>
<td>1.5</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
<td>32</td>
<td>40</td>
</tr>
</tbody>
</table>

Fig. 2.2 illustrates the total rate of VWN versus $N_g$. For all mentioned sets, with increasing $N_g$, system rate is increased as expected from multi-user diversity gain. However, the performance gap of EPA and Alg. 1 decreases as $N_g$ increases. This is because at higher number of users, the probability to have the high value of CSI, e.g., high SNR (signal-to-noise ratio), is increased due to multi-user diversity gain. At high SNR regime, EPA approaches the optimal solution [48]. Consequently, the gap between Alg.1 and EPA vanishes.

From Fig. 2.2, the rate of different sets in Table 2.1 can be compared. Note that for convex optimization problems, when the feasibility set is expanding, the obtained objective function at the optimum value is increased [44]. For (2.3), when values of $R^\text{RSV}_g$, $K^\text{RSV}_g$ and $P^\text{RSV}_g$ are decreased, the feasibility set is expanded, i.e., there are more power vectors satisfying $C_1$, $C_2$, $C_3$ and $C_4$. 

Therefore, Set 1 is supposed to have the highest system rate since its feasibility set is the largest compared to the others, i.e., the constraints of C4 and C5 have the small values which is also demonstrated in Fig. 2.2.

However, Set 3 should have the lowest rate since the constraints of C4 and C5 have the largest values, leading to shrinking the feasibility set which is observable in Fig. 2.2. From Fig. 2.2, the rate of Set 2 falls between the performances of Set 1 and Set 3, since the values of C4 and C5 are in the middle between those for Set 1 and Set 3, respectively. It is interesting to note that the performance of Set 5 with relaxed version of constraint C4 approaches to that of Set 1, indicating that large values of $P_{rsv}^g$ and $K_{rsv}^g$ in C3 does not have profound impact on the system rate. Set 4 due to restricted constraint C4, has lower rate than Set 2 at low number of users. However, due to multiuser diversity gain with increasing $N_g$, the feasibility for C4 increases, and therefore, Set 4 has better performance than Set 2.

For a clear understanding of the effect of different values of $R_{rsv}^g$, $K_{rsv}^g$ and $P_{rsv}^g$ on feasibility condition of (2.3), in Fig. 2.3, we plot the outage probability of system defined as $Pr(outage) = Pr(C3' \cup C4')$, where

$$C3' : \sum_{n_g \in N_g} \sum_{k \in K} w_{n_g,k} \leq K_{rsv}^g \forall g \in G_1$$

and

$$C4' : \sum_{n_g \in N_g} \sum_{k \in K} w_{n_g,k} P_{n_g,k} \leq P_{rsv}^g \text{ or } \sum_{n_g \in N_g} \sum_{k \in K} R_{n_g,k} \leq R_{rsv}^g \forall g \in G_2,$$

respectively.

From Fig. 2.3, with increasing $P_{max}^\text{max}$, the outage probability is reduced for all sets in Table 2.1 as we expected from expanding feasibility region. Also, EPA has more outage probability compared to Alg. 1 for all sets and $P_{max}^\text{max}$. This is because EPA cannot adjust the amount of transmit power for sub-carriers to satisfy $R_{rsv}^g$ and $P_{rsv}^g$, indicating that EPA is not suitable for VWN. Fig. 2.3 points out that Pr(outage) of Set 3 has the largest value compared to the other sets. With decreased values of $R_{rsv}^g$, $K_{rsv}^g$ and $P_{rsv}^g$ for other sets, e.g., for Set 1, Pr(outage) is decreased while still non-zero for $P_{max}^\text{max} < 10$ dB, meaning that the system needs to adjust these values based on the CSIs of users in different slices. To avoid outage and bring system into feasible situation, we introduced an admission control policy for (2.3) as discussed in Section IV. The next subsection presents our simulation results for admission control policy for (2.3).
2.5.2 Performance of Admission Control Algorithm

Here, we investigate the performance of admission control policy for three scenarios: *Scenario 1*) in which all slices belong to $G_1$, *Scenario 2*) in which all slices belong to $G_2$, *Scenario 3*) where there exist slices from both $G_1$ and $G_2$. Also for clear presentation, we choose more restricted constraints as given in Table 2.2.

Fig. 2.2  Rate versus Number of Users in slices.
For the first scenario, we consider two slices for $G_1$ where $D_{n_1} \in [0.2, 0.6]$ and $D_{n_2} \in [0.5, 0.9]$ for slices $g_1$ and $g_2 \in G_1$, respectively. Fig. 2.4 plots $\varepsilon_{gr}$ versus $P_{rsv}^{rsv}$. Clearly, $\varepsilon_{g_2r}$ is higher than $\varepsilon_{g_1r}$, since the users of $g_2$ have larger distance to the BS, leading to the less average channel gains compared to the users of $g_1$. Therefore, more adjustment is required for $P_{g_2}^{rsv}$ as
Joint Resource Provisioning and Admission Control in VWN

Fig. 2.4  $\varepsilon_{g r}$ versus $R^\text{rsv}_g$ for Scenario 1.

compared to $R^\text{rsv}_{g_1}$, e.g., $\varepsilon_{g_1 r} = 0$ and $\varepsilon_{g_2 r} > 0$ for $R^\text{rsv}_{g_1}$ and $R^\text{rsv}_{g_2}$ less than 3 bit/s/Hz. With increasing $R^\text{rsv}_g$, due to transmit power limitation of BS, the $R^\text{rsv}_{g_1}$ is not feasible. Therefore, $R^\text{rsv}_{g_1}$ should also be adjusted and $\varepsilon_{g_1 r} > 0$. Fig. 2.5 illustrates the performance of Alg. 2 for Scenario 2. The network simulation is similar to Fig. 2.4 except that $g_1$ and $g_2 \in \mathcal{G}_2$, and $K^\text{rsv}_{g_1} = 32$. From Fig. 2.5, when $\alpha_{g_1} = \alpha_{g_2} = 1$, i.e., two slices have equal priority, most of the sub-carriers are allocated to the users of slice $g_1$ which are close to BS in order to maximize the utility function of (2.8) and $\varepsilon_{g_1 k} \approx 0$ while $\varepsilon_{g_2 k} \gg 1$ for slice $g_2$. When $\alpha_{g_1} = 1$ and $\alpha_{g_2} = 10$, i.e., higher priority for $g_2$, $\varepsilon_{g_2 k} \approx 0$ and $\varepsilon_{g_1 k} \gg 1$. In other words, due to the higher priority of $g_2$, more sub-carriers are allocated to the $g_2$ despite the fact that users of $g_2$ have poor CSI. This is because the penalty for decreasing $K^\text{rsv}_{g_2}$ is increased in the objective function of (2.8) when $\alpha_{g_2} \gg 1$. The results of Fig. 2.5 demonstrate how the priority factor can change the results of admission control algorithm.


2 Joint Resource Provisioning and Admission Control in VWN

Fig. 2.5 $\varepsilon_{gk}$ versus $K_r^{rsv}$ and $\alpha_g$ for Scenario 2.

Table 2.3 $\varepsilon_{gr}$, $\varepsilon_{gp}$ and $\varepsilon_{gp}$ with $\alpha_{g1} = \alpha_{g2} = 1$ for Scenario 3

<table>
<thead>
<tr>
<th>Cases</th>
<th>Admission Control Parameters</th>
<th>Set 1</th>
<th>Set 2</th>
<th>Set 3</th>
<th>Set 4</th>
<th>Set 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case A</td>
<td>$\varepsilon_{gr}$ (bps/Hz)</td>
<td>0.18</td>
<td>0.62</td>
<td>1.23</td>
<td>1.05</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>$\varepsilon_{gp}$ (dB)</td>
<td>0</td>
<td>0</td>
<td>1.94</td>
<td>0</td>
<td>2.21</td>
</tr>
<tr>
<td></td>
<td>$\varepsilon_{gp}$ (dB)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Case B</td>
<td>$\varepsilon_{gr}$ (bps/Hz)</td>
<td>0</td>
<td>0.013</td>
<td>0.02</td>
<td>0</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>$\varepsilon_{gp}$ (dB)</td>
<td>3.0</td>
<td>3.81</td>
<td>8.43</td>
<td>3.189</td>
<td>5.21</td>
</tr>
<tr>
<td></td>
<td>$\varepsilon_{gp}$ (dB)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
In Table 2.3, performance of admission control policy for two cases is investigated:

- **Case A)** when the users of $g_1 \in G_1$ are located at the cell boundary, e.g., $D_{n_{g_1}} \in [0.85, 0.9]$ while the users of $g_2 \in G_2$ are close to the BS, e.g., $D_{n_{g_2}} \in [0.3, 0.35]$;

- **Case B)** when the users of $g_1 \in G_1$ are close to the BS, e.g., $D_{n_{g_1}} \in [0.3, 0.35]$, and the users of $g_2 \in G_2$ are located on the cell boundary, e.g., $D_{n_{g_2}} \in [0.85, 0.9]$.

For two simulation cases, both slices have same priority, i.e., $\alpha_{g_1} = \alpha_{g_2} = 1$.

Results of Table 2.3 reveal that $\varepsilon_{g_1r}$ of Case A has non-zero values for all sets due to their poor CSI. However, $\varepsilon_{g_1r}$ has large values for Set 3 and Set 4 due to the high value of $R_{g_1}^{rsv}$ while it has low values for other sets. For users of $g_2$, $\varepsilon_{g_2r} = 0$ since the CSI is relatively good. For Set 3 and Set 5, $\varepsilon_{g_2p}$ and $\varepsilon_{g_2k}$ are non-zero due to restricted constraints in C3. For Case B, $\varepsilon_{g_1r}$ approaches to zero, however, $\varepsilon_{g_2p}$ and $\varepsilon_{g_2k}$ have the higher value compared to Case A. This is because the users of $g_2$ experience high attenuation for Case B. From the network perspective, to reach a higher rate and keep the resource provisioning problem feasible, $P_{g_2}^{rsv}$ and $K_{g_2}^{rsv}$ should be adjusted. Also, decreasing $P_{g_2}^{rsv}$ effects the network feasibility more significantly compared to $K_{g_2}^{rsv}$, e.g., $\varepsilon_{g_2p} = 10.25$ and $\varepsilon_{g_2k} = 1$ for Set 3.

### Table 2.4 $\varepsilon_{gr}, \varepsilon_{gk}$ and $\varepsilon_{gp}$ for Different Sets in Scenario 3

<table>
<thead>
<tr>
<th>$\alpha_{g_1} = 3$, $\alpha_{g_2} = 1$</th>
<th>Set 1</th>
<th>Set 2</th>
<th>Set 3</th>
<th>Set 4</th>
<th>Set 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_{gr}$ (bps/Hz)</td>
<td>0.028</td>
<td>0.42</td>
<td>0.97</td>
<td>0.87</td>
<td>0.03</td>
</tr>
<tr>
<td>$\varepsilon_{gp}$ (dB)</td>
<td>0</td>
<td>1.48</td>
<td>10.25</td>
<td>0.15</td>
<td>7.3</td>
</tr>
<tr>
<td>$\varepsilon_{gk}$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

To demonstrate effects of priority factors on the performance of admission control policy, Case A with $\alpha_{g_1} = 3$ and $\alpha_{g_2} = 1$ is considered. For this case, simulation results are summarized in Table 2.4. By increasing priority of $g_1$, it is clear that the values of $\varepsilon_{g_1r}$ are reduced as compared to the simulation results of Table 2.3 for all sets. This is because higher value of $\alpha_{g_1}$ increases the penalty in utility function in (2.8). Therefore, admission control policy tries to maintain $R_{g_1}^{rsv}$ while with adjusting $\varepsilon_{g_2p}$, it holds the feasibility of resource provisioning problem. The results
of Table 2.4 are in line with results of Fig. 2.5, which further confirm that admission control policy has an effective performance with considering the priority and users’ CSI to preserve the feasibility.

2.6 Concluding Remarks

In this chapter, we proposed a joint resource provisioning and admission control policy for OFDMA-based VWNs to maintain the requirements of two groups of slices: 1) rate-based slices with minimum required rate, and 2) resources-based slices with minimum required power and sub-carriers. We propose efficient algorithms with reasonable computational complexity to solve them. The simulation results confirm the efficiency of proposed resource provisioning compared to other approaches. Furthermore, results demonstrate the ability of the admission control algorithm to avoid outage by efficiently adjusting the minimum slice requirements based on CSI and priority of slices.
Chapter 3

Resource Provisioning with Massive MIMO in VWN

This chapter proposes a dynamic resource provisioning scheme for an OFDMA-based VWN, where one BS is equipped with a large number of antennas and serves users belonging to multiple service providers via different slices. In particular, to maximize a sum-utility while maintaining a minimum rate per slice, a joint power, sub-carrier, and antenna allocation problem is presented for both perfect and imperfect channel knowledge cases. Subsequently, relaxation and variable transformation are applied to develop a simple and efficient algorithm to solve the formulated non-convex, combinational optimization problem. Simulations are performed to demonstrate the effectiveness of proposed algorithm. The obtained results show the network performance for different system conditions and confirm the benefits of applying a large number of antennas in this setup\(^1\).

3.1 Introduction

Wireless network virtualization is a promising approach to provide service customization for next-generation wireless networks. In VWNs, limited wireless resources (e.g., power and spectrum) are shared among different groups of users referred to as slices. Each slice may have a different set of QoS requirements, which calls for effective resource provisioning algorithms in VWNs to maximize the sum utility of all slices, while holding the minimum required rate of each slice.

\(^1\)Parts of this chapter have been published in [49].
However, due to the random nature of wireless channels, power and spectrum limitations, there is always a non-zero probability that the minimum required rate cannot be satisfied. Thus, this will cause an infeasibility problem for resource provisioning and lead to poor network performance. In this chapter, to expand the feasibility region, we take advantage of the large spatial degree of freedom introduced by massive MIMO technology. In massive MIMO, each BS is equipped with a large number of antennas serving multiple single-antenna users.

The massive MIMO setup scales up the MIMO multiplexing gain, provided that the exact CSI of all users is available at the BS [50]. However, extracting the precise values of CSI requires ideal orthogonal pilot signals between different BSs, which is practically infeasible [50]. In this chapter, taking this issue into account, we consider two cases: 1) perfect CSI, where it is ideally assumed to have all the precise values of CSI; and 2) imperfect CSI where the effect of non-orthogonal pilots is considered in the achievable rate.

Subsequently, we formulate a resource provisioning problem for the up-link transmission in a VWN. The objective is to maximize the sum utility of all slices subject to the minimum required rate of each slice and transmit power of each user within each slice. We introduce a new utility function representing the difference between the total achieved user rate and its corresponding costs for allocated power and antennas. The considered pricing mechanisms for the allocated power and antennas enable an effective control of inter-slice interference and available antennas in VWN, respectively. To the best of our knowledge, there exists no related work in the context of VWN with massive MIMO.

As the formulated problem is both non-convex and combinatorial, it suffers from high computational complexity. To develop an efficient algorithm, we apply variable transformations and constraint relaxations. Simulation results reveal that the values of power and antenna pricing variables have a significant impact on the total achieved user throughput.

The rest of this chapter is organized as follows. Section 3.2 introduces the system model and problem formulations. Section 3.3 provides a solution for the formulated resource provisioning problem, followed by Section 3.4 where the simulation results are presented. Section 3.5 provides concluding remarks.
3.2 Network Model and Problem Formulation

We consider the up-link transmission in an OFDMA VWN where a BS with $M$ antennas serves a set of slices $\mathcal{G} = \{1, \ldots, G\}$. Each slice $g \in \mathcal{G}$ has a set of single-antenna users denoted by $\mathcal{N}_g = \{1, \ldots, N_g\}$ and requires a minimum rate $R_g^{\text{sys}}$. It should be noted that $N = \sum_{g \in \mathcal{G}} N_g$ and $N \ll M$. Consider $\mathbf{M} = [\mathbf{M}_1, \ldots, \mathbf{M}_G]$ as the allocated antenna vector for all slices where $\mathbf{M}_g = \left[\mathbf{M}_{n_g}\right]_{n_g=1}^{N_g}$, $\mathbf{M}_{n_g} = [M_{n_g,1}, \ldots, M_{n_g,k}, \ldots, M_{n_g,K}]$, and $M_{n_g,k}$ is the number of antennas allocated to user $n_g$ on sub-carrier $k$. Let $\mathbf{w} = [\mathbf{w}_1, \ldots, \mathbf{w}_G]$ be the sub-carrier assignment vector for all slices where $\mathbf{w}_g = [\mathbf{w}_{n_g}]_{n_g=1}^{N_g}$, $\mathbf{w}_{n_g} = [w_{n_g,1}, \ldots, w_{n_g,k}, \ldots, w_{n_g,K}]$, and $w_{n_g,k} = 1$ indicates that sub-carrier $k$ is allocated to user $n_g$ and otherwise $w_{n_g,k} = 0$.

In this setup, let $\mathbf{h}_{n_g,k} \in \mathbb{C}^{1 \times M_{n_g,k}}$ be the channel vector of user $n_g$ on sub-carrier $k$, where $h_{n_g,k,m_{n_g,k}}$ is the channel coefficient of user $n_g$ on sub-carrier $k$ and antenna $m_{n_g,k}$. More specifically, $h_{n_g,k,m_{n_g,k}} = g_{n_g,k,m_{n_g,k}} \sqrt{d_{n_g}}$ where $g_{n_g,k,m_{n_g,k}}$ represents the small-scale fading coefficient with variance of 1, and $d_{n_g}$ denotes the large-scale fading coefficient of user $n_g$ on sub-carrier $k$. Note that $d_{n_g}$ includes both path loss and shadowing [50]. Practically, the channel coefficients are estimated by the BS based on the uplink pilots with duration $\tau$ at the specific part of the coherence interval of $T_c$ [51]. In the perfect CSI case, the up-link received sample at BS after using the linear detector from user $n_g \in \mathcal{N}_g$ on sub-carrier $k$ is [51]

$$ y_{n_g,k}^{\text{Perf}} = \sqrt{p_{n_g,k}} \mathbf{h}_{n_g,k} \mathbf{f}_{n_g,k}^{\text{Perf}} x_{n_g,k} + I_{n_g,k} + \mathbf{z}_{n_g,k} \mathbf{f}_{n_g,k}^{\text{Perf}}, $$

(3.1)

where $x_{n_g,k}$ and $\mathbf{f}_{n_g,k}^{\text{Perf}} \in \mathbb{C}^{M_{n_g,k} \times 1}$ represent the transmit symbol and the precoding vector of user $n_g$ on sub-carrier $k$, respectively. Moreover, $p_{n_g,k}$ is the transmit power of user $n_g$ on sub-carrier $k$ and $\mathbf{z}_{n_g,k}$ is a vector of additive white Gaussian noise (AWGN) at the BS with zero mean and power spectral density $\sigma$, which is assumed to be independent and identically distributed (i.i.d.) for all users over all sub-carriers. For notational simplicity, we normalize $\sigma$ to 1 and thus $p_{n_g,k}$ is the transmit SNR and dimensionless. $I_{n_g,k} = \sum_{g \in \mathcal{G}} \sum_{n_g' \neq n_g} \sqrt{T_{n_g',k}} \mathbf{h}_{n_g',k} \mathbf{f}_{n_g',k}^{\text{Perf}} x_{n_g',k}$ is the interference from other users to user $n_g$. In the imperfect CSI case, due to pilot contamination error in the channel estimation and the linear detector, the received signal from user $n_g$ on sub-carrier $k$ is [51]

$$ y_{n_g,k}^{\text{Imperf}} = \sqrt{p_{n_g,k}} \tilde{\mathbf{h}}_{n_g,k} \mathbf{f}_{n_g,k}^{\text{Imperf}} x_{n_g,k} + \delta_{n_g,k}(\mathbf{P}) $$

(3.2)

where $\tilde{\mathbf{h}}_{n_g,k}$ is the estimated channel vector including the contamination error, $\mathbf{P} = [\mathbf{P}_1, \ldots, \mathbf{P}_G]$. 

is the allocated power vector of all slices in which \( \mathbf{P}_g = [\mathbf{P}_{ng}]_{n_g=1}^{N_g} \) and \( \mathbf{P}_{ng} = [p_{ng,1}, \ldots, p_{ng,k}] \), \( \mathbf{f}_{ng,k}^\text{Imperf} \in \mathbb{C}^{M_{ng,k} \times 1} \) is the precoding vector in the imperfect CSI case, and \( \delta_{ng,k}(\mathbf{P}) \) is the function of contamination error, interference and noise. Using maximum ratio combining (MRC) detector with \( M \to \infty \), the rate of user \( n_g \) on sub-carrier \( k \) is derived as follow.

**Proof.** For mutually independent \( 1 \times n \) vectors \( \mathbf{p} = [p_1, \ldots, p_n] \) and \( \mathbf{q} = [q_1, \ldots, q_n] \) whose elements are i.i.d. zero-mean, unity-variance random variables (RVs), it can be shown by the law of large numbers that \( \lim_{n \to \infty} \frac{1}{n} \mathbf{p} \mathbf{p}^\text{H} \xrightarrow{a.s.} 1 \) and \( \lim_{n \to \infty} \frac{1}{n} \mathbf{p} \mathbf{q}^\text{H} \xrightarrow{a.s.} 0 \) where \( \xrightarrow{a.s.} \) denotes the almost sure convergence.

For **perfect CSI** using MRC, \( \mathbf{f}_{ng,k} = \mathbf{h}_{ng,k}^\text{H} \) [51], and the SINR of user \( n_g \) on sub-carrier \( k \) is

\[
\gamma_{ng,k}^\text{Perf} = p_{ng,k} \| \mathbf{h}_{ng,k} \|^4 / (\| I_{ng,k} \|^2 + \| \mathbf{h}_{ng,k} \|^2).
\]

According to the law of large numbers for large \( M_{ng,k} \), \( \| \mathbf{h}_{ng,k} \|^2 \to M_{ng,k} d_{ng} \) and \( \| I_{ng,k} \|^2 \to 0 \). Thus rate for perfect CSI case is

\[
R_{ng,k}^\text{Perf} = \log_2 \left( 1 + p_{ng,k} d_{ng} M_{ng,k} \right).
\]

For **imperfect CSI** in (3.2),

\[
\delta_{ng,k}(\mathbf{P}) = \sum_{g \in \mathcal{G}} \sum_{n_g \neq n_g \in \mathcal{G}} \sqrt{p_{ng,k}^\text{Pilot} d_{ng}} \mathbf{h}_{ng,k} \mathbf{f}_{ng,k}^\text{Imperf} x_{ng,k} - \sqrt{p_{ng,k}} \sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{G}} \mathbf{e}_{ng,k} \mathbf{f}_{ng,k}^\text{Imperf} x_{ng,k} + \mathbf{z}_{ng,k} \mathbf{f}_{ng,k}^\text{Imperf},
\]

where \( \mathbf{e}_{ng,k} = \mathbf{h}_{ng,k} - \mathbf{h}_{ng,k} \) whose elements are RVs with zero mean and variance \( \frac{p_{ng,k}^\text{Pilot} d_{ng}}{p_{ng,k}^\text{Pilot} d_{ng} + 1} \) and \( p_{ng,k} = \tau p_{ng,k} \) [51]. With MMSE-based channel estimation, the elements of \( \mathbf{h}_{ng,k} \) are i.i.d. RVs with zero mean and variance \( \frac{p_{ng,k}^\text{Pilot} d_{ng}}{\tau p_{ng,k}^\text{Pilot} d_{ng} + 1} \) [51]. With MRC precoder, \( \mathbf{f}_{ng,k}^\text{Imperf} = \mathbf{h}_{ng,k}^\text{H} \). Since \( \mathbf{e}_{ng,k} \) and \( \mathbf{h}_{ng,k} \) are independent of \( \mathbf{h}_{ng,k} \) and \( \mathbf{f}_{ng,k}^\text{Imperf} \), the first term of \( \delta_{ng,k}(\mathbf{P}) \) is zero and second term is equal to \( \sqrt{p_{ng,k}^\text{Pilot} \mathbf{e}_{ng,k} \mathbf{f}_{ng,k}^\text{Imperf} x_{ng,k}} \). Thus, the SINR for the imperfect CSI case is

\[
\gamma_{ng,k}^\text{Imperf} = p_{ng,k} \| \mathbf{h}_{ng,k} \|^4 / (p_{ng,k} \| \mathbf{e}_{ng,k} \| \mathbf{h}_{ng,k} \|^2 + \| \mathbf{h}_{ng,k} \|^2)
\]

\[
= M_{ng,k}^2 p_{ng,k} \left( \frac{p_{ng,k}^\text{Pilot} d_{ng}}{p_{ng,k}^\text{Pilot} d_{ng} + 1} \right)^2.
\]
By some mathematical manipulations, \( \gamma_{n_g,k}^{\text{Imperf}} = \frac{M_{n_g,k} P_{\text{Pilot}} n_g, k p_{n_g,k}^2 d_{n_g}^2}{p_{n_g,k} d_{n_g} + P_{\text{Pilot}} n_g, k d_{n_g} + 1} \). Now, by substituting \( p_{n_g,k}^\text{Pilot} = \tau p_{n_g,k} \) and considering \( p_{n_g,k} = \rho_{n_g,k} \sqrt{M_{n_g,k}} \) [51], we have

\[
\gamma_{n_g,k}^{\text{Imperf}} = \frac{\tau \rho_{n_g,k}^2 d_{n_g}^2}{1 + (1 + \tau) \rho_{n_g,k} d_{n_g} / \sqrt{M_{n_g,k}}}. \]

When \( M_{n_g,k} \to \infty \), we have \( \gamma_{n_g,k}^{\text{Imperf}} = \tau \rho_{n_g,k}^2 d_{n_g}^2 = \tau p_{n_g,k}^2 d_{n_g}^2 M_{n_g} \). Thus,

\[
R_{n_g,k} = \frac{T_c - \tau}{T_c} \log_2 \left( 1 + \tau p_{n_g,k}^2 d_{n_g}^2 M_{n_g} \right),
\]

where \((T_c - \tau)/T_c\) is the fraction of transmission frame for sending the actual data [51].

From the above results, we obtain

\[
R_{n_g,k} = \begin{cases} 
\log_2 \left( 1 + p_{n_g,k} d_{n_g} M_{n_g,k} \right) & \text{for perfect CSI,} \\
\frac{T_c - \tau}{T_c} \log_2 \left( 1 + \tau p_{n_g,k}^2 d_{n_g}^2 M_{n_g,k} \right) & \text{for imperfect CSI,}
\end{cases}
\] (3.3)

and the total achieved rate of user \( n_g \) is \( R_{n_g}(P, w, M) = \sum_{k \in K} w_{n_g,k} R_{n_g,k} \). Now, we define utility function of slice \( g \) as

\[
U_g(P, w, M) = \sum_{n_g \in N_g} R_{n_g}(P, w, M) - c_g^M \sum_{n_g \in N_g} \sum_{k \in K} w_{n_g,k} M_{n_g,k} - c_g^P \sum_{n_g \in N_g} \sum_{k \in K} w_{n_g,k} p_{n_g,k},
\]

where \( c_g^M \) and \( c_g^P \) are pricing factors for the number of allocated antennas and the transmit power of slice \( g \), respectively. Hence, the resource provisioning problem can be written as

\[
\max_{P, w, M} \sum_{g \in G} U_g(P, w, M),
\] (3.4)

subject to the following constraints C1 – C4.

C1: Exclusive sub-carrier allocation in OFDMA implies:

\[
w_{n_g,k} \in \{0, 1\} \quad \text{and} \quad \sum_{g \in G} \sum_{n_g \in N_g} w_{n_g,k} \leq 1, \quad \forall k \in K.
\]
C2: Transmit power limitation for each user \( n_g \) implies:
\[
\sum_{k \in K} w_{n_g,k} p_{n_g,k} \leq P_{n_g}^{\max}, \quad \forall n_g \in N_g, \quad \forall g \in G,
\]
where \( P_{n_g}^{\max} \) is the maximum transmit power of user \( n_g \).

C3: Minimum required rate, \( R_{g}^{sv} \), of slice \( g \in G \) implies:
\[
\sum_{n_g \in N_g} R_{n_g} = R_{g}^{sv}, \quad \forall g \in G.
\]

C4: Control on the number of antennas allocated to each slice to preserve the fairness between slices and to improve the energy efficiency for the VWNs can be expressed as [50],
\[
\sum_{n_g \in N_g} \sum_{k \in K} w_{n_g,k} M_{n_g,k} \in \{ M_{g}^{\min}, M_{g}^{\min+1}, \ldots, M_{g}^{\max} \},
\]
for each slice \( g \in G \) where \( M_{g}^{\min} \) and \( M_{g}^{\max} \) are the minimum number of reserved antennas and the maximum allowable number of allocated antennas for slice \( g \), respectively.

Since (3.4) has combinatorial properties, it involves high computational complexity. Next, we propose an efficient algorithm to solve this problem by applying variable transformations and constraint relaxation.

### 3.3 Proposed Algorithm

To obtain a low-complexity solution to (3.4), we relax the conditions C1, C3, and C4, by considering \( w_{n_g,k} \) as a continuous variable in the interval \([0, 1]\) and the numbers of allocated antennas to be non-negative real-valued. In the new definition, \( w_{n_g,k} \) indicates the fraction of time that sub-carrier \( k \) is assigned to user \( n_g \) for a specific transmission frame. Furthermore, we define two new variables: \( x_{n_g,k} = w_{n_g,k} p_{n_g,k} \) and \( y_{n_g,k} = w_{n_g,k} M_{n_g,k} \). For \( M_{n_g,k} \gg 1 \), \( R_{n_g} \) can be approximated and represented as
\[
\tilde{R}_{n_g} \approx \begin{cases} 
\sum_{k \in K} w_{n_g,k} \log_2 \left( \frac{x_{n_g,k} y_{n_g,k} d_{n_g}}{w_{n_g,k}} \right), & \text{for perfect CSI,} \\
\frac{T_c - \tau}{T_c} \sum_{k \in K} w_{n_g,k} \log_2 \left( \frac{\tau x_{n_g,k}^2 y_{n_g,k} d_{n_g}^2}{w_{n_g,k}^3} \right), & \text{for imperfect CSI,}
\end{cases}
\]
which are convex functions with respect to $x_{n_g,k}$, $y_{n_g,k}$, and $w_{n_g,k}$ [44, 52]. Consequently, conditions C1-C4 are changed to the following $\tilde{C}_1$ – $\tilde{C}_4$, respectively,

\begin{align*}
\tilde{C}_1 & : \ w_{n_g,k} \in [0, 1] \quad \text{and} \quad \sum_{g \in G} \sum_{n_g \in N_g} w_{n_g,k} \leq 1, \ \forall k \in K, \\
\tilde{C}_2 & : \ \sum_{k \in K} x_{n_g,k} \leq P_{n_g}^\text{max}, \quad \forall n_g \in N_g, \ \forall g \in G, \\
\tilde{C}_3 & : \ \sum_{n_g \in N_g} \tilde{R}_{n_g} \geq R_g^\text{rv}, \quad \forall g \in G, \\
\tilde{C}_4 & : \ M_g^\text{min} \leq \sum_{n_g \in N_g} \sum_{k \in K} y_{n_g,k} \leq M_g^\text{max}, \ \forall g \in G,
\end{align*}

and (3.4) is transformed to

\begin{equation}
\max_{x, w, y} \sum_{g \in G} \tilde{U}_g(x, w, y), \tag{3.5}
\end{equation}

subject to : $\tilde{C}_1$ – $\tilde{C}_4$,

where $x$ and $y$ are vectors of all $x_{n_g,k}$ and $y_{n_g,k}$, and

\[ \tilde{U}_g(x, w, y) = \sum_{n_g \in N_g} \tilde{R}_{n_g}(x, w, y) - c_g^M \sum_{n_g \in N_g} \sum_{k \in K} y_{n_g,k} - c_g^P \sum_{n_g \in N_g} \sum_{k \in K} x_{n_g,k}. \]

Since (3.5) involves continuous variables and convex functions, we can solve it with the following Lagrange function.

\begin{equation}
L(x, w, y, \lambda_{n_g}, \phi_g, \theta_g, \psi_g) = \sum_{g \in G} \tilde{U}_g + \sum_{n_g \in N_g} \lambda_{n_g} \left( \sum_{k \in K} x_{n_g,k} - P_{n_g}^\text{max} \right) + \sum_{g \in G} \phi_g \left( R_g^\text{rv} - \sum_{n_g \in N_g} \tilde{R}_{n_g} \right) + \sum_{g \in G} \theta_g \left( M_g^\text{min} - \sum_{n_g \in N_g} \sum_{k \in K} y_{n_g,k} \right) + \sum_{g \in G} \psi_g \left( \sum_{n_g \in N_g} \sum_{k \in K} y_{n_g,k} - M_g^\text{max} \right), \tag{3.6}
\end{equation}

where $\lambda_{n_g}, \phi_g, \theta_g$ and $\psi_g$ are the Lagrange multipliers for $\tilde{C}_2$, $\tilde{C}_3$ and $\tilde{C}_4$. Now, we propose an iterative approach to solve the dual problem of (3.6), in the following algorithm.

**Algorithm 1:**

**Initialization:**

Set $w^*(l = 0) = 1$, $P_{n_g}^*(l = 0) = P_{n_g}^\text{max} / K$, and $M_g^*(l = 0) = M_g^\text{max}$ for all $n_g \in N_g$ and $g \in G$. Initialize $\lambda(l = 0) = 1$, $\phi(l = 0)$, $\psi(l = 0)$ and $\theta(l = 0)$. 

Repeat:

- Update dual variables, $\lambda_n, \phi_g, \theta_g$ and $\psi_g$, by gradient descent method for all $g \in G$ where 
  \[ x^+ = \max\{x, 0\} \]
  \[
  \lambda_n(l + 1) = \left[ \lambda_n(l) + \delta \lambda_n \frac{\partial \mathcal{L}}{\partial \lambda_n} \right]^+ , \forall n \in \mathcal{N},
  \]
  \[
  \phi_g(l + 1) = \left[ \phi_g(l) + \delta \phi_g \frac{\partial \mathcal{L}}{\partial \phi_g} \right]^+ , \forall g \in \mathcal{G},
  \]
  \[
  \theta_g(l + 1) = \left[ \theta_g(l) + \delta \theta_g \frac{\partial \mathcal{L}}{\partial \theta_g} \right]^+ , \forall g \in \mathcal{G},
  \]
  \[
  \psi_g(l + 1) = \left[ \psi_g(l) + \delta \psi_g \frac{\partial \mathcal{L}}{\partial \psi_g} \right]^+ , \forall g \in \mathcal{G}.
  \]

- Using the above updated parameters for $(l + 1)$, compute $p_{n_g,k}^*(l + 1)$ and $M_{n_g,k}^*(l + 1)$ for all $n_g \in \mathcal{N}_g$, and $k \in \mathcal{K}$.

<table>
<thead>
<tr>
<th>For $l + 1$</th>
<th>Perfect CSI case</th>
<th>Imperfect CSI case</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{n_g,k}^*$</td>
<td>$\left[ \frac{1 + \phi_g}{\ln(2)(\lambda_{n_g} + c_{G}n_g)} \right]<em>0 P</em>{n_g}^{\max}$</td>
<td>$\left[ \frac{(T_c - \tau)}{T_c} \frac{2(1 + \phi_g)}{\ln(2)(\lambda_{n_g} + c_{G}n_g)} \right]<em>0 P</em>{n_g}^{\max}$</td>
</tr>
<tr>
<td>$M_{n_g,k}^*$</td>
<td>$\left[ \frac{1 + \phi_g}{\ln(2)(\psi_g - \theta_g + c_{M}M_g)} \right]<em>0 M</em>{n_g}^{\max}$</td>
<td>$\left[ \frac{(T_c - \tau)}{T_c} \frac{1 + \phi_g}{\ln(2)(\psi_g - \theta_g + c_{M}M_g)} \right]<em>0 M</em>{n_g}^{\max}$</td>
</tr>
</tbody>
</table>

where $[x]_a^b = \min\{b, \max\{x, a\}\}$.

- Perform sub-carrier allocation for all $n_g \in \mathcal{N}_g$ and for all $k \in \mathcal{K}$, where by applying KKT conditions [45, 53],
  \[
  \frac{\partial \mathcal{L}}{\partial w_{n_g,k}} = (1 + \phi_g) \begin{cases} \log_2(p_{n_g,k}d_{n_g}M_{n_g,k}) - \frac{2}{\ln(2)} , & \text{perfect CSI} \\ \frac{(T_c - \tau)}{T_c} \left( \log_2(p_{n_g,k}^2d_{n_g}^2M_{n_g,k}) - \frac{3}{\ln(2)} \right) , & \text{imperfect CSI} \end{cases}
  \]

  and consequently, the optimal sub-carrier allocation $w_{n_g,k}^*(l + 1)$ can be derived as in [52], i.e., $w_{n_g,k}^*(l + 1) = 1$, if $\frac{\partial \mathcal{L}}{\partial w_{n_g,k}} > 0$, $w_{n_g,k}^*(l + 1) = 0$, otherwise.

- Set $l = l + 1$

Until $||P(l + 1) - P(l)|| \leq \varepsilon$ or $l \geq l_{\max}$. 
Finally, the optimum integer \( M_{n_g, k}^* \) is selected as \( \lfloor M_{n_g, k}(l) \rfloor \) where \( \lfloor x \rfloor \) denotes the largest integer less than or equal to the value of \( x \).

### 3.4 Numerical Results and Discussions

![Figure 3.1](image-url)

**Fig. 3.1** Total rate vs. \( M_g^{\text{max}} \)

To evaluate the performance of VWN with the proposed algorithm, we consider one BS serving two slices (i.e., \( G = 2 \)) with \( N_1 = N_2 = 4 \). Furthermore, we assume \( K = 64 \) and \( P_{n_g}^{\text{max}} = 0 \) dB. After precoding, channel is modeled as \( ||h_{n_g, k, m_{n_g, k}}||^2 = d_{n_g} = D_{n_g}^{-\beta} \) where \( \beta = 3 \) is the path loss exponent and \( D_{n_g} \in [0.2, 0.6] \) is the distance of user \( n_g \) to BS. Furthermore, we set \( \varepsilon = 10^{-4} \) for Algorithm 1 and \( R_{1\text{sv}} = R_{2\text{sv}} = 2 \) bps/Hz unless otherwise stated. \( T_c \) is selected based on the parameters in [51] and simulations are based on the ratio of \( \tau/T_c \). All the results are presented in terms of average rate over 100 channel realizations based on random locations of the users. When \( R_{g}^{\text{sv}} \) does not hold for a channel realization, the total rate is set to zero. Also, to demonstrate the effect of \( c_g^M \) and \( c_g^P \), we consider three sets with different values of \( c_g^M \) and \( c_g^P \). For
Set 1, we have $c_g^M = c_g^P = 0$ which means that the utility of each user is equal to its rate. For Set 2, we have $c_g^M = 0.07$ and $c_g^P = 1$. Set 3 imposes higher restrictions to use antennas and power by employing $c_g^M = 0.09$ and $c_g^P = 2$.

Fig. 3.1 shows that the total rate is a non-decreasing function of $M_g^{\text{max}}$ in both perfect and imperfect CSI cases where $\frac{\tau}{T_c} = 0.3$. This can be explained by the fact that the feasibility regions are expanded as $M_g^{\text{max}}$ increases. However the utility function is formulated as a decreasing function of $c_g^P$ and $c_g^M$, so that by increasing the costs from Set 1 to Set 3, the total rate of VWN is decreased. As a result, the increase in the achieved rate is reduced for large values of $M_g^{\text{max}}$, especially for high $c_g^P$ and/or $c_g^M$, e.g., for $M_g^{\text{max}} > 75$, the total rate for Sets 2 and 3 is almost unchanged. Furthermore, due to the channel estimation errors, the performance with imperfect CSI is worse than that with perfect CSI for all sets, as expected.

To demonstrate the effects of $R_g^{\text{BSV}}$, Fig. 3.2 plots the total rate versus $R_g^{\text{BSV}}$ for Set 1, where $\frac{\tau}{T_c} = 0.3$, $M_1^{\text{max}} = M_2^{\text{max}} = 100$, and $R_1^{\text{BSV}} = R_2^{\text{BSV}} = R_g^{\text{BSV}}$. It can be observed that in both perfect and imperfect CSI cases, the total rate decreases with increasing $R_g^{\text{BSV}}$. This is because increasing $R_g^{\text{BSV}}$ shrinks the feasibility region for convex optimization problem (3.4) and thus reduces the optimal value of objective function [44].
To study further the effect of \( \tau \) on the performance of VWN, we focus on Set 1 with \( M_g^{\text{max}} = \{200, 150\} \) and plot the rate versus \( \frac{\tau}{T_c} \) in Fig. 3.3. Obviously, when \( 0.1 \leq \frac{\tau}{T_c} \leq 0.3 \), the rate increases with increasing \( \frac{\tau}{T_c} \) benefiting from the more accurate CSI estimation. However, for \( 0.3 \leq \frac{\tau}{T_c} \leq 0.9 \), the spectral efficiency of VWN is decreased with increasing \( \frac{\tau}{T_c} \) due to the fact that most of the transmission time wastes for pilot signals. Therefore, considering the optimum values of \( \frac{\tau}{T_c} \) to reach the best performance of VWN is essential. Fig. 3.3 also highlights that increasing \( M_g^{\text{max}} \) can increase the total rate for any \( \tau \) due to the feasibility region expansion.

To further study the effects of pricing on the total rate of VWN, in Fig. 3.4, the total rate is plotted versus \( c_g^M \) and \( c_g^P \). Obviously, the total rate is decreased with increasing \( c_g^M \) and \( c_g^P \). However, the effect of increasing \( c_g^M \) is much more profound compared to that of \( c_g^P \). This is because in massive MIMO, we have \( M_g^{\text{max}} \gg P_n^{\text{max}} \), e.g., compare \( M_g^{\text{max}} > 100 \) and \( P_n^{\text{max}} = 0 \) dB. Figures 3.3 and 3.4 highlight that allocating dynamic and optimal pricing values for VWN and optimum value of \( \tau/T_c \) are essential to reach the best performance which are left for the future work.

![Fig. 3.3 Total rate vs. \( \tau/T_c \)](image-url)
3.5 Concluding Remarks

Utility-based resource provisioning for massive MIMO-based VWN was investigated in this chapter for perfect and imperfect CSI scenarios. We proposed an efficient iterative algorithm to solve the developed resource allocation problems. Via simulation results, the effects of power and antenna pricing mechanisms on the performance of VWN were investigated.
Chapter 4

Resource Provisioning with Stable Queue Control for VWN

This chapter investigates a dynamic resource provisioning with queue stability in VWNs. Aiming to maximize the total average rate of a VWN over a transmission frame, a dynamic resource provisioning policy is proposed. In this context, a minimum average required rate of each slice and a stable-queue constraint of a VWN are preserved. Based on a Lyapunov drift-plus-penalty algorithm and the variable transformation techniques, an iterative algorithm is proposed for joint power and sub-carrier allocation. The simulation results confirm the effectiveness of proposed algorithm in maintaining the stable queue states of a VWN.

4.1 Introduction

The recent trend of service-oriented demands of mobile users calls for more efficient and modular wireless networks. The concept of wireless virtualization is a solution to provide flexibility and service customization in the existing networks via sharing the limited wireless resources (e.g., power and sub-carriers) among various slices. Generally, each slice has its own set of users and supports diverse QoS requirements for offered services.

Considering the dynamic features of wireless channels (e.g., fading and shadowing) and limited resources in a VWN, an effective, dynamic resource provisioning is required to satisfy the PHY-layer-based QoS requirements of each slice such as minimum required rate of each slice,

1Parts of this chapter have been presented in [54].
subject to VWN limitations such as transmit power at BSs. Recently, this problem has received a great deal of attention. For instance, in [28], a resource management scheme is studied by introducing rate-based and resource-based slices requiring minimum rate and physical resources, respectively. In [39], an auction game is applied for the interaction of slices, network operator, and users. To guarantee the minimum rate of slices under fading of wireless channels, an admission control policy is proposed in [40] where the minimum rate of each slice is adjusted based on its users’ CSI. To extend the feasibility condition of VWN for supporting diverse QoS, [49] borrowed the concept of massive MIMO. In [55], an elastic resource allocation problem is combined with time and space for an OFDMA system.

In addition to the PHY layer parameters considered in the above-mentioned works, it is very important to consider random traffic arrivals and the limited queue backlogs to reach reliable service support for each user. To address this issue, we propose the dynamic resource provisioning algorithm based on both PHY layer parameters (i.e., CSI) and the MAC layer parameters (i.e., QSI). The set objective of this optimization problem is to maximize the total rate of VWNs subject to an average queue stability condition of each user and average minimum required rate of each slice over one complete transmission frame including multiple time slots.

The proposed optimization problem of this chapter is challenging to solve because of its cross-layer nature and both PHY and MAC constraints. To tackle this, we apply the Lyapunov drift-plus-penalty algorithm [56] to transform the average rate dynamic problem over one transmission frame into one time slot optimal problem. Lyapunov drift approach and drift-plus-penalty algorithms have been widely applied for the resource allocation problem of conventional wireless networks by considering limited queue length or maximum delay requirement [22, 23, 57–62]. However, to the best of our knowledge, there is no other related works that apply this framework to solve the dynamic resource allocation problem in VWN.

In the down-link OFDMA-based transmission scenario, the optimization problem of each time slot includes both sub-carrier assignment and power allocation variables. Therefore, proposing the efficient algorithm to get rid of high computational complexity is of high importance. To this end, we first apply the technique of variable relaxations and transformations [52] to transform non-convex problem into a related convex one. Afterwards, we propose a two-level iterative algorithm to solve the queue-aware dynamic resource provisioning: 1) outer loop, where queue parameters are updated based on the achieved rate of slices, and 2) inner loop, where power and sub-carriers are derived by applying KKT conditions.

The rest of this chapter is organized as follows. Section 4.2 introduces the system model and
problem formulations, followed by Section 4.3, which provides detailed solution for the problem via Lyapunov drift-plus-penalty approach and eventually proposes the iterative algorithm and its performance analysis. Section 4.4 presents the simulation results and Section 4.5 concludes the chapter.

4.2 Network Model and Dynamic Resource Provisioning Problem

We consider a slotted down-link OFDMA-based VWN which provides the coverage to a specific region by one BS. In this area, VWN serves a set of $G = \{1, \ldots, G\}$ slices. Each slice $g \in G$ serves a group of users denoted by $N_g = \{1, \ldots, N_g\}$. The total number of users is $N = \sum_{g \in G} N_g$.

The network operates in slotted transmission time $T = \{1, \ldots, t, \ldots, T_{\text{max}}\}$. The total bandwidth of $B$ Hz is shared between users of all slices through OFDMA scheme, which is divided into a set of $K = \{1, \ldots, K\}$ sub-carriers. The bandwidth of each sub-carrier, $B_k = B/K$, is assumed to be small compared to the coherent bandwidth $B_c$ of the wireless channel. Based on this assumption, consider $h_{ng,k}^t$ as the channel power gain of user $n_g$ on sub-carrier $k \in K$ in time slot $t \in T$. Also, consider $p_{ng,k}^t$ as the allocated power to user $n_g$ on sub-carrier $k$ in time slot $t$ and $w_{ng,k}^t = \{0, 1\}$ as the sub-carrier assignment variable where $w_{ng,k}^t = 1$ if the sub-carrier $k$ is allocated to user $n_g$ at time slot $t$, otherwise $w_{ng,k}^t = 0$. In this setup, the total transmit power at the BS is limited to the maximum power $P_{\text{max}}$ in each time slot, i.e.,

$$C1: \sum_{g \in G} \sum_{n_g \in N_g} \sum_{k \in K} w_{ng,k}^t p_{ng,k}^t \leq P_{\text{max}}.$$

Furthermore, from OFDMA limitation, each sub-carrier is assigned exclusively to one user where

$$C2: \sum_{g \in G} \sum_{n_g \in N_g} w_{ng,k}^t \leq 1 \text{ and } w_{ng,k}^t \in \{0, 1\}, \forall k \in K.$$

Without loss of generality, let $\sigma$ be the noise power in each sub-carrier. In this scenario, the rate of user $n_g \in N_g$ at time slot $t$ is

$$R_{ng}^t (p_{ng}, w_{ng}) = \sum_{k \in K} w_{ng,k}^t \log_2 \left(1 + \frac{p_{ng,k}^t h_{ng,k}^t}{\sigma}\right),$$
where $\mathbf{P}_{n_g} = [p_{n_g,k}^t]_{k=1}^K$, $\mathbf{w}_{n_g} = [w_{n_g,k}^t]_{k=1}^K$ and 

$$\bar{R}_{n_g} = \lim_{T \to \infty} \frac{1}{T} \sum_{t \in T} R_{n_g}^t(\mathbf{P}_{n_g}, \mathbf{w}_{n_g})$$

is the average rate of user $n_g$ over a transmission frame. Now, the minimum required average rate of slice $g \in \mathcal{G}$ over one transmission frame, $\bar{R}_g^{\text{sv}}$ is presented as

C3: \[
\sum_{n_g \in \mathcal{N}_g} R_{n_g} \geq R_g^{\text{sv}}, \forall g \in \mathcal{G}.
\]

Let $\alpha_{n_g}^t$ and $Q_{n_g}^t$ respectively be the i.i.d. packet arrival process and queue length for user $n_g$ in time slot $t$. Furthermore, let $l$ be the size of a packet in bits per Hz per slot. Therefore, the queue dynamics for user $n_g$ can be mathematically represented as

$$Q_{n_g}^{t+1} = \max \left[ Q_{n_g}^t - R_{n_g}^t (\mathbf{P}_{n_g}, \mathbf{w}_{n_g}) T_t, 0 \right] + \alpha_{n_g}^t, \tag{4.1}$$

where $T_t$ is the duration of each time slot $t \in \mathcal{T}$. In this setup, network has a queue stability condition if and only if the queue of each user is stable. The queue of each user is stable when “it has a bounded time-average backlog”, i.e., $\limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}[Q(t)] \leq \infty$ [63], which can be explained as

C4: \[
\mu_{n_g} \leq \bar{R}_{n_g},
\]

where $\mu_{n_g}$ is the average packet arrival rate over one transmission frame. Thus, with the objective to maximize the average rate of users over all slices, the optimization problem can be formulated as

$$\max_{\mathbf{P}, \mathbf{w}} \sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} \bar{R}_{n_g}, \tag{4.2}$$

subject to C1 - C4,

where $\mathbf{P}$ and $\mathbf{w}$ are the vectors of all $\mathbf{P}_{n_g}$ and $\mathbf{w}_{n_g}$ for all $n_g \in \mathcal{N}_g$ and $g \in \mathcal{G}$, respectively.

Note that (4.2) is a cross-layer optimization problem which aims to maximize the average rate of system, while considering the queue dynamics of MAC layer and CSI of physical layer. The
multidimensional nature and the discrete variables make this problem hard to be solved in this form. In the next section, we resolve these issues via variable relaxation and Lyapunov drift.

4.3 Dynamic Resource Provisioning Algorithm

In the first step, by relaxing the integer variable $w_{n_g,k}$, we transform the constraint $C2$ to

$$\tilde{C2} : \sum_{g \in G} \sum_{n_g \in N_g} w_{n_g,k}^t \leq 1 \text{ and } w_{n_g,k}^t \in [0, 1], \forall k \in K.$$ 

In the context of new definition, $w_{n_g,k}^t$ represents the portion of a time slot $t$ for which sub-carrier $k$ is assigned to user $n_g$. Since by definition, $R_{n_g}^t(P_{n_g}, w_{n_g})$ is not convex w.r.t. $P_{n_g}$ and $w_{n_g}$, we consider another variable transformation i.e., $x_{n_g,k}^t = w_{n_g,k}^t P_{n_g,k}^t$ to change $R_{n_g}^t(P, w)$ into a convex function

$$\tilde{R}_{n_g}^t(x, w) = \sum_{k \in K} w_{n_g,k}^t \log_2 \left(1 + \frac{x_{n_g,k}^t h_{n_g,k}^t}{w_{n_g,k}^t \sigma} \right),$$

similar to the general class of functions $f(x, y) = y \log(1 + x/y)$ [52]. The problem (4.2), under the new definitions of variables, can be rewritten as

$$\max_{x, w} \sum_{g \in G} \sum_{n_g \in N_g} \tilde{R}_{n_g}$$

subject to $\tilde{C1}$, $\tilde{C2}$, $\tilde{C3}$, and $\tilde{C4}$,

where $\tilde{R}_{n_g} = \lim_{T \to \infty} \frac{1}{T} \sum_{t \in T} \tilde{R}_{n_g}^t(x, w)$ is the time average rate with new definitions of variables,

$$\tilde{C1} : \sum_{g \in G} \sum_{n_g \in N_g} \sum_{k \in K} x_{n_g,k}^t \leq P_{n_g}^{\text{max}}, \tilde{C3} : \sum_{n_g \in N_g} \tilde{R}_{n_g} \geq \bar{R}_{g}^{\text{rvs}} \text{ and } \tilde{C4} : \mu_{n_g} \leq \tilde{R}_{n_g},$$

for all $n_g \in N_g$ and $g \in G$.

Since (4.3) belongs to general stochastic optimization problem, we apply a simple and elegant theory of Lyapunov drift optimization frameworks to convert the average based optimization problem into the single time slot optimization problem which can be solved based on the instantaneous values of CSI, QSI and arrival rate of each user, resulting in much less computational
complexity [63]. From the quadratic Lyapunov function, the optimization problem in (4.3) can be converted into the following optimization problem

\[
\min_{x, w} - V (\sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} \tilde{R}_{n_g}^t) + \sum_{g \in \mathcal{G}} Z_g^t (\tilde{R}_{g}^t - \sum_{n_g \in \mathcal{N}_g} \tilde{R}_{n_g}^t) + \sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} Q_{n_g}^t (\alpha_{n_g}^t - \tilde{R}_{n_g}^t) \tag{4.4}
\]

subject to: \( \tilde{C}_1 - \tilde{C}_2 \),

where \( V > 0 \) and \( Z^t_g \) are trade-off parameter and virtual queue for constraint \( \tilde{C}_3 \), respectively in Lyapunov drift-plus-penalty algorithm. The algorithm to solve (4.4) is presented in Table 4.1. This algorithm contains two major steps: 1) update the queue and virtual queue related parameters, i.e., \( Q^t_{n_g} \) and \( Z^t_g \), based on Q1 and Q2 in Table 4.1, respectively; 2) update the optimal power and sub-carrier allocation vectors.

To derive the optimal power and sub-carrier allocation, we can apply the Lagrange dual decomposition to (4.4) because of its convex nature. The Lagrange function related to (4.4) is

\[
\mathcal{L}(x, w, \lambda, \rho) = - V (\sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} \tilde{R}_{n_g}^t) + \sum_{g \in \mathcal{G}} Z_g^t (\tilde{R}_{g}^t - \sum_{n_g \in \mathcal{N}_g} \tilde{R}_{n_g}^t) + \sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} Q_{n_g}^t (\alpha_{n_g}^t - \tilde{R}_{n_g}^t) + \sum_{k \in \mathcal{K}} \rho_k (\sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} w_{n_g, k}^t - 1) + \lambda (P_{\text{max}} - \sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} \sum_{k \in \mathcal{K}} x_{n_g, k}^t).
\]

From KKT conditions related to (4.4), we obtain the following solution for optimal power allocation

\[
P_{n_g, k}^* = \left[ \frac{V + Z_g^t + Q_{n_g}^t}{\lambda \ln(2)} - \frac{\sigma^2}{h_{n_g, k}} \right] P_{\text{max}}.
\]

For sub-carriers, we obtain the following necessary condition for \( w_{n_g, k} \), [45]

\[
\frac{\partial \mathcal{L}(x, w, \lambda, \rho)}{\partial w_{n_g, k}^*} \left\{ \begin{array}{ll}
< 0, & w_{n_g, k}^* = 0 \\
= 0, & 0 < w_{n_g, k}^* < 1 \quad \forall k \in \mathcal{K}, n_g \in \mathcal{N}_g, \\
> 0, & w_{n_g, k}^* = 1
\end{array} \right.
\]
Table 4.1 Dynamic Resource Provisioning Algorithm

Algorithm 1 QoS-aware Resource Provisioning Algorithm

Initialization: Initialize $V$, $Z = [Z_1, \cdots, Z_G]$ and $Q = Q_{n_g, n_g} \forall g \in \mathcal{G}$, $T_{\text{max}}$ and $i_{\text{max}}$, $\varepsilon = 10^{-3}$.

Step 1: For $t = 1, \cdots, T_{\text{max}}$
- Set $w^*(i = 0) = 1$, $P^*_n (i = 0) = P_{\text{max}} / K$, $\forall n_g \in \mathcal{N}_g, g \in \mathcal{G}$.
- Set arrival rate vector $\alpha = [\alpha_{n_g, n_g}]_{n_g=1}^{N_g} \forall g \in \mathcal{G}$, according to a Poisson random variable.

Step 2: Single shot loop iteration $i$:
- $\lambda(i) = \left[\lambda(i-1) + \delta_n \frac{\partial \mathcal{L}(\mathbf{x}, \mathbf{w}, \lambda, \rho)}{\partial \mathcal{L}} \right]^+$.
- Update power and sub-carrier according to (4.6) and (4.7),
  - Select $n_g, k = \max[s_{n_g, k}^t]$ and set $w_{n_g, k}^t = 1$, otherwise $w_{n_g, k}^t = 0$.
  - When $||(\lambda(i) - \lambda(i-1))|| \leq \varepsilon$, or $i > i_{\text{max}}$, stop, otherwise go to Step 2.

Update:
- $Q_{n_g}^{t+1} = \max[\langle Q_{n_g}^t - R_{n_g}^t, T_t \rangle, 0] + \alpha_{n_g}^t, \forall n_g \in \mathcal{N}_g, \forall g \in \mathcal{G}$.
- $Z_g^{t+1} = \max[Z_g^t + R_{\text{RSV}} - \sum_{n_g \in \mathcal{N}_g} R_{n_g}^t, 0], \forall g \in \mathcal{G}$.
- When $t \geq T_{\text{max}}$, Stop,
- Otherwise, set $t = t + 1$, goto Step 1.

where

\[
\frac{\partial \mathcal{L}(\mathbf{x}, \mathbf{w}, \lambda, \rho)}{\partial \mathcal{L}}(\mathbf{x}, \mathbf{w}, \lambda, \rho) = \rho_k - (V + Z_g^t + Q_{n_g}^t) \left( \log_2(1 + \gamma_{n_g, k}^t) - \frac{\gamma_{n_g, k}^t}{(1 + \gamma_{n_g, k}^t) \ln(2)} \right), \forall g \in \mathcal{G}.
\]

Therefore, we reach to

\[
s_{n_g, k}^t = (V + Z_g^t + Q_{n_g}^t) \left( \log_2(1 + \gamma_{n_g, k}^t) - \frac{\gamma_{n_g, k}^t}{(1 + \gamma_{n_g, k}^t) \ln(2)} \right), \forall g \in \mathcal{G}, \tag{4.7}
\]

where $\gamma_{n_g, k}^t = \frac{x_{n_g, k}^t h_{n_g, k}^t}{\sigma w_{n_g, k}^t}$. The sub-carrier $k$ is allocated to user $n_g$, i.e., $w_{n_g, k}^t = 1$ for which $s_{n_g, k}^t$ is maximum otherwise $w_{n_g, k}^t = 0$. In the inner loop of algorithm in Table 4.1, the power and sub-carriers are allocated according to current CSI in the time slot $t$ for fixed queues, i.e., $Q_{n_g}$ for all users and $Z_g$ for all slices. Then, $Q_{n_g}$ and $Z_g$ are updated according to the allocated power and sub-carriers in the outer loop until $t = T_{\text{max}}$ where $T_{\text{max}}$ is a maximum number of time slots.

Via this algorithm, the solution for time-average-based problem (4.3) is derived from (4.4) which is instantaneously solved at each time slot based on the CSI values of this slot and without considering the average over other time slots. Now, the important question is to declare the
relationship between solutions of (4.3) and (4.4). This relation is discussed in the following theorem.

**Theorem 1:** The Lyapunov quadratic approximation of (4.3) satisfies all the constraints in (4.2) and the obtained rate is less than $R_{\text{opt}}$ by a gap of $C/V$, which decreases with increasing $V$ and $C > 0$ derived based on [56].

**Proof.** Due to power limitation, for this setup, there exist two finite constants $C_1$ and $C_2$ which satisfy

$$\frac{1}{2} \sum_{g \in G} \mathbb{E} \left\{ \sum_{n_g \in N_g} \tilde{R}_{n_g}^t - \tilde{R}_g^{\text{sv}} \right\}^2 \leq C_1,$$

$$\frac{1}{2} \sum_{g \in G} \sum_{n_g \in N_g} \mathbb{E} \{ \mu_{n_g} - \tilde{R}_{n_g}^t \}^2 \leq C_2.$$

Assuming $V \geq 0$ and $Z_{g}^1 = 0, \forall g \in G$, all the desired constraints in (4.3) are satisfied, i.e.,

$$\limsup_{T \to \infty} \frac{1}{T} \sum_{t \in T} \mathbb{E} \left\{ \sum_{n_g \in N_g} \tilde{R}_{n_g}^t - \tilde{R}_g^{\text{sv}} \right\} \leq 0, \forall g \in G,$$

$$\limsup_{T \to \infty} \frac{1}{T} \sum_{t \in T} \mathbb{E} \{ \mu_{n_g} - \tilde{R}_{n_g}^t \} \leq 0, \forall n_g \in N_g, \forall g \in G.$$

Also, if $R_{\text{opt}}$ is the maximum optimal rate, then the average expectation of objective function in (4.3) satisfies

$$\frac{1}{T} \sum_{t \in T} \mathbb{E} \left\{ \sum_{g \in G} \sum_{n_g \in N_g} \tilde{R}_{n_g}^t \right\} \leq R_{\text{opt}} + \frac{C}{V}, \forall t \in \mathcal{T},$$

where $C = C_1 + C_2$. Furthermore, for all time slots $t \in \mathcal{T}$, we have

$$\frac{1}{T} \sum_{t \in T} \mathbb{E} \left\{ \sum_{n_g \in N_g} \tilde{R}_{n_g}^t - \tilde{R}_g^{\text{sv}} \right\} \leq \mathbb{E} \{ Z_g [T + 1]/T \}, \forall g \in G,$$

$$\frac{1}{T} \sum_{t \in T} \mathbb{E} \{ \mu_{n_g} - \tilde{R}_{n_g}^t \} \leq \mathbb{E} \{ Q_{n_g} [T + 1]/T \}, \forall n_g \in N_g, \forall g \in G.$$
and

\[
\sum_{g \in \mathcal{G}} \mathbb{E}\{Z_g[T+1]^2/T^2\} + \sum_{g \in \mathcal{G}} \sum_{n_g \in N_g} \mathbb{E}\{Q_{n_g}[T+1]^2/T^2\} \leq \frac{1}{T}[2C + 2(R^{\text{opt}} - \theta)],
\]  

(4.9)

where \( \theta \) is a finite constant such that \( \mathbb{E}\{\sum_{g \in \mathcal{G}} \sum_{n_g \in N_g} \tilde{R}_{n_g}^t\} \geq \theta, \forall t \in \mathcal{T} \). Such \( \theta \) exists because \( \tilde{R}_{n_g}^t \) is lower bounded [56].

The above statement declares that via the Lyapunov drift approach, when (4.2) is feasible, the solution of (4.4) satisfies the constraints of (4.2), and it is within \( O(1/V) \) of the optimum value of (4.2). Therefore, the time-average based solution can be derived by solution of optimization problem in each frame. The difference between solutions of (4.2) and (4.4), i.e., \( C/V \), can be made arbitrarily small by increasing the value of parameter \( V \). However, increasing the value of \( V \) affects the convergence time, as shown by (4.8) and (4.9). Consequently, \( V \) is a trade-off parameter for convergence time and optimality of the proposed algorithm in Table 4.1 [56].

### 4.4 Numerical Results and Discussions

The simulation setup consists of one BS with \( K = 128 \) sub-carriers which serves \( G = 3 \) slices. Each slice \( g \in \mathcal{G} \) contains \( N_g = 3 \) users. The locations of all the users from different slices are fixed in the range of distance \( D_{n_g} \in \{0.1, 0.6\} \) as shown in the Fig. 4.1. Furthermore, under Rayleigh fading assumption, the channel power gain of user \( n_g \) on sub-carrier \( k \) is modeled as \( h_{n_g,k} = \mathcal{X}_{n_g,k}/D_{n_g}^\beta \), where \( \beta = 3 \) is the path-loss exponent and \( \mathcal{X}_{n_g,k} \) is the exponential random variable for user \( n_g \) on sub-carrier \( k \) with mean one. We set \( P_{\text{max}} = 5 \) dB, \( \sigma = 1 \), \( \tilde{R}_{n_g}^{\text{sv}} = 0.5 \) bps/Hz, and average Poisson packet arrival rate \( \mu_{n_g} = \mu = 1 \) packet/slot with each packet of size \( l = 1 \) bits/Hz/slot, unless otherwise stated. All the simulation results are generated with \( T_{\text{max}} = 5000 \) time slots with each time slot length \( T_t = 1 \) ms for the bandwidth \( B_k = 200 \) KHz. Fig. 4.2 shows the total rate versus \( V \) for various values of \( P_{\text{max}} \). It is observed that the total rate is increased with increasing \( V \). It is expected because the proposed optimization problem (4.2) is solved via Lyapunov drift-plus-penalty approach, which gives asymptotically optimal solution in terms of average rate over time as \( V \) approaches infinity [56]. In other words, the gap between the obtained average rate and optimal average rate reduces to zero with increasing \( V \).

Next, we study the effect of packet size arriving in users’ queue on average total rate. Fig.4.3 illustrates the decrease in total rate with increasing packet size \( l \). This is because the feasibility re-
Fig. 4.1 Locations of users of three slices.

Fig. 4.2 Total Rate versus $V$ and $P_{\text{max}}$. 
region shrinks with increasing \( l \) due to constraint C4 in problem (4.2) which decreases the obtained achievable rate [44]. Furthermore, since increasing \( R_{g}^{\text{HSV}} \) also shrinks the feasibility region due to constraint C3, the total rate decreases with increasing \( R_{g}^{\text{HSV}} \) as depicted in Fig.4.4. However, the
total average rate increases with increasing $V$ for both Fig.4.3 and Fig.4.4, which further confirms our observations in Fig.4.2.

To further demonstrate the convergence of Alg. 1 in terms of queue stability, in Fig.4.5, we plot the queue lengths of farthest users $n_1$ and $m_1$ from slice 1 in addition to user $n_2$ from slice 2 at each time slot. It is observed that the queue lengths of all users remains bounded as time slots increases. In other words, network queue stability condition holds for the VWN, confirming the effectiveness of the proposed algorithm.

### 4.5 Concluding Remarks

In this chapter, we proposed a QoS-oriented dynamic resource provisioning policy for OFDMA-based VWNs which simultaneously satisfies the condition of stable average queue lengths of the users of all slices and the minimum reserved rate of each slice. The solution was proposed by transforming the cross-layer problem into PHY layer problem via Lyapunov drift-plus-penalty approach. Finally, the performed simulations under different system parameters confirmed the effectiveness of the proposed algorithm.
Chapter 5

Delay-Aware and Energy-Efficient Resource Allocation in VWN

This chapter proposes a delay-aware resource provisioning policy for VWNs to minimize the total average transmit power while holding the minimum required average rate of each slice and maximum average packet transmission delay for each user. The proposed cross-layer optimization problem is inherently non-convex and has high computational complexity. To develop an efficient solution, we first transform cross-layer dependent constraints to the physical layer dependent ones. Afterwards, we apply different convexification techniques using variable transformation and relaxation, and propose an iterative algorithm to reach an optimal solution. Simulation results illustrate the effects of the required average packet transmission delay and minimum average slice rate on the total transmission power in VWNs.

5.1 Introduction

Wireless network virtualization has been recently considered as a promising approach to enhance spectrum efficiency via sharing infrastructures among different slices serving their own specific sets of users [65]. Successful sharing requires proper isolation among slices to prevent harmful effects of user activity in one slice on the other slices [3].

Various resource provisioning policies in VWNs have been proposed to provide this isolation either by static resource allocation or dynamic throughput reservations of slices [28]. In [66], the

\[^1\]Parts of this chapter have been presented in [64].
authors propose a Karnaugh-map-like online embedding algorithm for VWN to handle network requests. In [39], the concept of game theory is used to provide dynamic interactions between slices and network operators. An opportunistic spectrum sharing method in wireless virtualization with multiple physical networks is proposed in [67]. The resource allocation schemes in VWNs to maximize the total rate under various QoS requirements of slices are investigated in [40, 49] for OFDMA and massive MIMO-based VWNs.

While the above works focus on the physical-layer parameters to provide the isolation among slices, it is essential to consider the traffic characteristics of end-users, e.g., arrival rate and tolerable delay, to ensure high-quality end-user experience in a VWN. This issue calls for a new constraint related to maximum delay of each user in a resource allocation problem. Additionally, energy efficiency should be considered as an important objective of VWNs for the next generation of wireless networks [3]. In this Chapter, we propose a cross-layer resource provisioning policies suitable for VWNs to minimize the power consumption in consideration of traffic arrival rate and tolerable delay while satisfying the dynamic slice isolation.

We consider an up-link transmission of an OFDMA-based VWN. In this setup, we propose a delay-aware resource allocation that minimizes the total average power of VWNs while holding the minimum average rate of each slice for isolation under the limit of maximum delay in packet transmission. Due to these two constraints, the formulated resource allocation problem has an inherent cross-layer, as well as non-convex nature, and suffers from high computational complexity. To reach a tractable formulation and capture the meaning of tolerable delay from end-users perspectives, we resort to the concept of effective capacity [68–70]. By replacing the delay-aware constraints with more tractable formulations and applying relaxation and transformation techniques, we convexify the formulated problem and develop an efficient iterative algorithm for a solution. Through simulations, we investigate the effects of different parameters on the transmit power of VWN for two user-location scenarios: cell-center and cell-boundary. We show how packet size, arrival rate of users and maximum tolerable delay can affect the total transmit power of VWN.

The rest of this chapter is organized as follows. Section 5.2 introduces the system model and problem formulation. Section 5.3 develops the proposed iterative algorithm for slice provisioning. Section 5.4 presents the simulation results and Section 5.5 concludes the chapter.
5 Delay-Aware and Energy-Efficient Resource Allocation in VWN

5.2 Network Model and Problem Formulation

We consider an up-link OFDMA transmission of a single-cell VWN based on OFDMA scheme, where the BS is virtualized to support a set of $G = \{1, \cdots, G\}$ slices. Each slice $g \in G$ serves a set of $N_g = \{1, \cdots, N_g\}$ users and $N = \sum_{g \in G} N_g$ is the total number of users in the VWN. The total wireless channel bandwidth $B$ of wireless channel is equally divided into $K = \{1, \cdots, K\}$ set of OFDMA sub-carriers. The bandwidth $B_k = B / K$ of each sub-carrier $k \in K$ is assumed to be less than the coherence bandwidth $B_c$. Therefore, the link from user $n_g$ to BS on sub-carrier $k$ exhibits flat fading with channel power gain $h_{n_g,k}$. It is also assumed that the overall channel power gain vector $h = [h_{n_g,k}]_{\forall n_g, \forall g, \forall k}$ has a known stationary and ergodic and known cumulative distribution function (cdf) [71, 72].

Let $w_{n_g,k}$ denote the sub-carrier assignment indicator, where $w_{n_g,k} = 1$ indicates that the sub-carrier $k$ is assigned to user $n_g$, otherwise $w_{n_g,k} = 0$. Then, the OFDMA exclusive sub-carrier assignment policy can be expressed as the following constraint

$$C1: \sum_{g \in G} \sum_{n_g \in N_g} w_{n_g,k} \leq 1 \text{ and } w_{n_g,k} \in \{0, 1\}, \forall k \in K.$$ 

If $p_{n_g,k}$ is the power allocated to user $n_g$ on sub-carrier $k$ in time slot $t$, the corresponding achievable rate is $R_{n_g,k} = \log_2(1 + \frac{p_{n_g,k} h_{n_g,k}}{\sigma})$, where $\sigma$ is the noise power in each sub-carrier and users. Consequently, the total achievable rate of user $n_g$ becomes $R_{n_g}(P, w) = \sum_{k \in K} R_{n_g,k}$, where $P = [p_{n_g,k}]_{\forall n_g, \forall g, \forall k}$ and $w = [w_{n_g,k}]_{\forall n_g, \forall g, \forall k}$ are the allocated power and sub-carrier vector to all users of all slices. To maintain isolation among slices and offer reliable quality-of-service to each user in all slices, we consider following constraints:

- **Slice-Isolation constraints VWN**: by guaranteeing the minimum average rate $R_{rsv}^{g}$ i.e.,

$$C2: E_h \left\{ \sum_{n_g \in N_g} R_{n_g}(P, w) \right\} \geq R_{rsv}^{g}, \forall g \in G,$$

where operator $E_h \{\cdot\}$ represents the expectation over random vector $h$.

- **User QoS constraints**: by keeping the average traffic delay for each user $n_g$ for all $g \in G$
is kept below the predefined threshold, $\mu_{ng}$, i.e., [63]

$$C3: \mu_{ng} E_h \{Q_{ng}\} \leq T_{ng}, \ \forall n_g \in \mathcal{N}_g, \ \forall g \in \mathcal{G},$$

where $Q_{ng}$ and $\mu_{ng}$ denote the queue length and average packet arrival rate at the queue of user $n_g$, respectively.

With considering energy efficiency for VWN under the above two types of constraints, the overall optimization problem can be written as

$$\begin{aligned}
\min_{\mathbf{P}, \mathbf{w}} & \quad E_h \left\{ \sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} \sum_{k \in \mathcal{K}} w_{n_g,k} P_{n_g,k} \right\} \\
\text{subject to} & \quad C1-C3
\end{aligned}$$

(5.1)

subject to C1-C3.

Note that the proposed resource provisioning problem (5.1) contains discrete and continuous variables such as $\mathbf{w}$ and $\mathbf{P}$, and is inherently non-convex optimization problem [21]. Additionally, the constraints from both physical layer (C1 and C2) and MAC layer (C3) add multi dimensional complexity to the optimization problem.

### 5.3 Proposed Algorithm

The constraint C3 contributes to the major complexity of (5.1) due to its relation with cross-layer parameters, i.e., $Q_{ng}$ and $\mu_{ng}$. As a first step to simplify (5.1), we need to find the equivalence of C3 in terms of $\mathbf{P}$ and $\mathbf{w}$ instead of $Q_{ng}$ and $\mu_{ng}$. From the context of effective capacity, C3 can be rewritten as [68]

$$\tilde{C3} : \quad E_h [R_{ng}(\mathbf{P}, \mathbf{w})] \geq Z_{ng}, \ \forall n_g \in \mathcal{N}_g, \ \forall g \in \mathcal{G}$$

where

$$Z_{ng} = \frac{(2T_{ng}\mu_{ng} + 2) + \sqrt{(2T_{ng}\mu_{ng} + 2)^2 - 8T_{ng}\mu_{ng} L_{ng}}}{4T_{ng} L_{ng}}$$

and $L_{ng}$ is the average packet size at the queue of user $n_g$.

In the next step, to get rid of combinatorial structure of problem, we apply the relaxation and re-transformation techniques to (5.1) where we consider $w_{n_g,k} \in [0, 1]$, representing the sub-carrier assignment for the fraction of a time slot [73]. Based on new transformations, C1 is
changed to

\[ \tilde{C}_1 : \sum_{g \in G} \sum_{n_g \in N_g} w_{n_g,k} \leq 1 \text{ and } w_{n_g,k} \in [0, 1], \forall k \in K. \]

Furthermore, in order to convexify (5.1), we consider a new variable \( x_{n_g,k} \) to transform the rate as

\[ \tilde{R}_{n_g}(x, w) = \sum_{k \in K} w_{n_g,k} \log_2 \left( 1 + \frac{x_{n_g,k} h_{n_g,k}}{\sigma w_{n_g,k}} \right) \]

where \( x \) is the vector of \( x_{n_g,k} \) for all users and sub-carriers. The above transformed throughput belongs to the general class of convex function \( f(x, y) = x \log(1 + y/x) \) [21, 52]. Therefore, \( C_2 \) and \( C_3 \) become

\[ \tilde{C}_2 : \mathbb{E}_h \left\{ \sum_{n_g \in N_g} \tilde{R}_{n_g}(x, w) \right\} \geq R_{\text{rsv}}^g, \forall g \in G \text{ and } \]

\[ \tilde{C}_3 : \mathbb{E}_h \left[ \tilde{R}_{n_g}(x, w) \right] \geq Z_{n_g}, \forall n_g \in N_g, \forall g \in G, \]

respectively. Consequently, (5.1) can be rewritten as

\[ \min_{x, w} \mathbb{E}_h \left\{ \sum_{g \in G} \sum_{n_g \in N_g} \sum_{k \in K} x_{n_g,k} \right\} \]

subject to \( \tilde{C}_1 - \tilde{C}_3 \)

Now, since (5.2) is a convex problem, it can be solved via Lagrange dual method [74, 75] and the optimal solution can be achieved for this scenario. The Lagrange function of (5.2) is

\[ \mathcal{L}(\phi_g, \zeta_{n_g}, \rho_k, x, w) = -\mathbb{E}_h \left\{ \sum_{g \in G} \sum_{n_g \in N_g} \sum_{k \in K} x_{n_g,k} \right\} + \sum_{g \in G} \phi_g \left( \mathbb{E}_h \left\{ \sum_{n_g \in N_g} \tilde{R}_{n_g} \right\} - R_{\text{rsv}}^g \right) \]

\[ + \sum_{g \in G} \sum_{n_g \in N_g} \zeta_{n_g} \left( \tilde{R}_{n_g} - Z_{n_g} \right) + \sum_{k \in K} \rho_k \left( 1 - \sum_{g \in G} \sum_{n_g \in N_g} w_{n_g,k} \right), \]

where \( \rho_k \) for all \( k \in K \), \( \phi_g \) for all \( g \in G \) and \( \zeta_{n_g} \) for all \( n_g \in N_g \) are the positive Lagrange
variables of $\tilde{C}_1$, $\tilde{C}_2$ and $\tilde{C}_3$, respectively. Let $\rho$, $\phi$ and $\zeta$ be the vectors of the corresponding Lagrange variables $\rho_k$, $\phi_g$ and $\zeta_n$, respectively. Now, the Dual function related to (5.3) is

$$D(\phi, \zeta, \rho) = \max_{x,w} L(\phi, \zeta, \rho, x, w),$$

and, consequently, the dual problem is

$$\min_{\phi, \zeta, \rho} D(\phi, \zeta, \rho) \quad (5.4)$$

subject to: $\tilde{C}_1 - \tilde{C}_3$.

Due to the convexity of (5.2), the duality gap is zero, i.e., the solution of dual problem is equivalent to the solution of primal problem. By applying the KKT condition to (5.4), the optimal power for user $n_g$ on sub-carrier $k$, $p^*_{n_g,k}$, is

$$p^*_{n_g,k} = \left[ \frac{\phi_g + \zeta_{n_g}}{\ln(2)} - \frac{\sigma}{h_{n_g,k}} \right] p_{\text{max}} 0,$$

(5.5)

where $[x]^a_b = \max\{\min\{x, a\}, b\}$. From KKT conditions, for optimal sub-carrier allocation $w^*_{n_g,k}$, we have,

$$w^*_{n_g,k} = \begin{cases} 0, & \partial L(\phi_g, \zeta_{n_g}, \rho_k, x, w) \frac{\partial}{\partial w^*_{n_g,k}} < 0, \\ \in [0, 1], & \partial L(\phi_g, \zeta_{n_g}, \rho_k, x, w) \frac{\partial}{\partial w^*_{n_g,k}} = 0, \\ 1, & \partial L(\phi_g, \zeta_{n_g}, \rho_k, x, w) \frac{\partial}{\partial w^*_{n_g,k}} > 0. \end{cases}$$

(5.6)

where

$$\frac{\partial L(\phi_g, \zeta_{n_g}, \rho_k, x, w)}{\partial w^*_{n_g,k}} = (\phi_g + \zeta_{n_g}) \left( \log_2(1 + \gamma_{n_g,k}) - \frac{\gamma_{n_g,k}}{(1 + \gamma_{n_g,k}) \ln(2)} \right),$$

and $\gamma_{n_g,k} = \frac{x_{n_g,k}h_{n_g,k}}{\sigma w_{n_g,k}}$. In order to hold the exclusive sub-carrier allocation of OFDMA, the sub-carrier $k$ is allocated to user which satisfy the followings

$$w^*_{n_g,k} = \begin{cases} 1, & n'_g = \max_{n_g \in N_g} \max_{\gamma_g \in G} \frac{\partial L(\phi_g, \zeta_{n_g}, \rho_k, x, w)}{\partial w^*_{n_g,k}}, \\ 0, & n_g \neq n'_g. \end{cases}$$

(5.6)
The iterative algorithm to derive the optimal sub-carrier and power algorithm for (5.2), summarized in Algorithm 2, consists of two phases:

1) **Off-line Phase**: According to a known channel distribution information (CDI) of users, generate different CSIs of VWN. Based on these samples, via gradient descent method, the Lagrange variables $\phi$ and $\zeta$ are updated from

$$
\phi^*_g(j) = \phi_g(j - 1) + \delta_{\phi_g}(\frac{\partial L}{\partial \phi_g})
$$

and

$$
\forall g \in G, \quad \zeta^*_n_g(j) = \zeta_n_g(j - 1) + \delta_{\zeta_n_g}(\frac{\partial L}{\partial \zeta_n_g}) \quad \forall n_g \in N_g \quad \text{and} \quad \forall g \in G, \quad \text{where} \quad 0 < \delta_{\phi_g} \ll 1 \quad \text{and} \quad 0 < \delta_{\zeta_n_g} \ll 1 \quad \text{are small positive step sizes for} \quad \phi_g \quad \text{and} \quad \zeta_n_g, \quad \text{respectively. The update functions can be equivalently rewritten as}
$$

$$
\phi^*_g(j) = \phi_g(j - 1) + \delta_{\phi_g}(E_h \left\{ \sum_{n_g \in N_g} \tilde{R}_{n_g}(x(j - 1), w(j - 1)) \right\} - R_g^{\text{sv}}),
$$

and

$$
\zeta^*_n_g(j) = \zeta_n_g(j - 1) + \delta_{\zeta_n_g}(E_h \left\{ \tilde{R}_{n_g}(x(j - 1), w(j - 1)) \right\} - Z_{n_g}).
$$

The iterative process will be terminated if

$$
\|\phi^*_g(j) - \phi_g(j - 1)\| \leq \varepsilon_1, \quad \text{and} \quad \|\zeta^*_n_g(j) - \zeta_n_g(j - 1)\| \leq \varepsilon_2
$$

where $0 < \varepsilon_1 \ll 1$ and $0 < \varepsilon_2 \ll 1$.

2) **On-line Phase**: Power and sub-carrier are allocated according to (6.9) and (6.10) at each time slot, based on the values derived from off-line phase for fixed CDIs.

When CDI is changed, **Off-line Phase** is executed to update $\phi$ and $\zeta$ [76].

### 5.4 Numerical Results and Discussions

In this section, we evaluate the performance of proposed algorithm via simulation results. We consider a single BS serving $G = 2$ slices with $K = 64$ OFDMA sub-carriers. Each slice has $N_1 = N_2 = 2$ users and $N = 4$ in VWN. Each slot length $T$ is normalized to one. Furthermore,
**Algorithm 2**: Slice Provisioning Algorithm

**Off-line Phase**:

**Initialization**: Set arbitrary values for $j = 0$, $\phi(j = 0)$, $\zeta(j = 0)$, $\delta_{\phi g}(j = 0)$, $\delta_{\zeta ng}(j = 0)$, and $j_{\text{max}} \gg 1$ as the maximum number of iteration for offline phase.

**Repeat**: $j = j + 1$

Update $\phi(j)$ and $\zeta(j)$ according to (5.7) and (5.8).

**Until** (5.9) holds or $j \geq j_{\text{max}}$.

**On-line Phase**:

Update $P$ and $w$ according to (6.9) and (6.10), respectively.

If CDI changes, go to **Off-line Phase**, otherwise, continue **On-line Phase** for next time instance.

---

we set maximum average delay of packet transmission at each user’s queue $T_{n g} = 0.5$, which is normalized to time slot, minimum reserved rate $R_{r sv g_1} = R_{r sv g_2} = 1.0$ (bps/Hz) for slices $g_1$ and $g_2$, and packet arrival process as Poisson process with average 5 packets/slot, i.e., $\mu = \mu_{n g} = 5$ packets/slot for all $n g$ unless otherwise stated. The channel gain fading follows Rayleigh distribution, i.e., $h_{n g,k} = \mathcal{X}D_{n g}^{-\beta}$, where $\beta = 3$ is path loss exponent and $D_{n g}$ is distance of users from BS, and $\mathcal{X}$ is exponentially distributed with mean of one. The simulations are performed for 1000 on-line time slots and off-line parameters are up-dated after every 25 slots. For off-line phase, we set $\varepsilon_1 = \varepsilon_2 = 10^{-3}$.

<table>
<thead>
<tr>
<th>$R_{r sv g}$</th>
<th>$L_{n g}$</th>
<th>2.50 b/Hz</th>
<th>1.50 b/Hz</th>
<th>0.25 b/Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0 bps/Hz</td>
<td>Set 1</td>
<td>Set 2</td>
<td>Set 3</td>
<td></td>
</tr>
<tr>
<td>1.0 bps/Hz</td>
<td>Set 4</td>
<td>Set 5</td>
<td>Set 6</td>
<td></td>
</tr>
<tr>
<td>0.5 bps/Hz</td>
<td>Set 7</td>
<td>Set 8</td>
<td>Set 9</td>
<td></td>
</tr>
</tbody>
</table>

Clearly, by increasing $R_{r sv g}$ and decreasing $T_{n g}$, the transmit power should be increased to hold the related constraints. To understand better their effects on the allocated power, performance of the proposed approaches, we consider two different user-location cases: **Case 1** with users close to cell-center, i.e., $D_{n g} = \{0.35, 0.45\}$; and **Case 2** with users close to cell-boundary, i.e., $D_{n g} = \{0.55, 0.65\}$.

The different sets of system parameters, $R_{r sv g}$ and $L_{n g}$, are summarized in Table 5.1. These two parameters determine the lower bound of $\tilde{C}_2$ and $\tilde{C}_3$, respectively. By increasing these two
parameters, VWN needs more transmit power to satisfy the slice-isolation and/or QoS constraints. Based on our definition, Set 1 has the most stringent constraints while Set 9 has the loosest constraints. Specifically, Sets 1-3 have the strictest $\tilde{C}_2$ with $R_{g1}^{sv} = R_{g2}^{sv} = 2.0$ while Sets 7-9 have relaxed $\tilde{C}_2$ with $R_{g1}^{sv} = 0.5$. Sets 1, 4 and 7 have a larger packet size $L_{n_g} = 2.5$ b/Hz leading to higher $Z_{n_g}$ than Sets 3, 6 and 9 with $L_{n_g} = 0.5$ b/Hz.

Figs. 5.1(a) and 5.1(b) depict the transmit power versus $T_{n_g}$ for Case 1 and Case 2, respec-
Both figures show that for sets 3, 6 and 9 with smallest average packet size $L_{n_g}$, the variations in total transmission power with respect to $T_{n_g}$ is negligible in comparison to the other sets with higher values of $L_{n_g}$. This is because for the small packet size, $T_{n_g}$ does not have a strict effect on $Z_{n_g}$, and hence, the total transmit power is robust against variation in $T_{n_g}$, while for large $L_{n_g}$ (i.e., Sets 1, 4, and 7), $T_{n_g}$ has a strong effect on $Z_{n_g}$, leading to higher transmit power. Furthermore, it is also observed that with increasing $R_y^{sv}$, the transmit power increases to
achieve the required minimum rate of slices. The transmit power of Case 2 is larger than Case 1 to compensate the path-loss.

Figs. 5.2(a) and 5.2(b) depict the transmit power versus $\mu_{n_g}$ for Case 1 and Case 2, respectively, where $R_{g1}^{rvw}$ and $R_{g2}^{rvw}$ are set to 1.0 bps/Hz and $T_{n_g} = 0.5$. Both figures show that with increasing $\mu_{n_g}$, the total transmit power is increased for $L_{n_g} > 0.25$. However, for $L_{n_g} \leq 0.25$, $\mu_{n_g}$ does not have considerable affect on the transmit power. This is because for large values of packet size ($L_{n_g} > 0.25$), $Z_{n_g}$ increases with increasing $\mu_{n_g}$ and hence increases the strictness of constraint $\tilde{C}_3$. As a result, total transmit power requirements of users increases to satisfy the feasibility of $\tilde{C}_3$. On the other side, for small values of packet size ($L_{n_g} \leq 0.25$), $Z_{n_g}$ does not increases significantly with increasing $\mu_{n_g}$ and hence $\tilde{C}_3$ is not affected much. Consequently, for small packet size, increasing $\mu_{n_g}$, negligibly affect the total power transmission of users. Therefore, when $\tilde{C}_3$ is dominant constraint among the other constraints of considered resource allocation problem, increasing $\mu_{n_g}$ affect the feasibility region of resource allocation problem, and hence, the total transmit power is increased. Whereas, for small packet size, $\tilde{C}_3$ is no longer dominate as compared to other constraints of the considered problem and the increment of $\mu_{n_g}$ does not affect the feasibility region.

In summary, Figs. 5.1-5.2 show the effect of channel attenuation, i.e., the users with lower channel gains need higher transmit power. In case that the transmit power of users is limited, the admission control policy, e.g., [40], is required to maintain the performance of networks. Otherwise, even with maximum transmit power by users, none of QoS of users and isolation factor between slices can be satisfied.

Finally, we study the convergence of Lagrange variables in off-line phase for 1000 CSI samples generated with prior CDI information. Fig. 5.3 shows the values of Lagrange multipliers $\phi_g$ (for $\tilde{C}_2$) and $\zeta_{n_g}$ (for $\tilde{C}_3$) for all $g \in G$ and $n_g \in N_g$ versus off-line iterations with $L_{n_g} = 2.5$ b/Hz. All the Lagrange variables converge to a constant value with-in 60 iterations, indicating the effectiveness of Alg 2.

### 5.5 Concluding Remarks

In this work, we proposed a delay-aware resource provisioning policy for VWNs to minimize total transmit power while considering minimum average rates of each slice and maximum average packet transmission delay of each user. Via effective capacity, first, we transformed all the constraints to PHY-layer dependent constraints, and afterwards, convexified the formulated
cross-layer resource allocation problem with relaxation techniques. Finally an iterative algorithm was developed to solve the proposed delay-aware and power-efficient, joint power and sub-carrier allocation problem. The effects of various system parameters on the total power of VWN were investigated via the simulations.
Chapter 6

Energy-Efficient Robust Resource Provisioning in VWN

This chapter proposes a robust resource allocation approach in VWNs to address the uncertainty in CSI at the BS due to estimation error and mobility of users. In this set-up, the resources of an OFDMA-based wireless network are shared among different slices where the minimum reserved rate is considered as the QoS requirement of each slice. We formulate the robust resource allocation problem against the worst-case CSI uncertainty, aiming to maximize the overall energy efficiency (EE) of VWN in terms of a newly defined slice utility function. Uncertain CSI is modeled as the sum of its true estimated value and an error assumed to be bounded in a specific uncertainty region. The formulated problem suffers from two major issues: computational complexity and energy-efficiency degradation due to the considered error in the maximum extent. To deal with these issues, we consider a specific form of uncertainty region to solve the robust resource allocation problem via an iterative algorithm. The simulation results demonstrate the effectiveness of the proposed algorithms\(^1\).

6.1 Introduction

Wireless network virtualization is a promising paradigm to improve the spectrum efficiency and enable service customization among slices belonging to different service providers via introducing abstraction and modularity in wireless networks \([2, 3, 32]\). In a single-cell VWN, dif-

\(^1\)Parts of this chapter have been presented in [77].
different slices can share physical network resources (e.g., BS) and wireless resources (e.g., subcarriers and power) where each slice comprises a set of users, and has its own QoS requirements. Due to the diverse QoS requirements of slices and wireless resource limitations, resource provisioning among slices is challenging and essential, which has drawn a lot of attentions recently [28, 33, 39–42, 49].

Generally, the resource provisioning problems considered in [28, 33, 39–42, 49] are based on a common assumption that the accurate CSI of all users of different slices to a BS is available. Also, some of these works consider the total throughput of a VWN as an objective function, e.g., [33, 40, 42, 49]. However, due to users’ mobility, the stochastic nature of wireless channels and delay in feedback channels, perfect CSI knowledge may not be available in practice. Besides, considering a utility function to investigate the energy efficiency is of high importance for wireless networks [78]. In this chapter, we aim to focus on these two issues as follows. We first introduce a utility function for each slice based on its total rate and its cost of transmit power. We then show that how this per-slice utility function can increase the energy efficiency of VWN.

To immunize the performance of VWN against the uncertainty in the CSI values, we apply the worst-case optimization theory, which has been widely applied in the resource allocation in wireless networks [24–27]. In this context, the uncertain parameter is modeled as an estimated value plus an error that is modeled as a bounded value in the specific region and the performance of network is maximized under the worst condition of error. It is well-known that the worst-case approach is capable of preserving the instantaneous VWN performance against the uncertain parameters, while it suffers from high computational complexity and total throughput reduction due to its conservative view of worst-case error [25, 26].

To deal with the above mentioned issues, we resort to the moderate version of robust optimization theory in which the error is assumed to have joint bounded and stochastic nature [26]. Also, instead of maximizing the throughput under the worst-case error condition, the throughput outage probability is preserved below a predefined threshold. By selecting the appropriate uncertainty region, as well as variable-relaxation and transformation techniques, we convexify the resource allocation problem and propose an efficient two-level iterative solution algorithm.

Simulation results verify the energy efficiency of the proposed robust resource allocation algorithm in VWNs, based on the slice utility function. Specifically, they show how important the cost factor of the utility function is in controlling the EE factor of VWN.

The rest of this chapter is organized as follows. Section 6.2 introduces the system model and problem formulations, followed by Section 6.3, where a solution to the robust problem and an
iterative algorithm are proposed. In Section 6.4, the simulation results are presented. Finally, Section 6.5 concludes the chapter.

## 6.2 Network Model and Problem Formulation

We consider the down-link transmission of OFDMA-based VWN with one central BS serving a set of slices, i.e., $G = \{1, \cdots, G\}$, in which each slice $g \in G$ requires a minimum reserved rate $R_g^{rsv}$. Furthermore, each slice $g \in G$ has a set of users, i.e., $N_g = \{1, \cdots, N_g\}$, where $N_g$ is the total number of users in slice $g$ and $N = \sum_{g \in G} N_g$ represents the total number of users in VWN. Considering the OFDMA scheme, the total bandwidth $B$ is equally divided into a set of sub-carriers, i.e., $K = \{1, \cdots, K\}$, where each sub-carrier bandwidth $B_k = B/K$ is assumed to be small compared to the coherent bandwidth $B_c$ of wireless channel. Thereby, the channel gain $h_{n_g,k}$ of user $n_g$ on sub-carrier $k$ exhibits flat fading.

Let $w_{n_g,k} \in \{0, 1\}$ be the sub-carrier allocation indicator for user $n_g$ on sub-carrier $k$, where $w_{n_g,k} = 1$ indicates that sub-carrier $k$ is assigned to user $n_g$, and otherwise $w_{n_g,k} = 0$. Via exclusive orthogonal sub-carrier assignment imposed by OFDMA implementation issue, we have

$$C1: \sum_{g \in G} \sum_{n_g \in N_g} w_{n_g,k} \leq 1, \quad \forall k \in K,$$

which means that each sub-carrier is allocated to maximum one user. Due to the transmit power limitation of BS, we have

$$C2: \sum_{g \in G} \sum_{n_g \in N_g} \sum_{k \in K} w_{n_g,k} p_{n_g,k} \leq P_{\text{max}},$$

where $p_{n_g,k}$ and $P_{\text{max}}$ are the allocated power to user $n_g$ over sub-carrier $k$ and maximum transmit power of BS, respectively. Therefore, the rate of user $n_g \in N_g$ is

$$R_{n_g}(P, w) = \sum_{k \in K} w_{n_g,k} \log_2 \left( 1 + \frac{p_{n_g,k} h_{n_g,k}}{\sigma} \right),$$

where $P = [p_{n_g,k}]_{n_g,g,k}$ and $w = [w_{n_g,k}]_{n_g,g,k}$ are the allocated power vector and the sub-carrier assignment vector of all users, respectively. The constraint on the minimum rate reserved for
each slice $g \in \mathcal{G}$ is represented as

$$C_3 : \sum_{n_g \in \mathcal{N}_g} R_{n_g}(P, w) \geq R_g^{rsv}, \quad \forall g \in \mathcal{G}.$$ 

For energy-efficient design of VWN, we consider the following slice utility function, $\forall g \in \mathcal{G}$,

$$U_g(P, w) = \sum_{n_g \in \mathcal{N}_g} R_{n_g}(P, w) - C_g^{E} \sum_{n_g \in \mathcal{N}_g} \sum_{k \in \mathcal{K}} w_{n_g,k} P_{n_g,k},$$

where the energy-cost coefficient of slice $g \in \mathcal{G}$, $C_g^{E}$ provides the trade-off between its achieved throughput and its power consumption. Aiming to maximize the sum utility of all slices, while satisfying the minimum required slice rates, the nominal VWN optimization problem is

$$\max_{P, w} \sum_{g \in \mathcal{G}} U_g(P, w), \quad (6.1)$$

subject to: $C_1 - C_3$.

In (6.1), perfect CSI knowledge is assumed. However, in practice, due to delay in feedback channel, user mobility, and error in the estimation, such CSI knowledge can be imperfect. To deal with this issue, we consider the uncertainty in CSI at the BS and introduce a robust counterpart of the above resource allocation problem.

The imperfect CSI is modeled as the sum of its estimated value and an additive error i.e.,

$$\tilde{h}_{n_g} = \bar{h}_{n_g} + \hat{h}_{n_g}, \quad \forall n_g \in \mathcal{N}_g, \ g \in \mathcal{G},$$

where $\tilde{h}_{n_g} = [h_{n_g,k}]_{\forall k}$ is the $1 \times K$ uncertain CSI vector, and, $\bar{h}_{n_g} = [\bar{h}_{n_g,k}]_{\forall k}$ and $\hat{h}_{n_g} = [\hat{h}_{n_g,k}]_{\forall k}$ are, respectively, the $1 \times K$ estimated CSI and error vectors of user $n_g$. In the context of worst-case robust optimization, the errors on the estimated values are trapped in the bounded region, called uncertainty region, defined as

$$\mathcal{E}_{n_g} = \{\tilde{h}_{n_g} \mid \|\tilde{h}_{n_g} - \bar{h}_{n_g}\| \leq \epsilon_{n_g}\}, \forall n_g \in \mathcal{N}_g, \forall g \in \mathcal{G},$$

where $\epsilon_{n_g} \geq 0$ is the uncertainty bound, assumed to be small, and $\|\mathbf{x}\|$ denotes the norm function of vector $\mathbf{x}$ [76].

The effect of uncertainty on $\tilde{h}_{n_g}$ can be represented by a new vector of variables in the
throughput of each user. Let \( \hat{R}_{ng} \) denotes the throughput of user \( ng \) in the robust resource allocation, which depends on \( h = [\tilde{h}_{ng}]_{g,n} \). When the uncertainty region shrinks to zero (i.e., \( \epsilon_{ng} = 0 \)), the total throughput of the nominal and robust optimization problems are identical i.e.,

\[
R_{ng}(P, w) = \hat{R}_{ng}(P, w, h)|_{\epsilon_{ng}=0}, \quad \forall ng \in \mathcal{N}_g.
\]

The objective of the worst-case approach is to find the optimal transmit power and sub-carrier allocation for each user that optimize their total throughput under the worst condition of error in the uncertainty region. In this approach, the robust VWN resource allocation problem based on (6.1) becomes [26]

\[
\max_{P, w} \sum_{g \in \mathcal{G}} \tilde{U}_g(P, w, h),
\]

subject to : C1 – C3,

where \( \tilde{U}_g(P, w, h) \) is the robust counter part of the utility \( U_g \), mathematically expressed as

\[
\tilde{U}_g(P, w, h) = \sum_{ng \in \mathcal{N}_g} \min_{\tilde{h}_{ng} \in \tilde{\mathcal{E}}_{ng}} \hat{R}_{ng}(P, w, h) - C_E \sum_{k \in \mathcal{K}} w_{ng,k} p_{ng,k}
\]

In general, solving the robust counterpart (6.2) involves high computational complexity, because, in addition to the inherent computational complexity from (6.1), it has a new set of optimization variables with uncertain parameters, i.e., \( \tilde{h}_{ng} \). To reduce the computational complexity of (6.2), we treat each \( \tilde{h}_{ng,k} \) as a bounded random variable. Then, we demonstrate how the inner minimization over \( \tilde{h}_{ng} \) is solved. Interestingly, we will also show that the proposed reformulation provides a trade-off between performance and robustness.

### 6.3 Robust EE Resource Provisioning Algorithm

The direct way to solve (6.2) is to obtain the inner minimization analytically, and then, solve the outer maximization, either numerically or analytically [76]. In the following, we will show how the inner and outer optimization problems can be solved.
6.3.1 Inner Optimization Problem

The inner optimization problem of (6.2) is

$$\min_{\tilde{h}_{ng} \in \mathcal{E}_{ng}} \hat{R}_{ng}(P, w, \hat{h}), \quad \forall n_g \in \mathcal{N}_g, \; g \in \mathcal{G}. \quad (6.3)$$

For general definition of norm function of $\mathcal{E}_{ng}$, a closed-form expression of $\tilde{h}_{ng}$ cannot be obtained for a given values of $P$ and $w$ for (6.3). To simplify (6.3), following the same argument as in [76, 79, 80], we assume that $\tilde{h}_{ng,k}$ for all $n_g \in \mathcal{N}_g$ and $k \in \mathcal{K}$ are i.i.d. random variables with the probability distribution function (pdf) of $f(\hat{h}_{ng,k})$. In this case, the uncertainty region is transformed into $\hat{h}_{ng,k} \in [-\varepsilon_{ng,k}, \varepsilon_{ng,k}]$, where $\varepsilon_{ng,k}$ is the bound of uncertainty region for user $n_g$ on sub-carrier $k$. Now, by utilizing the pdf of $\hat{h}_{ng,k}$, the inner optimization problem of (6.3) is transformed into [79]

$$\min_t \sum_{k \in \mathcal{K}} w_{ng,k} t_{ng,k}, \quad (6.4)$$

subject to:

$$C4: \quad \Pr \left\{ \log_2 \left( 1 + \frac{p_{ng,k} h_{ng,k}}{\sigma} \right) < t_{ng,k} \right\} > \eta_{ng,k},$$

$$C5: \quad \hat{h}_{ng,k} \in [-\varepsilon_{ng,k}, \varepsilon_{ng,k}], \quad \forall k \in \mathcal{K}, \; \forall n_g \in \mathcal{N}_g,$$

where $t = [t_{ng,k}]_{\forall n_g,g,k}$ and $t_{ng,k} \geq 0$ is an auxiliary variable for this transformation. Also, $0 < \eta_{ng,k} < 1$ is the probability factor against the uncertain parameters. $C4$ can be simplified to $t_{ng,k} > \log_2(1 + \frac{p_{ng,k} F^{-1}(\eta_{ng,k})}{\sigma})$ for all $k \in \mathcal{K}$ and $n_g \in \mathcal{N}_g$. If $F(\hat{h}_{ng,k})$ has a uniform distribution over the interval $[-\varepsilon_{ng,k}, \varepsilon_{ng,k}]$, we have $F^{-1}(\eta_{ng,k}) = 2\varepsilon_{ng,k} \eta_{ng,k} + \tilde{h}_{ng,k} - \varepsilon_{ng,k}$. Therefore, the solution of (6.4) for all $n_g$ and $k$ is

$$\tilde{h}_{ng,k} = 2\varepsilon_{ng,k} \eta_{ng,k} + \tilde{h}_{ng,k} - \varepsilon_{ng,k}, \quad (6.5)$$

for all $k \in \mathcal{K}$ and $n_g \in \mathcal{N}_g$. From (6.5), for $0.5 \leq \eta_{ng,k} < 1$, $\tilde{h}_{ng,k} \leq \hat{h}_{ng,k}$, and is in-line with the concept of worst-case robust optimization, in which the error is at its own maximum extent. For this, we will focus on this case in the rest of this chapter.
The throughput of each user with uncertainty can be rewritten as

$$\hat{R}_{ng}(P, w) = \sum_{k \in K} w_{ng,k} \log_2 (1 + \frac{p_{ng,k} \tilde{h}_{ng,k}}{\sigma})$$

for all $ng \in N_g$ and $g \in G$. Therefore, (6.2) is simplified to

$$\max_{P, w} \sum_{g \in G} \hat{U}_g(P, w),$$

subject to: $C1 - C2$ and $\hat{C}3$,

where

$$\hat{U}_g(P, w) = \sum_{ng \in N_g} \hat{R}_{ng}(P, w) - C^{E}_g \sum_{ng \in N_g} \sum_{k \in K} w_{ng,k} p_{ng,k}$$

and

$$\hat{C}3: \sum_{ng \in N_g} \hat{R}_{ng}(P, w) \geq R^{sv}_g, \forall g \in G.$$
Furthermore, we consider a new variable $x_{n_g,k} = w_{n_g,k}P_{n_g,k}$, which transforms $\tilde{R}_{n_g}(P, w)$ to

$$\tilde{R}_{n_g}(x, w) = \sum_{k \in \mathcal{K}} w_{n_g,k} \log_2(1 + \frac{x_{n_g,k} \tilde{h}_{n_g,k}}{\sigma w_{n_g,k}}), \forall n_g \in \mathcal{N}_g.$$ 

Therefore, the utility function is simplified to

$$\tilde{U}_g(x, w) = \sum_{n_g \in \mathcal{N}_g} \tilde{R}_{n_g}(x, w) - C^E_g \sum_{n_g \in \mathcal{N}_g} \sum_{k \in \mathcal{K}} x_{n_g,k}.$$ 

In this context, $C_2$ and $\tilde{C}_3$ are transformed into

$$\tilde{C}_2 : \sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} \sum_{k \in \mathcal{K}} x_{n_g,k} \leq P^{\text{max}} \quad \text{and}$$

$$\tilde{C}_3 : \sum_{n_g \in \mathcal{N}_g} \tilde{R}_{n_g}(x, w) \geq R^{\text{rsv}}_g, \forall g \in \mathcal{G},$$

respectively. Since $\tilde{R}_{n_g}(x, w)$ belongs to the class of convex functions represented as $f(x, y) = x \log_2(1 + \frac{y}{x}) \forall x, y \geq 0$ [43]. Therefore, the convexified robust counterpart of (6.1) is

$$\max_{x, w} \sum_{g \in \mathcal{G}} \tilde{U}_g(x, w),$$  \hspace{1cm} (6.7)

subject to : $\tilde{C}_1, \tilde{C}_2$ and $\tilde{C}_3$.

Now, we can solve (6.7) by solving the dual optimization problem and applying KKT conditions. Let $\rho_k, \lambda$ and $\phi_g$ represent the Lagrange multipliers for constraints $\tilde{C}_1, \tilde{C}_2$ and $\tilde{C}_3$, respectively. Therefore, the Lagrange function for (6.7) is

$$L(w, x, \lambda, \phi, \rho) = -\sum_{g \in \mathcal{G}} \tilde{U}_g + \lambda(\sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} \sum_{k \in \mathcal{K}} x_{n_g,k} - P^{\text{max}})$$  \hspace{1cm} (6.8)

$$+ \sum_{g \in \mathcal{G}} \phi_g(R^{\text{rsv}}_g - \sum_{n_g \in \mathcal{N}_g} \tilde{R}_{n_g}) + \sum_{k \in \mathcal{K}} \rho_k(\sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} w_{n_g,k} - 1).$$

Applying KKT conditions to (6.8), we obtain the optimal power solution of (6.7) $\forall k \in \mathcal{K}$,
Algorithm 3: Robust Slice Provisioning

Initialization: Set \( w^*(l = 0) = 1 \), \( P^*_{n_g}(l = 0) = P_{\text{max}} / K \), \( \forall n_g \in N_g, g \in G, l = 0, l_{\text{max}}, \)

\( i_{\text{max}} \) and \( 0 < \zeta_m \ll 1 \) for \( m = \{1, 2, 3\} \).

OL: Repeat \( l = l + 1 \):

\[
\lambda(l) = \left[ \lambda(l - 1) + \delta \frac{\partial \zeta}{\partial \lambda} \right]^+, \\
\phi_g(l) = \left[ \phi_g(l - 1) + \delta \frac{\partial \phi_g}{\partial \phi_g} \right]^+, \quad \forall g \in G.
\]

IL: Repeat \( i = i + 1 \):

Update \( P^*(i) \) according to (6.9).

Update \( w^*(i) \) according to (6.10).

Until \( (||P(i) - P(i - 1)|| \leq \zeta_1) \) or \( i > i_{\text{max}} \).

Until \( (||\phi_g(l) - \phi_g(l - 1)|| \leq \zeta_2 \) and \( ||\lambda(l) - \lambda(l - 1)|| \leq \zeta_3 \) or \( l > l_{\text{max}} \)).

\( n_g \in N_g \) and \( g \in G \) as

\[
p_{n_g,k} = \left[ \frac{1 + \phi_g}{\ln(2)(\lambda + CE_g)} - \frac{\sigma}{\hat{h}_{n_g,k}} \right]^{P_{\text{max}}}_{0}. \quad (6.9)
\]

In order to obtain the solution for sub-carrier allocation, we obtain the following necessary condition for \( w_{n_g,k} \) for all \( k \in K \) and \( n_g \in N_g \)

\[
w_{n_g,k}^* = \begin{cases} 
0, & \frac{\partial L(w, x, \lambda, \phi, \rho)}{\partial w_{n_g,k}} < 0, \\
[0, 1], & \frac{\partial L(w, x, \lambda, \phi, \rho)}{\partial w_{n_g,k}} = 0, \\
1, & \frac{\partial L(w, x, \lambda, \phi, \rho)}{\partial w_{n_g,k}} > 0,
\end{cases}
\]

where [45]

\[
\frac{\partial L(w, x, \lambda, \phi, \rho)}{\partial w_{n_g,k}^*} = (1 + \phi_g) \left( \log_2(1 + \gamma_{n_g,k}) - \frac{\gamma_{n_g,k}}{(1 + \gamma_{n_g,k}) \ln(2)} \right).
\]

For holding the exclusive sub-carrier allocation of OFDMA, the sub-carrier \( k \) is allocated to user which satisfy the followings

\[
w_{n_g,k}^* = \begin{cases} 
1, & n_g' = \max_{n_g \neq n_g} \frac{\partial L(w, x, \lambda, \phi, \rho)}{\partial w_{n_g,k}^*}, \\
0, & n_g \neq n_g',
\end{cases} \quad (6.10)
\]
The iterative algorithm to allocate the optimal power and sub-carrier with uncertain CSI is presented in Algorithm 3. It starts with initialization of variables followed by an outer loop where Lagrange variables $\lambda$ and $\phi_g$ are updated for all $g \in G$ via gradient method, where $0 < \delta_x \ll 1$ is the step size for Lagrange variable $x$. In the inner loop, the power and sub-carriers are computed from the updated values of Lagrange variables. The iterative processes are stopped when power and sub-carrier converge to the constant values.

### 6.4 Numerical Results and Discussions

In this section, we investigate the proposed solution of the resource provisioning problem (6.2) via simulation results. For simulation settings, we consider two slices $g_1$ and $g_2 \in G$ where BS has $K = 64$ sub-carriers and $\sigma = 1$. The CSI is derived from Rayleigh fading distribution, modeled as $h_{n_g,k} = \mathcal{X}D_{n_g}^{-\beta}$, where $\beta = 4$ is the path loss exponent, $\mathcal{X}$ is exponential random variable with mean one, and $D_{n_g}$ is the distance of user $n_g$ from BS. For all the simulations, we set the minimum reserved rate $R_{g}^{\text{rsv}} = 1.0$ bps/Hz for each slice, $\eta_{n_g,k} = 0.9$, $C_g^{\text{E}} = 3.0$, $P_{\text{max}} = 15$ dB, and $\varepsilon_{n_g,k} = \varepsilon = 0.3$, $\forall n_g \in N_g$ and $\forall k \in K$ unless otherwise stated. For the simulations in Figs. 6.1 and 6.2, all the users of slices are randomly located in the range of distance $D_{n_g} \in \{0.2, 0.5\}$. All the plotted results are obtained from the average of over 100 CSI realizations. To demonstrate the results, we define the energy efficiency (EE) factor as $\text{EE} = \sum_{g \in G} \sum_{n_g \in N_g} R_{n_g}(P, \mathbf{w})/(\sum_{g \in G} \sum_{n_g \in N_g} \sum_{k \in K} P_{n_g,k} + P_c)$ where $P_c = -10$ dB is the constant signal processing power required at the BS [81].

Fig. 6.1 illustrates the total EE factor versus number of users $N$ for different values of $P_{\text{max}}$. The EE increases with increasing $N$ for all considered $P_{\text{max}}$ due to the multi-user diversity gain, which increases the total rate leading to higher energy efficiency. From Fig. 6.2, increasing $K$ also increases the EE factor. This is because that VWN has more options to assign sub-carriers with better channel gains to the users with increasing $K$. Thus, the total rate of VWN and hence EE would be increased. Figs. 6.1 and 6.2 indicate that the total EE reduces with increasing $P_{\text{max}}$. This happens because higher power cannot help to increase throughput (due to the tradeoff between throughput and power cost in the defined utility) as much as required to compensate the power increase in the denominator of EE factor.

To analyze the behavior of EE with respect to $\varepsilon_{n_g,k}$, we consider two scenarios based on the locations of users, 1) **Users at the cell-edge** or low-SNR scenario where $D_{n_g} \in [0.4, 0.6]$ for all $n_g$, and 2) **Users at the cell-center** or high-SNR scenario where $D_{n_g} \in [0.2, 0.3]$ for all $n_g$. It
Fig. 6.1 EE factor versus number of users $N$

Fig. 6.2 EE versus number of subcarriers $K$
Fig. 6.3  EE factor versus $\varepsilon$ and $C_g^E$ for cell-center users scenario

Fig. 6.4  EE factor versus $\varepsilon$ and $C_g^E$ for cell-edge users scenario
can be observed that increasing $\varepsilon$ decreases the EE factor due to the conservative feature of the worst-case approach, where error is considered to the maximum extent. From both Figs. 6.3 and 6.4, increasing $C^E_g$ increases the EE factor since for higher value of price, the VWN consumes less power. Consequently, EE is increased. Moreover, EE factor in Fig. 6.3 is higher than that in Fig. 6.4. This is because when users are located at the boarder of the cell, the rate of VWN decreases because of limited transmit power and large scale fading. Consequently, EE factor decreases.

6.5 Concluding Remarks

In this chapter, we proposed the robust resource provisioning policy for OFDMA-based VWNs, aiming to maximize the total energy efficiency of a network while satisfying the minimum rate requirements of all the slices. The non-convex problem was transformed into a convex one by applying the appropriate selection of uncertainty region, variable transformations and relaxations. Based on the solution of the convexified problem, an iterative algorithm was developed. Via simulation results, the effects of system parameters, including error in CSI, on total EE factor of VWN was investigated.
Chapter 7

Conclusion

7.1 Summary

This thesis has proposed resource provisioning policies for VWNs to support different QoS requirements of slices. In this context, various algorithms have been proposed to satisfy the QoS requirements of slices, such as minimum reserved rates and resources of slices, as well as stable queue state and maximum packet transmission delay.

In Chapter 2, we have considered the QoS requirements of two groups of slices in VWN: rate-based groups, which requires minimum reserved rates and resource-based groups, which requires a minimum number of sub-carriers and fraction of total transmission power. In this context, an optimal power and sub-carrier allocation algorithm for slices has been developed. Via simulation results, we have observed that with increasingly stringent QoS requirements of slices under the limited wireless resources, the probability of service outage due to infeasibility increases in VWNs. To deal with this issue, a dynamic admission control policy has been introduced, where the minimum requirements of slices have been adjusted according to a slice priority factor and channel conditions. The simulation results have successfully demonstrated that admission control variables dynamically check the infeasible QoS parameters of slices while considering the slice priority factor. Therefore, an effective admission control in VWNs is essential to not only prevent the over-utilization of wireless resources but also to achieve slice isolation.

Chapter 3 has introduced a resource provisioning scheme for massive MIMO aided VWN, where a massive MIMO setup has been utilized to improve the QoS of VWN via expanding its feasibility region. Simulation results have successfully demonstrated the expansion of the feasible region of VWNs due to a massive MIMO setup via supporting higher rate requirements.
of slices. Furthermore, results have shown that the maximum rate value is observed for a specific range of pilot duration in imperfect CSI scenarios.

Chapter 4 has addressed the issue of performance degradation due to random arrival of data in users’ queues with limited size in realistic VWNs. Specifically, a stable queue constraint of each user has been considered in addition to average reserved rates of slices in the proposed cross-layer resource provisioning problem. The simulation results have shown that the feasibility region of VWNs shrinks with increasing packet size of users and reserved rates of slices because of limited VWN resources. This indicates that the performance of VWN with queue stability constraint degrades when network has larger packet size or strict slice isolation requirements. Moreover, the results have demonstrated the bounded states of users’ queues over a complete transmission frame, indicating the stable queue state of the system.

In Chapter 5, we have introduced another cross-layer resource provisioning policy to improve end-users’ experience in energy-efficient VWNs. In this context, a slice provisioning algorithm for power and sub-carrier allocation is developed to satisfy the maximum average transmission delay requirements of users in VWNs. Simulation results have demonstrated that the stringent QoS requirements of slices in VWNs, such as delay-based requirements, are power expensive to satisfy due to the random wireless channel and limited wireless resources. In other words, with increasing strictness in delay-based constraints, VWNs require higher transmission power to operate in the feasible region.

Finally, Chapter 6 has introduced a robust resource provisioning policy for energy-efficient VWNs. By considering the energy-cost coefficients in the utility function and worst case robust optimization theory, we have developed a robust slice provisioning algorithm for energy-efficient VWNs. The simulation results have demonstrated the positive effects of multi-user diversity on energy efficiency of VWNs. Moreover, it has been observed that the energy efficiency in VWNs suffers with increasing error in CSI estimation, improving when users are located close to the BS.

## 7.2 Suggested Future Works

As noted in the thesis, wireless virtualization is a new concept and requires a great deal of work. Although, this thesis has addressed various questions related to the QoS based resource provisioning policies in VWN, there are still issues, which require further investigation in future works as discussed below.

- From the numerical results of Chapter 3, it can be observed that for the imperfect CSI case,
the rate is maximized for a specific value of pilot duration. It will be interesting to find the optimal value of pilot duration to maximize the overall rate of the system through a resource provisioning problem in VWNs.

- In Chapter 5, the packet transmission delay constraint is an average constraint over a transmission frame. Although challenging, it will be good to study the effects of instantaneous packet transmission delay requirements in the resource provisioning problem for VWNs in future works.

- All the resource provisioning problems proposed in this thesis are designed for a single-cellular VWNs. Future works can consider the multi-cellular VWN scenarios and investigate the effect of inter-cell interferences on resource provisioning policies. For instance, a joint BS, power and sub-carrier allocation policy is an interesting problem for multicellular VWN.

- As wireless virtualization offers a lot of flexibility when integrated with cloud-RAN, an interesting resource provisioning policy for cloud-RAN based VWN can be investigated in future works. The problem of user allocations to the cloud-RAN units (base band processing units and remote radio heads) under the constraints of limited front haul capacity can be investigated.
References


References


