ACTA

C

TECHNICA

Pradeep Chathuranga Weeraddana

OPTIMIZATION TECHNIQUES FOR RADIO RESOURCE MANAGEMENT IN WIRELESS COMMUNICATION NETWORKS

UNIVERSITY OF OULU,
FACULTY OF TECHNOLOGY,
DEPARTMENT OF COMMUNICATIONS ENGINEERING;
CENTRE FOR WIRELESS COMMUNICATIONS;
INFOTECH OULU

PRADEEP CHATHURANGA WEERADDANA

OPTIMIZATION TECHNIQUES FOR RADIO RESOURCE MANAGEMENT IN WIRELESS COMMUNICATION NETWORKS

Academic dissertation to be presented, with the assent of the Faculty of Technology of the University of Oulu, for public defence in Auditorium IT115, Linnanmaa, on 2 December 2011, at 9 a.m.

Weeraddana, Pradeep Chathuranga, Optimization techniques for radio resource management in wireless communication networks.
University of Oulu, Faculty of Technology, Department of Communications Engineering; Centre for Wireless Communications; Infotech Oulu, P.O. Box 4500, FI-90014 University of Oulu, Finland
Acta Univ. Oul. C 402, 2011

Abstract

The application of optimization techniques for resource management in wireless communication networks is considered in this thesis. It is understood that a wide variety of resource management problems of recent interest, including power/rate control, link scheduling, cross-layer control, network utility maximization, beamformer design of multiple-input multiple-output networks, and many others are directly or indirectly reliant on the general weighted sum-rate maximization (WSRMax) problem. Thus, in this dissertation a greater emphasis is placed on the WSRMax problem, which is known to be NP-hard.

A general method, based on the *branch and bound technique*, is developed, which solves globally the nonconvex WSRMax problem with an optimality certificate. Efficient analytic bounding techniques are derived as well. More broadly, the proposed method is not restricted to WSRMax. It can also be used to maximize *any* system performance metric, which is Lipschitz continuous and increasing on signal-to-interference-plus-noise ratio. The method can be used to find the optimum performance of any network design method, which relies on WSRMax, and therefore it is also useful for evaluating the performance loss encountered by any heuristic algorithm. The considered link-interference model is general enough to accommodate a wide range of network topologies with various node capabilities, such as singlepacket transmission, multipacket transmission, simultaneous transmission and reception, and many others.

Since global methods become slow in large-scale problems, fast local optimization methods for the WSRMax problem are also developed. First, a general multicommodity, multichannel wireless multihop network where all receivers perform singleuser detection is considered. Algorithms based on *homotopy methods* and *complementary geometric programming* are developed for WSRMax. They are able to exploit efficiently the available multichannel diversity. The proposed algorithm, based on homotopy methods, handles efficiently the self interference problem that arises when a node transmits and receives simultaneously in the same frequency band. This is very important, since the use of supplementary combinatorial constraints to prevent simultaneous transmissions and receptions of any node is circumvented. In addition, the algorithm together with the considered interference model, provide a mechanism for evaluating the gains when the network nodes employ self interference cancelation techniques with different degrees of accuracy. Next, a similar multicommodity wireless multihop network is considered, but all receivers perform multiuser detection. Solutions for the WSRMax problem are obtained by imposing additional constraints, such as that only one node can transmit to others at a time or that only one node can receive from others at a time. The WSRMax problem of downlink OFDMA systems is also considered. A fast algorithm based on *primal decomposition techniques* is developed to jointly optimize the multiuser subcarrier assignment and power allocation to maximize the weighted sum-rate (WSR). Numerical results show that the proposed algorithm converges faster than Lagrange relaxation based methods.

Finally, a distributed algorithm for WSRMax is derived in multiple-input single-output multicell downlink systems. The proposed method is based on classical *primal decomposition methods* and *subgradient methods*. It does not rely on zero forcing beamforming or high signal-to-interference-plus-noise ratio approximation like many other distributed variants. The algorithm essentially involves coordinating many local subproblems (one for each base station) to resolve the inter-cell interference such that the WSR is maximized. The numerical results show that significant gains can be achieved by only a small amount of message passing between the coordinating base stations, though the global optimality of the solution cannot be guaranteed.

Keywords: distributed optimization methods, global (nonconvex) optimization methods, mathematical optimization, radio resource management, weighted sum-rate maximization

Weeraddana, Pradeep Chathuranga, Optimointitekniikoita radioresurssien hallintaan langattomissa tiedonsiirtoverkoissa.
Oulun yliopisto, Teknillinen tiedekunta, Tietoliikennetekniikan osasto, Centre for Wireless Communications; Infotech Oulu, PL 4500, 90014 Oulun yliopisto
Acta Univ. Oul. C 402, 2011

Tiivistelmä

Tässä työssä tutkitaan optimointimenetelmien käyttöä resurssienhallintaan langattomissa tiedonsiirtover-koissa. Monet ajankohtaiset resurssienhallintaongelmat, kuten esimerkiksi tehonsäätö, datanopeuden säätö, radiolinkkien ajastus, protokollakerrosten välinen optimointi, verkon hyötyfunktion maksimointi ja keilan-muodostus moniantenniverkoissa, liittyvät joko suoraan tai epäsuorasti painotetun summadatanopeuden maksimointiongelmaan (weighted sum-rate maximization, WSRMax). Tästä syystä tämä työ keskittyy eri-tyisesti WSRMax-ongelmaan, joka on tunnetusti NP-kova.

Työssä kehitetään yleinen *branch and bound* -tekniikkaan perustuva menetelmä, joka ratkaisee epäkon-veksin WSRMax-ongelman globaalisti ja tuottaa todistuksen ratkaisun optimaalisuudesta. Työssä johdetaan myös tehokkaita analyyttisiä suorituskykyrajojen laskentatekniikoita. Ehdotetun menetelmän käyttö ei rajoitu vain WSRMax-ongelmaan, vaan sitä voidaan soveltaa *minkä tahansa* suorituskykymetriikan maksi-mointiin, kunhan se on Lipschitz-jatkuva ja kasvava signaali-häiriö-plus-kohinasuhteen funktiona. Mene-telmää voidaan käyttää minkä tahansa WSRMax-ongelmaan perustuvan verkkosuunnittelumenetelmän optimaalisen suorituskyvyn määrittämiseen, ja siksi sitä voidaan hyödyntää myös minkä tahansa heuristi-sen algoritmin aiheuttaman suorituskykytappion arvioimiseen. Tutkittava linkki-häiriömalli on riittävän yleinen monien erilaisten verkkotopologioiden ja verkkosolmujen kyvykkyyksien mallintamiseen, kuten esimerkiksi yhden tai useamman datapaketin siirtoon sekä yhtäaikaiseen lähetykseen ja vastaanottoon.

Koska globaalit menetelmät ovat hitaita suurien ongelmien ratkaisussa, työssä kehitetään WSRMax-ongelmalle myös nopeita paikallisia optimointimenetelmiä. Ensiksi käsitellään yleistä useaa eri yhteyspal-velua tukevaa monikanavaista langatonta monihyppyverkkoa, jossa kaikki vastaanottimet suorittavat yhden käyttäjän ilmaisun, ja kehitetään algoritmeja, joiden perustana ovat *homotopiamenetelmät* ja *komplemen-taarinen geometrinen optimointi*. Ne hyödyntävät tehokkaasti saatavilla olevan monikanavadiversiteetin. Esitetty homotopiamenetelmiin perustuva algoritmi käsittelee tehokkaasti itsehäiriöongelman, joka syntyy, kun laite lähettää ja vastaanottaa samanaikaisesti samalla taajuuskaistalla. Tämä on tärkeää, koska näin voi-daan välttää lisäehtojen käyttö yhtäaikaisen lähetyksen ja vastaanoton estämiseksi. Lisäksi algoritmi yhdes-sä tutkittavan häiriömallin kanssa auttaa arvioimaan, paljonko etua saadaan, kun laitteet käyttävät itsehäiri-ön poistomenetelmiä erilaisilla tarkkuuksilla. Seuraavaksi tutkitaan vastaavaa langatonta monihyppyverk-koa, jossa kaikki vastaanottimet suorittavat monen käyttäjän ilmaisun. Ratkaisuja WSRMax-ongelmalle saadaan asettamalla lisäehtoja, kuten että vain yksi lähetin kerrallaan voi lähettää tai että vain yksi vastaan-otin kerrallaan voi vastaanottaa. Edelleen tutkitaan WSRMax-ongelmaa laskevalla siirtotiellä OFDMA-jär-jestelmässä, ja johdetaan *primaalihajotelmaan* perustuva nopea algoritmi, joka yhteisoptimoi monen käyt-täjän alikantoaalto- ja tehoallokaation maksimoiden painotetun summadatanopeuden. Numeeriset tulokset osoittavat, että esitetty algoritmi suppenee nopeammin kuin Lagrangen relaksaatioon perustuvat menetel-mät.

Lopuksi johdetaan hajautettu algoritmi WSRMax-ongelmalle monisoluisissa moniantennilähetystä käyttävissä järjestelmissä laskevaa siirtotietä varten. Esitetty menetelmä perustuu klassisiin *primaalihajo-telma-* ja *aligradienttimenetelmiin*. Se ei turvaudu nollaanpakotus-keilanmuodostukseen tai korkean signaa-li-häiriö-plus-kohinasuhteen approksimaatioon, kuten monet muut hajautetut muunnelmat. Algoritmi koor-dinoi monta paikallista aliongelmaa (yhden kutakin tukiasemaa kohti) ratkaistakseen solujen välisen häiri-ön siten, että WSR maksimoituu. Numeeriset tulokset osoittavat, että merkittävää etua saadaan jo vähäisel-lä yhdessä toimivien tukiasemien välisellä viestinvaihdolla, vaikka globaalisti optimaalista ratkaisua ei voi-dakaan taata.

Asiasanat: globaalit (epäkonveksit) optimointimenetelmät, hajautetut optimointimenetelmät, matemaattinen optimointi, painotetun summadatanopeuden maksimointi, radioresurssien hallinta

To my wife

Preface

Abbreviations

Roman-letter notations

a_{jc}	Channel gain from the BS to the jth user in subcarrier c in a single cell OFDMA downlink system
$\mathbf{a}_i$	Closest vertex to $\boldsymbol{\gamma}_{\min}$ (in the ith dimension) of the axis aligned rectangle $\mathcal{Q}$
$\bar{\mathbf{a}}_i$	Closest vertex to $\boldsymbol{\gamma}_{\min}$ (in the ith dimension) of the axis aligned rectangle $\bar{\mathcal{Q}}^\star$
$\mathbf{A}(\cdot)$	Square matrix obtained by using $\mathbf{B}(\boldsymbol{\gamma})$ and $\mathbf{G}$, i.e., $\mathbf{A}(\boldsymbol{\gamma}) = \mathbf{I} - \mathbf{B}(\boldsymbol{\gamma})\mathbf{G}$
$\mathcal{A}$	Set of link pairs (i, j) for which the transmitter of the ith link and the receiver of the jth link coincide
b_{jc}	Normalized power gain from the BS to the jth user in subcarrier c in a single cell OFDMA downlink system
$\mathbf{b}(\cdot)$	Vector obtained by using $\mathbf{B}(\boldsymbol{\gamma})$ and σ^2, i.e., $\mathbf{b}(\boldsymbol{\gamma}) = \sigma^2 \mathbf{B}(\boldsymbol{\gamma})\mathbf{1}$
B	Dynamic cross-layer control algorithm parameter used to characterize the performance of the algorithm
$\mathbf{B}(\cdot)$	Diagonal matric containing the normalized SINR values
$\bar{\mathbf{B}}_i(\cdot)$	Square matrix obtained from $\mathbf{B}(\boldsymbol{\gamma})$ by eliminating the ith row and the ith column
$\mathcal{B}_k$	Set of axis-aligned rectangles at the kth iteration of the BB algorithm for WSRMax
c_{ij}	Small-scale fading coefficient between the transmitter of link i to the receiver of link j; an argument t can be used to indicate the time slot or the fading realization index
$\mathbf{c}_{jl}$	Vector of small-scale fading coefficients between the transmitter of data stream j and the receiver of data stream l in a multicell MIMO Downlink system
C	Number of orthogonal channels or subcarriers
C1	Condition 1 requited to ensure the convergence of the BB algorithm for WSRMax

C2	Condition 2 requited to ensure the convergence of the BB algorithm for WSRMax
$\mathcal{C}$	Set of orthogonal channels or subcarriers
$\mathcal{C}_j$	Set of subcarriers allocated to the jth user
$\acute{\mathcal{C}}$	Single fading realization, which is a set of arbitrarily generated fading coefficients; an argument t can be used to indicate the time slot or the fading realization index
d_0	Far field reference distance used in modeling the power gain between distinct nodes in general wireless networks
d_{ij}	Distance from the transmitter of the ith link (data stream) to the receiver of the jth link (data stream)
d_s	Destination node of commodity s
D	Lipschitz constant of function f_0
D_0	A common distance used in modeling the SNR operating point in a wireless network
D	Node distance matrix
$\mathcal{D}$	Any subset of $L-$dimensional nonnegative orthant
$\mathcal{D}(n)$	Set of links which are decoded at node n
$f_0(\cdot)$	Function used to represent the $(-)$ WSR in the case of singlecast networks
$f_n(\cdot)$	Function used to represent the solution of the nth subproblem (or BS optimization problem) in the algorithm for WSRMax in a multicell downlink system
$f_c^{\star}(\cdot)$	Function used to represent the solution of cth subproblem in the algorithm for WSRMax in the OFDMA downlink
$\check{f}_n(\cdot)$	Function used to establish an upper bound on $f_n(\cdot)$
$\tilde{f}(\cdot)$	Extended value extension of function $f_0(\cdot)$
$\tilde{f}_0(\cdot)$	Function used to represent the $(-)$ WSR in the case of multicast networks
g	Residual self-interference gain after a certain self interference cancelation technique was employed at the network's nodes
$g_{ijc}(t)$	Interference coefficient from the transmitter of ith link (data stream) to the receiver of jth link (data stream) in channel c during time slot t if $i \neq j$; the power gain of ith link (data stream) in channel c during time slot t if $i = j$; the time slot index t is

	dropped sometimes for simplicity; in the single-channel case the channel index c is dropped
$g_l(\cdot)$	Objective function of the optimization problem, which is used to compute the improved lower bound function of the BB algorithm for WSRMax in singlecast networks
$\bar{g}_l(\cdot)$	Objective function of the auxiliary optimization problem, which is used to compute the improved lower bound function of the BB algorithm for WSRMax in singlecast networks
$\mathbf{G}$	Cross-coupling matrix
$\mathbf{G}(t)$	Interference coefficient and power gain matrix; in a single-channel case $[\mathbf{G}(t)]_{i,j} = g_{ij}(t)$
$\bar{\mathbf{G}}_i$	Square matrix obtained from $\mathbf{G}$ by eliminating the ith row and the ith column
$\mathcal{G}$	Set of achievable SINR vectors in a singlecast network
$\breve{\mathcal{G}}$	Set obtained from $\mathcal{G}$ by eliminating certain dimensions of $\mathcal{G}$
$\tilde{\mathcal{G}}$	Set of achievable SINR vectors in a multicast network
$h(\cdot)$	Function used to illustrate the behavior of the algorithm for WSRMax in the OFDMA downlink
$h_{ijc}(t)$	Channel gain from the transmitter of link i to the receiver of link j in channel c during time slot t; the time slot index t is dropped sometimes for simplicity; in a single-channel case the channel index c is dropped
$\mathbf{h}_{jl}$	Channel matrix between transmitter of data stream j and the receiver of data stream l in a multicell MIMO downlink system
$\mathcal{H}$	Set used in the derivation of an upper bound on $f_n(\cdot)$
$\breve{\mathcal{H}}$	Set used in the derivation of an upper bound on $f_n(\cdot)$
(i,j)	Ordered pair of distinct nodes i and j
$\mathcal{I}(n)$	Set of links incoming to the nth node
J	Number of users
L	Number of links or data streams
L_k	Lower bound for $(-)$ WSR at the kth iteration of the BB algorithm for WSRMax
$\mathcal{L}$	Set of links or data streams
$\mathcal{L}(n)$	Set of data streams transmitted by nth BS

$\mathcal{L}_{\text{int}}$	Set of data streams that are subject to out-of-cell interference in a multicell MIMO downlink system
$\mathcal{L}_{\text{int}}(n)$	Set of data streams for which base station n acts as an out-of-cell interferer in a multicell MIMO downlink system
$\mathcal{L}_{\text{local}}(n)$	Subset of data streams transmitted by nth BS, which are not interfered by any out-of-cell interference in a multicell MIMO downlink system
M_n	Number of multicast transmissions from the nth node
n_{rand}	Number of arbitrarily chosen initialization points in the parallel version of the algorithm for WSRMax in the OFDMA downlink
N	Number of nodes (in the case of general networks) or base stations (in the case of cellular networks)
N_l	Power spectral density of the noise at the receiver of link l
$\mathcal{N}$	Set of nodes (in the case of general networks) or base stations (in the case of cellular networks)
$\mathcal{N}_{\text{int}}(l)$	Set of out-of-cell interfering BSs of the receiver node of the lth data stream in a multicell MIMO downlink system; the set is determined by the distance between BSs and the receiver node l
$O(\cdot)$	Big O notation
$\mathcal{O}(n)$	Set of links outgoing from the nth node
$\mathcal{O}^m(n)$	Set of links associated with the mth multicast transmission of the nth node
$p_{lc}(t)$	Power allocated to the lth link (data stream) in channel c during time slot t in the case of multichannel singlecast networks; the time slot index t is dropped sometimes for simplicity; in a single-channel case the channel index c is dropped
$p_0^{\max}$	Transmit sum power constraint common to all nodes
p_n^m	Power allocated to the mth multicast transmission of the nth node in the case of multicast networks
$p_n^{\max}$	Transmit sum power constraint of the nth node
$\mathbf{p}$	Vector of p_l or p_n^m, i.e., $\mathbf{p} = (p_1, \ldots, p_L)$ for singlecast networks and $\mathbf{p} = (p_n^m)_{n \in \mathcal{T}, m=1,\ldots,M_n}$ for multicast networks; an argument t is used sometimes to indicate the time index
$\bar{\mathbf{p}}_i$	Vector obtained from $\mathbf{p}$ by eliminating the ith component
P_ξ	Probability of missing the global optimum; a performance metric

14

	used to compare different algorithms for WSRMax in the OFDMA downlink
$\mathbf{P}(t)$	Overall power allocation matrix during time slot t in the case of multichannel singlecast networks; $[\mathbf{P}(t)]_{l,c} = p_{lc}(t)$
$\mathcal{P}$	Feasible set of the subproblems of the general WSRMax in the OFDMA downlink
$\bar{\mathcal{P}}$	Feasible set of the master problem of the general WSRMax in the OFDMA downlink
q_{jc}	Power allocated at the BS to the jth user in subcarrier c
$q_n^s(t)$	Backlog of commodity s data stored at the nth node during time slot t
$\bar{q}_c$	Power allocated at the BS for subcarrier c
$\mathcal{Q}$	$L-$dimensional axis-aligned rectangle inside $\mathcal{Q}_{\mathrm{init}}$
$\tilde{\mathcal{Q}}$	Axis-aligned rectangle inside $\tilde{\mathcal{Q}}_{\mathrm{init}}$
$\mathcal{Q}_{\mathrm{init}}$	$L-$dimensional axis-aligned rectangle, which is used in the initialization of the BB algorithm for WSRMax in singlecast networks
$\tilde{\mathcal{Q}}_{\mathrm{init}}$	Axis-aligned rectangle, which is used in the initialization of the BB algorithm for WSRMax in multicast networks
$\bar{\mathcal{Q}}^\star$	Smallest $L-$dimensional axis-aligned rectangle, which is used to compute the improved lower bound function of the BB algorithm for WSRMax in singlecast networks
$r_l(t)$	Achievable rate of the lth link (data stream) during time slot t; the time index t is dropped sometimes for simplicity
r_n^m	Maximum achievable rate of all links in the mth multicast transmission of the nth node
$\mathbf{r}$	Vector of r_l or r_n^m used in the generic formulation of WSRMax problem; in the case of singlecast networks, $\mathbf{r} = (r_1, \ldots, r_L)$ and in the case of multicast networks, $\mathbf{r} = (r_n^m)_{n \in \mathcal{T}, m=1,\ldots,M_n}$
$R_n^{\max}$	Dynamic cross-layer control algorithm parameter used to control the burstiness of data delivered to the network layer at node n
$\mathcal{R}$	Set of receiver nodes
$\mathcal{R}^{\mathrm{AVE-MC}}(\cdot)$	Average rate region in the case of multicast networks
$\mathcal{R}^{\mathrm{AVE-SC}}(\cdot)$	Average rate region in the case of singlecast networks
$\mathcal{R}^{\mathrm{INS-MC}}(\cdot)$	Instantaneous rate region in the case of multicast networks
$\mathcal{R}^{\mathrm{INS-SC}}(\cdot)$	Instantaneous rate region in the case of singlecast networks

$\mathcal{R}^{\mathrm{DIR-MC}}(\cdot)$	Directly achievable rate region in the case of multicast networks
$\mathcal{R}^{\mathrm{DIR-SC}}(\cdot)$	Directly achievable rate region in the case of singlecast networks
$\mathcal{R}^{\mathrm{SIC}}(\cdot)$	Achievable rate region in a network with multiuser detectors at nodes
$\bar{\mathcal{R}}$	Set used to represent a set of achievable link rates in the generic formulation of the WSRMax problem
s_l	Information symbol associated with the lth data stream
S	Number of commodities
$\mathcal{S}_n$	Set of commodities which arrives exogenously at the nth node
$\acute{T}$	Number of fading realizations or time slots considered in averaging
$\mathcal{T}$	Set of transmitter nodes
$u_i(\cdot)$	Constraint function used in the convex reformulation of the WSRMax problem over the capacity region of scalar broadcast channel
$u_n^s(\cdot)$	Nondecreasing utility function representing the "reward" received by sending data commodity s from node n to the destination node of commodity s at a given long term average rate
U_k	Upper bound for $(-)$ WSR at the kth iteration of BB algorithm for WSRMax
$\mathbf{U}$	Upper triangular matrix used in the convex reformulation of the WSRMax problem over the capacity region of scalar broadcast channel
v_{lc}	Auxiliary variable used to reformulate the WSRMax problem as a CGP
$\mathbf{v}_l$	Transmit beamformer associated to the lth data stream
$\acute{\mathbf{v}}_i$	ith component of $\acute{\mathcal{V}}$
V	Dynamic cross-layer control algorithm parameter used to characterize the performance of the algorithm
$\acute{\mathcal{V}}$	Set of vertices of the outer polyblock approximation that can be used to further improve $\phi_{\mathrm{lb}}^{\mathrm{Imp}}(\cdot)$
w_{nl}	Out-of-cell interference power from nth BS to the receiver of data stream l in a multicell MIMO downlink system
$\mathbf{w}$	Vector of w_{nl} in a multicell MIMO downlink system, i.e., $\mathbf{w} = (w_{nl})_{n \in \mathcal{N}, l \in \mathcal{L}_{\mathrm{int}}(n)}$
W_c	Bandwidth of channel c

16

x_{jc}	Transmitted signal from the BS to the jth user in subcarrier c in a single cell OFDMA downlink system
$\mathbf{x}_n$	Transmitted signal vector from nth BS in a multicell MIMO downlink system
$x_n^s(t)$	Amount of data commodity s admitted in the network at the nth node during time slot t
$\bar{x}_n^s$	long term average rate of data commodity s admitted in the network at the nth node
y_{jc}	Received signal by the jth user in subcarrier c in a single cell OFDMA downlink system
y_l	Received signal by the receiver node of data stream l in a multicell MIMO downlink system
z_{jc}	Received noise by the jth user in subcarrier c in a single cell OFDMA downlink system
z_l	Received noise by the receiver node of data stream l in a multicell MIMO downlink system

Greek-letter notations

$\beta_l(t)$	Positive weight associated with the lth link (data stream, user) during time slot t in the case of singlecast networks; the time slot index t is dropped sometimes for simplicity
β_n^m	Positive weight associated with the mth multicast transmission of the nth node in the case of multicast networks
$\boldsymbol{\beta}$	Vector of β_l or β_n^m used in the generic formulation of the WSRMax problem; in the case of singlecast networks, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_L)$ and in the case of multicast networks, $\boldsymbol{\beta} = (\beta_n^m)_{n \in \mathcal{T}, m=1,\dots,M_n}$
γ_{lc}	(Auxiliary) variable associated to the SINR of the lth link (data stream) in channel c in the case of multichannel singlecast networks; in single-channel case the channel index c is dropped
$\gamma_{l,\max}$	lth component of $\boldsymbol{\gamma}_{\max}$
$\gamma_{l,\min}$	lth component of $\boldsymbol{\gamma}_{\min}$
γ_n^m	Auxiliary variable associated to the SINRs of links belonging to the mth multicast transmission of the nth node in a multicast network

$\hat{\gamma}_{lc}$	Component at the lth row the cth column of $\hat{\boldsymbol{\gamma}}$, i.e., $[\hat{\boldsymbol{\gamma}}]_{l,c}$; in a single-channel case the channel index c is dropped; a superscript is used sometimes to indicate the algorithm iteration index
$\bar{\gamma}_l^{\star}$	lth component of $\bar{\boldsymbol{\gamma}}^{\star}$
$\boldsymbol{\gamma}$	Vector of γ_l or γ_n^m, i.e., $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_L)$ for singlecast networks and $\boldsymbol{\gamma} = (\gamma_n^m)_{n \in \mathcal{T}, m=1,\ldots,M_n}$ for multicast networks
$\boldsymbol{\gamma}_{\max}$	Maximum element of the axis-aligned rectangle $\mathcal{Q}$ w.r.t the generalized inequality $\preceq_{\mathbb{R}_+^n}$
$\boldsymbol{\gamma}_{\min}$	Minimum element of the axis-aligned rectangle $\mathcal{Q}$ (as well as $\bar{\mathcal{Q}}^{\star}$) w.r.t the generalized inequality $\preceq_{\mathbb{R}_+^n}$
$\breve{\boldsymbol{\gamma}}_{\min}$	Vector obtained from $\boldsymbol{\gamma}_{\min}$ by eliminating certain components of $\boldsymbol{\gamma}_{\min}$
$\bar{\boldsymbol{\gamma}}_{\min,i}$	Vector obtained from $\boldsymbol{\gamma}_{\min}$ by eliminating the ith component
$\bar{\boldsymbol{\gamma}}^{\star}$	Maximum element of the axis-aligned rectangle $\bar{\mathcal{Q}}^{\star}$ w.r.t the generalized inequality $\preceq_{\mathbb{R}_+^n}$
$\hat{\boldsymbol{\gamma}}$	Initial SINR guess for GP, CGP, or SP based algorithms; a superscript is used sometimes to indicate the algorithm iteration index
ΔD_{WSR}	Average normalized weighted-sum-rate deviation; a performance metric used to compare different algorithms for WSRMax in the OFDMA downlink
ϵ	Accuracy required for the BB method for WSRMax
ε	Accuracy required for CGP based algorithms
η	Path loss exponent used in modeling the power gain between distinct nodes in general wireless networks
$\lambda_{\max}$	Scalar used in the dual decomposition based algorithm for WSRMax in the OFDMA downlink
$\lambda_{\min}$	Scalar used in the dual decomposition based algorithm for WSRMax in the OFDMA downlink
$\boldsymbol{\Lambda}$	Network layer capacity region
μ	Interference coupling index used in modeling the power gain between distinct nodes in simple bipartite networks
ξ	Accuracy used in the definition of P_{ξ}
$\boldsymbol{\pi}_n$	Arbitrary permutation of links in $\mathcal{D}(n)$, which determines the decoding and the cancellation order of signals received at node n

$\rho(\cdot)$	Spectral radius of a given square matrix
$\boldsymbol{\rho}_n$	Specific permutation of links in $\mathcal{O}(n)$
$\boldsymbol{\varrho}_n$	Specific permutation of links in $\mathcal{I}(n)$
σ^2	Power of thermal noise at a receiver
σ_l^2	Power of thermal noise at the receiver of the lth link (data stream)
$\phi_{\text{lb}}(\cdot)$	Lower bound function of the BB algorithm for WSRMax
$\phi_{\text{lb}}^{\text{Basic}}(\cdot)$	Basic lower bound function of the BB algorithm for WSRMax
$\phi_{\text{lb}}^{\text{Imp}}(\cdot)$	Improved lower bound function of the BB algorithm for WSRMax
$\phi_{\text{ub}}(\cdot)$	Upper bound function of the BB algorithm for WSRMax
$\phi_{\text{ub}}^{\text{Basic}}(\cdot)$	Basic upper bound function of the BB algorithm for WSRMax
$\phi_{\text{ub}}^{\text{Imp}}(\cdot)$	Improved upper bound function of the BB algorithm for WSRMax
$\phi_{\text{ub}}^{\text{ImpCGP}}(\cdot)$	Improved upper bound function of the BB algorithm for WSRMax obtained by using CGP
$\phi_{\text{min}}(\cdot)$	Function that returns the optimum $(-)$ WSR over a given axis-aligned rectangle $\mathcal{Q}$

Mathematical Operator notations and symbols

$\text{cond}(\mathcal{Q})$	Condition number of the axis-aligned rectangle $\mathcal{Q}$
$\deg(n)$	Degree of the nth node, i.e., the sum of incoming and outgoing links at the nth node
$\text{rec}(l)$	Receiver node of the lth link (data stream)
$\text{size}(\mathcal{Q})$	Maximum half length of the sides of the axis-aligned rectangle $\mathcal{Q}$
$\text{SINR}_n^{ml}(\cdot)$	SINR of the lth link belongs to the mth multicast transmission of the nth node in a multicast network
$\text{tran}(l)$	Transmitter node of the lth link (data stream)
$\text{vol}(\mathcal{Q})$	Volume of the axis-aligned rectangle $\mathcal{Q}$
$\text{conv}(\mathcal{X})$	Convex hull of the set $\mathcal{X}$
$\text{diag}(\mathbf{x})$	Diagonal matrix with the elements of vector $\mathbf{x}$ on the main diagonal
$\text{E}\{\cdot\}$	Expectation
$\inf(\cdot)$	Largest lower bound (infimum)
$\log(\cdot)$	Logarithm in base 2
$\ln(\cdot)$	Natural logarithm
$\max(\cdot)$	Maximum
$\min(\cdot)$	Minimum

Prob$\{\cdot\}$	Probability of the event
rank($\mathbf{X}$)	Rank of matrix $\mathbf{X}$
vec($\cdot$)	Vec-operator; if $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_n]$, then vec($\mathbf{X}$) = $[\mathbf{x}_1^{\mathrm{T}}, \ldots, \mathbf{x}_n^{\mathrm{T}}]^{\mathrm{T}}$
$\mathbf{e}_i$	ith standard unit vector
$\mathbf{I}$	Identity matrix; size of the matrix is implicit
$\mathbf{1}$	Vector of 1s, i.e., $(1, \ldots, 1)$; size of the vector is implicit
$|x|$	Absolute value of the complex number x
$(x)^+$	Positive part of scalar x, i.e., $\max(0, x)$
$\|\mathbf{x}\|_2$	ℓ_2-norm of complex vector $\mathbf{x}$
$[\mathbf{x}]_n$	nth component of vector $\mathbf{x}$
$\mathbf{X}^{\mathrm{H}}$	Conjugate transpose (Hermtian) of matrix $\mathbf{X}$
$\mathbf{X}^{\mathrm{T}}$	Transpose of matrix $\mathbf{X}$
$\mathbf{X}^{-1}$	Inverse of matrix $\mathbf{X}$
$[\mathbf{X}]_{i,j}$	Element at the ith row and the jth column of matrix $\mathbf{X}$
$|\mathcal{X}|$	Cardinality of set $\mathcal{X}$
$\nabla f(\mathbf{x})$	Gradient of function f at $\mathbf{x}$
$\nabla^2 f(\mathbf{x})$	Hessian matrix of function f at $\mathbf{x}$
$\mathbb{C}$	Set of complex numbers
$\mathbb{C}^n$	Set of complex n-vectors
$\mathbb{C}^{m \times n}$	Set of complex $m \times n$ matrices
$\mathbb{R}, \mathbb{R}_+, \mathbb{R}_{++}$	Set of real, nonnegative real, and positive real numbers
$\mathbb{R}_+^n$	Cone of nonnegative, real n-vectors (the set of nonnegative and real n-vectors)
$\mathbb{R}_+^{m \times n}$	Set of nonnegative, real $m \times n$ matrices
$\mathbb{R}_{++}^n$	Set of positive, real n-vectors
$\mathbb{R}_{++}^{m \times n}$	Set of positive, real $m \times n$ matrices
$\mathbb{R}^n$	Set of real n-vectors
$\mathbb{R}^{m \times n}$	Set of real $m \times n$ matrices
$\mathcal{CN}(m, K)$	Circularly symmetric complex Gaussian distribution with mean m and variance $K/2$ per dimension
$\mathcal{CN}(\mathbf{m}, \mathbf{K})$	Circularly symmetric complex Gaussian vector distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{K}$
$\mathcal{U}(a, b)$	Uniform distribution with mean $(a + b)/2$
$(\cdot)^{\star}$	Solution of an optimization problem
$\sim$	Distributed according to

$\geq$	Greater than or equal operator; between real matrices, it represents componentwise inequality

Acronyms

BB	Branch and Bound
BLSLA	Base Line Single Link Activation
BS	Base Station
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CGP	Complementary Geometric Programming
CPU	Central Processing Unit
DC	Difference of Convex
DSL	Digital Subscriber Lines
FDMA	Frequency Division Multiple Access
GP	Geometric Programming
IC	Interference Channel
LB_{Basic}	Lower Bound (Basic)
LB_{Imp}	Lower Bound (Improved)
LHS	Left-Hand Side
LP	Linear Programming
MC	MultiCast
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
MWM	MaxWeight Matching
NUM	Network Utility Maximization
OFDMA	Orthogonal Frequency Division Multiple Access
RA	Resource Allocation
RHS	Right-Hand Side
SC	SingleCast
SIC	Successive Interference Cancellation
SDMA	Space Division Multiple Access
SINR	Signal-To-Interference-Plus-Noise Ratio
SNR	Signal-To-Noise Ratio
SOCP	Second-Order Cone Programming
SP	Signomial Programming
s.t.	Such That

TDMA	Time Division Multiple Access
UB$_{\text{Basic}}$	Upper Bound (Basic)
UB$_{\text{Imp}}$	Upper Bound (Improved)
UB$_{\text{ImpCGP}}$	Upper Bound (Improved by using CGP)
w.r.t	With Respect To
WSR	Weighted Sum-Rate
WSRMax	Weighted Sum-Rate Maximization

Contents

1   Introduction

During the past few years, there has been an explosive growth of the wireless mobile community and its accelerated demand for additional functionalities and services, including multimedia applications, real-time-gaming applications, wireless internet access for realtime video and music, mobile social networks, and many others [1, 2]. Unfortunately, realizing such a growth in demand is indeed challenging in the presence of scarce/expensive radio resources, such as spectrum, and inescapable constraints, such as channel capacity, delay requirements, quality of service requirements, interference requirements, and many others. Therefore, sophisticated radio resource management strategies for wireless communication networks are of paramount importance and are to be designed carefully. To handle such problems, mathematical optimization is an increasingly important tool, which provides general frameworks and systematic guidelines.

Many resource management problems of recent interest in wireless communication can be posed in the framework of mathematical optimization [3–8]. The focus of this dissertation is to apply optimization techniques for resource management in wireless communication networks. In particular, a greater emphasis is placed on the general weighted sum-rate maximization (WSRMax) problem for a set of interfering links. The generic WSRMax problem is

$$\text{maximize} \quad \boldsymbol{\beta}^{\mathrm{T}}\mathbf{r}$$
$$\text{subject to} \quad \mathbf{r} \in \bar{\mathcal{R}} \;,$$

with the variable $\mathbf{r}$, i.e., the vector of achievable rates, where $\boldsymbol{\beta}$ is the vector of positive weights and $\bar{\mathcal{R}}$ is a set of achievable link rates, which depends on factors such as transmission strategy, reception strategy, available radio resources, noise, and many others. In general, the WSRMax problem above is NP-hard.

1.1   Motivation

Among various resource management policies, the WSRMax for an arbitrary set of interfering links plays a central role in many network control and optimization methods. For example, the problem is encountered in power/rate allocation in wireless, as well as in wireline networks [9–11], MaxWeight link

scheduling in multihop wireless networks [12], finding achievable rate regions of singlecast/multicast wireless networks [13, 14], joint optimization of transmit beamforming patterns, transmit powers, and link activations in multiple-input multiple-output (MIMO) networks [15], the resource allocation (RA) subproblem in various cross-layer control policies [16–18] and network utility maximization (NUM) [19], and dynamic power and subcarrier assignment in orthogonal frequency division multiple access (OFDMA) networks [20, 21], among others. Thus, WSRMax appears to be a central component in many network design problems.

Unfortunately, the general WSRMax problem is not yet amendable to a convex formulation [22]. In fact, it is NP-hard [23]. Therefore, we must rely on global optimization approaches [24, 25] for computing an exact solution of the WSRMax problem. Such global solution methods are increasingly important since they can be used to provide performance benchmarks by back-substituting them into any network design method, which relies on WSRMax. They are also very useful for evaluating the performance loss encountered by any heuristic algorithm for the WSRMax problem.

Though global methods find the solution of the WSRMax problem, they are typically slow. Even small problems, with a few tens of variables, can take a very long time to solve WSRMax. Therefore, it is natural to seek suboptimal algorithms for WSRMax that are efficient enough, and still close to optimal; the compromise is optimality [3]. Such algorithms for the WSRMax problem are of central importance since they can be fast and widely applicable in large-scale network control and optimization methods.

Due to the explosion of problem size and the signal overhead required in centralized network control and optimization methods, it is highly desirable to develop decentralized variants of those algorithms. Therefore, finding distributed methods for the WSRMax problem is of crucial importance from a theoretical, as well as from a practical perspective for decentralized implementation of many network control and optimization methods, such as [12, 18].

1.2   Review of earlier and parallel work

Interference is inherent in wireless networks when multiple transmitters and receivers operate over a shared medium, e.g., in spatial-TDMA networks [26]

28

or code division multiple access networks [9]. A similar kind of interference also arises in wireline networks due to electromagnetic coupling between the transmitted signals over wires which are closely bundled, e.g., in digital subscriber lines (DSL) [11]. Due to interference, the achievable rates on different links are interdependent, i.e., the achievable rate of a particular link depends on the powers allocated to all other links. In general, this coupling makes many network control and optimization problems extremely difficult to solve [22, 23]. In many such problems, we see that the WSRMax problem holds an essential role. In the sequel, we first discuss diverse application domains where the WSRMax problem arises directly or indirectly.

1.2.1 Diverse application domains of WSRMax

Network utility maximization (NUM)

In the late nineties, Kelly et. al. [27, 28] introduced the concept of NUM for fairness control in wireline networks. It was shown therein that maximizing the sum-rate under the fairness constraint is equivalent to maximizing certain network utility functions and different network utility functions can be mapped to different fairness criteria. Many aspects of the NUM concept in the case of wireless networks have been substantially discussed in [19, 29–32]. In this context, the WSRMax problem appears as a part of the Lagrange dual problem of the overall NUM problem [33–37].

Cross-layer control policies for wireless networks

A number of papers discuss variants of cross-layer control policies, such as [16–18, 38–48]. Many of these policies are essentially identical. It has been shown that an optimal cross-layer control policy, which achieves data rates arbitrarily close to the optimal operating point, can be decomposed into three subproblems that are normally associated with different network layers. Specifically, *flow control* resides at the transport layer, *routing and in-node scheduling* [1] resides at the network layer, and *resource allocation* (or RA) is usually associated with the

---

[1]*in-node scheduling* refers to selecting the appropriate commodity and it is not to be confused with the links scheduling mechanism which is handled by the resource allocation subproblem [17].

medium access control and physical layers [44]. The first two subproblems are convex optimization problems and can be solved relatively easily. It turns out that under reasonably mild assumptions, the RA subproblem can be cast as a general WSRMax problem over the instantaneous achievable rate region [16, 17]. The weights of the links are given by the differential backlogs and the policy resembles the well known *backpressure algorithm* introduced by Tassiulas and Ephremides in [12, 49] and further extended in [16] to dynamic networks with power control.

MaxWeight link scheduling for wireless networks

Maximum weighted link scheduling for wireless networks [12, 38–41, 49–53] is a place, in which the problem of WSRMax is directly used. Note that, for networks with fixed link capacities, the maximum weighted link scheduling problem reduces to the classical maximum weighted matching problem and can be solved in polynomial time [38, 53, 54]. However, no solution is known for the general case when the link rates depend on the power allocation of all other links.

Power/rate control policies for wireless networks

We see sometimes that the WSRMax problem is directly used as the basis for the power/rate control policy in wireless, as well as in wireline networks [9–11]. For example, in DSL networks, there is considerable research on resource management policies, which rely directly on the WSRMax problem for multiuser spectrum balancing [55–66]. Direct application of WSRMax as an optimization criterion can also been seen extensively in joint power control and subcarrier assignment algorithms for OFDMA networks [20, 21, 67–71].

Resource management in MIMO networks

There are also a number of resource management algorithms in multiuser MIMO networks, which rely on the problem of WSRMax. For example, the methods proposed in [15, 72–74] rely on WSRMax for joint design of linear transmit and receive beamformers. In addition, many other references have applied WSRMax directly as an optimization criterion for beamformer design in MIMO networks,

including [74–80].

Finding achievable rate regions in wireless networks

In multiuser systems many users share the same network resources, e.g., time, frequency, codes, space, etc. Thus, there is naturally a tradeoff between the achievable rates of the users. In other words, one may require to reduce its rate if another user wants a higher rate. In such multiuser systems, the achievable rate regions are, of course, important since they characterize the tradeoff achievable by any resource management policy [9, 10]. As noted in [81], the rate regions are convex; by invoking a time sharing argument, one can always assume that the rate region is convex [9]. Therefore, any boundary point of the rate region can be obtained by using the solution of an WSRMax problem for some weights.

1.2.2 Global methods for WSRMax in wireless networks

Since the general WSRMax problem is NP-hard [23], it is natural to rely on global optimization approaches [24, 25] for computing an exact solution. One straightforward approach is based on exhaustive search in the variable space [55]. The main disadvantage of this approach is the prohibitively expensive computational complexity, even in the case of very small problems. A better approach is to apply branch and bound techniques [24, 25, 82], which essentially implement the exhaustive search in an intelligent manner; see [57–60, 62] and the author's contributions [83–88]. Branch and bound methods based on difference of convex functions (DC) programming [24] have been proposed in [57–59] to solve (a subclass of) WSRMax. Although DC programming is the core of their algorithms, it also limits the generality of their method to the problems, in which the objective function cannot easily be expressed as a DC [24]. For example, in the case of multicast wireless networks, expressing the objective function as a DC cannot be easily accomplished, even when Shannon's formula is used to express the achievable link rates. Another branch and bound method has been used in [60] in the context of DSL bit loading, where the search space is discretized in advance. As a result of discretization, this method does not allow a complete control of the accuracy of the solution. An alternative optimal method was proposed in [62], where the WSRMax problem is cast as a generalized linear

fractional program [89] and solved via a polyblock algorithm [48]. The method works well for small scale problems, but as pointed out in [25, Ch. 2, pp. 40-41] and [89, Sec. 6.3], it may show much slower convergence than branch and bound methods as the problem size increases. A special form of the WSRMax problem is presented in [8, p. 78][61], where the problem data and the constraints must obey certain properties and, consequently, the problem can be reduced to a convex formulation. However, these required properties correspond to very unlikely events in wireless/wireline networks, and therefore the method has a very limited applicability.

1.2.3 Local methods for WSRMax in wireless networks

Indeed, the worst case computational complexity for solving the general WSRMax problem by applying global optimization approaches [24, 25, 82] can increase more than polynomially with the number of variables. As a result, these methods are prohibitively expensive, even for off line optimization of moderate size networks. Therefore, certainly, the problem of WSRMax deserves efficient algorithms, which even though suboptimal, perform well in practice.

Several approximations have been proposed for the case when all links in the network operate in certain signal-to-interference-plus-noise ratio (SINR) regions. For example, the assumption that the achievable rate is a linear function of the SINR (i.e., low SINR region) is widely used in the ultra-wide-band systems [90–92]. Other references, which provide solutions for the power and rate control in low SINR regions include [39, 93, 94]. The high SINR region is treated in [34, 95, 96]. However, at the optimal operating point different links correspond to different SINR regions, which is usually the case with multihop networks. Therefore, all methods mentioned above that are based on either the low or the high SINR assumption can fail to solve the general problem.

One promising method is to cast the WSRMax problem into a signomial program (SP) formulation [5, Sec. 9] or into a complementary geometric program (CGP) [97, 98], where a suboptimal solution can be obtained quite efficiently; we can readily convert an SP to a CGP and vice versa [98, Sec. 2.2.5]. Applications of SP and CGP, or closely related solution methods, have been demonstrated in various signal processing and digital communications problems, e.g., [15, 64, 66, 72, 73, 98–100] and the author's contributions [101, 102]. There are a number of

other important papers proposing suboptimal solution methods for the WSRMax problem, such as [103] proposed by the author and [38, 40, 53, 63, 65, 74–80], among others.

Though the suboptimal methods mentioned above, including SP/CGP based algorithms, can perform reasonably well in many cases [2], it is worth pointing out that not all of them can handle the general WSRMax problem. The reason is the self-interference problem, which arises when a node transmits and receives simultaneously in the same frequency band. Since there is a huge imbalance between the transmitted signal power and the received signal power of nodes, the transmitted signal strength is typically few orders of magnitude larger than the received signal strength. Thus, when a node transmits and receives simultaneously in the same channel, the useful signal at the receiver of the incoming link is overwhelmed by the transmitted signal of the node itself. As a result, the SINR values at the incoming link of a node that simultaneously transmits in the same channel is very small. Therefore, the self-interference problem plays a central role in WSRMax in general wireless networks; see, e.g., the author's papers [104, 105].

Thus, in the case of general multihop wireless networks, the WSRMax problem must also cope with the self-interference problem. Under such circumstances SP/CGP cannot be directly applicable even to obtain a better suboptimal solution, since initialization of the algorithms is critical. One approach to dealing with self interference consists of adding supplementary combinatorial constraints, which prevent any node in the network from transmitting and receiving simultaneously [43, 51, 106–111]. This is sometimes called the node-exclusive interference model; only subsets of mutually exclusive links can simultaneously be activated in order to avoid the large self interference encountered if a node transmits and receives in the same frequency band. Of course, such approaches based on the node-exclusive interference model induce a combinatorial nature for the WSRMax problem in general. The combinatorial nature is circumvented in the authors contributions [112, 113], where *homotopy methods* (or continuation methods) [114] together with complementary geometric programming [97] are adopted to derive efficient algorithms for the general WSRMax problem. Here, the term "efficient" can mean faster convergence, or convergence to a point with

---

[2]e.g., when a node does not transmits and receives simultaneously in the same frequency band

better objective value.

Along with the emergence of future wireless technologies (e.g., 3G-LTE, WIMAX, IMT-A) [115–118], which rely heavily on OFDMA based physical layer specifications, there has also been a lot of research on designing resource management algorithms for OFDMA networks; see, e.g., [20, 21, 67–71, 119–121, 121–132]. In general, this process requires solving combinatorial optimization problems and the complexity of the problem increases exponentially with the number of subcarriers. Nevertheless, specific algorithms designed to power control and dynamic subcarriers assignment to the users, based on WSRMax, include the author's contributions [129–133] and [20, 21, 67–71], among others.

The general WSRMax problem has been addressed in [20] to characterize the frequency division multiple access (FDMA) capacity region for a broadcast channel. Due to the nontractability of the original problem, a modified convex problem formulation, FDMA-TDMA (time division multiple access), was proposed. The authors of [20] also considered algorithms to obtain optimal and suboptimal solutions to a particular variation of the original problem, where the total power is evenly divided among the used set of subcarriers. A Lagrangian relaxation based method has been proposed in [21]. Here, a bisection search method was used to update the dual variable until the algorithm converges. Due to the nonconvexity of the optimization problem the optimality of the algorithm is, of course, not guaranteed. Computationally efficient algorithms for maximizing the sum-rate have been developed in [67, 68]. Although these algorithms are optimal for sum-rate maximization, they are not applicable to solving the WSRMax problem. The reason is that the properties exploited for solving the sum-rate maximization problem are destroyed when there are arbitrary weights. A suboptimal method for characterizing the achievable rate region of the two-user FDMA channel have been presented in [69]. A number of papers also proposed suboptimal methods for variants on the WSRMax problem, such as [70, 71].

1.2.4   Distributed WSRMax in wireless networks

The emergence of large scale communication networks, as well as accompanying network control and optimization methods with huge signalling overheads triggered a considerable body of recent research on developing distributed

34

algorithms for resource management, see [4, 134, 135] and the references therein. Such distributed algorithms rely only on local observations and are carried out with limited access to global information. These algorithms essentially involve coordinating many local subproblems to find a solution to a large global problem. It is worth emphasizing that the convexity of the problems is crucial in determining the behavior of the distributed algorithms [4, Ch. 9]. For example, in the case of nonconvex problems such algorithms need not converge, and if it does converge, it need not converge to an optimal point, which is the case with the WSRMax problem. Nevertheless, finding even a suboptimal but distributed method for WSRMax is crucial for deploying distributively many network control and optimization methods, e.g., [12, 18, 32, 38, 41, 50–53], which rely on WSRMax.

Distributed implementation of the WSRMax problem has been investigated in [63, 64, 66, 136] in the context of DSL networks. However, in the case of cellular systems with multiple transmit antennas, the decision variables space is, of course, larger, e.g., joint optimization of transmit beamforming patterns, transmit powers, and link activations is required. Therefore, designing efficient distributed methods for WSRMax is a more challenging task, due to the extensive amount of message passing required to resolve the coupling between decision variables. In the sequel, we limit ourselves to basic, but still very important, results that develop distributed coordinated algorithms for resource management in networks with multiple antennas.

Several distributed methods for WSRMax in multiple-input and single-output (MISO) cellular networks have been proposed in [137–141] and in the author's contributions [142, 143]. Specifically, in [137] a two-user MISO interference channel (IC) is considered and a distributed algorithm is derived by using the commonly used high SINR approximation [34]. Moreover, another approximation, which relies on zero forcing (ZF) beamforming is introduced in [137] to handle networks with many MISO ICs. The methods proposed in [138–140] derived the necessary (but not sufficient) solution structure of the WSRMax problem and used it as the basis for their distributed solution. However, many parameters must be selected heuristically to construct a potential distributed solution and there is, in general no systematic method for finding those parameters. In particular, the algorithms in [138, 139] are designed for systems with limited backhaul signaling resources. Thus, [138, 139] do not consider any iterative base station

(BS) coordination mechanism to resolve the out-of-cell interference coupling. Even though the method proposed in [140] relies on stringent requirements on the message passing between BSs during each iteration of the algorithm, their results show that BS coordination can provide considerable gains compared to uncoordinated methods. An inexact cooperate descent algorithm for the case where each BS is serving only one cell edge user has been proposed in [141]. The method proposed in [144] is designed for sum-rate maximization and uses high SINR approximation. The method is very similar to the one proposed in [137] for WSRMax problem. A cooperative beamforming algorithm is proposed in [145] for MISO IC, where each BS can transmit only to a single user. Their proposed method employs an iterative BS coordination mechanism to resolve the out-of-cell interference coupling. However, the convexity properties exploited for distribution of the problem are destroyed when more than one user is served by any BS. Thus, their proposed method is not directly applicable to the WSRMax problem.

Algorithms based on game theory are found in [146–151]. Their proposed methods are restricted to interference channels, e.g., MISO IC, MIMO IC. In addition, the methods often require the coordination between receiver nodes and the transmitter nodes during algorithm's iterations. Therefore, even in a infrastructure based network (e.g., cellular network), communication overhead may be significantly noticeable.

Many optimization criteria other than the weighted sum-rate have been considered in the references [144, 145, 152–158] to distributively optimize the system resources (e.g., beamforming patterns, transmit powers, etc.) in multiantenna cellular networks. In particular, the references [152–155] used the characterization of the Pareto boundary of the MISO interference channel [159] as the basis for their distributed methods. The proposed methods do not employ any BS coordination mechanism to resolve the out-of-cell interference coupling. The simple power control methods considered can be unreliable, especially when the degree of freedom at the transmitters is not sufficient to perform ZF. In [156–158] distributed algorithms have been derived to minimize a total (weighted) transmitted power, or the maximum per antenna power across the BSs, subject to SINR constraints at the user terminals. Many related solution methodologies, though not distributed, can be found in [160–162] and references therein.

1.3   Aims and the outline of the thesis

The aim of this thesis is to apply optimization theory and techniques for developing global, fast local, as well as distributed solution methods to important and challenging problems that arise in radio resource management in wireless communication networks. Specifically, a greater emphasis is placed on the general WSRMax problem, which plays a central role in diverse application domains, such as NUM, cross-layer design, link scheduling, power/rate control, MIMO beamformer design, and finding achievable rate regions, among others.

Chapter 2, the results of which have been documented in [83, 85–87], proposes a solution method, based on the branch and bound technique, which solves globally the nonconvex WSRMax problem with an optimality certificate. Efficient analytic bounding techniques are introduced and their impact on convergence is numerically evaluated. The considered link-interference model is general enough to model a wide range of network topologies with various node capabilities, e.g. single- or multipacket transmission (or reception), simultaneous transmission and reception. Diverse application domains of WSRMax are considered in the numerical results, including cross-layer network utility maximization and maximum weighted link scheduling for multihop wireless networks, as well as finding achievable rate regions for singlecast/multicast wireless networks.

Chapter 3, the results of which have been presented in [101–105, 112, 113], proposes fast suboptimal algorithms for the WSRMax problem in multicommodity, multichannel wireless networks. First, the case where all receivers perform singleuser detection [3] is considered and algorithms are derived by applying complementary geometric programming and homotopy methods. Here we analyze the quantitative impact of gains that can be achieved at the network layer in terms of end-to-end rates and network congestion, by incorporating the proposed algorithms within a general cross-layer utility maximization framework. In addition, we apply them in evaluating the gains achievable at the network layer when the network nodes employ self interference cancelation techniques with different degrees of accuracy. Finally, a case where all receivers perform multiuser detection is considered and solutions are obtained by imposing additional constraints, such as that only one node can transmit to others at a time or that

---

[3]i.e., a receiver decodes each of its intended signals by treating all other interfering signals as noise.

only one node can receive from others at a time.

Chapter 4 is devoted to developing algorithms for the WSRMax problem in downlink OFDMA systems; the results are presented in [129, 130, 132, 133]. A low-complexity suboptimal power control and subcarrier assignment algorithm is proposed for the WSRMax problem. The algorithm is based on an a primal decomposition based method. The original, nonconvex optimization problem is split into a number of subproblems (one subproblem for each subcarrier) and a master problem. The subproblems, which can be solved independently are coordinated to compute an approximate solution for the master problem. Numerical results are provided to compare the performance of the proposed algorithm to Lagrange relaxation based suboptimal methods as well as to the optimal exhaustive search based method.

Chapter 5, the results of which have been documented in [142, 143], considers the WSRMax problem in a multicell downlink system. We derive a distributed algorithm based on primal decomposition and subgradient methods. The key idea is to break the original, nonconvex problem into a number of subproblems (one for each BS) and a master problem. Each BS optimizes locally its own decision variables (i.e., beamformers' directions and power allocation) by using an iterative ascent algorithm. The BS optimizations are coordinated to find an approximate solution for the master problem, which resolves the out-of-cell interference. Numerical results are provided to observe the the behavior of the algorithm under different degrees of coordination between the cooperating BS.

Chapter 6 concludes the thesis. The main results are summarized and some open problems for future research are pointed out.

1.4   The author's contribution to the publications

The thesis is based, in part, on five journal papers [83, 104, 112, 129, 142] and thirteen conference papers [85–87, 101–103, 105, 113, 130–133, 143]. The first four journal papers [83, 104, 112, 129] have already been published and the last one [142] is under revision. Nevertheless, a short conference version of [142] has been published recently; see [143].

The author had the main responsibility for carrying out the analysis, developing the simulation software, generating the numerical results, and writing all the journal papers [83, 104, 112, 129, 142]. Other authors provided invaluable

comments, criticism, and support during the process. The support from Dr. Marian Codreanu was priceless; he was involved in deep technical discussions, helped discussing the structures of the papers and rephrasing certain parts of the papers, which were, indeed, important in improving the clarity of the articles.

The conference papers [85–87, 101–103, 105, 113, 130–133, 143] are mainly based on the journal articles [83, 104, 112, 129, 142]. The author had the main responsibility for preparing all the conference papers, except [87, 102] for which Dr. Marian Codreanu took the responsibility of manuscript preparation.

In addition to the papers above, the author contributed to one other journal paper [84] and a conference paper [88]. The articles [84, 88] extended the BB algorithm for WSRMax developed in [83] further to handle MISO wireless networks. The author was actively involved in the discussions during the process. Moreover, the author provided the main simulation software developed in [83] on which the additional software components to handle the MISO case could be coded.

2 A branch and bound method for WSRMax

The main contribution of this chapter is to provide a branch and bound method for solving globally the general WSRMax problem for a set of interfering links. At each step, the algorithm computes upper and lower bounds for the optimal value. The algorithm terminates when the difference between the upper and the lower bounds is within a pre-specified accuracy level. Efficient analytic bounding techniques are introduced and their impact on the convergence is numerically evaluated. The considered link-interference model is general enough to model a wide range of network topologies with various node capabilities, e.g., single- or multipacket transmission (or reception), simultaneous transmission and reception. In contrast to the previously proposed branch and bound based techniques [57–59], our method does not rely on the convertibility of the problem into a DC problem. Therefore, our proposed method applies to a broader class of WSRMax problems (e.g., WSRMax in multicast wireless networks). Moreover, the method proposed here is not restricted to WSRMax; it can also be used to maximize *any* system performance metric that can be expressed as a Lipschitz continuous and increasing function of SINR values.

Our proposed branch and bound method shows some analogy to the one proposed in [60] in terms of the initial search domain and the basic bounding techniques. However, the two methods are fundamentally different in terms of branching techniques, as the algorithm proposed in [60] is designed specifically to search over a discrete space whilst our method is optimized for a continuous search space. We also provide improved bounding techniques which substantially improve the convergence speed of the algorithm.

Given its generality, the proposed algorithm can be adapted to address a wide range of network control and optimization problems. Performance benchmarks for various network topologies can be obtained by back-substituting it into any network design method which relies on WSRMax. Several applications, including cross-layer network utility maximization and maximum weighted link scheduling for multihop wireless networks, as well as finding achievable rate regions for singlecast/multicast wireless networks, are presented. As suboptimal but low-complex algorithms are typically used in practice, our algorithm can

also be used for evaluating their performance loss.

2.1   System model and problem formulation

The network considered consists of a collection of nodes which can send, receive, and relay data across a set of links. The set of all nodes is denoted by $\mathcal{N}$ and we label the nodes with the integer values $n = 1, \ldots, N$. A link is represented as an ordered pair (i, j) of distinct nodes. The set of all links is denoted by $\mathcal{L}$ and we label the links with the integer values $l = 1, \ldots, L$. We define tran(l) as the transmitter node of link l, and rec(l) as the receiver node of link l. The existence of a link $l \in \mathcal{L}$ implies that a direct transmission is possible from node tran(l) to node rec(l). Note that, in the most general case, $\mathcal{L}$ may consist of a combination of wireless and wireline links, e.g., in the case of hybrid networks. We define $\mathcal{O}(n)$ as the set of links that are outgoing from node n, and $\mathcal{I}(n)$ as the set of links that are incoming to node n. Furthermore, we denote the set of transmitter nodes by $\mathcal{T}$ and the set of receiver nodes by $\mathcal{R}$, i.e., $\mathcal{T} = \{n \in \mathcal{N}|\mathcal{O}(n) \neq \emptyset\}$ and $\mathcal{R} = \{n \in \mathcal{N}|\mathcal{I}(n) \neq \emptyset\}$.

The model above covers a wide range of network topologies from very simple ones to more complicated ones, as shown in Figure 2.1. A particular class of network topologies is the one for which the set of transmitters $\mathcal{T}$ and the set of receivers $\mathcal{R}$ are disjoint and we refer to these networks as *bipartite* networks. Figures 2.1(a) and 2.1(b) show two examples of bipartite networks. In Figure 2.1(a) each transmitter node has only one outgoing link and each receiving node has only one incoming link, i.e., $|\mathcal{O}(n)| = 1$ for all $n \in \mathcal{T}$ and $|\mathcal{I}(n)| = 1$ for all $n \in \mathcal{R}$. Borrowing terminology from graph theory, we say this network has *degree* one [4]. In contrast, the network shown in Figure 2.1(b) has degree three, since all nodes $n \in \{3, 7, 9\}$ have degree 3. A network for which $\mathcal{T} \cap \mathcal{R} \neq \emptyset$ is referred to as a *nonbipartite* network. Examples of nonbipartite networks are shown in Figures 2.1(c) and 2.1(d). Note that all bipartite networks are necessarily singlehop networks whilst the nonbipartite networks can be either singlehop [e.g., Figure 2.1(c)] or multihop [e.g., Figure 2.1(d)] networks.

---

[4]In graph theory, the degree of a vertex is the number of edges incident on it and the degree of a graph is the maximum degree of any vertex. By associating the network's nodes with vertices and the network's links with (oriented) edges, we say that the degree of node n is given by $\deg(n) = |\mathcal{I}(n)| + |\mathcal{O}(n)|$ and the degree of the network is given by $\max_{n \in \mathcal{N}} \deg(n)$.
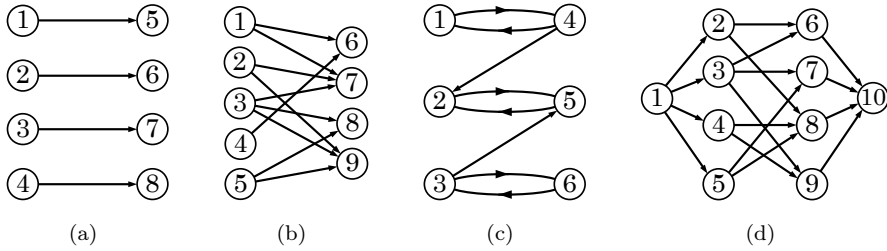
Fig 2.1. **Various network topologies: (a) Bipartite network,** $\mathcal{T} = \{1, 2, 3, 4\}$, $\mathcal{R} = \{5, 6, 7, 8\}$, **degree** 1; **(b) Bipartite network,** $\mathcal{T} = \{1, 2, 3, 4, 5\}$, $\mathcal{R} = \{6, 7, 8, 9\}$, **degree** 3; **(c) Nonbipartite singlehop network,** $\mathcal{T} = \mathcal{R} = \mathcal{N} = \{1, 2, 3, 4, 5, 6\}$, **degree** 3; **(d) Nonbipartite multihop network,** $\mathcal{T} = \{1, 2, \ldots, 9\}$, $\mathcal{R} = \{2, 3, \ldots, 10\}$, $\mathcal{T} \cap \mathcal{R} = \{2, 3, \ldots, 9\}$, **degree** 4, **[83] © 2011, IEEE.**

Furthermore, all networks with degree one are necessarily bipartite and all nonbipartite networks have degrees larger than one.

In general, depending on the complexity limitations and the transceiver techniques employed at different nodes of the network, some nodes may have restricted transmit and receive capabilities. For example, certain nodes may have only singlepacket receive and/or transmit capabilities [5] and some nodes may not be able to transmit and receive simultaneously. These limitations create subsets of mutually exclusive links and induce a combinatorial nature for the power and rate optimization in the case of networks with degree larger than one [38, 51, 163–167]. An example is the maximum weighted link scheduling for multihop wireless networks [12].

We assume that all links share a common channel and the interference is controlled via power allocation. We denote the channel gain from the transmitter of link i to the receiver of link j by h_{ij}. For any pair of distinct links $i \neq j$, we denote the interference coefficient from link i to link j by g_{ij}. In the case of nonadjacent links (i.e., links i and j do not have common nodes), g_{ij} represents the power of the interference signal at the receiver node of link j when one unit of power is allocated to the transmitter node of link i, i.e., $g_{ij} = |h_{ij}|^2$. When links i and j are adjacent, the value of g_{ij} also depends on the transmit and receive capabilities of the common node. Specifically, we set $g_{ij} = \infty$

---

[5]We say that a node has singlepacket receive capability if it can only receive from a single incoming link at a time. Similarly, we say that a node has singlepacket transmit capability if it can transmit only through a single outgoing link at a time.
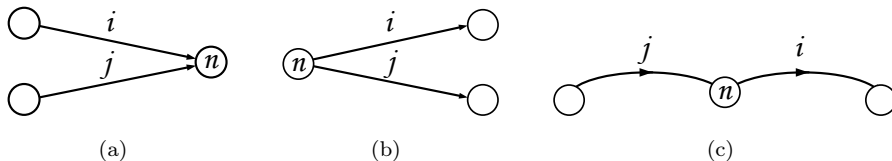
Fig 2.2. Choosing the value of interference coefficient in the case of adjacent links: (a) $i, j \in \mathcal{I}(n)$**,** $g_{ij} = g_{ji} = \infty$ **if node** n **has singlepacket receive capability or** $g_{ij} = |h_{ii}|^2$**,** $g_{ji} = |h_{jj}|^2$ **if node** n **has multipacket receive capability; (b)** $i, j \in \mathcal{O}(n)$**,** $g_{ij} = g_{ji} = \infty$ **if node** n **has singlepacket transmit capability or** $g_{ij} = |h_{jj}|^2$**,** $g_{ji} = |h_{ii}|^2$ **if node** n **has multipacket transmit capability; (c)** $i \in \mathcal{O}(n)$**,** $j \in \mathcal{I}(n)$**,** $g_{ij} = g_{ji} = \infty$ **if node** n **can not transmit and receive simultaneously or** $g_{ij} = |h_{ij}|^2$ **and** $g_{ji} = |h_{ji}|^2$ **if node** n **can transmit and receive simultaneously, [83] © 2011, IEEE.**

if links i and j are mutually exclusive and $g_{ij} = |h_{ij}|^2$ if links i and j can be simultaneously activated. Thus, $g_{ij} = g_{ji} = \infty$ for any pair of mutually exclusive links. Figure 2.2 illustrates three examples of choosing the value of the interference coefficient in the case of adjacent links. Note that in the case of nonbipartite networks, when $i \in \mathcal{O}(n)$ and $j \in \mathcal{I}(n)$, the term g_{ij} represents the power gain within the same node from its transmitter to its receiver, and is referred to as the self-interference coefficient [see Figure 2.2(c)]. In the case of wireless networks, these gains can be several orders of magnitude larger than the power gains between distinct nodes. References [168–171] discuss various self interference cancelations techniques that provide different degrees of accuracy. When such schemes are employed, g_{ij} models the residual self-interference coefficient after a certain (imperfect) self interference cancelation technique was performed.

It is worth noting that the interference model described previously can easily be extended to accommodate different multiple access techniques by appropriately reinterpreting the interference coefficients. For example, in the case of wireless CDMA networks, the interference coefficient g_{ij} would model the residual interference at the output of the despreading filter of node $\mathrm{rec}(j)$ [9]. Similarly, in the case of wireless SDMA networks, where nodes are equipped with multiple antennas, g_{ij} represents the equivalent interference coefficient measured at the output of the antenna combiner of node $\mathrm{rec}(j)$ [9]. Extensions to a multichannel scenario (e.g., FDMA or FDMA-SDMA networks) are also

possible by introducing multiple links between nodes, one link for each available spectral channel, and by setting $g_{ij} = 0$ if links i and j correspond to orthogonal channels. However, many such extensions are beyond the main scope of this thesis.

We consider the case where all receiver nodes are using *singleuser detection* (i.e., a receiver decodes each of its intended signals by treating all other interfering signals as noise) and assume that the achievable rate of link l is given by

$$r_l = \log\left(1 + \frac{g_{ll}p_l}{\sigma^2 + \sum_{j \neq l} g_{jl}p_j}\right) , \qquad (2.1)$$

where p_l is the power allocated to link l, σ^2 represents the power of the thermal noise at the receiver, and g_{ll} represents the power gain of link l, i.e. , $g_{ll} = |h_{ll}|^2$. The use of the Shannon formula [6] for the achievable rate in (2.1) is common practice (see, e.g., [9, 11]) but it must be noted that this is not strictly correct in the case of finite length packets. However, as the packet length increases, it is asymptotically correct.

Let us first consider the case of singlecast networks, where all links carry different information. Let β_l denote an arbitrary nonnegative number which represents the weight associated with link l. Assuming that the power allocation is subject to a maximum power constraint $\sum_{l \in \mathcal{O}(n)} p_l \leq p_n^{\max}$ for each transmitter node $n \in \mathcal{T}$ [7], the problem of weighted sum-rate maximization can be expressed as

$$
\begin{aligned}
\text{maximize} \quad & \sum_{l \in \mathcal{L}} \beta_l \log\left(1 + \frac{g_{ll}p_l}{\sigma^2 + \sum_{j \neq l} g_{jl}p_j}\right) \\
\text{subject to} \quad & \sum_{l \in \mathcal{O}(n)} p_l \leq p_n^{\max}, \ n \in \mathcal{T} \\
& p_l \geq 0, \ l \in \mathcal{L} ,
\end{aligned}
\qquad (2.2)
$$

where the variable is $(p_l)_{l \in \mathcal{L}}$.

In the case of multicast networks, a transmitter can simultaneously send common information to multiple receiver nodes. We consider the general

---

[6]The algorithm proposed in this section can be used for any other rate vs. SINR dependence. The only restriction is that the rate must be a nondecreasing and Lipschitz continuous function of SINR.

[7]For the sake of clarity we only consider the case of sum-power constraints for each transmitter node. However, supplementary sum-power constraints can be also handled by the proposed algorithm. For example, in the case of a cellular downlink employing the cooperation of several multiantenna base stations, sum-power constraints per subsets of nodes (one subset of nodes corresponds to a base station) should be also considered [73].
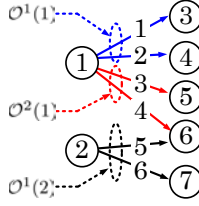
Fig 2.3. Multicast network: Different colors represents different multicast transmissions. $\mathcal{T} = \{1, 2\}$, $M_1 = 2$, $M_2 = 1$, $\mathcal{O}^1(1) = \{1, 2\}$, $\mathcal{O}^2(1) = \{3, 4\}$, **and** $\mathcal{O}^1(2) = \{5, 6\}$, **[83] © 2011, IEEE.**

case where each transmitter node can have several multicast transmissions. Thus, for each $n \in \mathcal{T}$ we partition $\mathcal{O}(n)$ into M_n disjoint subsets of links, i.e., $\mathcal{O}(n) = \cup_{m=1}^{M_n} \mathcal{O}^m(n)$, where M_n is the number of multicast transmissions from node n and the set $\mathcal{O}^m(n)$ contains all links associated with the mth multicast transmission of node n (see Figure 2.3). Let p_n^m and β_n^m be the power and the nonnegative weight allocated to the mth multicast transmission of node n. Moreover, let $\mathbf{p} = (p_n^m)_{n \in \mathcal{T}, m=1,\ldots,M_n}$ and denote the SINR of the lth link belongs to the mth multicast transmission of the nth node by $\text{SINR}_n^{ml}(\mathbf{p})$, where

$$\text{SINR}_n^{ml}(\mathbf{p}) = \frac{g_{ll} p_n^m}{\sigma^2 + \sum_{j \in \mathcal{T}, j \neq n} \sum_{k=1}^{M_j} p_j^k \max_{i \in \mathcal{O}^k(j)} g_{il} + \sum_{k=1, k \neq m}^{M_n} p_n^k \max_{i \in \mathcal{O}^k(n)} g_{il}}$$

$$\text{for all } n \in \mathcal{T}, \ m = 1, \ldots, M_n \ . \quad (2.3)$$

Clearly, for any link in the mth multicast transmission of node n, i.e., $l \in \mathcal{O}^m(n)$, interference at $\text{rec}(l)$ is created by the other multicast transmissions of node n itself and by multicast transmissions of other nodes. The $\max(\cdot)$ operator in the denominator of SINR expressions is used to impose mutually exclusive multicast transmissions, e.g., if node 6 in Figure 2.3 has singlepacket reception capability, then $\mathcal{O}^2(1)$ and $\mathcal{O}^1(2)$ are mutually exclusive.

Thus, by noting that the maximum rate achievable by all links in $\mathcal{O}^m(n)$ is given by $r_n^m = \min_{l \in \mathcal{O}^m(n)} r_l$, the weighted sum-rate maximization problem can be expressed as

$$\begin{aligned} \text{maximize} \quad & \sum_{n \in \mathcal{T}} \sum_{m=1}^{M_n} \beta_n^m \min_{l \in \mathcal{O}^m(n)} \log\left(1 + \text{SINR}_n^{ml}(\mathbf{p})\right) \\ \text{subject to} \quad & \sum_{m=1}^{M_n} p_n^m \leq p_n^{\max}, \ n \in \mathcal{T} \\ & p_n^m \geq 0, \ n \in \mathcal{T}, \ m = 1, \ldots, M_n \ , \end{aligned} \quad (2.4)$$

where the variable is $(p_n^m)_{n \in \mathcal{T}, m=1,\ldots,M_n}$.

46

2.2   Algorithm derivation

For the sake of clarity, let us first address the case of singlecast networks. Extension to multicast case is presented separately in Section 2.4. We start by equivalently reformulating the original problem (2.2) as minimization of a nonconvex function over an L-dimensional rectangle. Then, we describe our proposed algorithm based on a branch and bound technique [82] to minimize the nonconvex function over the L-dimensional rectangle.

By introducing auxiliary variables γ_l, $l \in \mathcal{L}$ we first reformulate problem (2.2) in the following equivalent form:

$$
\begin{aligned}
\text{minimize} \quad & \sum_{l \in \mathcal{L}} -\beta_l \log(1 + \gamma_l) \\
\text{subject to} \quad & \gamma_l \leq \frac{g_{ll} p_l}{\sigma^2 + \sum_{j \neq l} g_{jl} p_j}, \ l \in \mathcal{L} \\
& \sum_{l \in \mathcal{O}(n)} p_l \leq p_n^{\max}, \ n \in \mathcal{T} \\
& p_l \geq 0, \ l \in \mathcal{L} \ ,
\end{aligned}
\tag{2.5}
$$

where the variables are $(p_l)_{l \in \mathcal{L}}$ and $(\gamma_l)_{l \in \mathcal{L}}$. The equivalence between problems (2.2) and (2.5) follows from the monotone increasing property of the $\log(\cdot)$ function. Clearly, any feasible γ_l, $\in \mathcal{L}$ in problem (2.5) represents an achievable SINR value for link l. Let us denote the objective function of problem (2.5) by $f_0(\boldsymbol{\gamma}) = \sum_{l \in \mathcal{L}} -\beta_l \log(1 + \gamma_l)$ and the feasible set for the variables $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_L)$ (or the achievable SINR values) by $\mathcal{G}$, i.e.,

$$
\mathcal{G} = \left\{ \boldsymbol{\gamma} \ \middle| \
\begin{array}{ll}
\gamma_l \leq \frac{g_{ll} p_l}{\sigma^2 + \sum_{j \neq l} g_{jl} p_j}, & l \in \mathcal{L} \\
\sum_{l \in \mathcal{O}(n)} p_l \leq p_n^{\max}, & n \in \mathcal{T} \\
p_l \geq 0, & l \in \mathcal{L}
\end{array}
\right\} .
\tag{2.6}
$$

The optimal value of problem (2.5) can be expressed compactly as $t^\star = \inf_{\boldsymbol{\gamma} \in \mathcal{G}} f_0(\boldsymbol{\gamma})$.

For clarity, let us define a new function $\tilde{f} : \mathbb{R}_+^L \to \mathbb{R}$ as

$$
\tilde{f}(\boldsymbol{\gamma}) = \begin{cases} f_0(\boldsymbol{\gamma}) & \boldsymbol{\gamma} \in \mathcal{G} \\ 0 & \text{otherwise} \end{cases}
\tag{2.7}
$$

and note that for any $\mathcal{D} \subseteq \mathbb{R}_+^L$ such that $\mathcal{G} \subseteq \mathcal{D}$, we have

$$
\inf_{\boldsymbol{\gamma} \in \mathcal{D}} \tilde{f}(\boldsymbol{\gamma}) = \inf_{\boldsymbol{\gamma} \in \mathcal{G}} f_0(\boldsymbol{\gamma}) = t^\star \ ,
\tag{2.8}
$$

47

where the first equality follows from the fact that for any $\boldsymbol{\gamma} \in \mathbb{R}_+^L$ we have $f_0(\boldsymbol{\gamma}) \leq 0$. It is also worth noting that the function $\tilde{f}$ is nonconvex over $\mathcal{D}$ and f_0 is a global lower bound on $\tilde{f}$, i.e., $f_0(\boldsymbol{\gamma}) \leq \tilde{f}(\boldsymbol{\gamma})$ for all $\boldsymbol{\gamma} \in \mathcal{D}$.

Let us now define the L-dimensional rectangle

$$\mathcal{Q}_{\text{init}} = \left\{ \boldsymbol{\gamma} \middle| 0 \leq \gamma_l \leq \sigma^{-2} g_{ll} p_{\text{tran}(l)}^{\max}, \ l \in \mathcal{L} \right\} ,$$

which encloses the set of all achievable SINR values, i.e., $\mathcal{G} \subseteq \mathcal{Q}_{\text{init}}$. By using (2.8), it follows that $t^\star = \inf\limits_{\boldsymbol{\gamma} \in \mathcal{Q}_{\text{init}}} \tilde{f}(\boldsymbol{\gamma})$. Thus, we have reformulated problem (2.2) equivalently as a minimization of the nonconvex function $\tilde{f}$ over the rectangle $\mathcal{Q}_{\text{init}}$. In what follows we show how the branch and bound technique is used to minimize $\tilde{f}$ over $\mathcal{Q}_{\text{init}}$.

Let $\mathcal{Q}$ be a L-dimensional rectangle defined as

$$\mathcal{Q} = \left\{ \boldsymbol{\gamma} \middle| \gamma_{l,\min} \leq \gamma_l \leq \gamma_{l,\max}, \ l \in \mathcal{L} \right\} ,$$

where $\gamma_{l,\min}$ and $\gamma_{l,\max}$ are real numbers such that $\gamma_{l,\min} \leq \gamma_{l,\max}$ for all $l \in \mathcal{L}$. For any L-dimensional rectangle $\mathcal{Q} \subseteq \mathcal{Q}_{\text{init}}$, let us now define the following function:

$$\phi_{\min}(\mathcal{Q}) = \inf_{\boldsymbol{\gamma} \in \mathcal{Q}} \tilde{f}(\boldsymbol{\gamma}) . \tag{2.9}$$

It can be easily observed that

$$\phi_{\min}(\mathcal{Q}_{\text{init}}) = \inf_{\boldsymbol{\gamma} \in \mathcal{Q}_{\text{init}}} \tilde{f}(\boldsymbol{\gamma}) = t^\star . \tag{2.10}$$

The key idea of the branch and bound method is to generate a sequence of asymptotically tight upper and lower bounds for $\phi_{\min}(\mathcal{Q}_{\text{init}})$. At each iteration k, the lower bound L_k and the upper bound U_k are updated by partitioning $\mathcal{Q}_{\text{init}}$ into smaller rectangles. To ensure convergence, the bounds should become tight as the number of rectangles in the partition of $\mathcal{Q}_{\text{init}}$ grows. To do this, the branch and bound method uses two functions $\phi_{\text{ub}}(\mathcal{Q})$ and $\phi_{\text{lb}}(\mathcal{Q})$, defined for any rectangle $\mathcal{Q} \subseteq \mathcal{Q}_{\text{init}}$ such that the following conditions are satisfied [82].

C1 : The functions $\phi_{\text{lb}}(\mathcal{Q})$ and $\phi_{\text{ub}}(\mathcal{Q})$ compute a lower bound and an upper bound respectively on $\phi_{\min}(\mathcal{Q})$, i.e.,

$$\forall \mathcal{Q} \subseteq \mathcal{Q}_{\text{init}} \text{ we have } \phi_{\text{lb}}(\mathcal{Q}) \leq \phi_{\min}(\mathcal{Q}) \leq \phi_{\text{ub}}(\mathcal{Q}) . \tag{2.11}$$

C2 : As the maximum half length of the sides of $\mathcal{Q}$ $\left(\text{i.e., size}(\mathcal{Q})=\frac{1}{2}\max_{l\in\mathcal{L}}\{\gamma_{l,\max}-\right.$ $\left.\gamma_{l,\min}\}\right)$ goes to zero, the difference between the upper and lower bounds uniformly converges to zero, i.e.,

$$\forall \epsilon > 0 \; \exists \delta > 0 \;\; \text{s.t.} \;\; \forall \mathcal{Q} \subseteq \mathcal{Q}_{\text{init}}, \text{size}(\mathcal{Q}) \leq \delta \Rightarrow \phi_{\text{ub}}(\mathcal{Q}) - \phi_{\text{lb}}(\mathcal{Q}) \leq \epsilon \; . \quad (2.12)$$

For the sake of clarity, the definition and computation of ϕ_{lb} and ϕ_{ub} are described in Section 2.3. In the remainder of this section we will present the proposed branch and bound method in more detail.

Let ϵ be an a priori specified tolerance. The algorithm starts by computing $\phi_{\text{ub}}(\mathcal{Q}_{\text{init}})$ and $\phi_{\text{lb}}(\mathcal{Q}_{\text{init}})$. If $\phi_{\text{ub}}(\mathcal{Q}_{\text{init}}) - \phi_{\text{lb}}(\mathcal{Q}_{\text{init}}) \leq \epsilon$, the algorithm terminates and **C1** in (2.11) confirms that we have an upper bound $\phi_{\text{ub}}(\mathcal{Q}_{\text{init}})$, which is at most ϵ-away from the optimal value $t^\star$. Otherwise, we start partitioning $\mathcal{Q}_{\text{init}}$ into smaller rectangles. At the kth partitioning step, $\mathcal{Q}_{\text{init}}$ is split into k rectangles such that $\mathcal{Q}_{\text{init}} = \mathcal{Q}_1 \cup \mathcal{Q}_2 \cup \ldots \cup \mathcal{Q}_k$ and $\phi_{\text{ub}}(\mathcal{Q}_k)$ and $\phi_{\text{lb}}(\mathcal{Q}_k)$ are computed. Then the lower bound L_k and upper bound U_k are updated as follows:

$$L_k = \min_{i \in \{1,2,\ldots,k\}} \phi_{\text{lb}}(\mathcal{Q}_i) \leq \phi_{\min}(\mathcal{Q}_{\text{init}}) = t^\star \leq \min_{i \in \{1,2,\ldots,k\}} \phi_{\text{ub}}(\mathcal{Q}_i) = U_k \; . \quad (2.13)$$

Note that the lower bound L_k and the upper bound U_k are refined at each step and they represent the best lower and upper bounds obtained so far. If the difference between new bounds become smaller than ϵ, then the algorithm terminates. Otherwise, further partitioning of $\mathcal{Q}_{\text{init}}$ is required until the difference between U_k and L_k is less than ϵ. The condition **C2** in (2.12) ensures that, the difference $U_k - L_k$ eventually becomes smaller than ϵ for some finite k. The proposed algorithm based on the branch and bound method can be summarized as follows:

Algorithm 2.1. *Branch and bound method for WSRMax*

1. Initialization; given tolerance $\epsilon > 0$. Set $k = 1$, $\mathcal{B}_1 = \{\mathcal{Q}_{\text{init}}\}$, $U_1 = \phi_{\text{ub}}(\mathcal{Q}_{\text{init}})$, and $L_1 = \phi_{\text{lb}}(\mathcal{Q}_{\text{init}})$.
2. Stopping criterion; if $U_k - L_k > \epsilon$ go to step 3, otherwise STOP.
3. Branching;

    (a) pick $\mathcal{Q} \in \mathcal{B}_k$ for which $\phi_{\text{lb}}(\mathcal{Q}) = L_k$ and set $\mathcal{Q}_k = \mathcal{Q}$.

(b) split $\mathcal{Q}_k$ along one of its longest edge into $\mathcal{Q}_I$ and $\mathcal{Q}_{II}$.

(c) form $\mathcal{B}_{k+1}$ from $\mathcal{B}_k$ by removing $\mathcal{Q}_k$ and adding $\mathcal{Q}_I$ and $\mathcal{Q}_{II}$.

4. Bounding;

(a) set $U_{k+1} = \min_{\mathcal{Q} \in \mathcal{B}_{k+1}} \{\phi_{\mathrm{ub}}(\mathcal{Q})\}$.

(b) set $L_{k+1} = \min_{\mathcal{Q} \in \mathcal{B}_{k+1}} \{\phi_{\mathrm{lb}}(\mathcal{Q})\}$.

5. Pruning;

(a) pick all $\mathcal{Q} \in \mathcal{B}_{k+1}$ for which $\phi_{\mathrm{lb}}(\mathcal{Q}) \geq U_{k+1}$.

(b) update $\mathcal{B}_{k+1}$ by removing all $\mathcal{Q}$ obtained in the above step 5-(a).

6. Set $k = k + 1$ and go to step 2.

The first step initializes the algorithm and the upper and lower bounds are computed over the initial rectangle $\mathcal{Q}_{\mathrm{init}}$. The second step checks the difference between the best upper and lower bounds found so far [bounds U_k and L_k are given by (2.13)]. The algorithm repeats steps 3 to 6 until $U_k - L_k < \epsilon$.

Step 3 is the *branching* mechanism of the algorithm. Here we adopt the following branching rule: select from the current partition of $\mathcal{Q}_{\mathrm{init}}$ (i.e., $\mathcal{B}_k$) the rectangle with the smallest lower bound and split it in two smaller rectangles along its longest edge. Splitting the chosen rectangle along its longest edge ensures the convergence of the algorithm [82]. At step 4 the best upper bound U_k and the best lower bound L_k are updated according to (2.13).

Step 5 is used to eliminate (or prune) rectangles for which the lower bound is larger than the best upper bound found so far, since those rectangles can never contain a minimizer of the function $\tilde{f}$. Note that *pruning* does not affect the speed of the main algorithm since none of the rectangles that were pruned will be selected later in the branching step 3 for further splitting. The advantage of pruning is the release of the memory otherwise used for storing unnecessary rectangles.

2.2.1   Convergence of the branch and bound algorithm

In this section we show the convergence of the proposed branch and bound method for WSRMax (i.e., Algorithm 2.1) within a finite number of iterations.

Algorithm convergence is established by the following theorem.

Theorem 2.1. *If for any $\mathcal{Q} \subseteq \mathcal{Q}_{\text{init}}$ with $\mathcal{Q} = \{\boldsymbol{\gamma} \,|\, \gamma_{l,\text{min}} \leq \gamma_l \leq \gamma_{l,\text{max}}, \ l \in \mathcal{L}\}$, the functions $\phi_{\text{ub}}(\mathcal{Q})$ and $\phi_{\text{lb}}(\mathcal{Q})$ satisfy the conditions **C1** and **C2**, then Algorithm 2.1 converges in a finite number of iterations to a value arbitrarily close to $t^\star$, i.e., $\forall \epsilon > 0, \ \exists K > 0 \ s.t. \ U_K - t^\star \leq \epsilon$.*

Proof. The proof is similar to that provided in [82, 172] and it is provided here for the sake of completeness.

First note that there are k rectangles in the set $\mathcal{B}_k$ without pruning. Let the volume of rectangle $\mathcal{Q}_{\text{init}}$, denoted by $\text{vol}(\mathcal{Q}_{\text{init}})$. Thus, we have

$$\min_{\mathcal{Q} \in \mathcal{B}_k} \text{vol}(\mathcal{Q}) \leq \frac{\text{vol}(\mathcal{Q}_{\text{init}})}{k} . \tag{2.14}$$

Therefore, as k increases at least one rectangle in the partition become small.

Then it is required to show that the smaller $\text{vol}(\mathcal{Q})$ the smaller $\text{size}(\mathcal{Q})$. To do this, we first define the *condition number* of the rectangle

$$\mathcal{Q} = \{\boldsymbol{\gamma} \,|\, \gamma_{l,\text{min}} \leq \gamma_l \leq \gamma_{l,\text{max}}, \ l \in \mathcal{L}\}$$

as

$$\text{cond}(\mathcal{Q}) = \frac{\max_l(\gamma_{l,\text{max}} - \gamma_{l,\text{min}})}{\min_l(\gamma_{l,\text{max}} - \gamma_{l,\text{min}})} . \tag{2.15}$$

Note that the splitting rule we use, i.e., splitting the rectangle along its longest edge, always ensures that for any k and any rectangle $\mathcal{Q} \in \mathcal{B}_k$ [172, Lem. 1]

$$\text{cond}(\mathcal{Q}) \leq \max\{\text{cond}(\mathcal{Q}_{\text{init}}), 2\} . \tag{2.16}$$

Moreover, we have,

$$\text{vol}(\mathcal{Q}) = \prod_{l=1}^{L} (\gamma_{l,\text{max}} - \gamma_{l,\text{min}}) \tag{2.17}$$

$$\geq \max_l(\gamma_{l,\text{max}} - \gamma_{l,\text{min}}) \left(\min_l(\gamma_{l,\text{max}} - \gamma_{l,\text{min}})\right)^{L-1} \tag{2.18}$$

$$= \left(\max_l(\gamma_{l,\text{max}} - \gamma_{l,\text{min}})\right)^{L} \left(\frac{\min_l(\gamma_{l,\text{max}} - \gamma_{l,\text{min}})}{\max_l(\gamma_{l,\text{max}} - \gamma_{l,\text{min}})}\right)^{L-1} \tag{2.19}$$

$$= \frac{(2 \, \text{size}(\mathcal{Q}))^{L}}{(\text{cond}(\mathcal{Q}))^{L-1}} \tag{2.20}$$

$$\geq \left(\frac{2 \, \text{size}(\mathcal{Q})}{\text{cond}(\mathcal{Q})}\right)^{L} , \tag{2.21}$$

where (2.17)-(2.20) clearly followed by using straightforward manipulations and (2.21) follows by noting that $\text{cond}(\mathcal{Q}) \geq 1$. Thus, from (2.21) we have

$$\text{size}(\mathcal{Q}) \leq \frac{1}{2}\text{cond}(\mathcal{Q})\text{vol}(\mathcal{Q})^{1/L} \ . \tag{2.22}$$

By using (2.14), (2.16), and (2.22) we obtain the following relation:

$$\min_{\mathcal{Q}\in\mathcal{B}_k} \text{size}(\mathcal{Q}) \leq \frac{1}{2}\max\{\text{cond}(\mathcal{Q}_{\text{init}}),2\}\left(\frac{\text{vol}(\mathcal{Q}_{\text{init}})}{k}\right) \ . \tag{2.23}$$

We are now ready to show that there exists a positive integer K such that for any $\epsilon > 0$, $U_K - t^\star \leq \epsilon$. To see this, we select K as the maximum number of iterations such that

$$\frac{1}{2}\max\{\text{cond}(\mathcal{Q}_{\text{init}}),2\}\left(\frac{\text{vol}(\mathcal{Q}_{\text{init}})}{K}\right) \leq \delta \ . \tag{2.24}$$

Thus from (2.23), for some $\mathcal{Q}$ denoted as $\tilde{\mathcal{Q}}$, $\text{size}(\tilde{\mathcal{Q}}) \leq \delta$ and from **C2** [see (2.12)], we have

$$\phi_{\text{ub}}(\tilde{\mathcal{Q}}) - \phi_{\text{lb}}(\tilde{\mathcal{Q}}) \leq \epsilon \ . \tag{2.25}$$

However, note that $U_K \leq \phi_{\text{ub}}(\tilde{\mathcal{Q}})$ [since $U_K = \min_{k\in\{1,2,\ldots,K\}}\phi_{\text{ub}}(\mathcal{Q}_k)$] and $t^\star \geq \phi_{\text{lb}}(\tilde{\mathcal{Q}})$. Thus, $U_K - t^\star \leq \epsilon$ and the result follows. $\qquad\square$

2.3   Computation of upper and lower bounds

Note that the main challenge in designing a global optimization algorithm based on the branch and bound method is to find cheaply computable functions $\phi_{\text{ub}}(\mathcal{Q})$ and $\phi_{\text{lb}}(\mathcal{Q})$ such that the conditions given in (2.11) and (2.12) are satisfied. Basically, the essence of the branch and bound method is based on the fact that for any $\mathcal{Q} \subseteq \mathcal{Q}_{\text{init}}$, the bounds $\phi_{\text{ub}}(\mathcal{Q})$ and $\phi_{\text{lb}}(\mathcal{Q})$ are substantially easier to compute than the true minimum $\phi_{\text{min}}(\mathcal{Q})$ [82].

In this section we propose several candidates for $\phi_{\text{lb}}(\mathcal{Q})$ and $\phi_{\text{ub}}(\mathcal{Q})$ in Algorithm 2.1. First, we describe two basic lower and upper bound functions, prove that they satisfy the conditions **C1** and **C2** [see (2.11) and (2.12)] and present efficient methods for computing them. Computationally efficient better bounds are presented later in this section.

2.3.1 Basic lower and upper bounds

Recall that $\mathcal{Q} = \{\boldsymbol{\gamma} \,|\, \gamma_{l,\min} \leq \gamma_l \leq \gamma_{l,\max}, \ l \in \mathcal{L}\}$. We now define the functions $\phi_{\mathrm{lb}}^{\mathrm{Basic}}(\mathcal{Q})$ and $\phi_{\mathrm{ub}}^{\mathrm{Basic}}(\mathcal{Q})$ as

$$\phi_{\mathrm{lb}}^{\mathrm{Basic}}(\mathcal{Q}) = \begin{cases} f_0(\boldsymbol{\gamma}_{\max}) & \boldsymbol{\gamma}_{\min} \in \mathcal{G} \\ 0 & \text{otherwise} \ ; \end{cases} \tag{2.26}$$

$$\phi_{\mathrm{ub}}^{\mathrm{Basic}}(\mathcal{Q}) = \tilde{f}(\boldsymbol{\gamma}_{\min}) = \begin{cases} f_0(\boldsymbol{\gamma}_{\min}) & \boldsymbol{\gamma}_{\min} \in \mathcal{G} \\ 0 & \text{otherwise} \ , \end{cases} \tag{2.27}$$

where $\boldsymbol{\gamma}_{\max} = (\gamma_{1,\max}, \ldots, \gamma_{L,\max})$, $\boldsymbol{\gamma}_{\min} = (\gamma_{1,\min}, \ldots, \gamma_{L,\min})$, and $\mathcal{G}$ is defined in (2.6). Note that the most computationally expensive part of evaluating $\phi_{\mathrm{lb}}^{\mathrm{Basic}}(\mathcal{Q})$ and $\phi_{\mathrm{ub}}^{\mathrm{Basic}}(\mathcal{Q})$ is to check the condition $\boldsymbol{\gamma}_{\min} \in \mathcal{G}$. An efficient method for checking this condition is provided soon after the following important properties of functions $\phi_{\mathrm{lb}}^{\mathrm{Basic}}$ and $\phi_{\mathrm{ub}}^{\mathrm{Basic}}$ are established.

Lemma 2.1. *The functions $\phi_{\mathrm{lb}}^{\mathrm{Basic}}(\mathcal{Q})$ and $\phi_{\mathrm{ub}}^{\mathrm{Basic}}(\mathcal{Q})$ satisfy the condition $\boldsymbol{C1}$.*

Proof. In the case of $\boldsymbol{\gamma}_{\min} \notin \mathcal{G}$ we can easily see that $\phi_{\mathrm{lb}}^{\mathrm{Basic}}(\mathcal{Q}) = \phi_{\min}(\mathcal{Q}) = \phi_{\mathrm{ub}}^{\mathrm{Basic}}(\mathcal{Q}) = 0$, and therefore the inequalities in $\boldsymbol{C1}$ hold with equalities. In the case of $\boldsymbol{\gamma}_{\min} \in \mathcal{G}$ we notice that

$$\phi_{\min}(\mathcal{Q}) = \inf_{\boldsymbol{\gamma} \in \mathcal{Q}} \tilde{f}(\boldsymbol{\gamma}) \leq \tilde{f}(\boldsymbol{\gamma}_{\min}) = f_0(\boldsymbol{\gamma}_{\min}) = \phi_{\mathrm{ub}}^{\mathrm{Basic}}(\mathcal{Q}) \ . \tag{2.28}$$

The first equality follows from (2.9), the inequality follows since $\boldsymbol{\gamma}_{\min} \in \mathcal{Q}$, and the second equality follows from (2.7). Moreover, we have

$$\phi_{\min}(\mathcal{Q}) = \inf_{\boldsymbol{\gamma} \in \mathcal{Q}} \tilde{f}(\boldsymbol{\gamma}) \geq \inf_{\boldsymbol{\gamma} \in \mathcal{Q}} f_0(\boldsymbol{\gamma}) = f_0(\boldsymbol{\gamma}_{\max}) = \phi_{\mathrm{lb}}^{\mathrm{Basic}}(\mathcal{Q}) \ , \tag{2.29}$$

where the inequality follows from the fact that $\tilde{f}(\boldsymbol{\gamma}) \geq f_0(\boldsymbol{\gamma})$ and the second equality is from the fact that $\mathcal{Q}$ is a rectangle and $f_0(\boldsymbol{\gamma})$ is monotonically decreasing in each variable γ_l, $l \in \mathcal{L}$. From (2.28) and (2.29) we conclude that $\phi_{\mathrm{lb}}^{\mathrm{Basic}}(\mathcal{Q}) \leq \phi_{\min}(\mathcal{Q}) \leq \phi_{\mathrm{ub}}^{\mathrm{Basic}}(\mathcal{Q})$. $\qquad \square$

Lemma 2.2. *The functions $\phi_{\mathrm{lb}}^{\mathrm{Basic}}(\mathcal{Q})$ and $\phi_{\mathrm{ub}}^{\mathrm{Basic}}(\mathcal{Q})$ satisfy the condition $\boldsymbol{C2}$.*

Proof. We first show that the function $f_0(\boldsymbol{\gamma}) = \sum_{l \in \mathcal{L}} -\beta_l \log(1 + \gamma_l)$ is Lipschitz continuous on $\mathbb{R}_+^L$ with the constant $D = \sqrt{\sum_{l \in \mathcal{L}} \beta_l^2}$, i.e.,

$$|f_0(\boldsymbol{\mu}) - f_0(\boldsymbol{\nu})| \leq D \, \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_2 \tag{2.30}$$

for all $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathbb{R}_+^L$. We start by noting that $f_0(\boldsymbol{\gamma})$ is convex. Therefore, for all $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathbb{R}_+^L$ we have [3, Sec. 3.1.3]

$$f_0(\boldsymbol{\mu}) - f_0(\boldsymbol{\nu}) \leq \nabla f_0(\boldsymbol{\mu})^{\mathrm{T}}(\boldsymbol{\mu} - \boldsymbol{\nu}) \ . \tag{2.31}$$

Without loss of generality, we can assume that $f_0(\boldsymbol{\mu}) - f_0(\boldsymbol{\nu}) \geq 0$. Otherwise, we can obtain exactly the same results by interchanging $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ in (2.31), i.e., $f_0(\boldsymbol{\nu}) - f_0(\boldsymbol{\mu}) \leq \nabla f_0(\boldsymbol{\nu})^{\mathrm{T}}(\boldsymbol{\nu} - \boldsymbol{\mu})$. Thus, we have that

$$|f_0(\boldsymbol{\mu}) - f_0(\boldsymbol{\nu})| \leq \left| \nabla f_0(\boldsymbol{\mu})^{\mathrm{T}}(\boldsymbol{\mu} - \boldsymbol{\nu}) \right| \tag{2.32}$$

$$\leq ||\nabla f_0(\boldsymbol{\mu})||_2 \, ||(\boldsymbol{\mu} - \boldsymbol{\nu})||_2 \tag{2.33}$$

$$\leq \max_{\boldsymbol{\gamma} \in \mathbb{R}_+^L} ||\nabla f_0(\boldsymbol{\gamma})||_2 \, ||(\boldsymbol{\mu} - \boldsymbol{\nu})||_2 \tag{2.34}$$

$$= \max_{\boldsymbol{\gamma} \in \mathbb{R}_+^L} \sqrt{\sum_{l \in \mathcal{L}} \frac{\beta_l^2}{(1 + \gamma_l)^2}} \, ||(\boldsymbol{\mu} - \boldsymbol{\nu})||_2 \tag{2.35}$$

$$= D \, ||(\boldsymbol{\mu} - \boldsymbol{\nu})||_2 \ , \tag{2.36}$$

where (2.32) follows from (2.31), (2.33) follows from the Cauchy-Schwarz inequality, (2.34) follows from the maximization operation, (2.35) follows by noting that $[\nabla f_0(\gamma)]_l = \frac{\beta_l}{(1+\gamma_l)}$, $l \in \mathcal{L}$, and (2.36) follows by setting $\gamma_l = 0$ for all $l \in \mathcal{L}$.

Now we can write the following relations:

$$\phi_{\mathrm{ub}}^{\mathrm{Basic}}(\mathcal{Q}) - \phi_{\mathrm{lb}}^{\mathrm{Basic}}(\mathcal{Q}) \leq f_0(\boldsymbol{\gamma}_{\mathrm{min}}) - f_0(\boldsymbol{\gamma}_{\mathrm{max}}) \tag{2.37}$$

$$\leq D \, ||\boldsymbol{\gamma}_{\mathrm{min}} - \boldsymbol{\gamma}_{\mathrm{max}}||_2 \tag{2.38}$$

$$= D \big|\big| \sum_{l \in \mathcal{L}} (\gamma_{l,\mathrm{max}} - \gamma_{l,\mathrm{min}}) \mathbf{e}_l \big|\big|_2 \tag{2.39}$$

$$\leq D \sum_{l \in \mathcal{L}} (\gamma_{l,\mathrm{max}} - \gamma_{l,\mathrm{min}}) \tag{2.40}$$

$$\leq 2DL \, \mathrm{size}(\mathcal{Q}) \ . \tag{2.41}$$

The first inequality (2.37) follows from (2.26) and (2.27) by noting that f_0 is nonincreasing, (2.38) follows from (2.30), (2.39) follows clearly by noting that $\mathbf{e}_l$ is the lth standard unit vector, (2.40) follows from triangle inequality, and (2.41) follows from the definition of $\mathrm{size}(\mathcal{Q})$ (see **C2**). Thus, for any given $\epsilon > 0$, we can select δ such that $\delta \leq \epsilon/2DL$, which in turns implies that condition **C2** is satisfied. $\qquad\square$

In the sequel, we present a computationally efficient method of checking the condition $\boldsymbol{\gamma}_{\min} \in \mathcal{G}$ which is central in computing $\phi_{\text{lb}}^{\text{Basic}}(\mathcal{Q})$ and $\phi_{\text{ub}}^{\text{Basic}}(\mathcal{Q})$ efficiently. Without loss of generality, we can assume that $\boldsymbol{\gamma}_{\min} > \mathbf{0}$. Note that the method can be extended to the case, where there are links l for which $\gamma_{l,\min} = 0$ in a straightforward manner; then, checking the original condition $\boldsymbol{\gamma}_{\min} \in \mathcal{G}$ is equivalent to checking a modified condition $\breve{\boldsymbol{\gamma}}_{\min} \in \breve{\mathcal{G}}$, where $\breve{\boldsymbol{\gamma}}_{\min}$ and $\breve{\mathcal{G}}$ are obtained by eliminating the dimensions (or link indexes) for which $\gamma_{l,\min} = 0$ and thus, we have $\breve{\boldsymbol{\gamma}}_{\min} > \mathbf{0}$.

Let us first consider the first set of inequalities in the description of $\mathcal{G}$, i.e.,

$$\gamma_l \le \frac{g_{ll} p_l}{\sigma^2 + \sum_{j \ne l} g_{jl} p_j}, \; l \in \mathcal{L} \; . \tag{2.42}$$

Let $\mathbf{p} = (p_1, \ldots, p_L)$. By rearranging the terms, (2.42) can be equivalently expressed as [72, 173]

$$(\mathbf{I} - \mathbf{B}(\boldsymbol{\gamma})\mathbf{G}) \, \mathbf{p} \ge \sigma^2 \mathbf{B}(\boldsymbol{\gamma}) \mathbf{1} \; , \tag{2.43}$$

where the matrices $\mathbf{B}(\boldsymbol{\gamma}) \in \mathbb{R}_+^{L \times L}$ and $\mathbf{G} \in \mathbb{R}_+^{L \times L}$ are defined by

$$\mathbf{B}(\boldsymbol{\gamma}) = \text{diag}\left(\frac{\gamma_1}{g_{11}}, \ldots, \frac{\gamma_L}{g_{LL}}\right) \; ; \quad [\mathbf{G}]_{i,j} = \begin{cases} g_{ji} & i \ne j \\ 0 & \text{otherwise} \; . \end{cases} \tag{2.44}$$

For the notational simplicity, let

$$\mathbf{A}(\boldsymbol{\gamma}) = \mathbf{I} - \mathbf{B}(\boldsymbol{\gamma})\mathbf{G} \; \text{ and } \; \mathbf{b}(\boldsymbol{\gamma}) = \sigma^2 \mathbf{B}(\boldsymbol{\gamma}) \mathbf{1} \; . \tag{2.45}$$

Thus, (2.42) can be compactly expressed as $\mathbf{A}(\boldsymbol{\gamma})\mathbf{p} \ge \mathbf{b}(\boldsymbol{\gamma})$. Let us denote the spectral radius [174, p. 5] of matrix $\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}$ by $\rho(\mathbf{B}(\boldsymbol{\gamma})\mathbf{G})$. The following theorem helps us to check if $\boldsymbol{\gamma} \in \mathcal{G}$.

Theorem 2.2. *For any $\boldsymbol{\gamma} > \mathbf{0}$, the following implications hold:*

1. *$\rho(\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}) \ge 1 \; \Rightarrow \; \boldsymbol{\gamma} \notin \mathcal{G}$.*
2. *$\rho(\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}) < 1$ and $\sum_{l \in \mathcal{O}(n)} p_l \le p_n^{\max}$ for all $n \in \mathcal{T}$, where $\mathbf{p} = \mathbf{A}^{-1}(\boldsymbol{\gamma})\mathbf{b}(\boldsymbol{\gamma})$ $\Rightarrow \boldsymbol{\gamma} \in \mathcal{G}$.*
3. *$\rho(\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}) < 1$ and $\exists n \in \mathcal{T}$ s.t. $\sum_{l \in \mathcal{O}(n)} p_l > p_n^{\max}$, where $\mathbf{p} = \mathbf{A}^{-1}(\boldsymbol{\gamma})\mathbf{b}(\boldsymbol{\gamma}) \Rightarrow \boldsymbol{\gamma} \notin \mathcal{G}$.*

Proof. See Appendix 1. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ □

Based on Theorem 2.2, the condition $\boldsymbol{\gamma}_{\min} \in \mathcal{G}$ can be checked as follows:

Algorithm 2.2. *Checking for condition* $\boldsymbol{\gamma}_{\min} \in \mathcal{G}$

1. Construct $\mathbf{B}(\boldsymbol{\gamma}_{\min})$ and $\mathbf{G}$ according to (2.44).
2. If $\rho\left(\mathbf{B}(\boldsymbol{\gamma}_{\min})\mathbf{G}\right) \geq 1$, then $\boldsymbol{\gamma}_{\min} \notin \mathcal{G}$ and STOP. Otherwise, let
   $\mathbf{p} = \mathbf{A}^{-1}(\boldsymbol{\gamma}_{\min})\mathbf{b}(\boldsymbol{\gamma}_{\min})$.
3. If $\sum_{l \in \mathcal{O}(n)} p_l \leq p_n^{\max}$ for all $n \in \mathcal{T}$, then $\boldsymbol{\gamma}_{\min} \in \mathcal{G}$ and STOP. Otherwise, $\boldsymbol{\gamma}_{\min} \notin \mathcal{G}$ and STOP.


2.3.2 Improved lower and upper bounds

Finding tighter bounds is very important as they can substantially increase the convergence speed of Algorithm 2.1. By exploiting the monotonically nonincreasing property of f_0 [i.e., $\boldsymbol{\gamma}_1 \leq \boldsymbol{\gamma}_2 \Rightarrow f_0(\boldsymbol{\gamma}_1) \geq f_0(\boldsymbol{\gamma}_2)$], one improved lower bound and two improved upper bounds are proposed in this subsection. Efficient methods of computing them are provided as well.

Note that, in the case of $\boldsymbol{\gamma}_{\min} \notin \mathcal{G}$ [i.e., $\mathcal{Q} \cap \mathcal{G} = \emptyset$, see Figure 2.4(a)], $\tilde{f}(\boldsymbol{\gamma}) = 0$ for any $\boldsymbol{\gamma} \in \mathcal{Q}$. Thus, both the basic lower bound (2.26) and the basic upper bound (2.27) are trivially zero and no further improvement is possible since they are tight. Consequently, tighter bounds can be found only in the case $\boldsymbol{\gamma}_{\min} \in \mathcal{G}$ [i.e., $\mathcal{Q} \cap \mathcal{G} \neq \emptyset$, see Figure 2.4(b)]. Thus, we consider only this case in the sequel, unless otherwise specified.

Improved lower bound

Roughly speaking, a tighter lower bound can be obtained as follows. We first construct the smallest rectangle $\bar{\mathcal{Q}}^\star \subseteq \mathcal{Q}$, which encloses the intersection $\mathcal{Q} \cap \mathcal{G}$ [see Figure 2.4(b)]. Let us denote this rectangle as $\bar{\mathcal{Q}}^\star = \{\boldsymbol{\gamma} \,|\, \gamma_{l,\min} \leq \gamma_l \leq \bar{\gamma}_l^\star, \ l \in \mathcal{L}\}$. The improved lower bound is given by $f_0(\bar{\gamma}_1^\star, \ldots, \bar{\gamma}_L^\star)$ [8].

---

[8]Further improvement can be obtained by constructing an outer polyblock approximation [89] for $\bar{\mathcal{Q}}^\star \cap \mathcal{G}$ that lies inside $\bar{\mathcal{Q}}^\star$. If $\{\dot{\mathbf{v}}_i\}_{i \in \dot{\mathcal{V}}}$ are the proper vertices of the polyblock, it is easy to see that an improved bound is given by $\min_{i \in \dot{\mathcal{V}}} f_0(\dot{\mathbf{v}}_i)$. Though interesting, in this thesis we do not consider these possible extensions, which can be carried out in a straightforward manner. But we refer the reader to [25, Ch. 2, Sec. 7], where similar bound improving techniques are discussed in the context of (difference of) monotonic optimization problems.
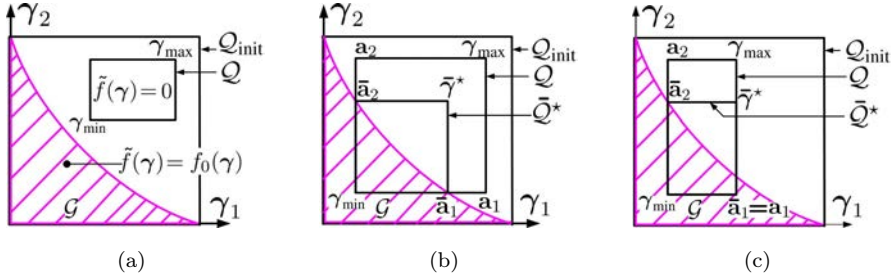
Fig 2.4. Illustration of the sets $\mathcal{G}$, $\mathcal{Q}_{\mathrm{init}}$, $\mathcal{Q}$, and $\bar{\mathcal{Q}}^{\star}$ in a $2-$dimensional space, [83] © 2011, IEEE.

Recall that $\mathcal{Q} = \{\gamma \,|\, \gamma_{l,\min} \leq \gamma_l \leq \gamma_{l,\max}, \; l \in \mathcal{L}\}$. For any $\mathcal{Q} \subseteq \mathcal{Q}_{\mathrm{init}}$, the improved lower bound can be formally expressed as

$$\phi_{\mathrm{lb}}^{\mathrm{Imp}}(\mathcal{Q}) = \begin{cases} f_0(\bar{\gamma}^{\star}) & \gamma_{\min} \in \mathcal{G} \\ 0 & \text{otherwise}, \end{cases} \tag{2.46}$$

where, $\bar{\gamma}^{\star} = (\bar{\gamma}_1^{\star}, \ldots, \bar{\gamma}_L^{\star})$ and $\bar{\gamma}_i^{\star}$ is the optimal value of the following optimization problem:

$$\begin{aligned} \text{maximize} \quad & \frac{g_{ii}p_i}{\sigma^2 + \sum_{j \neq i} g_{ji}p_j} \\ \text{subject to} \quad & \frac{g_{ii}p_i}{\sigma^2 + \sum_{j \neq i} g_{ji}p_j} \leq \gamma_{i,\max} \\ & \gamma_{l,\min} = \frac{g_{ll}p_l}{\sigma^2 + \sum_{j \neq l} g_{jl}p_j}, \; l \in \mathcal{L} \setminus \{i\} \\ & \sum_{l \in \mathcal{O}(n)} p_l \leq p_n^{\max}, \; n \in \mathcal{T} \\ & p_l \geq 0, \; l \in \mathcal{L}, \end{aligned} \tag{2.47}$$

where the variable is $(p_l)_{l \in \mathcal{L}}$. The first inequality constraint ensures that $\bar{\mathcal{Q}}^{\star} \subseteq \mathcal{Q}$, and it is active if and only if the corner point $\mathbf{a}_i = \gamma_{\min} + (\gamma_{i,\max} - \gamma_{i,\min})\mathbf{e}_i$ lies inside $\mathcal{G}$, i.e., $\mathbf{a}_i \in \mathcal{G}$ [see $\mathbf{a}_1$ in Figure 2.4(c)]. Therefore, when $\mathbf{a}_i \in \mathcal{G}$, $\bar{\gamma}_i^{\star} = \gamma_{i,\max}$. Otherwise (i.e., $\mathbf{a}_i \notin \mathcal{G}$), $\bar{\gamma}_i^{\star}$ is limited by the power constraints. In this case, the first constraint of problem (2.47) can be safely dropped and the resulting problem can be readily converted into a standard geometric program (GP) [3] so that the solution can be obtained numerically by using a GP solver, e.g., GGPLAB, GPPOSY, GPCVX [175]. However, it turns out that the particular structure of problem (2.47) allows us to analytically find the optimal value. This provides a more computationally efficient way to compute $\phi_{\mathrm{lb}}^{\mathrm{Imp}}(\mathcal{Q})$ without relying on a GP solver. This method is described soon after the following important property of $\phi_{\mathrm{lb}}^{\mathrm{Imp}}(\mathcal{Q})$ is established.

Lemma 2.3. *For any* $\mathcal{Q} \subseteq \mathcal{Q}_{\text{init}}$, *the lower bound* $\phi_{\text{lb}}^{\text{Imp}}(\mathcal{Q})$ *(2.46) is better than the basic lower bound* $\phi_{\text{lb}}^{\text{Basic}}(\mathcal{Q})$ *(2.26), i.e.,* $\phi_{\min}(\mathcal{Q}) \geq \phi_{\text{lb}}^{\text{Imp}}(\mathcal{Q}) \geq \phi_{\text{lb}}^{\text{Basic}}(\mathcal{Q})$.

Proof. If $\boldsymbol{\gamma}_{\min} \notin \mathcal{G}$, we have $\phi_{\min}(\mathcal{Q}) = \phi_{\text{lb}}^{\text{Imp}}(\mathcal{Q}) = \phi_{\text{lb}}^{\text{Basic}}(\mathcal{Q}) = 0$. Otherwise, i.e., when $\boldsymbol{\gamma}_{\min} \in \mathcal{G}$ we obtain

$$\phi_{\min}(\mathcal{Q}) = \inf_{\boldsymbol{\gamma} \in \mathcal{Q}} \tilde{f}(\boldsymbol{\gamma}) = \inf_{\boldsymbol{\gamma} \in \mathcal{G} \cap \mathcal{Q}} \tilde{f}(\boldsymbol{\gamma}) = \inf_{\boldsymbol{\gamma} \in \mathcal{G} \cap \mathcal{Q}} f_0(\boldsymbol{\gamma})$$
$$\geq f_0(\bar{\boldsymbol{\gamma}}^{\star}) = \phi_{\text{lb}}^{\text{Imp}}(\mathcal{Q}) \geq f_0(\boldsymbol{\gamma}_{\max}) = \phi_{\text{lb}}^{\text{Basic}}(\mathcal{Q}) , \quad (2.48)$$

where the first equality is from (2.9), the second equality follows from the fact that $\mathcal{G} \cap \mathcal{Q}$ is nonempty and $\tilde{f}(\boldsymbol{\gamma}) = 0$ for all $\boldsymbol{\gamma} \in \mathcal{Q} \setminus (\mathcal{G} \cap \mathcal{Q})$, the third equality follows from $\tilde{f}(\boldsymbol{\gamma}) = f_0(\boldsymbol{\gamma})$ for all $\boldsymbol{\gamma} \in \mathcal{G} \cap \mathcal{Q}$, the first inequality follows by noting that $\bar{\boldsymbol{\gamma}}^{\star} \geq \boldsymbol{\gamma}$ for all $\boldsymbol{\gamma} \in \mathcal{Q} \cap \mathcal{G}$ and f_0 is monotonically decreasing in each dimension, and the last inequality follows since $\boldsymbol{\gamma}_{\max} \geq \bar{\boldsymbol{\gamma}}^{\star}$ and f_0 is monotonically decreasing. $\square$

We describe now an efficient method to find $\bar{\gamma}_i^{\star}$ by solving problem (2.47) when $\boldsymbol{\gamma}_{\min} \in \mathcal{G}$ and $\mathbf{a}_i \notin \mathcal{G}$. We can assume without loss of generality that $\gamma_{l,\min} > 0$ for all $l \in \mathcal{L} \setminus \{i\}$; the proposed method can be extended to the case where there are links for which $\gamma_{l,\min} = 0$ for some $l \in \mathcal{L} \setminus \{i\}$. In such cases the original problem (2.47) is equivalent to a modified problem obtained by eliminating the dimensions $l \in \mathcal{L} \setminus \{i\}$ (i.e., link indexes) for which $\gamma_{l,\min} = 0$.

The proposed method can be summarized as follows. By using the equality constraints we eliminate the $L - 1$ variables $(p_l)_{l \in \mathcal{L} \setminus \{i\}}$ and transform problem (2.47) into a single variable optimization problem (with the variable p_i). This facilitates finding the optimal power $p_i^{\star}$ (and implicitly $\bar{\gamma}_i^{\star}$), in an efficient and straightforward manner.

For a detailed description of the above method it is useful to introduce a virtual network obtained from the original network by removing the ith link. Such a network is referred to as *reduced* network. For notational convenience let us define the following vectors and matrices associated to the reduced network: $\bar{\mathbf{p}}_i$ and $\boldsymbol{\gamma}_{\min,i}$ are obtained from $\mathbf{p}$ and $\boldsymbol{\gamma}_{\min}$ by removing the ith entries, i.e., $\bar{\mathbf{p}}_i = (p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_L)$ and $\bar{\boldsymbol{\gamma}}_{\min,i} = (\gamma_{1,\min}, \ldots, \gamma_{i-1,\min}, \gamma_{i+1,\min}, \ldots, \gamma_{L,\min})$; similarly, $\bar{\mathbf{B}}_i(\bar{\boldsymbol{\gamma}}_{\min,i})$ and $\bar{\mathbf{G}}_i$ are obtained from $\mathbf{B}(\boldsymbol{\gamma}_{\min})$ and $\mathbf{G}$ [see (2.44)] by removing the ith rows and the ith columns. It is important to note that if SINR

vector $\boldsymbol{\gamma}_{\min}$ is achievable in the original network then $\bar{\boldsymbol{\gamma}}_{\min,i}$ is also achievable in the reduced network.

Now we turn to problem (2.47). By rearranging the terms, the equality constraints can be expressed compactly as

$$[\mathbf{I} - \bar{\mathbf{B}}_i(\bar{\boldsymbol{\gamma}}_{\min,i})\bar{\mathbf{G}}_i]\bar{\mathbf{p}}_i + \mathbf{d}_i(\bar{\boldsymbol{\gamma}}_{\min,i})p_i = \sigma^2 \bar{\mathbf{B}}_i(\bar{\boldsymbol{\gamma}}_{\min,i})\mathbf{1} \ , \qquad (2.49)$$

where

$$\mathbf{d}_i(\bar{\boldsymbol{\gamma}}_{\min,i}) = -\left(\frac{g_{i1}}{g_{11}}\gamma_{1,\min}, \ldots, \frac{g_{ii-1}}{g_{i-1i-1}}\gamma_{i-1,\min}, \frac{g_{ii+1}}{g_{i+1i+1}}\gamma_{i+1,\min}, \ldots, \frac{g_{iL}}{g_{LL}}\gamma_{L,\min} \right) .$$

Similarly to (2.45), let us denote

$$\bar{\mathbf{A}}_i(\bar{\boldsymbol{\gamma}}_{\min,i}) = \mathbf{I} - \bar{\mathbf{B}}_i(\bar{\boldsymbol{\gamma}}_{\min,i})\bar{\mathbf{G}}_i \ ; \ \ \bar{\mathbf{b}}_i(\bar{\boldsymbol{\gamma}}_{\min,i}) = \sigma^2 \bar{\mathbf{B}}_i(\bar{\boldsymbol{\gamma}}_{\min,i})\mathbf{1} \qquad (2.50)$$

and rewrite (2.49) equivalently as

$$\bar{\mathbf{A}}_i(\bar{\boldsymbol{\gamma}}_{\min,i})\bar{\mathbf{p}}_i + \mathbf{d}_i(\bar{\boldsymbol{\gamma}}_{\min,i})p_i = \bar{\mathbf{b}}_i(\bar{\boldsymbol{\gamma}}_{\min,i}) \ . \qquad (2.51)$$

Since $\boldsymbol{\gamma}_{\min} \in \mathcal{G}$ it follows that the SINR vector $\bar{\boldsymbol{\gamma}}_{\min,i} > \mathbf{0}$ is achievable in the reduced network. Thus, Theorem 2.2 (applied to the reduced network) implies that the spectral radius of the matrix $\bar{\mathbf{B}}_i(\bar{\boldsymbol{\gamma}}_{\min,i})\bar{\mathbf{G}}_i$ is strictly smaller than one, i.e., $\rho(\bar{\mathbf{B}}_i(\bar{\boldsymbol{\gamma}}_{\min,i})\bar{\mathbf{G}}_i) < 1$. This, in turn, ensures that matrix $\bar{\mathbf{A}}_i(\bar{\boldsymbol{\gamma}}_{\min,i})$ is invertible and its inverse has nonnegative entries, i.e., $\bar{\mathbf{A}}_i^{-1}(\bar{\boldsymbol{\gamma}}_{\min,i}) \geq \mathbf{0}$ [174, Th. 2.5.3, items 2 and 17]. Therefore, we can parameterize all solutions of (2.49), using p_i as a free parameter [3, Sec. C.5, p. 681]. Thus, we obtain

$$\begin{bmatrix} \bar{\mathbf{p}}_i \\ p_i \end{bmatrix} = \begin{bmatrix} -\bar{\mathbf{A}}_i^{-1}(\bar{\boldsymbol{\gamma}}_{\min,i})\mathbf{d}_i(\bar{\boldsymbol{\gamma}}_{\min,i}) \\ 1 \end{bmatrix} p_i + \begin{bmatrix} \bar{\mathbf{A}}_i^{-1}(\bar{\boldsymbol{\gamma}}_{\min,i})\bar{\mathbf{b}}_i(\bar{\boldsymbol{\gamma}}_{\min,i}) \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} \bar{\mathbf{q}}_i \\ q_i \end{bmatrix} p_i + \begin{bmatrix} \bar{\mathbf{s}}_i \\ s_i \end{bmatrix} \ , \qquad (2.52)$$

where $q_i{=}1$, $s_i{=}0$, $\bar{\mathbf{q}}_i{=}-\bar{\mathbf{A}}_i^{-1}(\bar{\boldsymbol{\gamma}}_{\min,i})\mathbf{d}_i(\bar{\boldsymbol{\gamma}}_{\min,i})$, and $\bar{\mathbf{s}}_i{=}\bar{\mathbf{A}}_i^{-1}(\bar{\boldsymbol{\gamma}}_{\min,i})\bar{\mathbf{b}}_i(\bar{\boldsymbol{\gamma}}_{\min,i})$. The vectors $\bar{\mathbf{q}}_i$ and $\bar{\mathbf{s}}_i$ are introduced for notational simplicity and they have the following structure:

$$\bar{\mathbf{q}}_i = (q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_L) \ ; \ \ \bar{\mathbf{s}}_i = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_L) \ .$$

Furthermore, since $\bar{\mathbf{A}}_i^{-1}(\bar{\boldsymbol{\gamma}}_{\min,i}) \geq \mathbf{0}$ and by noting that $\mathbf{d}_i(\bar{\boldsymbol{\gamma}}_{\min,i}) \leq \mathbf{0}$ and $\mathbf{b}_i(\bar{\boldsymbol{\gamma}}_{\min,i}) \geq \mathbf{0}$ [see (2.50)], we can see that all entries in vectors $\bar{\mathbf{q}}_i$ and $\bar{\mathbf{s}}_i$ are

nonnegative, $\bar{\mathbf{q}}_i \geq \mathbf{0}$ and $\bar{\mathbf{s}}_i \geq \mathbf{0}$. Finally, we can rewrite parametrization (2.52) as

$$p_j = q_j p_i + s_j, \ j \in \mathcal{L} \ , \tag{2.53}$$

where $q_j \geq 0$, $s_j \geq 0$ for all $j \in \mathcal{L}$, and $q_i = 1$, $s_i = 0$.

Next we use parametrization (2.53) to convert problem (2.47) (with L power variables) into an equivalent one with a single power variable p_i. To do this, we first express the objective function of problem (2.47) $g_i(\mathbf{p})$ as a function of single variable p_i, i.e.,

$$g_i(\mathbf{p}) = \frac{g_{ii} p_i}{\sigma^2 + \sum\limits_{j \neq i} g_{ji} p_j} = \frac{g_{ii} p_i}{\sigma^2 + \sum\limits_{j \neq i} g_{ji} (q_j p_i + s_j)} = \bar{g}_i(p_i) \ . \tag{2.54}$$

The sum-power constraints of problem (2.47) (i.e., $\sum_{l \in \mathcal{O}(n)} p_l \leq p_n^{\max}$, $n \in \mathcal{T}$) can be expressed as

$$p_i \ \leq \ \frac{p_n^{\max} - \sum_{l \in \mathcal{O}(n)} s_l}{\sum_{l \in \mathcal{O}(n)} q_l}, \ n \in \mathcal{T} \ . \tag{2.55}$$

Furthermore, since $q_j \geq 0$, $s_j \geq 0$, all L nonnegativity power constraints of problem (2.47) can be replaced by $p_i \geq 0$, i.e., $p_i \geq 0$ in parametrization (2.53) implies that $p_j \geq 0$ for all $j \in \mathcal{L}$. Recall that we consider the nontrivial case $\mathbf{a}_i \notin \mathcal{G}$, and therefore the first inequality constraint of problem (2.47) can be safely dropped, and therefore problem (2.47) can be expressed equivalently as

$$
\begin{aligned}
\text{maximize} \quad & \bar{g}_i(p_i) \\
\text{subject to} \quad & p_i \leq \frac{p_n^{\max} - \sum_{l \in \mathcal{O}(n)} s_l}{\sum_{l \in \mathcal{O}(n)} q_l}, \ n \in \mathcal{T} \\
& p_i \geq 0 \ ,
\end{aligned}
\tag{2.56}
$$

where the variable is p_i. By recalling that $s_l \geq 0$ for all $\in \mathcal{L}$, it is easy to see that the first derivative of the objective function $\bar{g}_i(p_i)$ is strictly positive. Hence, the maximum $\bar{g}_i(p_i)$ can be found by increasing p_i until one power constraint become active. Thus, in the case of $\mathbf{a}_i \notin \mathcal{G}$, we have

$$p_i^\star = \min_{n \in \mathcal{T}} \frac{p_n^{\max} - \sum_{l \in \mathcal{O}(n)} s_l}{\sum_{l \in \mathcal{O}(n)} q_l} \tag{2.57}$$

and we can express the optimal $\bar{\gamma}_i^\star$ as $\bar{\gamma}_i^\star = \bar{g}_i(p_i^\star)$. Hence, the general solution of problem (2.47) can be expressed as

$$\bar{\gamma}_i^\star = \begin{cases} \gamma_{i,\max} & \mathbf{a}_i \in \mathcal{G} \\ \bar{g}_i(p_i^\star) & \text{otherwise} \ . \end{cases} \tag{2.58}$$

Note that, the proposed method for checking $\boldsymbol{\gamma}_{\min} \in \mathcal{G}$ (i.e., Algorithm 2.2) can be readily applied to check the condition $\mathbf{a}_i \in \mathcal{G}$ in (2.58) as well.

Improved upper bound

Based on monotonicity of f_0, L tighter upper bounds can be easily obtained by evaluating f_0 at the vertices of $\bar{\mathcal{Q}}^\star$ adjacent to $\boldsymbol{\gamma}_{\min}$. Specifically, they are given by $f_0(\bar{\mathbf{a}}_l)$, $l \in \mathcal{L}$, where $\bar{\mathbf{a}}_l = \boldsymbol{\gamma}_{\min} + (\bar{\gamma}_l^\star - \gamma_{l,\min})\mathbf{e}_l$ [see $\bar{\mathbf{a}}_1$ and $\bar{\mathbf{a}}_2$ in Figures 2.4(b) and 2.4(c)]. Note that the values $\bar{\gamma}_l^\star, \in \mathcal{L}$ have already been found for computing the improved lower bound $\phi_{\mathrm{lb}}^{\mathrm{Imp}}(\mathcal{Q})$ (2.46). Let $l^\star$ be the index of the vertex which provide the best (smallest) upper bound, i.e., $l^\star = \arg\min_{l \in \mathcal{L}} f_0(\bar{\mathbf{a}}_l)$. Thus, our first improved upper bound is given by

$$\phi_{\mathrm{ub}}^{\mathrm{Imp}}(\mathcal{Q}) = \begin{cases} f_0(\bar{\mathbf{a}}_{l^\star}) & \boldsymbol{\gamma}_{\min} \in \mathcal{G} \\ 0 & \text{otherwise} . \end{cases} \tag{2.59}$$

The following lemma ensures that $\phi_{\mathrm{ub}}^{\mathrm{Imp}}(\mathcal{Q})$ is tighter than the basic upper bound $\phi_{\mathrm{ub}}^{\mathrm{Basic}}(\mathcal{Q})$.

Lemma 2.4. *For any $\mathcal{Q} \subseteq \mathcal{Q}_{\mathrm{init}}$ and $\check{\boldsymbol{\gamma}} \in \mathcal{G} \cap \mathcal{Q}$ we have $\phi_{\min}(\mathcal{Q}) \leq f_0(\check{\boldsymbol{\gamma}}) \leq f_0(\boldsymbol{\gamma}_{\min}) = \phi_{\mathrm{lb}}^{\mathrm{Basic}}(\mathcal{Q})$.*

Proof. First note from (2.48) that, $\phi_{\min}(\mathcal{Q}) = \inf_{\boldsymbol{\gamma} \in \mathcal{G} \cap \mathcal{Q}} f_0(\boldsymbol{\gamma})$. Moreover, by noting that $\check{\boldsymbol{\gamma}} \in \mathcal{G} \cap \mathcal{Q}$, we have $\inf_{\boldsymbol{\gamma} \in \mathcal{G} \cap \mathcal{Q}} f_0(\boldsymbol{\gamma}) \leq f_0(\check{\boldsymbol{\gamma}})$ and since $\boldsymbol{\gamma}_{\min} \leq \check{\boldsymbol{\gamma}}$ and f_0 is monotonically decreasing in each dimension, we have $f_0(\check{\boldsymbol{\gamma}}) \leq f_0(\boldsymbol{\gamma}_{\min})$. Thus, we can combine these relations together and the result follows. $\square$

We can further improve the previously obtained bound by using efficient local optimization techniques. Specifically, we can use as an initial point $\boldsymbol{\gamma} = \bar{\mathbf{a}}_{l^\star}$ and (locally) minimize $f_0(\boldsymbol{\gamma})$ subject to $\boldsymbol{\gamma} \in \mathcal{G} \cap \mathcal{Q}$, i.e.,

$$\begin{aligned} \text{minimize} \quad & f_0(\boldsymbol{\gamma}) \\ \text{subject to} \quad & \boldsymbol{\gamma} \in \mathcal{G} \cap \mathcal{Q} , \end{aligned} \tag{2.60}$$

where the variable is $\boldsymbol{\gamma}$. Let us denote the obtained local optimum by $\boldsymbol{\gamma}_{\mathrm{ImpCGP}}$. Thus, our second improved upper bound is given by

$$\phi_{\mathrm{ub}}^{\mathrm{ImpCGP}}(\mathcal{Q}) = \begin{cases} f_0(\boldsymbol{\gamma}_{\mathrm{ImpCGP}}) & \boldsymbol{\gamma}_{\min} \in \mathcal{G} \\ 0 & \text{otherwise} . \end{cases} \tag{2.61}$$

One simple approach to efficiently compute $\boldsymbol{\gamma}_{\text{ImpCGP}}$ via complementary geometric programming (or CGP) [97] is presented in Appendix 2.

Since all improved bounds are tighter than the basic ones (see Lemma 2.3 and Lemma 2.4), any possible combination of a lower and an upper bound pair must also satisfy the conditions **C1** and **C2**. This ensures the convergence of the proposed Algorithm 2.1.

2.4   Extensions to multicast networks

In this section we consider the problem of WSRMax in multicast networks [i.e., problem (2.4)] and show how Algorithm 2.1 can be adapted to find the solution of problem (2.4). By noting the monotonically increasing property of $\log(\cdot)$ function, problem (2.4) can be expressed in the following equivalent form:

$$
\begin{aligned}
\text{maximize} \quad & \sum_{n\in\mathcal{T}}\sum_{m=1}^{M_n}\beta_n^m\log\big(1+\min_{l\in\mathcal{O}^m(n)}\text{SINR}_n^{ml}(\mathbf{p})\big) \\
\text{subject to} \quad & \sum_{m=1}^{M_n}p_n^m\leq p_n^{\max},\ n\in\mathcal{T} \\
& p_n^m\geq 0,\ n\in\mathcal{T},\ m=1,\dots,M_n\ ,
\end{aligned}
\tag{2.62}
$$

where the variable is $(p_n^m)_{n\in\mathcal{T},m=1,\dots,M_n}$. By introducing auxiliary variables $\gamma_n^m,\ n\in\mathcal{T}, m=1,\dots,M_n$, we can equivalently express problem (2.62) as

$$
\begin{aligned}
\text{minimize} \quad & \sum_{n\in\mathcal{T}}\sum_{m=1}^{M_n}-\beta_n^m\log(1+\gamma_n^m) \\
\text{subject to} \quad & \gamma_n^m\leq\text{SINR}_n^{ml}(\mathbf{p}),\ \begin{matrix} n\in\mathcal{T},\ m=1,\dots,M_n, \\ l\in\mathcal{O}^m(n) \end{matrix} \\
& \sum_{m=1}^{M_n}p_n^m\leq p_n^{\max},\ n\in\mathcal{T} \\
& p_n^m\geq 0,\ n\in\mathcal{T},\ m=1,\dots,M_n\ ,
\end{aligned}
\tag{2.63}
$$

where the variables are $(p_n^m)_{n\in\mathcal{T},m=1,\dots,M_n}$ and $(\gamma_n^m)_{n\in\mathcal{T},m=1,\dots,M_n}$. A close comparison of problems (2.63) and (2.5) reveals that they have a very similar structure. Therefore, the proposed branch and bound method (i.e., Algorithm 2.1) can be directly applied to solve problem (2.63) by redefining appropriately the following sets and functions.

1. $\boldsymbol{\gamma}=(\gamma_1,\dots,\gamma_L)$ is replaced by $\boldsymbol{\gamma}=(\gamma_n^m)_{n\in\mathcal{T},m=1,\dots,M_n}$.
2. $f_0(\boldsymbol{\gamma})$ is replaced by $\tilde{f}_0(\boldsymbol{\gamma})$, where $\tilde{f}_0(\boldsymbol{\gamma})=\sum_{n\in\mathcal{T}}\sum_{m=1}^{M_n}-\beta_n^m\log(1+\gamma_n^m)$.

3. $\mathcal{G}$ is replaced by $\tilde{\mathcal{G}}$, where

$$
\tilde{\mathcal{G}} = \left\{ \boldsymbol{\gamma} \left|
\begin{array}{ll}
\gamma_n^m \leq \mathrm{SINR}_n^{ml}(\mathbf{p}), & n \in \mathcal{T},\ m = 1, \ldots, M_n, \\
& l \in \mathcal{O}^m(n) \\
\sum_{m=1}^{M_n} p_n^m \leq p_n^{\max}, & n \in \mathcal{T} \\
p_n^m \geq 0, & n \in \mathcal{T},\ m = 1, \ldots, M_n
\end{array}
\right. \right\} .
$$

4. $\mathcal{Q}_{\mathrm{init}}$ is replaced by $\tilde{\mathcal{Q}}_{\mathrm{init}}$, where

$$
\tilde{\mathcal{Q}}_{\mathrm{init}} = \left\{ \boldsymbol{\gamma} \left| \ 0 \leq \gamma_n^m \leq \frac{\min\limits_{l \in \mathcal{O}^m(n)} g_{ll}}{\sigma^2}\ p_n^{\max},\ n \in \mathcal{T},\ m = 1, \ldots, M_n \right. \right\} .
$$

5. $\mathcal{Q}$ is replaced by $\tilde{\mathcal{Q}}$, where

$$
\tilde{\mathcal{Q}} = \left\{ \boldsymbol{\gamma} \left| \ \gamma_{n,\min}^m \leq \gamma_n^m \leq \gamma_{n,\max}^m,\ n \in \mathcal{T},\ m = 1, \ldots, M_n \right. \right\} .
$$

Note that the definitions of the lower and upper bound functions provided in the case of singlecast networks [i.e., (2.26), (2.27), (2.46), and (2.59)] are applicable in the case of multicast networks as well. However, instead of the proposed efficient methods based on M-matrix theory [174, p. 112] for checking $\boldsymbol{\gamma} \in \mathcal{G}$ (see Algorithm 2.2) and for evaluating $\bar{\gamma}_i^\star$ [see (2.58)], in the case of multicast networks, we have to rely on a linear programming (LP) or a GP solver.

2.5   Numerical examples

In this section we first compare the impact of the proposed lower bounds and upper bounds (Section 2.3) on the convergence of the proposed branch and bound method (Algorithm 2.1 in Section 2.2). Next, we provide various applications of Algorithm 2.1 and numerical examples for the considered applications. In summary, those applications include: sum-rate maximization in singlecast wireless networks, the problem of maximum weighted link scheduling for wireless multihop networks [12, Sec. III-B,V-A],[38, Sec. 4], cross-layer control policies for network utility maximization (NUM) in multihop wireless networks [17, Sec. 5], finding achievable rate regions in singlecast, as well as in multicast, wireless networks.

To simplify the presentation we use the abbreviations: $\mathrm{LB}_{\mathrm{Basic}}$ for the basic lower bound given in (2.26), $\mathrm{UB}_{\mathrm{Basic}}$ for the basic upper bound given in (2.27),
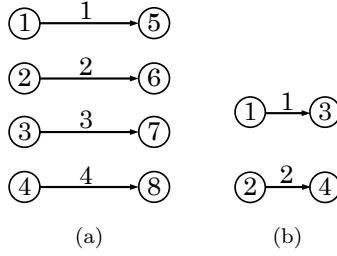
Fig 2.5. (a) Bipartite network, degree 1, $N = 8$, $L = 4$; **(b) Bipartite network, degree** 1, $N = 4$, $L = 2$, **[83] © 2011, IEEE.**

LB_{Imp} for the improved lower bound given in (2.46), UB_{Imp} for the improved upper bound given in (2.59), and $\text{UB}_{\text{ImpCGP}}$ for the improved upper bound given in (2.61).

## 2.5.1 Impact of different lower bounds and upper bounds on BB

To gain insight into the impact of the proposed lower and upper bounds on the convergence of Algorithm 2.1, we focus first on the problem of sum-rate maximization in a simple bipartite network of degree 1 [see Figure 2.5(a)]. The channel power gain between distinct nodes are modeled as

$$|h_{ij}|^2 = \mu^{|i-j|} c_{ij}, \ i, j \in \mathcal{L} \ , \tag{2.64}$$

where c_{ij}s are small-scale fading coefficients and the scalar $\mu \in [0, 1]$ is referred to as the interference coupling index, which parameterizes the interference between direct links. The fading coefficients are assumed to be exponentially distributed independent random variables to model Rayleigh fading. An arbitrarily generated set $\acute{\mathcal{C}}$ of fading coefficients, where $\acute{\mathcal{C}} = \{c_{ij} \mid i, j \in \mathcal{L}\}$ is referred to as a *single fading realization*; we use a discrete argument t sometimes, to indicate the fading realization index. For example $\acute{\mathcal{C}}(t)$ represents the tth fading realization. We define the signal-to-noise ratio (SNR) operating point as ($p_n^{\max} = p_0^{\max}$ for all $n \in \mathcal{T}$)

$$\text{SNR} = \frac{p_0^{\max}}{\sigma^2} \ . \tag{2.65}$$

We consider first the nonfading case, i.e., $c_{ij} = 1, i, j \in \mathcal{L}$, and the proposed

64
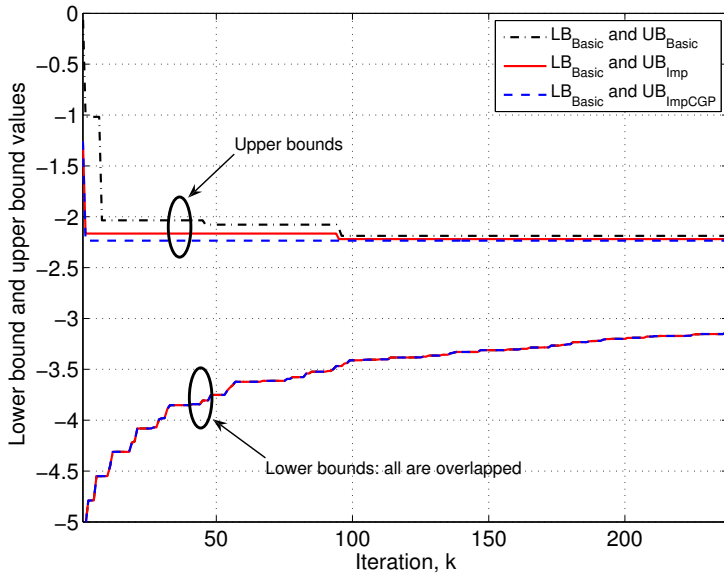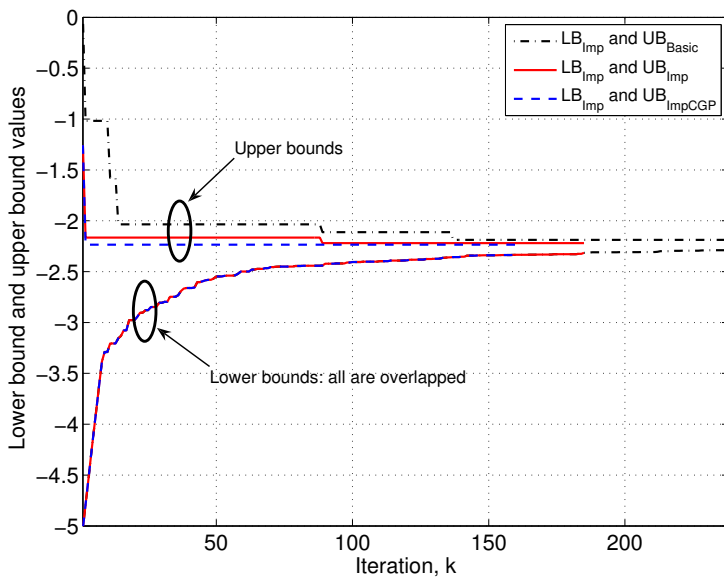
Fig 2.6. Evolution of lower and upper bounds: (a) Basic lower bound in conjunction with all upper bounds; (b) Improved lower bound in conjunction with all upper bounds, [83] © 2011, IEEE.

Algorithm 2.1 was run with all possible combinations of the proposed lower and upper bound pairs. Figure 2.6 shows the evolution of the upper and lower bounds for the optimal value of problem (2.5) [9] for SNR = 15dB, $\mu = 0.25$, and $\beta_l = 0.25$ for all $l \in \mathcal{L}$. Specifically in Figure 2.6(a), we used the basic lower bound LB_{Basic} in conjunction with all proposed upper bounds and in Figure 2.6(b) we used the improved lower bound LB_{Imp} in conjunction with all proposed upper bounds. The results show that the convergence speed of Algorithm 2.1 can be substantially increased by improving the lower bound whilst the tightness of the upper bound has a much reduced impact. Note that this is in general the behavior of a branch and bound method, where an approximative solution can be found relatively fast but certifying it typically takes a much larger number of iterations [82]. Note that in both Figure 2.6(a) and Figure 2.6(b) the evolution of lower bounds is independent of the upper bound used. This is due to the fact that in each iteration the branching mechanism depends only on the lower bound.

In order to provide a statistical description of the speed of convergence we turn to the fading case and run Algorithm 2.1 for a large number of fading realizations. For each one we store the number of iterations and the total CPU time required to find the optimal value of problem (2.5) within an accuracy of $\epsilon = 10^{-1}$ for SNR = 15dB, $\mu = 0.25$, and $\beta_l = 0.25$ for all $l \in \mathcal{L}$. Figure 2.7 shows the empirical cumulative distribution function (CDF) plots of the total number of iterations [Figure 2.7(a)] and the total CPU time [Figure 2.7(b)] for all possible combinations of lower and upper bounds pairs. Figure 2.7(a) shows that, irrespective of the upper bound we use, the improved lower bound LB_{Imp} provides a remarkable reduction in the total number of iterations when compared to LB_{Basic}. Results further show that, even though the improved upper bound UB_{ImpCGP} makes use of advanced optimization techniques, such as complementary geometric programming (see Algorithm 2.0.1, Appendix 2), the benefits from UB_{ImpCGP} over the improved upper bound UB_{Imp} is marginal in terms of the total number of iterations. In terms of the total CPU time [Figure 2.7(b)], significant improvements are often achieved by using the lower and upper bound pairs (LB_{Imp}, UB_{Imp}) and (LB_{Imp}, UB_{Basic}). Interestingly, the lower and upper bound pair (LB_{Imp}, UB_{ImpCGP}) performs very

---

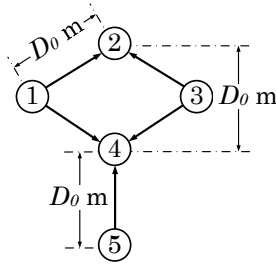[9]The optimal value of problem (2.5) is the negative of the optimal value of problem (2.2).
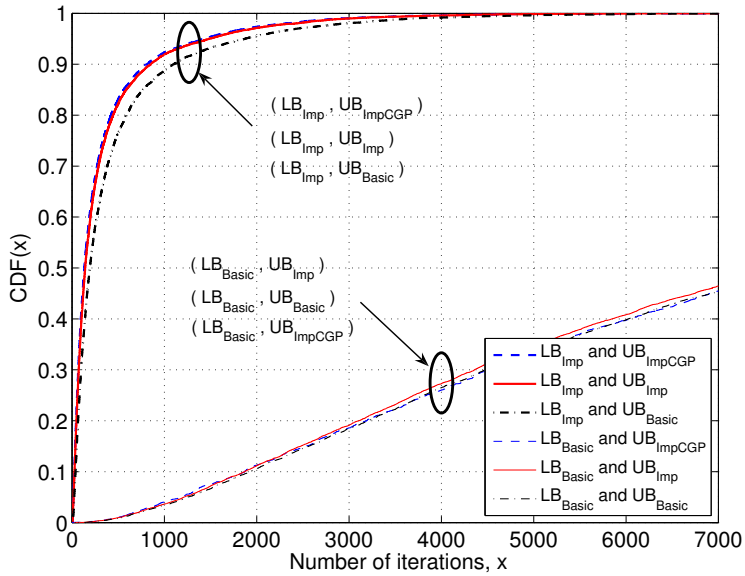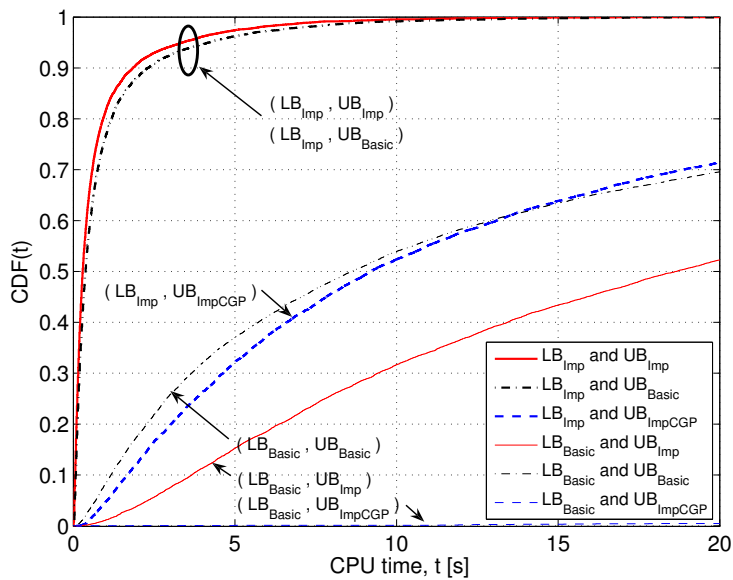
Fig 2.7. Empirical CDF plots of: (a) Total number of iterations; (b) Total CPU time, [83] © 2011, IEEE.

Fig 2.8. Bipartite network, degree 3, $N = 5$, $L = 5$, **[83] © 2011, IEEE.**

poorly. This behavior is due to the complexity of step 2 of Algorithm 2.0.1, where we have to rely on a GP solver.

Therefore, in all of the following numerical examples, Algorithm 2.1 is run with the lower and upper bound pair (LB_{Imp}, UB_{Imp}), unless otherwise specified.

2.5.2 Sum-rate maximization in singlecast wireless networks

Let us now consider the problem of sum-rate maximization in a bipartite singlecast network. To evaluate the benefits from multipacket transmit/receive capabilities of nodes, we chose a network setup with degree 3, as shown in Figure 2.8. The network is symmetric and the distances between nodes are chosen as shown in the figure. We assume an exponential path loss model, where the channel power gains between distinct nodes are given by

$$|h_{ij}|^2 = \left(\frac{d_{ij}}{d_0}\right)^{-\eta} c_{ij} \; , \tag{2.66}$$

where d_{ij} is the distance from the transmitter of link i to the receiver of link j, d_0 is the *far field reference distance* [176], η is the path loss exponent, and c_{ij} are defined similarly to their use in (2.64). Note that the interference coefficients g_{ij} s are chosen as we discussed in Section 2.1. The first term of (2.66) represents the path loss factor and the second term models Rayleigh small-scale fading. The SNR operating point is defined as ($p_n^{\max} = p_0^{\max}$ for all $n \in \mathcal{T}$)

$$\text{SNR} = \frac{p_0^{\max}}{\sigma^2} \left(\frac{D_0}{d_0}\right)^{-\eta} \; . \tag{2.67}$$

In the following simulations we set $D_0/d_0 = 10$ and $\eta = 4$.

68
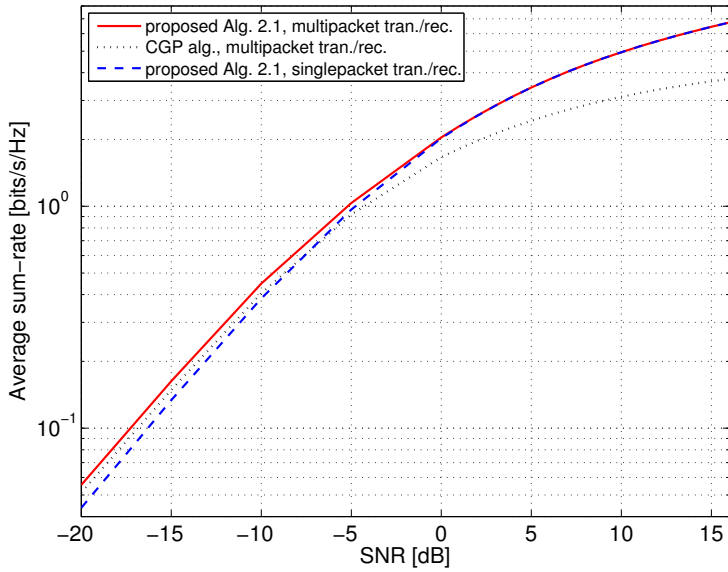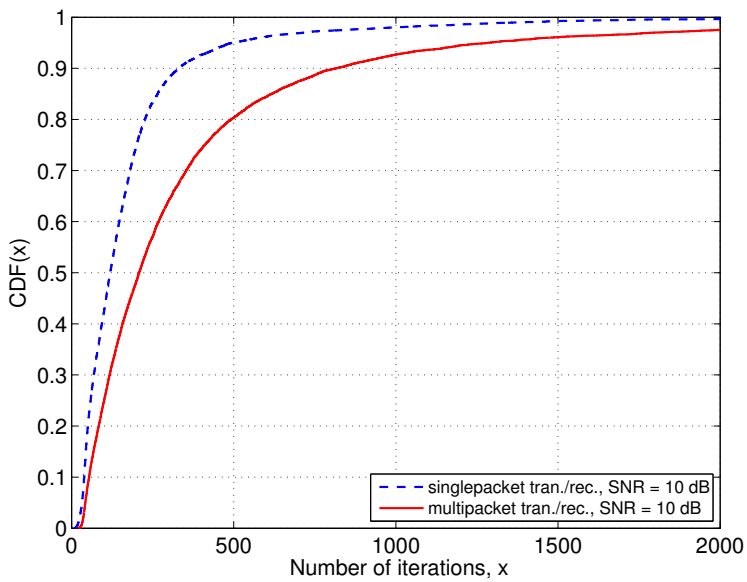
Fig 2.9. (a) Dependence of the average sum-rate on SNR; (b) Empirical CDF of the total number of iterations, [83] © 2011, IEEE.

Figure 2.9(a) shows the dependence of average sum-rate (i.e., $\beta_l = 1$ for all $l \in \mathcal{L}$) on the SNR. Results show that the average sum-rate, in the case of multipacket transmission/reception, is always better than, or equal to, the case of singlepacket transmission/reception and the performance gap increases as SNR decreases. However, as expected for practical SNR values, the benefits of multipacket transmission/reception are negligible when the receivers perform singleuser detection [9]. For comparison, we also plot the result obtained from a suboptimal solution method based on complementary geometric programming [5, 97, 98]. We refer to this suboptimal method as *CGP algorithm* for the rest of the section. Note that, CGP algorithm is equivalent to running Algorithm 2.0.1 (Appendix 2) with $\mathcal{Q} = \mathcal{Q}_{\text{init}}$ and a proper initialization $\hat{\gamma}$. Specifically, we found the initial $\hat{\gamma}_l$, $l \in \mathcal{L}$ according to (2.42) by using a uniform feasible power allocation, which will be referred to as uniform initialization in the rest of the section. Let us first focus on CGP performance in the case of multipacket transmission/reception. Results show that there is a significant performance loss due to the suboptimality of CGP algorithm, especially for SNR > 0dB. In the case of singlepacket transmission/reception, the average sum-rate that is obtained by using CGP algorithm is almost zero, irrespective of the SINR and not plotted in Figure 2.9(a) to preserve clarity. Results confirm that CGP algorithm cannot handle the huge imbalance between interference coefficient values [10].

Figure 2.9(b) shows the empirical CDF plots of the total number of iterations required to find the sum-rate by using Algorithm 2.1, which gives insight into the complexity of Algorithm 2.1. The plots are for the case of SNR = 10dB and $\epsilon = 10^{-3}$. Roughly speaking, results show that the total number of iterations required in the case of singlepacket transmission/reception is smaller compared to the case of multipacket transmission/reception.

2.5.3   Maxweight scheduling in multihop wireless networks

Next, we consider a multihop wireless network, where the nodes have only singlepacket transmit/receive capability and no node can transmit and receive simultaneously. In such setups the WSRMax problem is equivalent to the maxi-

---

[10]Recall from Figure 2.2(a) and 2.2(b) that, if nodes have singlepacket transmitter/receiver capabilities, then some of the interference coefficients are infinite.

(a)

(b)
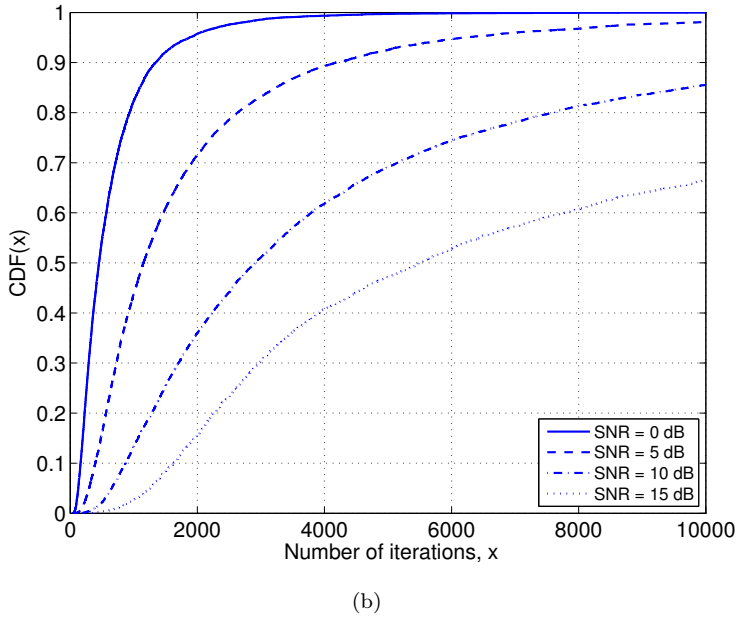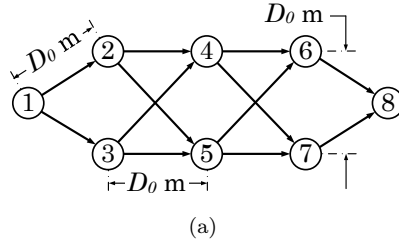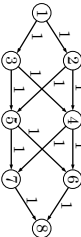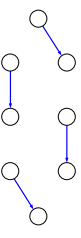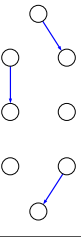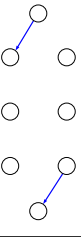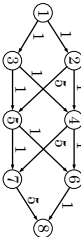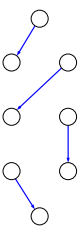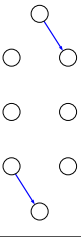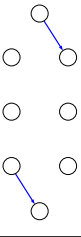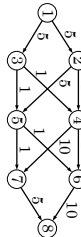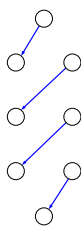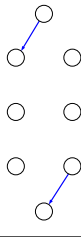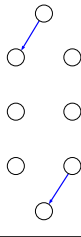
Fig 2.10. (a) Multihop network, $N = 8$, $L = 12$; (b) Empirical CDF of the total number of iterations, [83] © 2011, IEEE.

mum weighted matching [11] (MWM) problem [54]. Polynomial time algorithms are available for the problem in the case of fixed link rates [54],[38, Sec. 4.2]. To the best of our knowledge, there are no known solution methods for the MWM problem when the link rates depend on the power allocation of all other links. In such cases, it is worth noting that our proposed algorithm is able to find the MWM.

To show this, we use the symmetric multihop wireless network shown in Figure 2.10(a). The channel power gains, between nodes are given by (2.66) and

---

[11]Borrowing terminology from graph theory, a matching is a set of links, no two of which share a node [54].

Table 2.1. Maximum weighted matchings, [83] © 2011, IEEE.

Associated weights of the links	SNR [dB]	−10	0	5	10

72

the SNR operating point is given by (2.67). In the following simulations we set $D_0/d_0 = 10$ and $\eta = 4$.

Table 2.1 shows MWMs obtained for different link weights (see the left most column) and the SNR combinations. Here we consider a nonfading scenario (i.e., $c_{ij} = 1$, $i, j \in \mathcal{L}$) and an accuracy of $\epsilon = 10^{-10}$. Results show that the smaller the SNR, the larger the number of links that are activated simultaneously in the maximum weighted matching. This is intuitively expected since, at low SNR values, node transmission power is small, and therefore the interference generated is very small so that many links are activated simultaneously.

To gain some insight into the computational complexity of the algorithm we plot the CDF of the total number of iterations by running the algorithm for a large number of fading realizations.

Figure 2.10(b) shows the empirical CDF plots of the total number of iterations required to terminate Algorithm 2.1 (or to find the MWM). Plots are drawn for the cases of SNR = 0, 5, 10, and 15dB, $\beta_l = 1$ for all $l \in \mathcal{L}$, and $\epsilon = 10^{-2}$. Results show that the smaller the SNR, the smaller the total number of iterations required to find the MWM. For example, in the case of SNR = 0dB, with probability 0.9, the MWM is found in less than 1500 iterations. However, in the case of SNR = 5dB, with the same probability 0.9, the MWM is found in less than 4000 iterations.

2.5.4   Cross-layer control policies for NUM

In this section we specifically consider the problem of network utility maximization subject to stability constraints [17, Sec. 5]. Let us first revisit briefly the commodity description of the network. Exogenous data arrives at the source nodes and they are delivered to the destination nodes over several, possibly multihop, paths. The data is identified by their destinations, i.e., all data with the same destination are considered as a single commodity, regardless of their source. We label the commodities with integers $s = 1, \ldots, S$ ($S \leq N$). For every node, we define $\mathcal{S}_n \subseteq \{1, \ldots, S\}$ as the set of commodities, which can arrive exogenously at node n. The network is time slotted and at each source node, a set of flow controllers decides the amount of each commodity data admitted in every time slot in the network. Let $x_n^s(t)$ denote the amount of data commodity s admitted in the network at node n during time slot t. It is
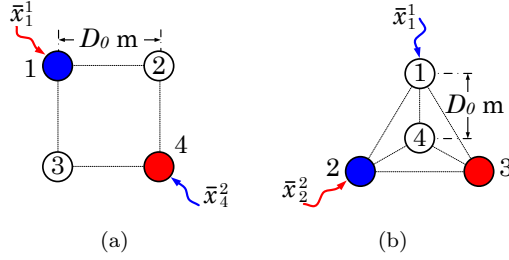
Fig 2.11. (a) Multihop network 1, $N = 4$, **fully connected,** $S = 2$; **(b) Multihop network** 2, $N = 4$, **fully connected,** $S = 2$, **[83] © 2011, IEEE.**

assumed that the data that is successfully delivered to its destination exits the network layer. Associated with each node-commodity pair $(n, s)_{s \in \mathcal{S}_n}$ we define a concave and nondecreasing *utility function* $u_n^s(y)$, representing the "reward" received by sending data of commodity s from node n to node d_s at a long term average rate of y [bits/slot]. Thus, the NUM problem under stability constraints can be formulated as [17, Sec. 5]

$$\begin{aligned} \text{maximize} \quad & \sum_{n \in \mathcal{N}} \sum_{s \in \mathcal{S}_n} u_n^s(y_n^s) \\ \text{subject to} \quad & (y_n^s)_{n \in \mathcal{N}, s \in \mathcal{S}_n} \in \mathbf{\Lambda} \ , \end{aligned} \tag{2.68}$$

where the variable is $(y_n^s)_{n \in \mathcal{N}, s \in \mathcal{S}_n}$ and $\mathbf{\Lambda}$ represents the *network layer capacity region* [17, Def. 3.7].

An arbitrarily close to optimal solutions for problem (2.68) is achieved by a cross-layer control policy, which consists of solving three subproblems: 1) flow control, 2) next-hop routing and in-node scheduling, and 3) RA, during each time slot [17]. The RA subproblem exactly resembles the WSRMax problem (2.2), where the weights are given by the maximum differential backlogs of network links [17]. Here, we implement the cross-layer control algorithm in [17] and, in the third step, we use our proposed Algorithm 2.1 to solve the RA subproblem. The cross-layer control algorithm is simulated for at least $\acute{T} = 10000$ time slots, and the average rates $\bar{x}_n^s$ are computed by averaging the last $t_0 = 3000$ time slots, i.e., $\bar{x}_n^s = 1/t_0 \sum_{t=\acute{T}-t_0}^{\acute{T}} x_n^s(t)$. We assume that the rates corresponding to all node-commodity pairs $(n, s)_{s \in \mathcal{S}_n}, n \in \mathcal{N}$ are subject to proportional fairness, and therefore we select the utility functions $u_n^s(y) = \log_e(y)$. For a detailed description of the cross-layer control policy [17] the reader may refer to Section 3.1.2.

Two fully connected multihop wireless network setups, as shown in Figure 2.11

are considered, where all nodes have multipacket transmit/receieve capability and no node can transmit and receive simultaneously. Each of the networks consist of four nodes (i.e., $N = 4$) and two commodities, which arrive exogenously at source nodes. In the case of the first network setup, shown in Figure 2.11(a), commodity 1 arrives exogenously at node 1, and is intended for node 4; commodity 2 arrives exogenously at node 4, and is intended for node 1. Nodes are located in a square grid such that the horizontal and the vertical distance between adjacent nodes are D_0 meters [m]. In the case of the second network setup, shown in Figure 2.11(b), commodity 1 arrives exogenously at node 1, and is intended for node 2; commodity 2 arrives exogenously at node 2, and is intended for node 3. Nodes are located such that three of them form an equilateral triangle and the fourth one is located at its center [see Figure 2.11(b)]. It is assumed that the distance from the middle node to any other is D_0 m. The channel power gains are given by (2.66) and SNR operating point is given by (2.67). We set $D_0/d_0 = 10$ and $\eta = 4$ in the following simulation.

Figure 2.12 shows the dependence of the average network layer sum-rate on the SNR for the considered network setups. As a reference, we first consider a suboptimal and more restrictive RA policy, where only one link can be activated during each time slot. This policy is called base line single link activation (BLSLA); BLSLA policy can be easily found and it consists of activating, during each time slot, only the link which achieves the maximum weighted rate. Other suboptimal RA policy is based on CGP algorithm (see Section 2.5.2). Specifically, we use two initialization methods for CGP algorithm: 1) the initial $\hat{\gamma}_l$, $l \in \mathcal{L}$ is found according to (2.42) by using BLSLA power allocation, 2) the uniform initialization, as discussed in Section 2.5.2.

Results show that the gains obtained by using Algorithm 2.1 are always larger compared to other suboptimal methods. The relative gains achieved by Algorithm 2.1 in the case of network setup 1 [Figure 2.12(a)] are more significant than in the case of network setup 2 [Figure 2.12(b)]. Results further show that, the suboptimal CGP algorithm is very sensitive to initialization. For example, in the case of uniform initialization, CGP algorithm performs extremely poorly compared to the case of BLSLA based initialization. Moreover, in the case of BLSLA based initialization, the suboptimal CGP algorithm can not perform beyond the limits that are achieved by simple BLSLA RA policy.
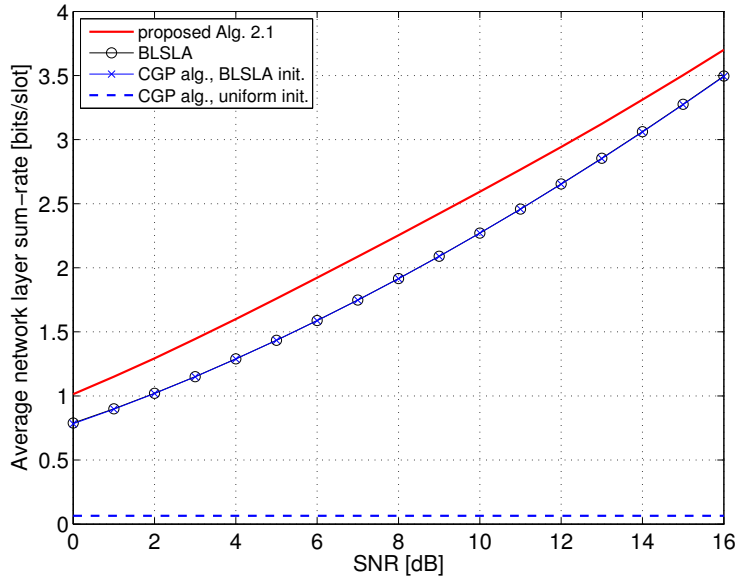
(a) Average network layer sum-rate $\bar{x}_1^1 + \bar{x}_4^2$
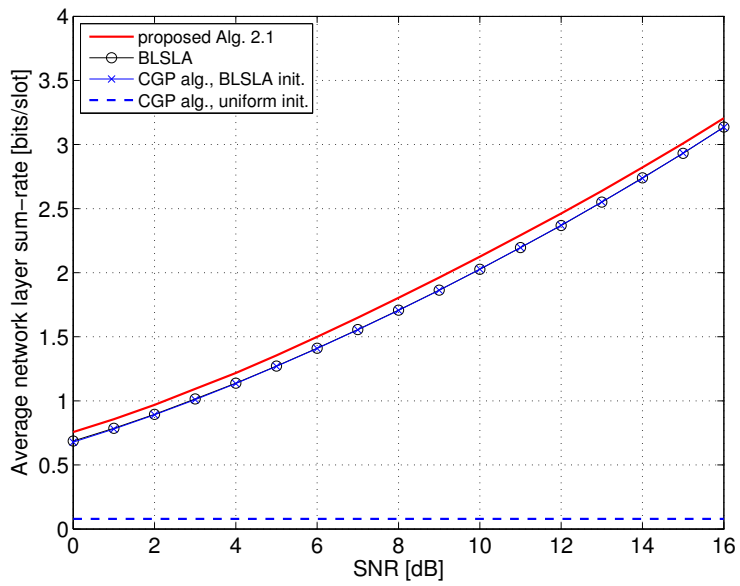
(b) Average network layer sum-rate $\bar{x}_1^1 + \bar{x}_2^2$

Fig 2.12. (a) Dependence of average network layer sum-rate on SNR for network 1; (b) Dependence of average network layer sum-rate on SNR for network 2, [83] © 2011, IEEE.

76

2.5.5 Achievable rate regions in singlecast wireless networks

In this section we illustrate how Algorithm 2.1 can be used to find the achievable rate region in singlecast wireless networks. Recall that we consider the case where all receiver nodes perform singleuser detection, and therefore the achievable rate regions we are referring to are different from the information theoretic capacity regions [177–179]. Note that the information theoretic capacity region is not known, even in the simple case of two interfering links [180].

To facilitate the graphical illustration, we consider a simple bipartite singlecast network of degree 1, as shown in Figure 2.5(b). The channel power gains are given by (2.64) and the SNR operating point is given by (2.65).

We start by defining the *directly achievable rate region*, the *instantaneous rate region*, and the *average rate region* for singlecast wireless networks. Let $\mathcal{R}^{\mathrm{DIR-SC}}(\mu, \acute{\mathcal{C}}(t), p_1^{\max}, p_2^{\max})$ denote the directly achievable rate region for a given interference coupling index μ, a given fading realization [12]

$$\acute{\mathcal{C}}(t) = \{c_{11}(t), c_{12}(t), c_{22}(t), c_{21}(t)\} , \qquad (2.69)$$

and maximum node transmission power $p_1^{\max}$ and $p_2^{\max}$, i.e.,

$$\mathcal{R}^{\mathrm{DIR-SC}}(\mu, \acute{\mathcal{C}}(t), p_1^{\max}, p_2^{\max}) \qquad (2.70)$$

$$= \left\{ (R_1, R_2) \left| \begin{array}{l} R_1 \leq \log\left(1 + \dfrac{c_{11}(t)p_1}{\sigma^2 + \mu c_{21}(t)p_2}\right) \\ R_2 \leq \log\left(1 + \dfrac{c_{22}(t)p_2}{\sigma^2 + \mu c_{12}(t)p_1}\right) \\ 0 \leq p_1 \leq p_1^{\max}, \quad 0 \leq p_2 \leq p_2^{\max} \end{array} \right. \right\} .$$

By invoking a time sharing argument, one can obtain the instantaneous rate region $\mathcal{R}^{\mathrm{INS-SC}}(\mu, \acute{\mathcal{C}}(t), p_1^{\max}, p_2^{\max})$; the convex hull of $\mathcal{R}^{\mathrm{DIR-SC}}(\mu, \acute{\mathcal{C}}(t), p_1^{\max}, p_2^{\max})$. That is,

$$\mathcal{R}^{\mathrm{INS-SC}}(\mu, \acute{\mathcal{C}}(t), p_1^{\max}, p_2^{\max}) = \mathrm{conv}\left\{ \mathcal{R}^{\mathrm{DIR-SC}}(\mu, \acute{\mathcal{C}}(t), p_1^{\max}, p_2^{\max}) \right\} ,$$

where $\mathrm{conv}\{\mathcal{R}\}$ denotes the convex hull of the set $\mathcal{R}$. As noted in [81], since the instantaneous rate region $\mathcal{R}^{\mathrm{INS-SC}}(\mu, \acute{\mathcal{C}}(t), p_1^{\max}, p_2^{\max})$ is convex, any boundary

---

[12]The argument t is used to indicate the fading realization index.

point of the rate region can be obtained by using the solution of an optimization problem in the form of (2.2) with $\beta_1 = \alpha$, $\beta_2 = (1 - \alpha)$ for some $\alpha \in [0, 1]$.

Finally, we define the average rate region $\mathcal{R}^{\text{AVE-SC}}(\mu, p_1^{\max}, p_2^{\max})$ for a given interference coupling index μ and a maximum node transmission power $p_1^{\max}$ and $p_2^{\max}$ as $\mathcal{R}^{\text{AVE-SC}}(\mu, p_1^{\max}, p_2^{\max}) = \frac{1}{\acute{T}} \sum_{t=1}^{\acute{T}} \mathcal{R}^{\text{INS-SC}}(\mu, \acute{\mathcal{C}}(t), p_1^{\max}, p_2^{\max})$, where addition and scalar multiplication of sets is used [13]. The nonnegative integer $\acute{T}$ is the total number of fading realizations we used in averaging. Note that, any boundary point $(R_1^{\text{b}}, R_2^{\text{b}})$ of $\mathcal{R}^{\text{AVE-SC}}(\mu, p_1^{\max}, p_2^{\max})$ is obtained by using the following steps for some $\alpha \in [0, 1]$: 1) solve problem (2.2) with $\beta_1 = \alpha$ and $\beta_2 = 1 - \alpha$ for $\acute{T}$ fading realizations, 2) for each fading realization $t \in \{1, \dots, \acute{T}\}$, evaluate the rate of link 1 and 2 denoted by $r_1(t), r_2(t)$ according to (2.1), and 3) average $r_1(t)$ and $r_2(t)$ over all $\acute{T}$ fading realizations to obtain $R_1^{\text{b}} = \frac{1}{\acute{T}} \sum_{t=1}^{\acute{T}} r_1(t)$ and $R_2^{\text{b}} = \frac{1}{\acute{T}} \sum_{t=1}^{\acute{T}} r_2(t)$.

Figure 2.13(a) shows $\mathcal{R}^{\text{INS-SC}}(\mu, \acute{\mathcal{C}}(t), p_1^{\max}, p_2^{\max})$, the instantaneous rate regions for different values of μ and for an arbitrary chosen fading realization in the case of SNR = 15dB. Specifically, the fading coefficients are $c_{11}(t) = 0.4185$, $c_{12}(t) = 0.3421$, $c_{22}(t) = 0.3700$, and $c_{21}(t) = 1.299$. As a reference, we also plot the directly achievable rate regions $\mathcal{R}^{\text{DIR-SC}}(\mu, \acute{\mathcal{C}}(t), p_1^{\max}, p_2^{\max})$ for all the scenarios considered. Note that the problem of finding any boundary point of $\mathcal{R}^{\text{DIR-SC}}(\mu, \acute{\mathcal{C}}(t), p_1^{\max}, p_2^{\max})$ can be easily cast as a GP, or as a problem of the form (2.47). Results show that the smaller the μ, the larger the rate regions. This is intuitively explained by noting that the smaller the μ, the smaller the interference coefficients, g_{ij} between links, and therefore the higher the rates. Results further show that, when $\mu \geq 0.2$, the directly achievable rate regions become nonconvex, whereas the instantaneous rate region is a triangle referred to as time division multiple access (TDMA) rate region, obtained by time sharing between the maximum rates of R_1 and R_2. Moreover, when $\mu < 0.2$, the instantaneous rate region expands beyond the TDMA rate region and for $\mu \leq 0.01$, the directly achievable rate region almost overlaps with the instantaneous rate region.

Figure 2.13(b) shows the average rate region $\mathcal{R}^{\text{AVE-SC}}(\mu, p_1^{\max}, p_2^{\max})$ for different values of μ in the case of SNR = 15dB. As a reference, we also plot the region obtained by using CGP algorithm to problem (2.2). Results show that

---

[13]For vector sets $\mathcal{A}_1$ and $\mathcal{A}_2$ and scalars α_1, α_2, the set $\alpha_1 \mathcal{A}_1 + \alpha_2 \mathcal{A}_2$ is defined as $\{\alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 \mid \mathbf{a}_1 \in \mathcal{A}_1, \mathbf{a}_2 \in \mathcal{A}_2\}$ [3, p. 38].
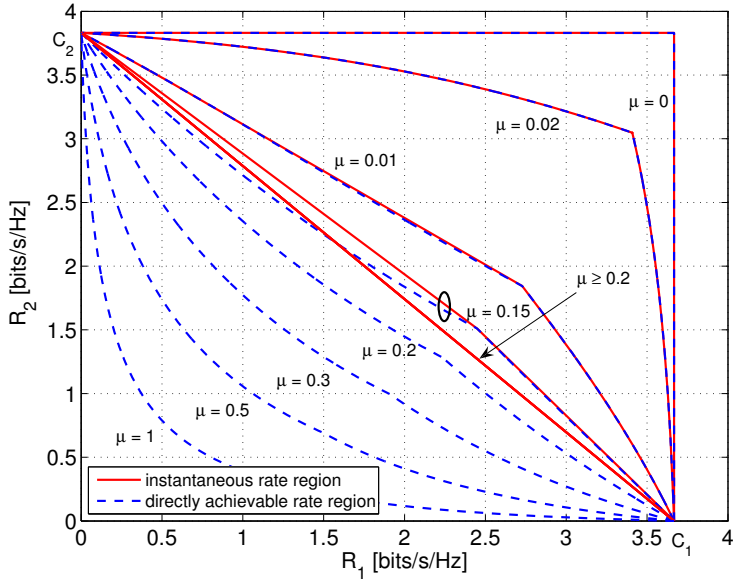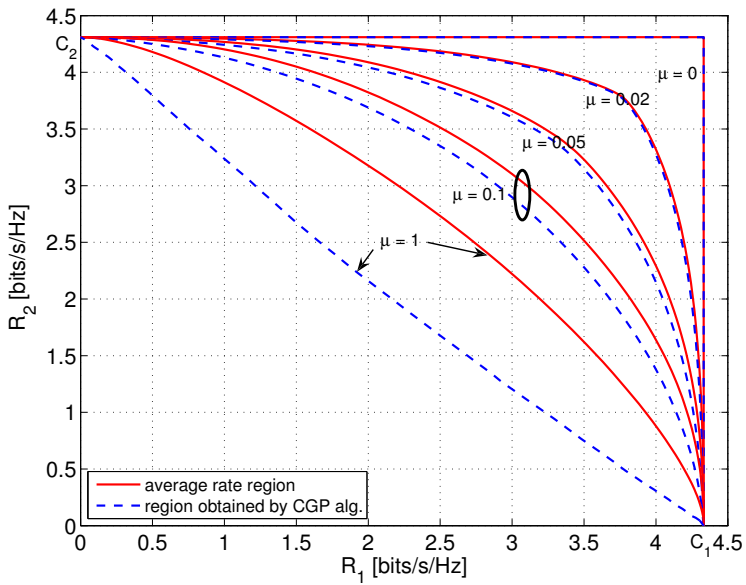
Fig 2.13. Rate regions: (a) Directly achievable and instantaneous rate regions; (b) Average rate regions, [83] © 2011, IEEE.

the region obtained by CGP algorithm is always worse than the average rate region. The gap in performance is more pronounced in the case of larger values of μ. Note that, even in the case of $\mu = 1$, the average rate region is bounded by a concave function with end points C_1 and C_2, although the corresponding instantaneous rate regions used in the averaging are triangles [see Figure 2.13(a)] in general. This phenomenon is due to the property of the set addition used in the definition of $\mathcal{R}^{\text{AVE}-\text{SC}}(\mu, p_1^{\max}, p_2^{\max})$. Results also show that the smaller the μ, the larger the average rate region.

2.5.6 Achievable rate regions in multicast wireless networks

We finally show the applicability of Algorithm 2.1 for finding the rate regions in a multicast wireless networks. A multicast with only two multicast transmissions [see Figure 2.14(a)] is considered for the sake of graphical illustration of the rate regions. Node 1 has common information to be sent to node 3 and 4, whereas node 2 has common information to be sent to node 3 and 5. We assume that node 3 has multipacket receiver capability. The channel power gains are given by (2.66) and SNR operating point is given by (2.67). Moreover, we set $D_0/d_0 = 10$ and $\eta = 4$.

As in the case of singlecast wireless networks, we first define the directly achievable rate region, instantaneous rate region, and the average rate region for multicast wireless networks. Particularized to the network setup in Figure 2.14(a), for a given set of interference coefficients / power gains

$$\acute{G}(t) = \{g_{11}(t), g_{22}(t), g_{33}(t), g_{44}(t), g_{14}(t), g_{32}(t)\} \tag{2.71}$$

and maximum node transmission power $p_1^{\max}$ and $p_2^{\max}$, the instantaneous rate region $\mathcal{R}^{\text{INS}-\text{MC}}(\acute{G}(t), p_1^{\max}, p_2^{\max})$ is defined as

$$\mathcal{R}^{\text{INS}-\text{MC}}(\acute{G}(t), p_1^{\max}, p_2^{\max}) = \text{conv}\left\{\mathcal{R}^{\text{DIR}-\text{MC}}(\acute{G}(t), p_1^{\max}, p_2^{\max})\right\}, \tag{2.72}$$

where $\mathcal{R}^{\text{DIR}-\text{MC}}(\acute{G}(t), p_1^{\max}, p_2^{\max})$ denotes the directly achievable rate region for
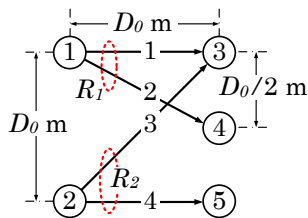
(a)
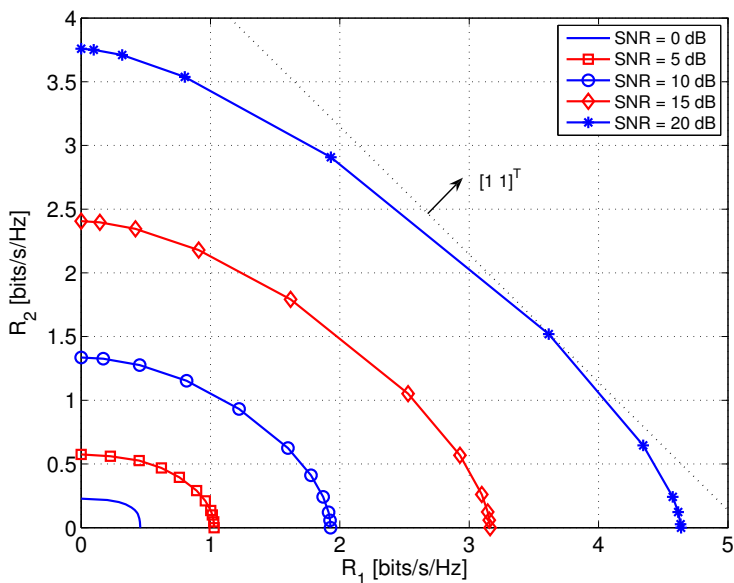
(b)

Fig 2.14. **(a) Multicast network,** $\mathcal{T} = \{1, 2\}$**,** $M_1 = 1$**,** $M_2 = 1$**,** $\mathcal{O}^1(1) = \{1, 2\}$**,**
$\mathcal{O}^1(2) = \{3, 4\}$**; (b) Average rate region, [83] © 2011, IEEE.**

multicast wireless networks, i.e.,

$$\mathcal{R}^{\mathrm{DIR-MC}}(\acute{G}(t), p_1^{\mathrm{max}}, p_2^{\mathrm{max}}) \tag{2.73}$$

$$= \left\{ (R_1, R_2) \left| \begin{array}{l} R_1 \leq \log\left(1 + \dfrac{g_{11}(t)p_1^1}{\sigma^2 + g_{33}(t)p_2^1}\right) \\[2mm] R_1 \leq \log\left(1 + \dfrac{g_{22}(t)p_1^1}{\sigma^2 + g_{32}(t)p_2^1}\right) \\[2mm] R_2 \leq \log\left(1 + \dfrac{g_{33}(t)p_2^1}{\sigma^2 + g_{11}(t)p_1^1}\right) \\[2mm] R_2 \leq \log\left(1 + \dfrac{g_{44}(t)p_2^1}{\sigma^2 + g_{14}(t)p_1^1}\right) \\[2mm] 0 \leq p_1^1 \leq p_1^{\mathrm{max}}, \quad 0 \leq p_2^1 \leq p_2^{\mathrm{max}} \end{array} \right. \right\} .$$

Finally, for a given maximum node transmission power p_1^{max} and p_2^{max}, the average rate region $\mathcal{R}^{\mathrm{AVE-MC}}(p_1^{\mathrm{max}}, p_2^{\mathrm{max}})$ is defined as $\mathcal{R}^{\mathrm{AVE-MC}}(p_1^{\mathrm{max}}, p_2^{\mathrm{max}}) = \frac{1}{T}\sum_{t=1}^{\acute{T}} \mathcal{R}^{\mathrm{INS-MC}}(\acute{G}(t), p_1^{\mathrm{max}}, p_2^{\mathrm{max}})$.

Figure 2.14(b) shows the average multicast rate region for different SNR values. Results show that, when the weights associated with rates R_1 and R_2 are the same, the resulting R_1 is always greater than R_2. For example, in the case of SNR = 20dB, we have $R_1 = 3.71$ bits/sec/Hz and $R_2 = 1.50$ bits/sec/Hz. Roughly speaking, this observation can be explained as follows: R_1 is determined by the rate of link 2 (the weakest of link 1 and 2), R_2 is determined by the rate of link 3 (the weakest of link 3 and 4) and rate of link 2 is larger than that of link 3 due to path losses.

2.6   Summary and discussion

We have considered the general WSRMax problem for a set of interfering links. In fact, this problem is NP-hard. A solution method, based on the branch and bound technique, has been proposed for solving the nonconvex WSRMax problem globally with an optimality certificate. Efficient and analytic bounds were proposed and their impact on convergence was numerically evaluated. The convergence speed of the proposed algorithm can be substantially increased by improving the lower bound, whilst the tightness of the upper bound has a much reduced impact. Numerical results showed that the proposed algorithm converged fairly fast in all considered setups. Nevertheless, since the problem is NP-hard, the worst case complexity can be exponential in the number of

variables. The considered link-interference model is fairly general so that it can model a wide range of network topologies with various node capabilities, such as single- or multipacket transmission (or reception) and simultaneous transmission and reception. Unlike other branch and bound based solution methods for WSRMax, our method does not require the problem to be convertible into a DC (difference of convex functions) problem. Therefore, the proposed method applies to a broader class of WSRMax problems (e.g., WSRMax in multicast wireless networks). Moreover, the method proposed can also be used to maximize *any* system performance metric that can be expressed as a Lipschitz continuous and increasing function of SINR values and is not restricted to WSRMax. Given its generality, the proposed algorithm can be adapted to address a wide range of network control and optimization problems. Performance benchmarks for various network topologies can be obtained by back-substituting it into any network design method which relies on WSRMax. Several applications, including cross-layer network utility maximization and maximum weighted link scheduling for multihop wireless networks, as well as finding achievable rate regions for singlecast/multicast wireless networks, have been presented. Since there are a number of suboptimal but low-complex algorithms are typically used in practice, the proposed algorithm can also be used for evaluating their performance loss.

3   Low-complexity algorithms for WSRMax

In this chapter we first develop efficient, low-complexity algorithms for the WSRMax problem in multicommodity, multichannel wireless networks by using homotopy methods [114] and complementary geometric program (or CGP) [97]. Our problem formulation is fairly general and it allows frequency reuse by activating multiple links in the same channel simultaneously. Here, the interference is solely resolved via power control. Furthermore, our formulation allows the possibility of exploiting multichannel diversity via dynamic power allocation across the available channels. The gains that can be achieved at upper layers in terms of end-to-end rates and network congestion are quantitatively analyzed by incorporating the proposed algorithms within Neely's cross-layer utility maximization framework [16, 17].

It is worth pointing out that the proposed algorithm, based on homotopy methods, also handles the self-interference problem in such a way that the combinatorial nature of the problem is circumvented. Here the imperfect self interference cancelation is modeled as a variable power gain from the transmitter to the receiver at all nodes. This simple model gives insight into the behavior of different network topologies when self interference cancellation is employed in network nodes. A similar approach can be used in a straightforward manner to model a wide range of network topologies with various node capabilities as well, e.g., singlepacket transmission, singlepacket reception, and many others. The proposed method can also be used to find the required level of accuracy for the self interference cancelation such that certain gains are achieved at the network layer. In addition, it provides a simple mechanism to evaluate the impact of scaling the distance between network nodes on the accuracy level of the self interference cancellation. Thus, from a network design perspective, the proposed method can be very useful.

Recall that WSRMax problem is NP-hard and we have to rely on exponentially complex global optimization techniques [24, 25, 82] to obtain the optimal solution. Nevertheless, the numerical results show that the proposed algorithms in this section perform close to global optimization methods. We further test our algorithms by carrying them out on large-scale problems, where global

optimization methods [55, 57, 60, 62, 83] cannot be used, due to prohibitive computational complexity. Results show that the proposed algorithms can provide significant gains at the network layer, in terms of end-to-end rates and network congestion, by exploiting efficiently the available multichannel diversity. We also evaluate the potential gains achievable at the network layer when the network nodes employ self interference cancelation techniques with different degrees of accuracy.

Finally, we consider different receiver capabilities and evaluate the effect of the use of multiuser detectors.

3.1 System model and problem formulation

3.1.1 Network model

The wireless network consists of a collection of nodes that can send, receive and relay data across wireless links. The set of all nodes is denoted by $\mathcal{N}$ and we label the nodes with the integer values $n = 1, \ldots, N$. A wireless link is represented as an ordered pair (i, j) of distinct nodes. The set of links is denoted by $\mathcal{L}$ and we label the links with the integer values $l = 1, \ldots, L$. We define $\mathrm{tran}(l)$ as the transmitter node of link l, and $\mathrm{rec}(l)$ as the receiver node of link l. The existence of a link $l \in \mathcal{L}$ implies that direct transmission is possible from $\mathrm{tran}(l)$ to $\mathrm{rec}(l)$. We assume that each node can be equipped with multiple transceivers, i.e., any node can simultaneously transmit to, or receive from, multiple nodes. We define $\mathcal{O}(n)$ as the set of links that are outgoing from node n, and $\mathcal{I}(n)$ as the set of links that are incoming to node n. Furthermore, we denote the set of transmitter nodes by $\mathcal{T}$ and the set of receiver nodes by $\mathcal{R}$, i.e., $\mathcal{T} = \{n \in \mathcal{N} | \mathcal{O}(n) \neq \emptyset\}$ and $\mathcal{R} = \{n \in \mathcal{N} | \mathcal{I}(n) \neq \emptyset\}$.

The network is assumed to operate in slotted time with slots normalized to integer values $t \in \{1, 2, 3, \ldots\}$. All wireless links are sharing a set $\mathcal{C}$ of orthogonal channels, labeled with integers $c = 1, \ldots, C$. When there are many channels which fade independently, at any one time there is a high probability that one of the channels will be strong. Thus, the main motivation for considering multiple channels is exploitation of the diversity that results from unequal links' behavior across a given wide band.

Let $h_{ijc}(t)$ denote the channel gain from the transmitter of link i to the receiver
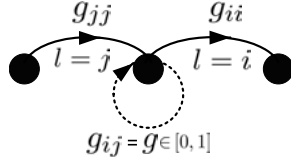
Fig 3.1. Choosing the value of interference coefficients, i.e., g_{ij} for $i \neq j$ and link power gains, i.e., g_{ii} and g_{jj} (channel index c and time index t are omitted for clarity): $\mathcal{A} = \{(i,j)\}$, $g_{ij} = g$, $g_{ji} = |h_{ji}|^2$, $g_{ii} = |h_{ii}|^2$, and $g_{jj} = |h_{jj}|^2$, [112] © 2011, IEEE.

of link j in channel c during time slot t. We assume that $h_{ijc}(t)$ are constant for the duration of a time slot and are independent and identically distributed over the time slots, links as well as over the channels. Let $g_{iic}(t)$ represent the power gain of link i in channel c during time slot t, i.e., $g_{iic}(t) = |h_{iic}(t)|^2$ (see Figure 3.1). For any pair of *distinct* links $i \neq j$, we denote the interference coefficient from link i to link j in channel c by $g_{ijc}(t)$. In the case of nonadjacent links (i.e., links i and j do not have common nodes), g_{ijc} represents the power of the interference signal at the receiver node of link j in channel c when one unit of power is allocated to the transmitter node of link i in channel c, i.e., $g_{ijc} = |h_{ijc}|^2$. When links i and j are adjacent, the value of g_{ijc} represents the power gain in channel c within the same node from its transmitter to its receiver, and is referred to as the self-interference coefficient (see Figure 3.1). For notational convenience let $\mathcal{A}$ denote the set of all link pairs (i,j) such that links i and j are adjacent. In other words, $\mathcal{A}$ represents the set of all link pairs (i,j) for which the transmitter of link i and the receiver of link j coincide, i.e., $\mathcal{A} = \{(i,j)_{i,j \in \mathcal{L}} | \text{ tran}(i) = \text{rec}(j)\}$ (see Figure 3.1). Specifically, for all $(i,j) \in \mathcal{A}$, we set $g_{ijc}(t) = g$ to model the residual self-interference gains after a certain self interference cancelation technique was employed at the network's nodes in channel c, where $g \in [0,1]$ is a scalar. We refer to g as the self-interference gain (see Figure 3.1). A value $g = 1$ means that no self interference cancelation technique is used and models the very large self interference that would affect the incoming links of a node if it simultaneously transmitted and received in the same channel. On the other hand, a value $g = 0$ corresponds to a perfect self interference cancelation. Note that, according to relative distances between network's nodes, $g_{ijc}(t)$ for all $(i,j) \in \mathcal{A}$ (i.e., the self-interference coefficients) can be several orders of magnitude larger than $g_{ijc}(t)$ for all $(i,j) \notin \mathcal{A}$. The

particular class of network topologies, for which $\mathcal{A} = \emptyset$ (i.e., $\mathcal{T} \cap \mathcal{R} = \emptyset$) is referred to as *bipartite* networks. On the other hand, the class of network topologies, for which $\mathcal{A} \neq \emptyset$ (i.e., $\mathcal{T} \cap \mathcal{R} \neq \emptyset$) is referred to as *nonbipartite* networks. Note that all multihop networks are necessarily nonbipartite.

In every time slot a network controller decides the power and rates allocated to each link in every channel. We denote the power allocated to each link l in channel c during time slot t by $p_{lc}(t)$. The power allocation is subject to a maximum power constraint $\sum_{c \in \mathcal{C}} \sum_{l \in \mathcal{O}(n)} p_{lc}(t) \leq p_n^{\max}$ for each node n.

We consider first the case where all receivers perform *singleuser detection*, i.e., any receiver decodes each of its intended signals by treating all other interfering signals as noise. Extensions to more advanced *multiuser detection* techniques will be addressed in Section 3.3. Suppose that the achievable rate of link l during time slot t is given by

$$r_l(t) = \sum_{c=1}^{C} W_c \log \left(1 + \frac{g_{llc}(t) p_{lc}(t)}{N_l W_c + \sum_{j \neq l} g_{jlc}(t) p_{jc}(t)} \right) \, , \qquad (3.1)$$

where W_c represents the bandwidth of channel c and N_l is the power spectral density of the noise at the receiver of link l. Note that for any link l, interference at rec(l) $\left(\text{i.e., the term } \sum_{j \neq l} g_{jlc}(t) p_{jc}(t)\right)$ is created by self transmissions $\left[\text{i.e., } \sum_{j \in \mathcal{O}(\text{rec}(l))} g_{jlc}(t) p_{jc}(t)\right]$, as well as by the other node transmissions $\left[\text{i.e., } \sum_{j \in \mathcal{L} \setminus \{\mathcal{O}(\text{rec}(l)) \cup \{l\}\}} g_{jlc}(t) p_{jc}(t)\right]$. To simplify the presentation, we assume in the sequel that all channels have equal bandwidths and that the noise power density is the same at all receivers (i.e., $W_c = W$ for all $c \in \mathcal{C}$ and $N_l = N_0$ for all $l \in \mathcal{L}$). The extension to the case of unequal bandwidths W_c and noise power spectral densities N_l is straightforward. Let $\sigma^2 = N_0 W$ denote the noise power, which is constant for all receivers in all channels. Furthermore, we denote by $\mathbf{P}(t) \in \mathbb{R}_+^{L \times C}$ the overall power allocation matrix, i.e., $p_{lc}(t) = [\mathbf{P}(t)]_{l,c}$. The use of the Shannon formula for the achievable rate in (3.1) is approximate in the case of finite length packets and is used to avoid the complexity of rate-power dependence in practical modulation and coding schemes. This is common practice but it must be noted that this is not strictly correct. However, as the packet length increases it is asymptotically correct.

88

3.1.2 Network utility maximization

Exogenous data arrive at the source nodes and they are delivered to the destination nodes over several, possibly multihop, paths. We identify the data by their destinations, i.e., all data with the same destination are considered as a single commodity, regardless of their source. Actually, our formulation also permits the anycast case, in which each packet exits the network as soon as any one of a particular destination set of nodes receives the packet successfully. We label the commodities with integers $s = 1, \ldots, S$ $(S \leq N)$ and the destination node of commodity s is denoted by d_s. For every node, we define $\mathcal{S}_n \subseteq \{1, \ldots, S\}$ as the set of commodities that can arrive exogenously at node n.

A network utility maximization, or NUM, framework similar to the one given in [17, Sec. 5.1] is considered. Specifically, exogenously arriving data is not directly admitted to the network layer. Instead, the exogenous data is first placed in the transport layer storage reservoirs. To avoid complications that may arise that are extraneous to our problem, we assume that all commodities have infinite demand at the transport layer. Nevertheless, the RA algorithms proposed in this section are still applicable when this assumption is relaxed. At each source node, a set of flow controllers decides the amount of each commodity data admitted during every time slot in the network. Let $x_n^s(t)$ denote the amount of data of commodity s admitted in the network at node n during time slot t. At the network layer, each node maintains a set of S internal queues for storing the current backlog (or unfinished work) of each commodity. Let $q_n^s(t)$ denote the current backlog of commodity s data stored at node n. We formally let $q_{d_s}^s(t) = 0$, i.e., it is assumed that data, which is successfully delivered to its destination, exits the network layer. Associated with each node-commodity pair $(n, s)_{s \in \mathcal{S}_n}$ we define a concave and nondecreasing *utility function* $u_n^s(y)$, representing the "reward" received by sending data of commodity s from node n to node d_s at a long term average rate of y [bits/slot].

The NUM problem under stability constraints can be formulated as [17, Sec. 5]

$$
\begin{aligned}
\text{maximize} \quad & \sum_{n \in \mathcal{N}} \sum_{s \in \mathcal{S}_n} u_n^s(y_n^s) \\
\text{subject to} \quad & (y_n^s)_{n \in \mathcal{N}, s \in \mathcal{S}_n} \in \mathbf{\Lambda} ,
\end{aligned}
\tag{3.2}
$$

where the variable is $(y_n^s)_{n \in \mathcal{N}, s \in \mathcal{S}_n}$ and $\mathbf{\Lambda}$ represents the *network layer capacity region*. In particular, the network layer capacity region $\mathbf{\Lambda}$ is the closure of the

set of all admissible arrival rate vectors that can be stably supported by the network, considering all possible strategies for choosing the control variables that affect routing, scheduling, and resource allocation (including those with perfect knowledge of future events) [17, p. 28].

A dynamic cross-layer control algorithm, which achieves a utility that is arbitrarily close to the optimal value of problem (3.2), has been introduced in [17, Sec. 5]. Specifically, the algorithm's performance can be characterized as follows:

$$
\sum_{n \in \mathcal{N}} \sum_{s \in \mathcal{S}_n} u_n^s(y^{\star s}_n) \; - \; \liminf_{T \to \infty} \sum_{n \in \mathcal{N}} \sum_{s \in \mathcal{S}_n} u_n^s \left(\frac{1}{T} \sum_{t=1:T} \mathrm{E}\{x_n^s(t)\} \right) \; \leq \; \frac{B}{V} \; , \quad (3.3)
$$

where $(y^{\star s}_n)_{n \in \mathcal{N}, s \in \mathcal{S}_n}$ is the optimal solution of problem (3.2), $B > 0$ is a well defined constant, and $V > 0$ is an algorithm parameter that can be used to control the tightness of the achieved utility to the optimal value [17, Sec. 5.2.1]. The details are extraneous to the central objective of this section. Particularized to our network model, in every time slot t, the algorithm performs the following steps:

Algorithm 3.1. *Dynamic cross-layer control algorithm [17, Sec. 5.2].*

1. Flow control; each node $n \in \mathcal{N}$ solves the following problem:

$$
\begin{aligned}
&\text{maximize} && \sum_{s \in \mathcal{S}_n} V u_n^s(x_n^s) - x_n^s q_n^s(t) \\
&\text{subject to} && \sum_{s \in \mathcal{S}_n} x_n^s \leq R_n^{\max}, \; x_n^s \geq 0 \; ,
\end{aligned} \quad (3.4)
$$

   where the variable is $(x_n^s)_{s \in \mathcal{S}_n}$. Set $(x_n^s(t) = x_n^s)_{s \in \mathcal{S}_n}$. The parameter $V > 0$ is a chosen parameter that affects the algorithm performance [see (3.3)] and $R_n^{\max} > 0$ is used to control the burstiness of data delivered to the network layer.

2. Routing and in-node scheduling; for each link l, let

$$
\begin{aligned}
\beta_l(t) &= \max_s \left\{ q_{\mathrm{tran}(l)}^s(t) - q_{\mathrm{rec}(l)}^s(t), 0 \right\} \\
c_l^\star(t) &= \arg\max_s \left\{ q_{\mathrm{tran}(l)}^s(t) - q_{\mathrm{rec}(l)}^s(t), 0 \right\} \; .
\end{aligned} \quad (3.5)
$$

   If $\beta_l(t) > 0$, the commodity that maximizes the differential backlog, i.e., $c_l^\star(t)$, is selected for potential routing over link l. This is the well known rule of next-hop transmission under the backpressure algorithm [12].

90

3. Resource allocation; the power allocation $\mathbf{P}(t)$ is given by $\mathbf{P}$ whose entries p_{lc} solve the following problem

$$
\begin{aligned}
\text{maximize} \quad & \sum_{l \in \mathcal{L}} \beta_l(t) \sum_{c \in \mathcal{C}} \log \left(1 + \frac{g_{llc}(t) p_{lc}}{\sigma^2 + \sum_{j \neq l} g_{jlc}(t) p_{jc}} \right) \\
\text{subject to} \quad & \sum_{c \in \mathcal{C}} \sum_{l \in \mathcal{O}(n)} p_{lc} \leq p_n^{\max}, \ n \in \mathcal{N} \\
& p_{lc} \geq 0, \ l \in \mathcal{L}, \ c \in \mathcal{C} \ .
\end{aligned}
\tag{3.6}
$$

Once the optimal power allocation $\mathbf{P}(t)$ is determined, compute rate allocation $r_l(t)$ for all $l \in \mathcal{L}$ by using (3.1). The resulting rate $r_l(t)$ is offered to the data of commodity $c_l^\star(t)$.

In the first step, each node n determines the amount of data of commodity s (i.e., $x_n^s(t)$ for all $s \in \mathcal{S}_n$) that are admitted in the network, based on the current backlogs (i.e., $q_n^s(t)$ for all $s \in \mathcal{S}_n$). In the second step, each node n computes β_l and the corresponding commodity $c_l^\star(t)$ for all $l \in \mathcal{O}(n)$. The commodity $c_l^\star(t)$ is selected for potential routing over link l during time slot t. Recall that *in-node scheduling* refers to selecting the appropriate commodity and it is not to be confused with the links scheduling mechanism, which is handled by the RA subproblem, i.e., step 3. The third step is the most difficult part of Algorithm 3.1, which computes the power allocation $\mathbf{P}(t)$ in each link l. Of course, the RA subproblem maximizes the sum of weighted rates, i.e., WSRMax. The solution $\mathbf{P}(t)$ determines implicitly the links/channels that should be activated in every time slot t. The power allocation $\mathbf{P}(t)$ is used to determine $r_l(t)$ [see (3.1)] and the resulting link rate $r_l(t)$ is offered to the data of commodity $c_l^\star(t)$. Since our main contribution resides in problem (3.6), extensive explanations of Algorithm 3.1 are avoided. However, we refer the reader to [17, Sec. 5] for more details.

3.2   Algorithm derivation: CGP and homotopy methods

In this section we focus on resource allocation problem (3.6). By using standard reformulation techniques, we first show that problem (3.6) is equivalent to a CGP [97]. Then we obtain a successive approximation algorithm for problem (3.6) in *bipartite* networks. Next we explain the challenges of the problem in

nonbipartite networks (e.g., multihop networks), due to the *self-interference problem*; when a node simultaneously transmits and receives in the same channel, its incoming links are affected by very large self interference levels. Finally, we propose a solution method based on *homotopy methods* [114], together with CGP, which circumvents the aforementioned difficulties.

3.2.1 CGP for WSRMax

Let us denote the negative of the objective function of problem (3.6) by $f_0(\mathbf{P})$. It can be expressed as

$$f_0(\mathbf{P}) = -\sum_{l\in\mathcal{L}}\sum_{c\in\mathcal{C}} \log\left(1 + \frac{g_{llc}p_{lc}}{\sigma^2 + \sum_{j\neq l} g_{jlc}p_{jc}}\right)^{\beta_l} \tag{3.7}$$

$$= \log \prod_{l\in\mathcal{L}}\prod_{c\in\mathcal{C}} \left(1 + \gamma_{lc}\right)^{-\beta_l}, \tag{3.8}$$

where the time index t was dropped for the sake of notational simplicity, and γ_{lc} represents the SINR of link l in channel c, i.e.,

$$\gamma_{lc} = \frac{g_{llc}p_{lc}}{\sigma^2 + \sum_{j\neq l} g_{jlc}p_{jc}}, \ l\in\mathcal{L}, \ c\in\mathcal{C} . \tag{3.9}$$

Since $\log(\cdot)$ is an increasing function, problem (3.6) can be reformulated equivalently as

$$\begin{aligned}
\text{minimize} \quad & \prod_{c\in\mathcal{C}}\prod_{l\in\mathcal{L}} \left(1 + \gamma_{lc}\right)^{-\beta_l} \\
\text{subject to} \quad & \gamma_{lc} = \frac{g_{llc}p_{lc}}{\sigma^2 + \sum_{j\neq l} g_{jlc}p_{jc}}, \ l\in\mathcal{L}, \ c\in\mathcal{C} \\
& \textstyle\sum_{c\in\mathcal{C}}\sum_{l\in\mathcal{O}(n)} p_{lc} \leq p_n^{\max}, \ n\in\mathcal{N} \\
& p_{lc} \geq 0, \ l\in\mathcal{L}, \ c\in\mathcal{C} ,
\end{aligned} \tag{3.10}$$

where the variables are $(p_{lc})_{l\in\mathcal{L},c\in\mathcal{C}}$ and $(\gamma_{lc})_{l\in\mathcal{L},c\in\mathcal{C}}$. Now we consider the related problem

$$\begin{aligned}
\text{minimize} \quad & \prod_{c\in\mathcal{C}}\prod_{l\in\mathcal{L}} \left(1 + \gamma_{lc}\right)^{-\beta_l} \\
\text{subject to} \quad & \gamma_{lc} \leq \frac{g_{llc}p_{lc}}{\sigma^2 + \sum_{j\neq l} g_{jlc}p_{jc}}, \ l\in\mathcal{L}, \ c\in\mathcal{C} \\
& \textstyle\sum_{c\in\mathcal{C}}\sum_{l\in\mathcal{O}(n)} p_{lc} \leq p_n^{\max}, \ n\in\mathcal{N} \\
& p_{lc} \geq 0, \ l\in\mathcal{L}, \ c\in\mathcal{C} ,
\end{aligned} \tag{3.11}$$

with the same variables $(p_{lc})_{l\in\mathcal{L},c\in\mathcal{C}}$ and $(\gamma_{lc})_{l\in\mathcal{L},c\in\mathcal{C}}$. Note that the equality constraints of problem (3.10) have been replaced with inequality constraints. We refer to these inequality constraints as SINR constraints for simplicity. Since the

objective function of problem (3.11) is decreasing in each γ_{lc}, we can guarantee that at any optimal solution of problem (3.11), the SINR constraints must be active. Therefore we solve problem (3.11) instead of problem (3.10).

Finally, by introducing the auxiliary variables $v_{lc} \leq 1 + \gamma_{lc}$ and rearranging the terms, problem (3.6) can be further reformulated as

$$
\begin{aligned}
\text{minimize} \quad & \prod_{c \in \mathcal{C}} \prod_{l \in \mathcal{L}} v_{lc}^{-\beta_l} \\
\text{subject to} \quad & v_{lc} \leq 1 + \gamma_{lc}, \ l \in \mathcal{L}, \ c \in \mathcal{C} \\
& \sigma^2 g_{llc}^{-1} p_{lc}^{-1} \gamma_{lc} + \sum_{j \neq l} g_{llc}^{-1} g_{jlc} p_{jc} p_{lc}^{-1} \gamma_{lc} \leq 1, \ l \in \mathcal{L}, \ c \in \mathcal{C} \quad (3.12) \\
& \sum_{c \in \mathcal{C}} \sum_{l \in \mathcal{O}(n)} \left(p_n^{\max} \right)^{-1} p_{lc} \leq 1, \ n \in \mathcal{N} \\
& p_{lc} \geq 0, \ l \in \mathcal{L}, \ c \in \mathcal{C} \ ,
\end{aligned}
$$

where the variables are $(p_{lc})_{l \in \mathcal{L}, c \in \mathcal{C}}$, $(\gamma_{lc})_{l \in \mathcal{L}, c \in \mathcal{C}}$, and $(v_{lc})_{l \in \mathcal{L}, c \in \mathcal{C}}$. Problem (3.12) can be identified as a CGP [97].

3.2.2 Successive approximation algorithm for WSRMax in bipartite networks

In this section we consider the case of bipartite networks. Recall from Section 3.1.1 that for such networks we have $\mathcal{A} = \emptyset$. By inspecting problem (3.12), we notice the following: 1) the objective is a monomial function [5, Sec. 2.1], 2) the right-hand side (RHS) terms of the first inequality constraints (i.e., $1 + \gamma_{lc}$) are posynomial functions, and 3) the left-hand side terms of all the inequality constraints are either monomial or posynomial functions. Note that if the RHS terms of the first inequality constraints were monomial functions (instead of posynomial ones), problem (3.12) would become a geometric program (or GP) in standard form. GPs can be reformulated as convex problems and they can be solved very efficiently, even for large scale problems [5, Sec. 2.5]. These observations suggest that by starting from an initial point, one can search for a close local optimum by solving a sequence of GPs, which locally approximate the original problem (3.12). At each step, the GP is obtained by replacing the posynomial functions in the RHS of the first inequality constraints with their best local monomial approximations near the the solution obtained at the previous step. The solution methods achieved by monomial approximations [5, 97] can be considered as a subset of a broader class of mathematical optimization

problems, known in mathematical literature as *inner approximation algorithms for nonconvex problems* [181]. The monomial approximation for the RHS terms of the first inequality constraints in problem (3.12) is described in the following lemma.

Lemma 3.1. *For any $\gamma > 0$, let $m(\gamma) = k\gamma^a$ be a monomial function used to approximate $s(\gamma) = 1 + \gamma$ near an arbitrary point $\hat{\gamma} > 0$. Then,*

1. the parameters a and k of the best monomial local approximation are given by

$$a = \hat{\gamma}(1 + \hat{\gamma})^{-1}, \quad k = \hat{\gamma}^{-a}(1 + \hat{\gamma}) . \tag{3.13}$$

2. $s(\gamma) \geq m(\gamma)$ for all $\gamma > 0$.

Proof. To show the first part we note the following: the monomial function m is the best local approximation of s near the point $\hat{\gamma}$ if

$$m(\hat{\gamma}) = s(\hat{\gamma}), \quad m'(\hat{\gamma}) = s'(\hat{\gamma}) . \tag{3.14}$$

By replacing the expressions of m and s in (3.14) we obtain the following system of equations:

$$\begin{cases} k\hat{\gamma}^a = 1 + \hat{\gamma} \\ ka\hat{\gamma}^{a-1} = 1 , \end{cases} \tag{3.15}$$

which has the solution given by (3.13).

The second part follows from (3.14) and noting that $s(\gamma)$ is affine and $m(\gamma)$ is concave on $\mathbb{R}_+$; concavity of $m(\gamma)$ follows from the fact that $k > 0$ and $0 < a < 1$ [3, Sec. 3.1.5]. $\square$

Figure 3.2 illustrates the monomial approximation given in Lemma 3.1 for different values of $\hat{\gamma}$. Note that the monomial approximation $m(\gamma)$ is reasonably close to $s(\gamma)$ for larger values of $\hat{\gamma}$. The approximation given in Lemma 3.1 turns out to be equivalent to the lower bound approximation used in [66, Sec. III-B] for dynamic spectrum management in digital subscriber lines.

Let us now turn to the GP obtained by using the local approximation given by Lemma 3.1. The posynomial functions $1 + \gamma_{lc}$ of the first inequality constraints of problem (3.12) are approximated near the point $\hat{\gamma}_{lc}$. Consequently the approximate inequality constraints become

$$v_{lc} \leq k_{lc}\gamma_{lc}^{a_{lc}}, \quad l \in \mathcal{L}, \ c \in \mathcal{C} , \tag{3.16}$$
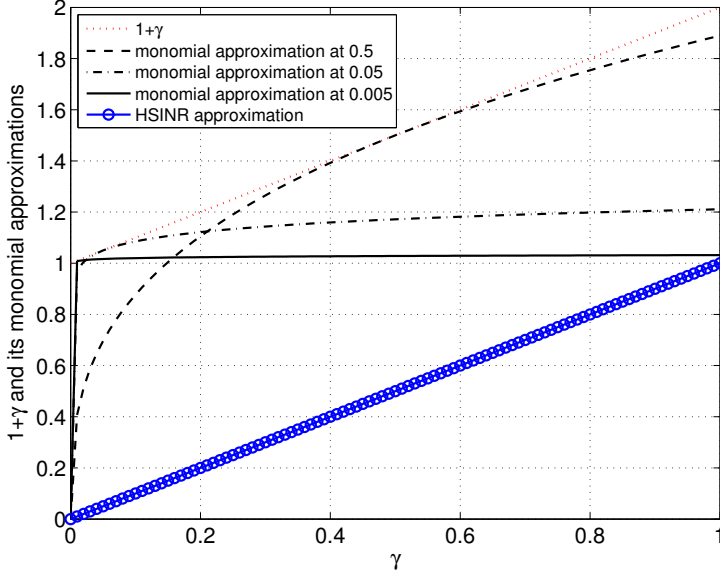
Fig 3.2. Monomial approximation given in Lemma 3.1 for $\hat{\gamma} = 0.5$**,** 0.05**, and** 0.005**.**

where a_{lc} and k_{lc} have the forms given in (3.13). Since the objective function of problem (3.12) is a decreasing function of v_{lc}, $l \in \mathcal{L}, c \in \mathcal{C}$, it can be easily verified that all of these modified inequality constraints will be active at the solution of the GP. Therefore, we can eliminate the auxiliary variables v_{lc} and rewrite the objective function of problem (3.12) as

$$\prod_{l \in \mathcal{L}} \prod_{c \in \mathcal{C}} v_{lc}^{-\beta_l} = \prod_{l \in \mathcal{L}} \prod_{c \in \mathcal{C}} k_{lc}^{-\beta_l} \gamma_{lc}^{-\beta_l a_{lc}} = K \prod_{l \in \mathcal{L}} \prod_{c \in \mathcal{C}} \gamma_{lc}^{-\beta_l \frac{\hat{\gamma}_{lc}}{1 + \hat{\gamma}_{lc}}} , \qquad (3.17)$$

where K is a multiplicative constant which does not affect the problem solution.

In the following subsections, we base our development on computationally efficient algorithms to obtain a suboptimal solution for problem (3.12). For notational convenience it is useful to define the overall SINR matrices $\boldsymbol{\gamma}, \hat{\boldsymbol{\gamma}} \in \mathbb{R}_+^{L \times C}$ as $[\boldsymbol{\gamma}]_{l,c} = \gamma_{lc}$ and $[\hat{\boldsymbol{\gamma}}]_{l,c} = \hat{\gamma}_{lc}$, respectively.

A very brief outline of the proposed successive approximation algorithm is as follows. It solves an approximated version of problem (3.12) in every iteration and the algorithm consists of repeating this step until convergence.

Algorithm 3.2. *Successive approximation algorithm for WSRMax*

1. Initialization; given tolerance $\varepsilon > 0$, a feasible power allocation $\mathbf{P}_0$. Set $i = 1$. The initial SINR guess $\hat{\boldsymbol{\gamma}}^{(i)}$ is given by (3.9).

2. Solving the GP;

$$
\begin{aligned}
\text{minimize} \quad & K^{(i)} \prod_{l \in \mathcal{L}} \prod_{c \in \mathcal{C}} \gamma_{lc}^{-\beta_l \frac{\hat{\gamma}_{lc}^{(i)}}{1+\hat{\gamma}_{lc}^{(i)}}} \\
\text{subject to} \quad & \alpha^{-1} \hat{\gamma}_{lc}^{(i)} \leq \gamma_{lc} \leq \alpha \hat{\gamma}_{lc}^{(i)}, \ l \in \mathcal{L}, \ c \in \mathcal{C} \\
& \sigma^2 g_{llc}^{-1} p_{lc}^{-1} \gamma_{lc} + \sum_{j \neq l} g_{llc}^{-1} g_{jlc} p_{jc} p_{lc}^{-1} \gamma_{lc} \leq 1, \ l \in \mathcal{L}, \ c \in \mathcal{C} \\
& \sum_{c \in \mathcal{C}} \sum_{l \in \mathcal{O}(n)} \left(p_n^{\max}\right)^{-1} p_{lc} \leq 1, \ n \in \mathcal{N} \ ,
\end{aligned}
\tag{3.18}
$$

with the variables $(p_{lc})_{l \in \mathcal{L}, c \in \mathcal{C}}$ and $(\gamma_{lc})_{l \in \mathcal{L}, c \in \mathcal{C}}$. Denote the solution by $(p_{lc}^{\star})_{l \in \mathcal{L}, c \in \mathcal{C}}$ and $(\gamma_{lc}^{\star})_{l \in \mathcal{L}, c \in \mathcal{C}}$.

3. Stopping criterion; if $\max_{(l,c) \in \mathcal{L} \times \mathcal{C}} \left| \gamma_{lc}^{\star} - \hat{\gamma}_{lc}^{(i)} \right| \leq \varepsilon$ STOP; otherwise go to step 4.

4. Set $i = i + 1$, $\left(\hat{\gamma}_{lc}^{(i)} = \gamma_{lc}^{\star}\right)_{l \in \mathcal{L}, c \in \mathcal{C}}$ and go to step 2.

The first step initializes the algorithm and an initial feasible SINR guess $\hat{\boldsymbol{\gamma}}^{(i)}$ is computed. For bipartite networks, there is no *self-interference problem*, and a simple uniform power allocation can be used.

The second step solves an equivalent GP approximation of problem (3.12) around the current guess $\hat{\boldsymbol{\gamma}}^{(i)}$ [see problem (3.18)]. Note that the auxiliary variables $(v_{lc})_{l \in \mathcal{L}, c \in \mathcal{C}}$ of problem (3.12) are eliminated and the objective function of problem (3.12) is replaced by using the monomial approximation at $\hat{\boldsymbol{\gamma}}^{(i)}$ given in (3.17); $K^{(i)}$ is a multiplicative constant which does not influence the solution of problem (3.18). These monomial approximations are sufficiently accurate only in the closer vicinity of the current guess $\hat{\boldsymbol{\gamma}}^{(i)}$. Therefore, the first set of inequality constraints are added to confine the domain of variables $\boldsymbol{\gamma}$ to a region around the current guess $\hat{\boldsymbol{\gamma}}^{(i)}$ [182]. The first set of inequality constraints of problem (3.18) are sometimes called trust region constraints [5, 182], which are not originally introduced in [97]. Therefore, Algorithm 3.2 is a slightly modified version of the solution method proposed in [97]. The parameter $\alpha > 1$ controls the desired approximation accuracy. However, it also influences the convergence

speed of Algorithm 3.2. At every step, each entry of the current SINR guess $\hat{\boldsymbol{\gamma}}^{(i)}$ can be increased or decreased at most by a factor α. Thus, a value of α close to 1 provides good accuracy for the monomial approximations at the cost of slower convergence speed, while a value much larger than 1 improves the convergence speed at the cost of reduced accuracy. In most practical cases, a fixed value $\alpha = 1.1$ offers a good speed/accuracy tradeoff [5]. Though we have trust region constraints for problem (3.18), it is not mandatary to include those here and Algorithm 3.2 can still be carried out.

The third step checks whether the SINRs $(\gamma_{lc}^{\star})_{l \in \mathcal{L}, c \in \mathcal{C}}$ obtained from the solution of problem (3.18) have been significantly changed compared to the entries of the current guess $\hat{\boldsymbol{\gamma}}^{(i)}$. If there are no substantial changes, then the algorithm terminates and the link rate $r_l(t) = \sum_{c=1}^{C} W \log(1 + \gamma_{lc}^{\star})$ is offered to the data of commodity $c_l^{\star}(t)$ [given by (3.5)]. Otherwise, the solution $(\gamma_{lc}^{\star})_{l \in \mathcal{L}, c \in \mathcal{C}}$ is taken as the current guess and the algorithm repeats steps 2 to 4 until convergence.

Note that the auxiliary variables $(v_{lc})_{l \in \mathcal{L}, c \in \mathcal{C}}$ were only used to reformulate problem (3.11) as a CGP [97] [i.e., problem (3.12)], but they do not appear in Algorithm 3.2. In fact, an identical algorithm results if, at each step, the objective function of problem (3.11) is locally approximated by a monomial function (see [15, Lem. 4.2.2]). This alternative derivation, known in optimization literature as *signomial programming* [5], is presented in Appendix 3. Careful comparisons reveal that the algorithm recently proposed in [100, p. 3034] is almost identical to our proposed Algorithm 3.2 for single channel case with no trust region constraints, i.e., $C = 1$ and $\alpha = \infty$.

The convergence of the Algorithm 3.2 to a Kuhn-Tucker solution of the original nonconvex problem (3.12) is guaranteed [181, Th. 1], since the algorithm falls into the broader class of mathematical optimization problems, *inner approximation algorithms for nonconvex problems* [181].

One interesting and important remark is that the objective function of the approximated problem (3.18) in each iteration i yields a upper bound on the objective function of the original problem (3.11), i.e.,

$$K^{(i)} \prod_{l \in \mathcal{L}} \prod_{c \in \mathcal{C}} \gamma_{lc}^{-\beta_l \frac{\hat{\gamma}_{lc}^{(i)}}{1+\hat{\gamma}_{lc}^{(i)}}} \geq \prod_{l \in \mathcal{L}} \prod_{c \in \mathcal{C}} \left(1 + \gamma_{lc}\right)^{-\beta_l} \tag{3.19}$$

for $(\gamma_{lc} > 0)_{l \in \mathcal{L}, \, c \in \mathcal{C}}$, with equality when $\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}}^{(i)}$. This follows directly from the second statement of Lemma 3.1. By using (3.19), we can show immediately that

Algorithm 3.2 is monotonically decreasing. The monotonicity of Algorithm 3.2 is established by the following theorem.

Theorem 3.1. *Let i and $i+1$ be any consecutive iterations of Algorithm 3.2 and $\hat{\boldsymbol{\gamma}}^{(i)}$ and $\hat{\boldsymbol{\gamma}}^{(i+1)}$ be the SINR guesses at the beginning of each iteration respectively. Then*

$$\prod_{l \in \mathcal{L}} \prod_{c \in \mathcal{C}} \left(1 + \hat{\gamma}_{lc}^{(i)}\right)^{-\beta_l} \geq \prod_{l \in \mathcal{L}} \prod_{c \in \mathcal{C}} \left(1 + \hat{\gamma}_{lc}^{(i+1)}\right)^{-\beta_l} . \tag{3.20}$$

Proof. To show this we write the following relations:

$$\prod_{l \in \mathcal{L}} \prod_{c \in \mathcal{C}} \left(1 + \hat{\gamma}_{lc}^{(i)}\right)^{-\beta_l} = K^{(i)} \prod_{l \in \mathcal{L}} \prod_{c \in \mathcal{C}} \left(\hat{\gamma}_{lc}^{(i)}\right)^{-\beta_l \frac{\hat{\gamma}_{lc}^{(i)}}{1+\hat{\gamma}_{lc}^{(i)}}} \tag{3.21}$$

$$\geq K^{(i)} \prod_{l \in \mathcal{L}} \prod_{c \in \mathcal{C}} \left(\hat{\gamma}_{lc}^{(i+1)}\right)^{-\beta_l \frac{\hat{\gamma}_{lc}^{(i)}}{1+\hat{\gamma}_{lc}^{(i)}}} \tag{3.22}$$

$$\geq \prod_{l \in \mathcal{L}} \prod_{c \in \mathcal{C}} \left(1 + \hat{\gamma}_{lc}^{(i+1)}\right)^{-\beta_l} , \tag{3.23}$$

where (3.21) follows from (3.19), (3.22) follows since $\hat{\boldsymbol{\gamma}}^{(i+1)}$ is the solution of problem (3.18), and (3.23) follows again from (3.19). $\qquad\square$

Therefore we see immediately that Algorithm 3.2 always yields a solution, which is at least as good as the one in the previous iteration. This is important in the context of practical implementations, since in practice, one can always stop the algorithm within few iterations, before it terminates.

3.2.3 The self-interference problem in nonbipartite networks

Let us now consider the nonbipartite networks; networks for which $\mathcal{A} \neq \emptyset$. In other words, the set of nodes cannot be divided into two distinct subsets, $\mathcal{T}$ and $\mathcal{R}$, i.e., $\mathcal{T} \cap \mathcal{R} \neq \emptyset$ (e.g., multihop wireless networks). For example see Figure 3.1 and Figure 3.3. For such network topologies, there is the *self-interference problem* and, consequently, the WSRMax problem must also cope with the self-interference problem. The difficulty comes from the fact that the self-interference gains g [14] can typically be a few orders of magnitude larger than the power gains

---

[14]Recall that $g_{ijc} = g$ for all $(i,j) \in \mathcal{A}$.
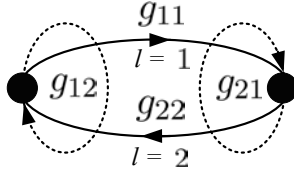
Fig 3.3. Two node network (channel index c and time index t are omitted for clarity): $\mathcal{A} = \{(1,2), (2,1)\}$, $g_{12} = 1$, $g_{21} = 1$, $g_{11} = |h_{11}|^2$, and $g_{22} = |h_{22}|^2$, [112] © 2011, IEEE.

between distinct network nodes $\{g_{jjc}\}_{j \in \mathcal{L}}$, e.g., when there is no self interference cancellation. Therefore there can be a huge imbalance between some entries of $\{g_{ijc}\}_{i,j \in \mathcal{L}}$, especially when g is large. Roughly speaking, this can destroy the smoothness of the functions associated with the WSRMax problem [e.g., the objective function of problem (3.6)] and can ruin the reliability and the efficiency of Algorithm 3.2 that solves it, at least suboptimally. In other words, there can be many highly suboptimal Kuhn-Tucker solutions for problem (3.12) at which Algorithm 3.2 can terminate by returning an undesirable suboptimal solution. Moreover, the SINR values at the incoming links of a node that simultaneously transmits in the same channel are very *small* and the convergence of Algorithm 3.2 can be very slow if it starts with an initial SINR guess $\hat{\gamma}$ containing entries with nearly zero values. These behaviors are reflected in the monomial approximations plotted in Figure 3.2. For example, when $\hat{\gamma}$ is smaller ($\hat{\gamma} = 0.005$), the monomial approximation $m(\gamma)$ shows abrupt changes at $\gamma = 0.005$ and remains constant at an undesirable level almost equal to 1 for all $\gamma > 0.005$. Therefore, the direct application of Algorithm 3.2 can perform very poorly and further improvements are necessary.

A standard way to deal with the self-interference problem consists of adding a supplementary combinatorial constraint into the WSRMax problem that does not allow any node in the network to transmit and receive simultaneously in the same channel [106, 107, 109]. We will refer to a power allocation, which satisfies this constraint as *admissible*. Note that this approach would require solving a power optimization problem (by using Algorithm 3.2) for each possible subset of links that can be simultaneously activated. This results inn a combinatorial nature for the WSRMax problem in the case of nonbipartite networks [38, 51, 163–167]. Of course, since the complexity of this approach grows exponentially with the number of links and number of channels, this solution method becomes quickly

impractical.

3.2.4 Successive approximation algorithm for WSRMax in nonbipartite networks: A homotopy method

To avoid difficulties pointed out in Section 3.2.3, we propose an algorithm inspired from *homotopy methods* [114] that can be traced back to late 80's; see [183] and the references therein. In fact, the well known interior-point methods [184],[3, Sec. 11] for convex optimization problems also fall into this general class of homotopy methods.

The underlying idea is to first introduce a parameterized problem that approximates the original problem (3.11). Specifically, we construct the parameterized problem from the original problem (3.11) by setting $g_{ijc} = g_v$ for all $(i,j) \in \mathcal{A}$, where $g_v \in \mathbb{R}_+$ is referred to as the homotopy parameter. Indeed, in the original problem (3.11) we have $g_{ijc} = g$ for all $(i,j) \in \mathcal{A}$. Note that the quality of the approximation improves as g_v reaches g; the true self-interference gain. Of course, when g_v is small (e.g., g_v and g_{jjc} are roughly in the same order), Algorithm 3.2 can be used reliably to find a suboptimal solution for the parameterized problem. Thus, a sequence of parameterized problems are solved, starting at a very small g_v and increasing the parameter g_v (thus the accuracy of the approximation) at each step until g_v reaches the true self-interference gain g. Moreover, in each step, when solving the parameterized problem for the current value of g_v, the initial guess for Algorithm 3.2 is obtained by using the solution (power) of the parameterized problem for the previous value of g_v.

The proposed algorithm based on homotopy methods can be summarized as follows:

Algorithm 3.3. *Successive approximation algorithm for WSRMax in the presence of self interferers*

1. Initialization; given an initial homotopy parameter $g_0 < g$, $\rho > 1$, a feasible power allocation $\mathbf{P}_0$. Let $g_v = g_0$, $\mathbf{P} = \mathbf{P}_0$.
2. Set $g_{ijc} = g_v$ for all $(i,j) \in \mathcal{A}$. Find the SINR guess $\hat{\boldsymbol{\gamma}}$ by using (3.9).
3. Solving parameterized problem; let $\hat{\boldsymbol{\gamma}}^{(1)} = \hat{\boldsymbol{\gamma}}$, perform steps 2 to 4 of Algorithm 3.2 until convergence to obtain the power $(p_{lc}^\star)_{l \in \mathcal{L}, c \in \mathcal{C}}$ and SINR

values $(\gamma_{lc}^{\star})_{l \in \mathcal{L}, c \in \mathcal{C}}$. Let $(p_{lc} = p_{lc}^{\star})_{l \in \mathcal{L}, c \in \mathcal{C}}$.

4. If $\exists (i, j) \in \mathcal{A}$ and $c \in \mathcal{C}$ such that $p_{ic} p_{jc} > 0$ (i.e., $\mathbf{P}$ is not admissible), then set $g_v = \min\{\rho g_v, g\}$ and go to step 5. Otherwise (i.e., $\mathbf{P}$ is admissible) STOP.

5. If $g_v < g$, go to step 2, otherwise STOP.

The first step initializes the algorithm; the homotopy parameter g_v is initialized by g_0, where g_0 is chosen in the same range of values as the power gains between distinct nodes. Specifically, in our simulations we select $g_0 = \max_{j \in \mathcal{L}, c \in \mathcal{C}} \{g_{jjc}\}$. Step 2 updates the problem data for the parameterized problem and a feasible SINR guess is computed. The third step finds a suboptimal solution for the parameterized problem. The algorithm terminates in step 4 if $\mathbf{P}$ is admissible (thus none of nodes in the network are transmitting and receiving simultaneously in the same channel). On the other hand, if $\mathbf{P}$ is not admissible, then the homotopy parameter g_v is increased. If g_v reaches its extreme allowed value (i.e., the actual self-interference gain value of g), the algorithm terminates. Otherwise (i.e., $g_v < g$), it returns to step 2 and continues. Terminating Algorithm 3.3 if the solution is *admissible* is intuitively obvious for the following reason. The data associated with the parameterized problem that is solved in step 3 of Algorithm 3.3 becomes independent of the homotopy parameter g_v, and therefore a further increase in g_v after having an *admissible* solution has no effect on the results. Our computational experience suggests that Algorithm 3.3 yields an *admissible* solution way before g_v reaches value g (e.g., in the case of no self interference cancellation, i.e., $g = 1$, an admissible power allocation is achieved in about $1 - 4$ iterations with $\rho = 2$).

Since Algorithm 3.3 runs a finite number of instances of Algorithm 3.2, its computational complexity *does not increase* more than polynomially with the problem size. Clearly, Algorithm 3.3 can converge to a Kuhn-Tucker solution of the last parameterized problem (the one just before the termination of Algorithm 3.3).

As a specific example to illustrate self interference, consider the simple network shown in Figure 3.3 and suppose that no self interference cancellation technique is employed at the network's nodes, i.e., $g = 1$. Here, $N = 2$, $L = 2$, and $C = 1$. Note that $\mathcal{A} = \{(1, 2), (2, 1)\}$ and let $\beta_1, \beta_2 \neq 0$. Suppose that $g_{12} \gg g_{22}$ and $g_{21} \gg g_{11}$, which is often the case due to path losses. Since

the gains $g_{12} = 1$ and $g_{21} = 1$ are very large compared to g_{22} and g_{11}, for any nonzero power allocation $p_1, p_2 = p_0$ the initial SINR guess $\hat{\gamma}_1, \hat{\gamma}_2$ will have nearly zero values. This results in difficulties in using Algorithm 3.2 directly. In Algorithm 3.3 this problem is circumvented by initializing the gains g_{12} and g_{21} by a parameter g_0 (e.g., $g_0 = \max\{g_{11}, g_{22}\}$) and executing Algorithm 3.2 repeatedly, increasing incrementally the parameter g_v until it reaches 1, the true values of g_{12} and g_{21}.

Regarding the complexity of the proposed algorithm we make the following remarks. The computational complexity of a GP depends on the number of variables and constraints, as well as on the sparsity pattern of the problem [5]. Unfortunately, it is difficult to quantify precisely the sparsity pattern, and therefore a general complexity analysis is not available. To give a rough idea, let us consider a fully connected network with $N = 9$ nodes and $C = 8$ channels. The number of variables in problem (3.18) is $2LC = 1152$ and the number of constraints is $3LC + N = 1737$, and it was solved in about 12 seconds on desktop computer. The number of iterations depends on the starting point, $p_n^{\max}$ and channel gains g_{ijc}, but typically Algorithm 3.2 required around 100 iterations to converge.

Nevertheless, with some slight modifications it is possible to dramatically decrease the average complexity per iteration, which is very important in the context of practical implementations. Two simple modifications are as follows:

1. Use a large values for the parameter α in Algorithm 3.2: as we discussed in Section 3.2.2, large α can improve the convergence speed of Algorithm 3.2 at the cost of reduced accuracy of the monomial approximation.

2. Eliminate (relatively) insignificant variables; we can eliminate the power variables p_{lc} and the associated SINR variables γ_{lc} from problem (3.18) when they have relatively very small contributions to the overall objective value of (3.18). Specifically, the exponent term $\beta_l \frac{\hat{\gamma}_{lc}^{(i)}}{1+\hat{\gamma}_{lc}^{(i)}}$ in the objective of (3.18) is evaluated for all $\in \mathcal{L}, c \in \mathcal{C}$ and if

$$\beta_l \frac{\hat{\gamma}_{lc}^{(i)}}{1 + \hat{\gamma}_{lc}^{(i)}} \ll \max_{\bar{l} \in \mathcal{L}, \bar{c} \in \mathcal{C}} \left(\beta_{\bar{l}} \frac{\hat{\gamma}_{\bar{l}\bar{c}}^{(i)}}{1 + \hat{\gamma}_{\bar{l}\bar{c}}^{(i)}} \right) \ ,$$

then p_{lc} s and the associated γ_{lc} s are eliminated in successive GPs.

3.2.5 Impact of scaling the node distances on the accuracy of self interference cancelation

Based on a simple exponential path loss model, in this section we discuss the impact of scaling the distance between network nodes on the accuracy level of the self interference cancellation.

For simplicity, we focus on the single-channel case (i.e., $C = 1$). Suppose an exponential path loss model, where the channel power gains $|h_{ij}(t)|^2$, between distinct nodes are given by

$$|h_{ij}(t)|^2 = \left(\frac{d_{ij}}{d_0}\right)^{-\eta} c_{ij}(t) . \tag{3.24}$$

Here d_{ij} is the distance from the transmitter of link i to the receiver of link j, d_0 is the *far field reference distance* [176], η is the path loss exponent, and $c_{ij}(t)$ are exponentially distributed random variables with unit mean, independent of the time slots and links. The first term of (3.24) represents the path loss factor and the second term models Rayleigh small-scale fading.

Suppose $p_n^{\max} = p_0^{\max}$ for all $n \in \mathcal{N}$. For all $l \in \mathcal{L}$ we define the SNR of link l as

$$\mathrm{SNR}_l = \frac{p_0^{\max}}{\sigma^2} \left(\frac{d_{ll}}{d_0}\right)^{-\eta} . \tag{3.25}$$

It represents the average SNR at $\mathrm{rec}(l)$ when $\mathrm{tran}(l)$ allocates all its transmission power to link l and all the other nodes are silent. Let $\mathbf{p}(t) \in \mathbb{R}_+^L$ denote the overall power allocation matrix, i.e., $p_l(t) = [\mathbf{p}(t)]_l$ (note that the channel index is dropped for simplicity, since $C = 1$).

Let us consider a network that is obtained from another one, by scaling the distance between distinct nodes and the maximum node transmission power such that all link's SNRs [see (3.25)] are conserved. In the sequel, we show that, in order to preserve the achievable rate region, the accuracy level of the self interference cancelation techniques must also be scaled appropriately.

We start by defining two matrices, which will be useful for later reference. Let $\mathbf{D} \in \mathbb{R}_+^{L \times L}$ denote the node distance matrix defined as $[\mathbf{D}]_{i,j} = d_{ij}$ and $\mathbf{G}(t) \in \mathbb{R}_+^{L \times L}$ denote the interference coefficient and the power gain matrix during time slot t, defined as $[\mathbf{G}(t)]_{i,j} = g_{ij}(t)$. The achievable rate region with singleuser detection at receivers for a given $\mathbf{G}(t)$ and a maximum node

transmission power $p_0^{\max}$ can be expressed as

$$\mathcal{R}(\mathbf{G}(t), p_0^{\max}) = \left\{ (r_1, \ldots, r_L) \left| \begin{array}{ll} r_l \leq \log\left(1 + \dfrac{g_{ll}(t)p_l}{\sigma^2 + \sum_{j \neq l} g_{jl}(t)p_j}\right), & l \in \mathcal{L} \\ \sum_{l \in \mathcal{O}(n)} p_l \leq p_0^{\max}, & n \in \mathcal{N} \\ p_l \geq 0, & l \in \mathcal{L} \end{array} \right. \right\}.$$

$$(3.26)$$

From (3.26), it follows that if the matrix $\mathbf{G}(t)$ is scaled by a factor of $1/\kappa$, and the maximum node transmission power $p_0^{\max}$ is scaled by a factor of κ, then the achievable rate region is unchanged, i.e.,

$$\mathcal{R}(\mathbf{G}(t), p_0^{\max}) = \mathcal{R}(\mathbf{G}(t)/\kappa, \kappa p_0^{\max}) . \qquad (3.27)$$

Let $\kappa = \theta^\eta$. According to the exponential path loss model given in (3.24), the scaling of $\mathbf{G}(t)$ by a factor of $1/\kappa$ (or $1/\theta^\eta$) is equivalent to the scaling of node distance matrix $\mathbf{D}$ by a factor of θ and the scaling of self-interference gains g by a factor of $1/\theta^\eta$. Therefore, with a slight abuse of notation, we rewrite (3.27) as

$$\mathcal{R}(\mathbf{D}, g, p_0^{\max}) = \mathcal{R}(\theta\mathbf{D}, g/\theta^\eta, \theta^\eta p_0^{\max}) . \qquad (3.28)$$

To interpret the relation in (3.28), we consider a network characterized by D, g, and $p_0^{\max}$. If we construct another network by scaling D by a factor of θ and by scaling $p_0^{\max}$ by a factor of θ^η, then to preserve the achievable rate region, the accuracy level of the self interference cancelation should be improved to g/θ^η. This is intuitively obvious since, the larger the distance between network nodes, the larger the power levels required to preserve the link SINRs, and therefore the higher the accuracy level required by the self interference cancelation techniques to remove the increased transmit power at nodes. Based on (3.28) we can establish similar equivalences in terms of network layer performance metrics as well. Roughly speaking, relation (3.28) suggests that in networks where the nodes are located far apart (e.g., cellular type of wireless networks), the accuracy of self interference cancellation is more stringent compared to that in networks where the nodes are located in close vicinity (e.g., a wireless network setup in an office).

3.3   Extensions to wireless networks with advanced transceivers

Until now the receiver structure was basically assumed to be equivalent to a bank of match filters, each of which attempts to decode one of the signals of interest at each node while treating the other signals as noise. This is a suboptimal detector structure that is commonly assumed. In this section, we investigate the possible gains achievable by using more advanced receiver structures. For clarity, we discuss first the single-channel case. The extension to the multichannel case is presented in Appendix 4. We assume that at every node $n \in \mathcal{N}$ the transmitter performs superposition coding over its outgoing links $\mathcal{O}(n)$, and the receiver decodes the signals of incoming links $\mathcal{I}(n)$ by using a multiuser receiver based on successive interference cancelation (SIC) strategy. One may of course assume other detector structure, including the optimum one that implements maximum likelihood. The largest set of achievable rates is obtained when the SIC receiver at every node $n \in \mathcal{N}$ is allowed to decode and cancel out the signals of all its incoming links $\mathcal{I}(n)$ and any subset of the remaining links in its complement set $\mathcal{L} \setminus \mathcal{I}(n)$. Let $\mathcal{D}(n)$ denote the set of links, which are decoded at the node n, i.e., $\mathcal{D}(n) = \mathcal{I}(n) \cup \mathcal{U}(n)$ for some $\mathcal{U}(n) \subseteq \mathcal{L} \setminus \mathcal{I}(n)$. Furthermore, let $\mathcal{R}^{\mathrm{SIC}}(\mathcal{D}(1), \ldots, \mathcal{D}(N), p_1^{\max}, \ldots, p_N^{\max})$ denote the achievable rate region for given $\mathcal{D}(1), \ldots, \mathcal{D}(N)$ and maximum node transmission power $p_1^{\max}, \ldots, p_N^{\max}$. We denote by $\mathcal{R}^{\mathrm{SIC}}(p_1^{\max}, \ldots, p_N^{\max})$ the achievable rate region, which is obtained as a union of all $\mathcal{R}^{\mathrm{SIC}}(\mathcal{D}(1), \ldots, \mathcal{D}(N), p_1^{\max}, \ldots, p_N^{\max})$ over all possible $2^{\sum_{n \in \mathcal{N}} (L - |\mathcal{I}(n)|)}$ combinations of sets $\mathcal{D}(1), \ldots, \mathcal{D}(N)$, i.e.,

$$\mathcal{R}^{\mathrm{SIC}}(p_1^{\max}, \ldots, p_N^{\max}) = \bigcup_{\mathcal{D}(1),\ldots,\mathcal{D}(N) | \forall n \in \mathcal{N} \; \exists \mathcal{U}(n) \subseteq \mathcal{L} \setminus \mathcal{I}(n) \text{ s.t. } \mathcal{D}(n) = \mathcal{I}(n) \cup \mathcal{U}(n)} \mathcal{R}^{\mathrm{SIC}}(\mathcal{D}(1), \ldots, \mathcal{D}(N), p_1^{\max}, \ldots, p_N^{\max}) \,. \quad (3.29)$$

The receiver of each node $n \in \mathcal{N}$ is allowed to perform SIC in its own order. Let $\boldsymbol{\pi}_n = \big(\pi_n(1), \ldots, \pi_n(|\mathcal{D}(n)|)\big)$ be an arbitrary permutation of the links in $\mathcal{D}(n)$, which describes the decoding and cancelation order at node n. Specifically, the signal of link $\pi_n(l)$ is decoded after all codewords of links $\pi_n(j)$, $j < l$ have been decoded and their contribution to the signal received at node n has been canceled. Thus, only the signals of the links $\pi_n(j)$, $j > l$ act as interference. The rate region $\mathcal{R}^{\mathrm{SIC}}(\mathcal{D}(1), \ldots, \mathcal{D}(N), p_1^{\max}, \ldots, p_N^{\max})$ is

obtained by considering all possible combinations of decoding orders for all nodes, i.e., all possible $\prod_{n \in \mathcal{N}} (|\mathcal{D}(n)| \, !)$ combinations $\boldsymbol{\pi} \stackrel{\triangle}{=} \boldsymbol{\pi}_1 \times \boldsymbol{\pi}_2 \times \ldots \times \boldsymbol{\pi}_N$. Thus, $\mathcal{R}^{\text{SIC}}(\mathcal{D}(1), \ldots, \mathcal{D}(N), p_1^{\max}, \ldots, p_N^{\max})$ can be expressed as

$$
\mathcal{R}^{\text{SIC}}(\mathcal{D}(1), \ldots, \mathcal{D}(N), p_1^{\max}, \ldots, p_N^{\max}) =
$$

$$
\bigcup_{\boldsymbol{\pi}} \left\{ (r_1, \ldots, r_L) \; \middle| \;
\begin{array}{l}
r_{\pi_n(l)} \leq \log \left(1 + \dfrac{G_{\pi_n(l)n}(t) \, p_{\pi_n(l)}}{\sigma^2 + \sum_{j>l} G_{\pi_n(j)n}(t) \, p_{\pi_n(j)}} \right), \\
\qquad \forall (n,l) \text{ s.t. } n \in \mathcal{N}, \; l \in \{1, \ldots, |\mathcal{D}(n)|\} \\
\sum_{l \in \mathcal{O}(n)} p_l \leq p_n^{\max}, \; n \in \mathcal{N} \\
p_l \geq 0, \; l \in \mathcal{L}
\end{array}
\right\} , \quad (3.30)
$$

where G_{ln}, $l \in \mathcal{L}$, $n \in \mathcal{N}$ represents the power gain from the transmitter of link l to the receiver at *node n*, and p_l represents the power allocated for the signal of link l. Clearly, the computational complexity experiences a formidable increase. Nevertheless, the RA subproblem at the third step of Dynamic Cross-Layer Control Algorithm 3.1 can be written as [15]

$$
\begin{array}{ll}
\text{maximize} & \sum_{l \in \mathcal{L}} \beta_l(t) r_l \\
\text{subject to} & (r_l)_{l \in \mathcal{L}} \in \mathcal{R}^{\text{SIC}}(p_1^{\max}, \ldots, p_N^{\max}) ,
\end{array}
\qquad (3.31)
$$

where the variable is $(r_l)_{l \in \mathcal{L}}$.

The combinatorial description of $\mathcal{R}^{\text{SIC}}(p_1^{\max}, \ldots, p_N^{\max})$ implies that solving problem (3.31) requires optimization over all possible combinations of decoding sets $\mathcal{D}(1), \ldots, \mathcal{D}(N)$ and decoding orders $\boldsymbol{\pi}$. This is intractable, even for off line optimization of moderate size networks. Therefore, in the following we propose two alternatives to find the solution of a more constrained version of problem (3.31) instead of solving problem (3.31) itself. The first alternative limits the access protocol so that only one node can transmit in all its outgoing links in each time slot. The second alternative adopts a similar view by assuming that only one node can receive from all its incoming links in each time slot. The main advantage of the alternatives above is their simplicity. As a result, a cheaply computable lower bound on the optimal value of problem (3.31) can be obtained. Moreover, these simple access protocols can be useful in practical applications with more advanced communication systems.

---

[15]Note that $\mathcal{R}^{\text{SIC}}(p_1^{\max}, \ldots, p_N^{\max})$ represents the set of *directly achievable* rates. By invoking a time sharing argument, one can extend the achievable rate region to the convex hull of $\mathcal{R}^{\text{SIC}}(p_1^{\max}, \ldots, p_N^{\max})$. However, this would not affect the optimal value of problem (3.31) because the objective function is linear [39].

3.3.1 Single node transmission case

By imposing the additional constraint that only one node can transmit during each slot, problem (3.31) is reduced to a problem where the optimal power and rate allocation can be computed via convex programming. Specifically, problem (3.31) is reduced to N WSRMax problems for the scalar broadcast channel, one for each possible transmitting node.

For any node $n \in \mathcal{N}$, let $\boldsymbol{\rho}_n = \big(\rho_n(1), \ldots, \rho_n(|\mathcal{O}(n)|)\big)$ be a permutation of the set of outgoing links $\mathcal{O}(n)$ such that

$$g_{\rho_n(1)\rho_n(1)}(t) \leq g_{\rho_n(2)\rho_n(2)}(t) \leq \cdots \leq g_{\rho_n(|\mathcal{O}(n)|)\rho_n(|\mathcal{O}(n)|)}(t) \;,$$

where $g_{ii}(t)$ denotes the power gain from the transmitter of link i to the receiver of link i during time slot t. Now we consider the case where node n is the transmitter. This results in a scalar Gaussian broadcast channel with $|\mathcal{O}(n)|$ users. Thus for all $i \in \{1, \ldots, |\mathcal{O}(n)|\}$, the optimal decoding and cancelation order at the receiver node of links $\rho_n(i)$ is specified by $\boldsymbol{\rho}_n$ [9, Sec. 6]. Specifically, the receiver of the link $\rho_n(i)$ decodes its own signal after all the codewords of links $\rho_n(j)$, $j < i$ have been decoded and their contribution to the received signal has been canceled. Thus, only the signals of the links $\rho_n(j)$, $j > i$ act as interference at the receiver of the link $\rho_n(i)$. Now we can rewrite problem (3.31) by using the capacity region descriptions of the scalar Gaussian broadcast channels [185] as

$$
\begin{aligned}
\text{maximize} \quad & \textstyle\sum_{l \in \mathcal{O}(n)} \beta_l r_l \\
\text{subject to} \quad & n \in \mathcal{N} \\
& r_{\rho_n(i)} \leq \log\left(1 + \frac{g_{\rho_n(i)\rho_n(i)} \; p_{\rho_n(i)}}{\sigma^2 + g_{\rho_n(i)\rho_n(i)} \sum_{j=i+1}^{|\mathcal{O}(n)|} \; p_{\rho_n(j)}}\right), \\
& \hspace{4cm} i \in \{1, \ldots, |\mathcal{O}(n)|\} \\
& \textstyle\sum_{l \in \mathcal{O}(n)} p_l \leq p_n^{\max} \\
& p_l \geq 0, \; l \in \mathcal{O}(n) \\
& p_l = 0, \; l \notin \mathcal{O}(n) \;,
\end{aligned}
\tag{3.32}
$$

where the variables are n, $(p_l)_{l \in \mathcal{L}}$, and $(r_l)_{l \in \mathcal{L}}$. Note that the time index t is dropped for notational convenience. The solution of problem (3.32) is obtained in two steps. First, we solve N independent subproblems (one subproblem for each possible transmitting node $n \in \mathcal{N}$). Then we select the solution of the subproblem with the largest objective value. The subproblem can be expressed

as

$$
\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^{|\mathcal{O}(n)|} \beta_{\rho_n(i)} r_{\rho_n(i)} \\
\text{subject to} \quad & r_{\rho_n(i)} = \log\left(1 + \frac{g_{\rho_n(i)\rho_n(i)}\, p_{\rho_n(i)}}{\sigma^2 + g_{\rho_n(i)\rho_n(i)} \sum_{j=i+1}^{|\mathcal{O}(n)|} p_{\rho_n(j)}}\right), \\
& \qquad\qquad\qquad\qquad\qquad\qquad\quad i \in \{1, \dots, |\mathcal{O}(n)|\} \\
& \sum_{l \in \mathcal{O}(n)} p_l \le p_n^{\max} \\
& p_l \ge 0, \ l \in \mathcal{O}(n)\ ,
\end{aligned}
\tag{3.33}
$$

where the variables are $(p_l)_{l \in \mathcal{O}(n)}$ and $(r_l)_{l \in \mathcal{O}(n)}$. Problem (3.33) represents the WSRMax over the capacity region of a scalar Gaussian broadcast channel [185, Sec. 2] with $|\mathcal{O}(n)|$ users. The *barrier method* [3, Sec. 11.3.1], or the explicit greedy method proposed in [185, Sec. 3.2], can be used to efficiently solve this problem. Here we use the barrier method and refer the reader to Appendix 5 for more details. Let $g^{(n)}$, $p_l^{(n)}$, and $r_l^{(n)}$ denote the optimal objective value and the corresponding optimal solution (i.e., power and rate) respectively. Then the rate/power relation can be expressed as

$$
r_{\rho_n(i)}^{(n)} = \log\left(1 + \frac{g_{\rho_n(i)\rho_n(i)}\, p_{\rho_n(i)}^{(n)}}{\sigma^2 + g_{\rho_n(i)\rho_n(i)} \sum_{j=i+1}^{|\mathcal{O}(n)|} p_{\rho_n(j)}^{(n)}}\right), \ i \in \{1, \dots, |\mathcal{O}(n)|\}
\tag{3.34}
$$

and the optimal solution of problem (3.32) is given by

$$
\begin{aligned}
n^\star &= \arg\max_{n \in \mathcal{N}} g^{(n)} \ ; \\
p_l^\star &= \begin{cases} p_l^{(n^\star)} & l \in \mathcal{O}(n^\star) \\ 0 & \text{otherwise} \ ; \end{cases} \\
r_l^\star &= \begin{cases} r_l^{(n^\star)} & l \in \mathcal{O}(n^\star) \\ 0 & \text{otherwise} \ . \end{cases}
\end{aligned}
\tag{3.35}
$$

3.3.2 Single node reception case

Here we consider the case where only one node can receive during each slot. As a result, the associated problem (3.31) is reduced to a simpler form where the optimal power and rate allocation can be computed very efficiently by considering N WSRMax problems for the Gaussian multiple access channel, one for each possible receiving node.

We start by considering the capacity region descriptions of the Gaussian multiaccess channel with $|\mathcal{I}(n)|$, $n \in \mathcal{N}$ users [179],[9, Sec. 6]. For any receiving

108

node $n \in \mathcal{N}$, the capacity region of a the $|\mathcal{I}(n)|$ user Gaussian multiaccess channel, with power constraints p_l, $l \in \mathcal{I}(n)$, is given by the set of rate vectors that lie in the intersection of the constraints

$$\sum_{l \in \mathcal{V}(n)} r_l \leq \log \left(1 + \frac{\sum_{l \in \mathcal{V}(n)} g_{ll} p_l}{\sigma^2} \right) \tag{3.36}$$

for every subset $\mathcal{V}(n) \subseteq \mathcal{I}(n)$. Thus, we can rewrite problem (3.31) as

$$\begin{aligned} \text{maximize} \quad & \sum_{l \in \mathcal{I}(n)} \beta_l r_l \\ \text{subject to} \quad & n \in \mathcal{N} \\ & \sum_{l \in \mathcal{V}(n)} r_l \leq \log \left(1 + \frac{\sum_{l \in \mathcal{V}(n)} g_{ll} p_l}{\sigma^2} \right), \; \mathcal{V}(n) \subseteq \mathcal{I}(n) \\ & 0 \leq p_l \leq p_{\mathrm{tran}(l)}^{\max}, \; l \in \mathcal{I}(n) \\ & p_l = 0, \; l \notin \mathcal{I}(n) \;, \end{aligned} \tag{3.37}$$

where the variables are n, $(p_l)_{l \in \mathcal{L}}$, and $(r_l)_{l \in \mathcal{L}}$. Again, the solution is obtained in two steps. First, we solve N independent subproblems (one subproblem for each possible receiving node $n \in \mathcal{N}$). Then we select the solution of the subproblem with the largest objective value. The subproblem has the form

$$\begin{aligned} \text{maximize} \quad & \sum_{l \in \mathcal{I}(n)} \beta_l r_l \\ \text{subject to} \quad & \sum_{l \in \mathcal{V}(n)} r_l \leq \log \left(1 + \frac{\sum_{l \in \mathcal{V}(n)} g_{ll} p_l}{\sigma^2} \right), \; \mathcal{V}(n) \subseteq \mathcal{I}(n) \\ & 0 \leq p_l \leq p_{\mathrm{tran}(l)}^{\max}, \; l \in \mathcal{I}(n) \;, \end{aligned} \tag{3.38}$$

where the variables are $(p_l)_{l \in \mathcal{I}(n)}$ and $(r_l)_{l \in \mathcal{I}(n)}$. Problem (3.38) is equivalent to the WSRMax over the capacity region of the Gaussian multiaccess channel with $|\mathcal{I}(n)|$ users [9, Sec. 6]. The solution is readily obtained by considering the *polymatroid* structure of the capacity region [179, Lem. 3.2]. Again we denote by $g^{(n)}$, $p_l^{(n)}$, and $r_l^{(n)}$ the optimal objective value and the optimal solution of problem (3.38) respectively. Thus, the solution of problem (3.38) can be written in closed form as $p_l^{(n)} = p_{\mathrm{tran}(l)}^{\max}$ for all $l \in \mathcal{I}(n)$ and

$$r_{\varrho_n(i)}^{(n)} = \log \left(1 + \frac{g_{\varrho_n(i)\varrho_n(i)} \, p_{\varrho_n(i)}^{(n)}}{\sigma^2 + \sum_{j=i+1}^{|\mathcal{I}(n)|} g_{\varrho_n(j)\varrho_n(j)} \, p_{\varrho_n(j)}^{(n)}} \right), \; i \in \{1, \ldots, |\mathcal{I}(n)|\} \;, \tag{3.39}$$

where $\varrho_n = (\varrho_n(1), \ldots, \varrho_n(|\mathcal{I}(n)|))$ is a permutation of the set of incoming links $\mathcal{I}(n)$ such that

$$\beta_{\varrho_n(1)} \leq \beta_{\varrho_n(2)} \leq \ldots \leq \beta_{\varrho_n(|\mathcal{I}(n)|)} \;. \tag{3.40}$$

One can in fact identify ϱ_n as the SIC order at the receiving node $n \in \mathcal{N}$. Finally, the optimal solution of problem (3.37) can be expressed as

$$
\begin{aligned}
n^\star &= \arg\max_{n \in \mathcal{N}} g^{(n)} \ ; \\
p_l^\star &= \begin{cases} p_l^{(n^\star)} & l \in \mathcal{I}(n^\star) \\ 0 & \text{otherwise} \ ; \end{cases} \\
r_l^\star &= \begin{cases} r_l^{(n^\star)} & l \in \mathcal{I}(n^\star) \\ 0 & \text{otherwise} \ . \end{cases}
\end{aligned} \tag{3.41}
$$

3.4   Numerical examples

In this section, we use the algorithms of the preceding sections to identify the solutions to the selected NUM problem and their properties, so as to get insight into network design and provisioning methods. Specifically, in every time slot t, the rate allocation at step 3 of the Dynamic Cross-Layer Control Algorithm (i.e., Algorithm 3.1, Section 3.1) is obtained by using the proposed algorithms for WSRMax described in Section 3.2 and Section 3.3.

We assume a block fading Rayleigh channel model where the channel gains are constant during each time slot and change independently from slot-to-slot. The small-scale fading components of the channel gains are assumed to be independent and identically distributed over the time slots, links, and channels. Recall that we consider equal power spectral density for all receivers, i.e., $N_l = N_0$ for all $l \in \mathcal{L}$ and equal channel bandwidths, i.e., $W_c = W$ for all $c \in \mathcal{C}$. Furthermore, the maximum power constraint is assumed to be the same for all nodes, i.e., $p_n^{\max} = p_0^{\max}$ for all $n \in \mathcal{N}$ (independent of the number of channels C). For a fair comparison between cases with different numbers of channels, we have assumed that the total available bandwidth is constant regardless of C, i.e., $\sum_{c=1}^{C} W_c = W_{\text{tot}}$. In all our simulations we have selected the total bandwidth to be normalized to one, i.e., $W_{\text{tot}} = 1$ Hz.

For comparing different algorithms, we consider the following two performance metrics: 1) *the average sum-rate* $\sum_{n \in \mathcal{N}} \sum_{s \in \mathcal{S}_n} \bar{x}_n^s$ and 2) *the average network congestion* $\sum_{n \in \mathcal{N}} \sum_{s=1}^{S} \bar{q}_n^s$. For each network instance, the Dynamic Cross-Layer Control Algorithm (i.e., Algorithm 3.1) is simulated for at least $\acute{T} = 10000$ time slots and the average rates $\bar{x}_n^s$ and queue sizes $\bar{q}_n^s$ are computed by averaging the last $t_0 = 3000$ time slots, i.e., $\bar{x}_n^s = 1/t_0 \sum_{t=\acute{T}-t_0}^{\acute{T}} x_n^s(t)$ and

$\bar{q}_n^s = 1/t_0 \sum_{t=\acute{T}-t_0}^{\acute{T}} q_n^s(t)$. We assume that the rates corresponding to all node-commodity pairs $(n,s)_{s \in \mathcal{S}_n}, n \in \mathcal{N}$ are subject to proportional fairness, and therefore we select the utility functions $u_n^s(y) = \log_e(y)$. In all considered setups, we selected $V = 100$ [see (3.4)] and the parameters $R_n^{\max}$ [see (3.4)] were chosen such that all conditions presented in [44, Sec. III-D] were satisfied.

We start with a simple network instance (Section 3.4.1); a bipartite network where no self interferers exist (i.e., $\mathcal{A} = \emptyset$) and the proposed successive approximation algorithm, Algorithm 3.2 (Section 3.2.2) is used in resource allocation. The associated results show important consequences on upper layers due to the proposed successive approximation algorithm. We then consider more general networks (Section 3.4.2), with the presence of self interferers (i.e., $\mathcal{A} \neq \emptyset$) and no self interference cancellation at network's nodes (i.e., $g = 1$). Here Algorithm 3.3 (Section 3.2.4) is used in resource allocation. The gains achievable at the network layer, due to different degrees of the self interference cancelation performed at the network nodes, are investigated quantitatively in Section 3.4.3. By changing g in the interval $[0,1]$, the results are able to capture the effect of self interference cancelation performed with different levels of accuracy. Finally, we look at the multiuser receiver scenario, again using the same network instance as in Section 3.4.2. The associated results (Section 3.4.4) show impacts in upper layer performance due to advanced receiver architectures.

3.4.1   NUM for bipartite networks with singleuser detection at receivers

A bipartite network, as shown in Figure 3.4, is considered. There are $N = 8$ nodes, $L = 4$ links, and $S = 4$ commodities. One distinct commodity arrives exogenously at every node n from the subset $\{1,2,3,4\} \subseteq \mathcal{N}$. Without loss of generality we assume that the nodes and commodities are labeled such that commodity i arrives at node i for any $i \in \{1,2,3,4\}$. The destination nodes are specified by the following commodity-destination node pairs $(s,d_s) \in \{(1,5),(2,6),(3,7),(4,8)\}$.

The channel power gains between distinct nodes are given by

$$|h_{ijc}(t)|^2 = \mu^{|i-j|} c_{ijc}(t), \ i,j \in \mathcal{L}, \ c \in \mathcal{C} , \tag{3.42}$$

where $c_{ijc}(t)$ are exponentially distributed independent random variables with unit mean used to model Rayleigh small-scale fading and the scalar $\mu \in [0,1]$ is
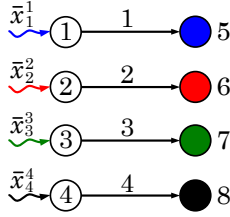
Fig 3.4. Bipartite wireless network with $N = 8$ nodes, $L = 4$ links, and $S = 4$ commodities, [112] © 2011, IEEE.

referred to as the interference coupling index, which parameterizes the interference between direct links. For example, if $\mu = 0$, transmissions of links are interference free. The interference between transmissions increases as the parameter μ grows. Similar channel gain models for bipartite networks has also been used in [186]. Of course, this simple hypothetical model provides useful insights into the performance of proposed algorithms in bipartite networks (e.g., cellular networks). We define the SNR operating point as

$$\text{SNR} = \frac{p_0^{\max}}{N_0 W_{\text{tot}}} \ . \tag{3.43}$$

Figure 3.5 shows the dependence of the average sum-rate [Figure 3.5(a)] and the average network congestion [Figure 3.5(b)] on the interference coupling index μ for our proposed Algorithm 3.2 and for the optimal base line single link activation (or BLSLA) policy [16]. We consider the single-channel case $C = 1$ operating at three different SNR values 2, 8, and 16dB. The initial power allocation $\mathbf{P}_0$ for Algorithm 3.2 is chosen such that $[\mathbf{P}_0]_{l,1} = p_0^{\max}$ unless otherwise specified. Here we can make several observations. First, the proposed Algorithm 3.2 provides substantial gains, both in the average sum-rate, as well as in the average network congestion, especially for small and medium values of the interference coupling index. The gains diminish as interference between direct links become significant. This is intuitively expected since for large SNR values the BLSLA policy becomes optimal when the interference coupling index μ approaches 1. It is interesting to note that at small SNR values the network can

---

[16] A channel access policy where, during each time slot, only one link is activated in each channel is called BLSLA policy. Finding the optimal BLSLA policy that solves problem (3.6) is a combinatorial problem with exponential complexity in C. Thus, it quickly becomes intractable, even for moderate values of C. However, for the case $C = 1$ the optimal BLSLA policy can be easily found and it consists of activating, during each time slot, only the link that achieves the maximum weighted rate.
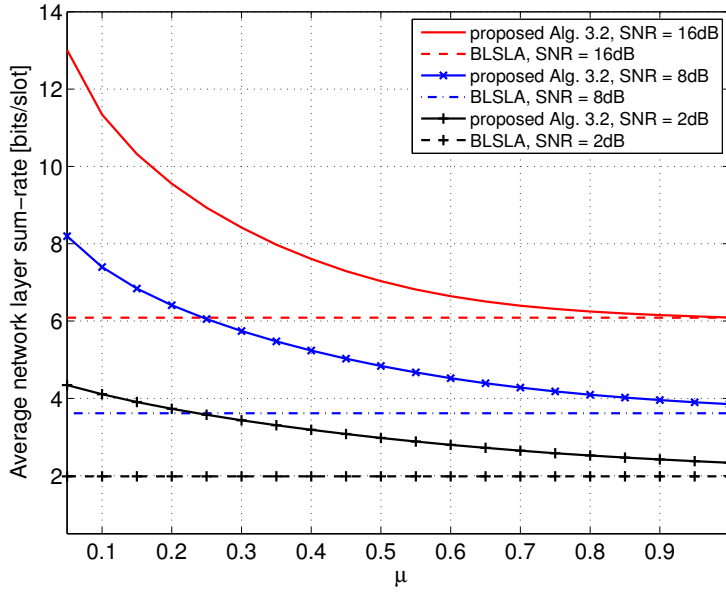
(a) Average network layer sum-rate $\sum_{s=1}^{4} \bar{x}_s^s$
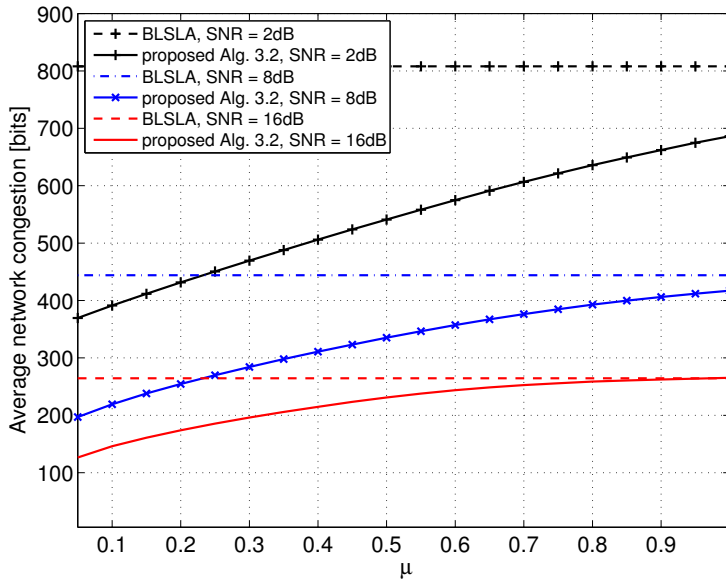
(b) Average network congestion $\sum_{s=1}^{4} \bar{q}_s^s$

Fig 3.5. Dependence of the average sum-rate and the average network congestion on the interference coupling index μ**;** $C = 1$ **and** $\mathrm{SNR} = 2, 8, 16$**dB, [112] © 2011, IEEE.**

still benefit from scheduling multiple links per slot, even for the case $\mu = 1$. This gain comes from the fact that the channel gains between interfering links are also affected by fading. Thus, links that experience low instantaneous interference levels can be simultaneously scheduled. Results suggest that, especially for small and medium values of the interference coupling index, the proposed solution method often yields designs that are far superior to those obtained by BLSLA.

Figure 3.6 shows the dependence of the average sum-rate [Figure 3.6(a)] and of the average network congestion [Figure 3.6(b)] on the number of iterations of Algorithm 3.2. We consider the single-channel case $C = 1$ with interference coupling index $\mu = 0.5$ and SNR values 2, 8, and 16dB. To facilitate faster convergence, Algorithm 3.2 is run without considering the trust region constraints; to do this, we can simply set the parameter α in Algorithm 3.2 to a very large positive number, e.g., $\alpha = 10^{100}$ [see problem (3.18)]. As a reference, we consider the optimal BLSLA policy. Results show that the incremental benefits are very significant for the first few iterations and are marginal later. For example, in the case of SNR $= 16$dB, when the numbers of iterations changes from 1 to 3, the improvement in the average sum-rate is around 18.1%, whereas when it changes from 7 to 9, the improvement is around 0.30%. Therefore, by running Algorithm 3.2 for a few iterations (e.g., 5 iterations) we can yield performance levels which are almost indistinguishable from those that would have been obtained by running Algorithm 3.2 until it terminates (see the stopping criterion in step 3). This observation can be very useful in practice, since we can terminate Algorithm 3.2 when the incremental improvements between consecutive iterations become negligible.

Figure 3.7 shows the dependence of the average sum-rate [Figure 3.7(a)] and of the average network congestion [Figure 3.7(b)] on the SNR for Algorithm 3.2 and optimal BLSLA policy. We have considered the case $C = 1$ and $\mu = 0.3$. For comparison, we also plot the results due to a commonly used high SINR approximation [34] where the achievable rates $\log(1 + \gamma_{lc})$ are approximated by $\log(\gamma_{lc})$. In particular, the objective function of problem (3.11) is approximated by $\prod_{c \in \mathcal{C}} \prod_{l \in \mathcal{L}} \gamma_{lc}^{-\beta_l}$. Recall that γ_{lc} represents the SINR of link l in channel c and β_l represents the differential backlog of link l. This results in a convex approximation (i.e., a GP) of problem (3.11). One should not confuse high SINR with high SNR, since those are fundamentally different and a high SNR value does not ensure high SINR values in all links. Results show that, when compared

114

(a) Average network layer sum-rate $\sum_{s=1}^{4} \bar{x}_s^s$
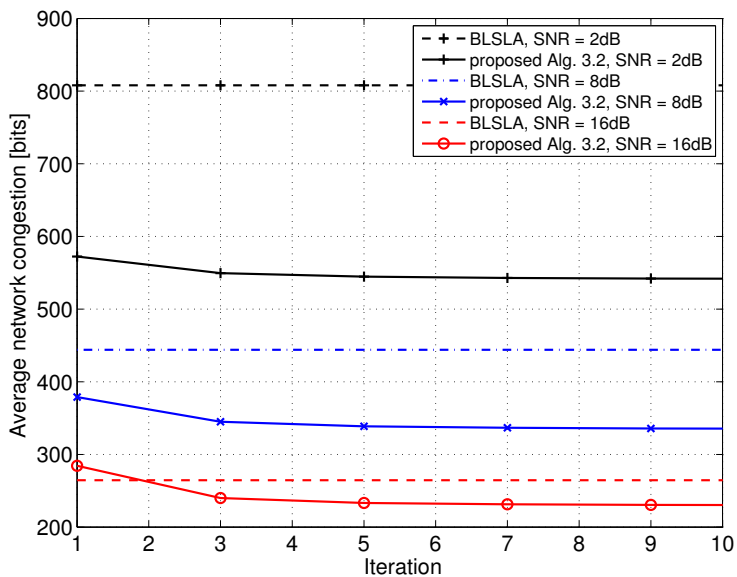
(b) Average network congestion $\sum_{s=1}^{4} \bar{q}_s^s$

Fig 3.6. Dependence of the average sum-rate and the average network congestion on the iteration; $\mu = 0.5$, $C = 1$ **and** $\mathrm{SNR} = 2, 8, 16$**dB, [112] © 2011, IEEE.**

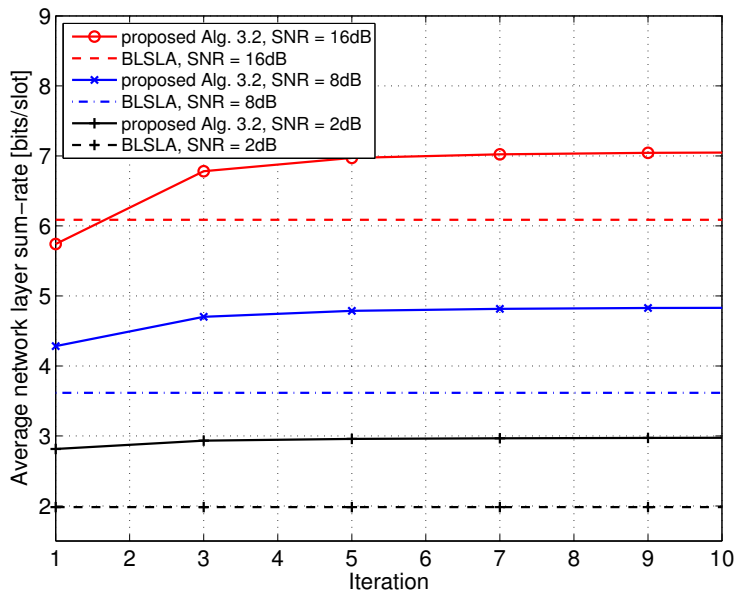(a) Average network layer sum-rate $\sum_{s=1}^{4} \bar{x}_s^s$

(b) Average network congestion $\sum_{s=1}^{4} \bar{q}_s^s$

Fig 3.7. Dependence of the average sum-rate and the average network congestion on the SNR; $C = 1$ **and** $\mu = 0.3$**, [112] © 2011, IEEE.**

with other methods, RA based on Algorithm 3.2 offers larger average sum-rate as well as reduced average network congestion. The relative gains of Algorithm 3.2 reduce, compared to the BLSLA at high SNR, e.g., the relative gain offered by proposed Algorithm 3.2 in the average sum-rate changes from 40% to 17% [Figure 3.7(a)] and the relative gain in the average network congestion changes from 23% to 15% [Figure 3.7(b)] when the SNR value is increased from $\gamma = 16$dB to $\gamma = 24$dB respectively. This observation is consistent with the fact that at high SNR it is very likely the optimal RA has a BLSLA structure. As a result, at the optimal RA, different links correspond to different SINR regions, and therefore the high SINR approximation is, of course, unreasonable and suffers a large penalty, especially at high SNR values. This poor performance is qualitatively consistent with intuition: the solution obtained by employing high SINR approximation in RA must contain all nonzero entries (i.e., nonzero γ_{lc}) to drive the approximated objective (i.e., $\prod_{c \in \mathcal{C}} \prod_{l \in \mathcal{L}} \gamma_{lc}^{-\beta_l}$) into a nonzero value, and therefore never yields a solution to the form of BLSLA.

Figure 3.8 shows the dependence of the average sum-rate [Figure 3.8(a)] and of the average network congestion [Figure 3.8(b)] on the number of channels C for Algorithm 3.2. We consider the case SNR $= 16$dB and $\mu = 0.3$ and the initial power allocation $\mathbf{P}_0$ for Algorithm 3.2 is simply chosen such that $[\mathbf{P}_0]_{l,c} = p_0^{\max}/C$. The plots illustrate that increasing the number of channels will yield better performance in both the average sum-rate and the average network congestion (e.g., when the number of channels C changes from 1 to 8, the improvement in the average sum-rate and the reduction in average network congestion is around 12% and 12.4% respectively). We stress that the benefits are solely achieved by opportunistically exploiting the available multichannel diversity in the network via the proposed Algorithm 3.2 without any supplementary bandwidth or power consumption. Moreover, the incremental benefits are very significant for small C, e.g., when the number of channels C changes from 1 to 2, the improvement in the average sum-rate is around 6%, whereas when C changes from 7 to 8, the improvement is around 0.25%. The plots give much insight into *why* multichannel designs are important and beneficial compared to the single-channel counterpart.
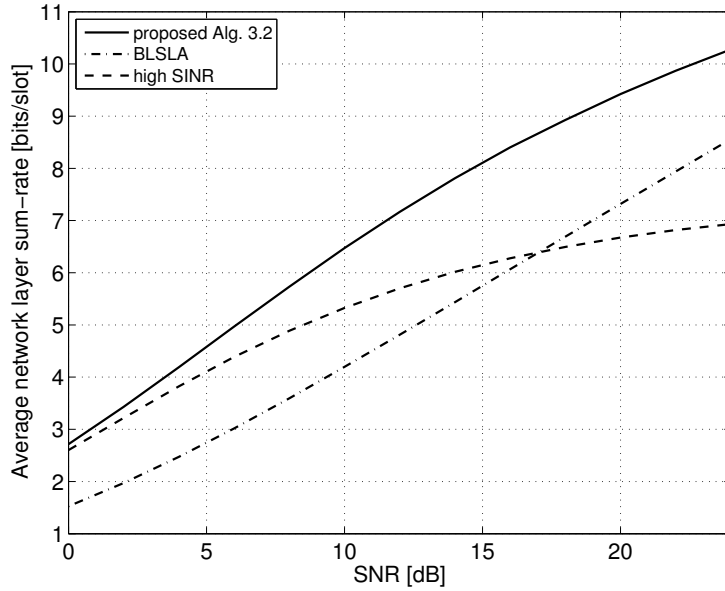
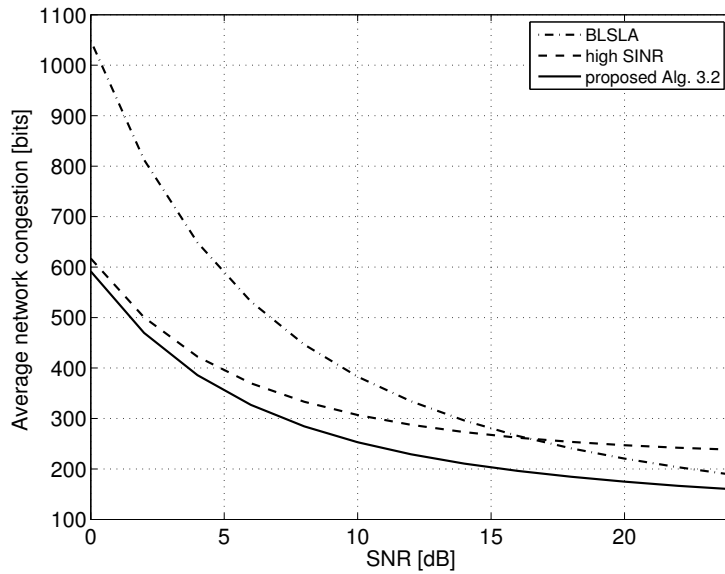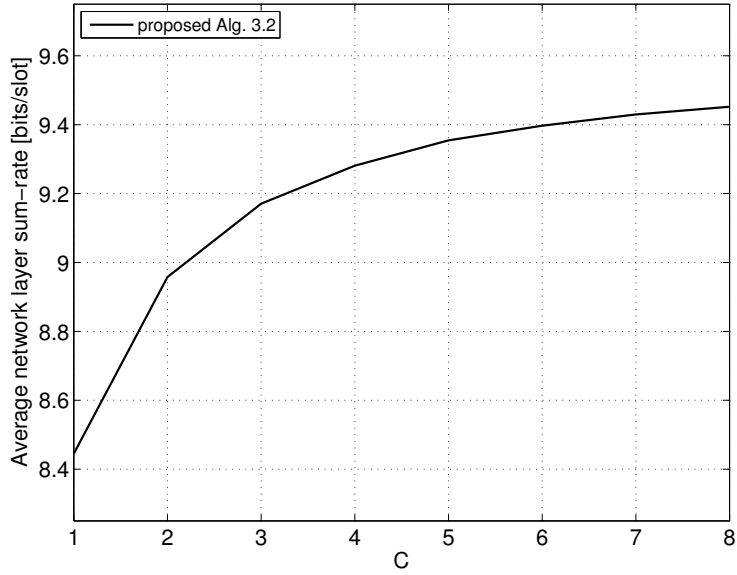(a) Average network layer sum-rate $\sum_{s=1}^{4} \bar{x}_s^s$

(b) Average network congestion $\sum_{s=1}^{4} \bar{q}_s^s$

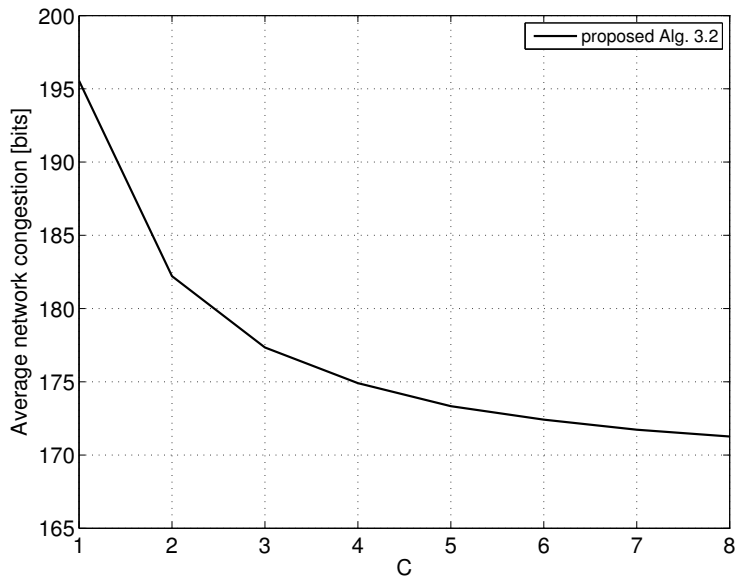**Fig 3.8. Dependence of the average sum-rate and the average network congestion on the number of channels C; $\mathrm{SNR} = 16$dB and $\mu = 0.3$, [112] © 2011, IEEE.**

3.4.2 NUM for nonbipartite networks with singleuser detection at receivers

First, two small fully connected multihop wireless network setups, which are identical to the once shown in Figure 2.11 are considered.

We assume an exponential path loss model; the channel power gains $|h_{ijc}(t)|^2$ between distinct nodes are given by

$$|h_{ijc}(t)|^2 = \left(\frac{d_{ij}}{d_0}\right)^{-\eta} c_{ijc}(t) \; , \tag{3.44}$$

where d_{ij} is the distance from the transmitter of link i to the receiver of link j, d_0 is the *far field reference distance* [176], η is the path loss exponent, and $c_{ijc}(t)$ are exponentially distributed random variables with unit mean, independent over the time slots, links, and channels. The first term of (3.44) represents the path loss factor and the second term models Rayleigh small-scale fading. The SNR operating point is defined as

$$\text{SNR} = \frac{p_0^{\text{max}}}{N_0 W_{\text{tot}}} \cdot \left(\frac{D_0}{d_0}\right)^{-\eta} \; . \tag{3.45}$$

In the following simulations we set $D_0/d_0 = 10$ and $\eta = 4$.

Figure 3.9 shows the dependence of the average network layer sum-rate on the SNR for the considered network setups, where we use $C = 1$. As a benchmark, we first consider the branch and bound algorithm proposed in Section 2.2 to optimally solve the RA subproblem. It should be stressed that the optimality of the algorithm proposed in Section 2.2 is achieved at the expense of prohibitive computational complexity, even in the case of very small problem instances. We then consider the optimal BLSLA policy and Algorithm 3.3 with two initialization methods: 1) Uniform initialization and 2) BLSLA based initialization. In the case of uniform initialization the initial power allocation $\mathbf{P}_0$ is chosen such that $[\mathbf{P}_0]_{l,1} = p_0^{\text{max}}/(|\mathcal{O}_{\text{tran}(l)}|)$. In the case of BLSLA based initialization the initial power allocation $\mathbf{P}_0$ is chosen such that $[\mathbf{P}_0]_{l^\star,1} : [\mathbf{P}_0]_{j,1} = P : 1$ for all $j \in \mathcal{L}$, $j \neq l^\star$ where $l^\star$ is the index of the active link obtained based on the optimal BLSLA policy and $P \gg 1$ is a real number. For comparison, we also plot the results for Algorithm 3.2 with uniform and BLSLA initializations.

Results show that the performance of Algorithm 3.3 is very close to the optimal branch and bound algorithm. Specifically, Algorithm 3.3 with BLSLA
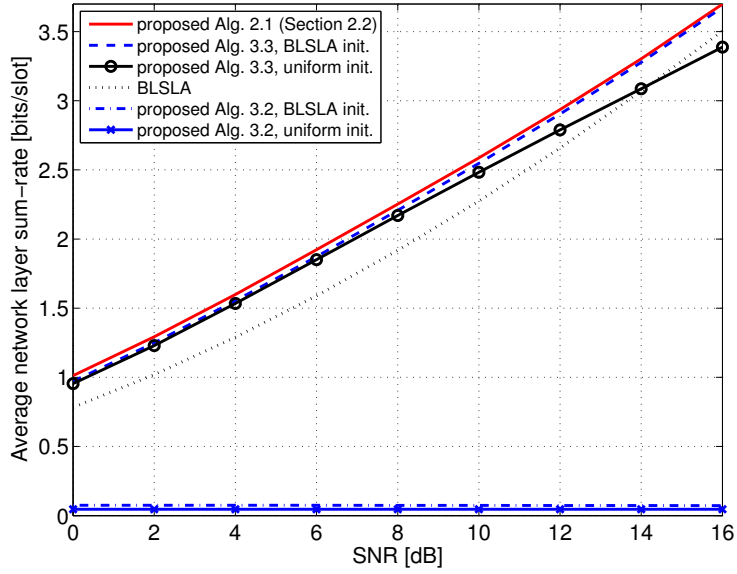
(a) Average network layer sum-rate $\bar{x}_1^1 + \bar{x}_4^2$
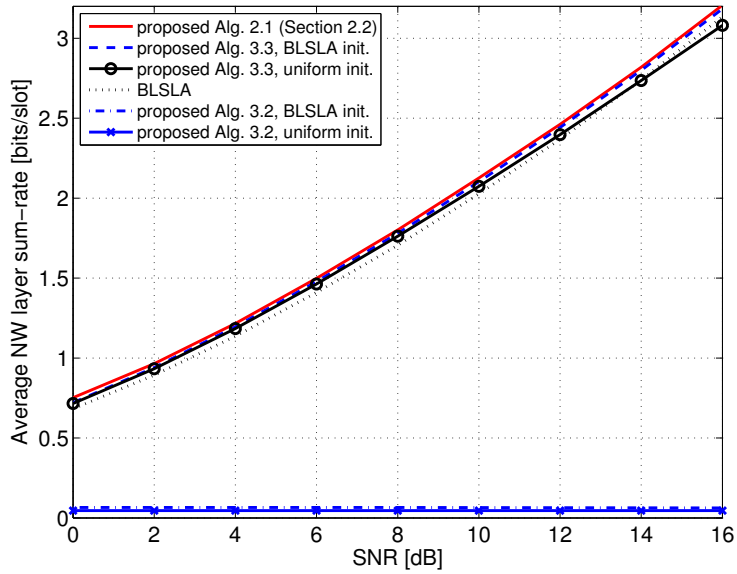
(b) Average network layer sum-rate $\bar{x}_1^1 + \bar{x}_2^2$

Fig 3.9. (a) Dependence of the average network layer sum-rate on SNR for network 1; (b) Dependence of the average network layer sum-rate on SNR for network 2, [112] © 2011, IEEE.
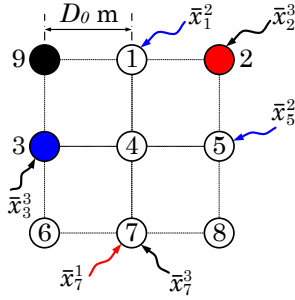
Fig 3.10. Multihop wireless network with $N = 9$ nodes and $S = 3$ commodities, [112] © 2011, IEEE.

initialization is almost indistinguishable from the optimum and at least as good as the optimal BLSLA, for all considered cases. In contrast, Algorithm 3.3 with uniform initialization exhibits significant deviations from both the optimal branch and bound algorithm and BLSLA, specially at high SNR values. This behavior is not surprising since Algorithm 3.3 is a local method for the nonconvex problem (3.6). Therefore, the initialization point of the algorithm can influence the resulting solution [3, Sec. 1.4.1]. Nevertheless, a carefully selected initialization point can improve the performance of Algorithm 3.3 to very close to the optimum. For example, at high SNR values, the performance of Algorithm 3.3 with BLSLA initialization is almost identical to the optimum, whereas the performance with uniform initialization deviates a bit from the optimum. It is important to remark that at low and moderate values of SNR, results derived from Algorithm 3.3 are not significantly affected by the initialization method. Results also convince that, in the presence of self interferers, Algorithm 3.2 cannot perform well and it can converge to a very bad suboptimal point, as we pointed out in Section 3.2.4. Therefore, even though the computational complexity of Algorithm 3.3 *does not increase* more than polynomially with the problem size, results show that Algorithm 3.3 with a proper initialization performs close to the optimum.

Next, a larger network; a fully connected multihop, multicommodity wireless network as shown in Figure 3.10 is considered. There are $N = 9$ nodes and $S = 3$ commodities. The commodities arrive exogenously at different nodes in the network as described in Table 3.1. Thus we have $\mathcal{S}_1 = \{2\}, \mathcal{S}_2 = \{3\}, \mathcal{S}_3 = \{3\}, \mathcal{S}_5 = \{2\}, \mathcal{S}_7 = \{1, 3\}$, and $\mathcal{S}_i = \emptyset$ for all $i \in \{4, 6, 8, 9\}$. The nodes are located in a rectangular grid such that the horizontal and vertical distances between adjacent nodes are D_0 m. The channel power gains, between nodes, are

Table 3.1. Network commodities, destination nodes, and source nodes, [112] ⓒ 2011, IEEE.

Commodity (s)	Destination node (d_s)	Source nodes
1	2	7
2	3	$1, 5$
3	9	$2, 3, 7$

given by (3.44) and the SNR operating point is given by (3.45). Moreover, we set $D_0/d_0 = 10$ and $\eta = 4$.

Figure 3.11 shows the dependence of the average sum-rate [Figure 3.11(a)] and of the average network congestion [Figure 3.11(b)] on the SNR for several algorithms, where we use $C = 1$. First we have considered the optimal BLSLA policy and Algorithm 3.3 with the two initialization methods: 1) Uniform initialization and 2) BLSLA based initialization (the same initializations as used when plotting Figure 3.9). For comparison, we also plot the results for the low-complex approaches where the set of nodes $\mathcal{N}$ is first partitioned into two disjoint subsets, the set of transmitting nodes $\mathcal{T}$ and the set of receiving nodes $\mathcal{R}$ and then Algorithm 3.2 and high SINR approximation are used in RA. The partitioning of the set of nodes $\mathcal{N}$ into two disjoint subsets is performed using two simple methods: 1) random partitioning and 2) greedy partitioning based on differential backlogs. In random partitioning, each node is allocated either to $\mathcal{T}$ or to $\mathcal{R}$ with equal probabilities. Greedy partitioning is performed as follows. We start with an empty set of links $\bar{\mathcal{L}} = \emptyset$. At each step, the link $l^\star$ from the set $\mathcal{L} \setminus \bar{\mathcal{L}}$ which has the largest differential backlog β_l (i.e., $l^\star = \arg\max_{l \in \mathcal{L} \setminus \bar{\mathcal{L}}} \beta_l$) is added to the set $\bar{\mathcal{L}}$. Then all links outgoing from $\mathrm{rec}(l^\star)$ and all links incoming to $\mathrm{tran}(l^\star)$ are deleted from $\mathcal{L}$. This procedure continues until there are no links left in $\mathcal{L} \setminus \bar{\mathcal{L}}$. The sets $\mathcal{T}$ and $\mathcal{R}$ can be found as $\mathcal{T} = \{\mathrm{tran}(l)|l \in \bar{\mathcal{L}}\}$ and $\mathcal{R} = \{\mathrm{rec}(l)|l \in \bar{\mathcal{L}}\}$.

From Figure 3.11 we make the following observations. First, Algorithm 3.3 with BLSLA based initialization yields results better than any other counterpart. In contrast, Algorithm 3.3 with uniform initialization shows significant deviations from the BLSLA solution at high SNR, especially in the terms of average sum-rate [Figure 3.11(a)]. Moreover, it is important to observe again that at low and moderate values of SNR, results derived from Algorithm 3.3 are not
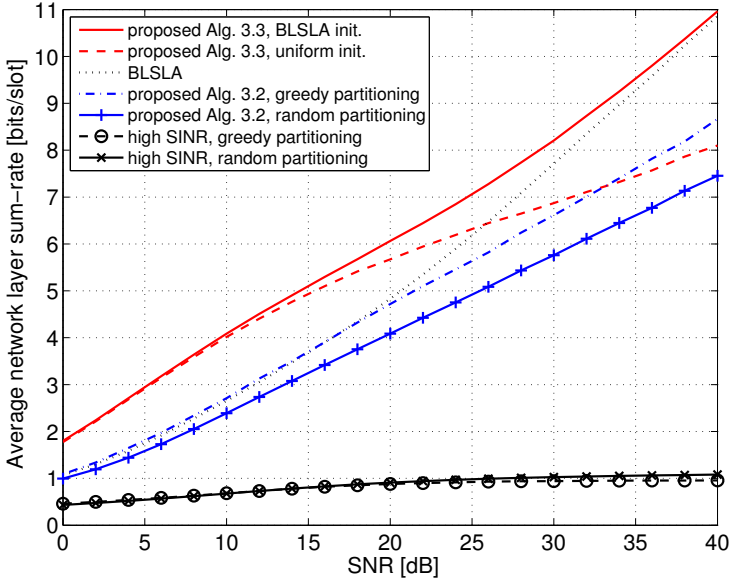
(a) Average network layer sum-rate $\sum_{n=1}^{9} \sum_{s \in \mathcal{S}_n} \bar{x}_n^s$
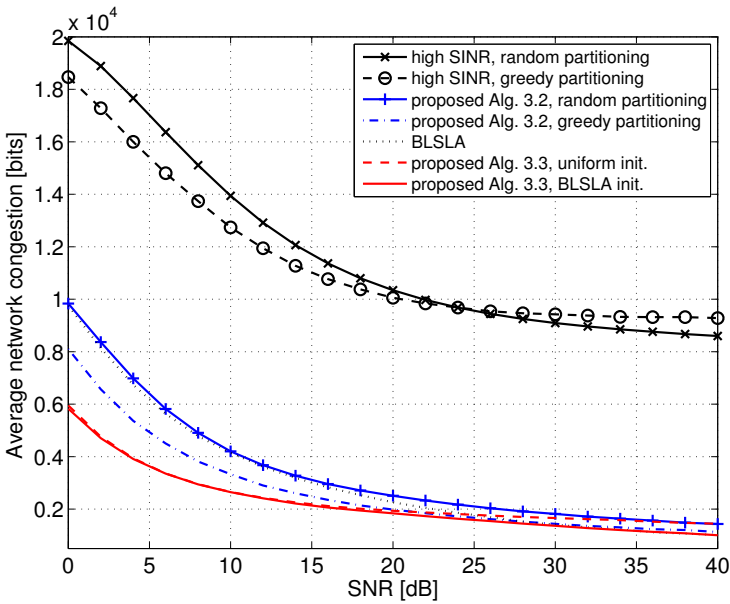
(b) Average network congestion $\sum_{n=1}^{9} \sum_{s=1}^{3} \bar{q}_n^s$

Fig 3.11. Dependence of the average sum-rate and the average network conges-tion on the SNR; $C = 1$, [112] © 2011, IEEE.

substantially affected by the initialization method. These observations are almost the same as those we saw in Figure 3.9. We also observe that Algorithm 3.3 with a proper initialization, can significantly outperform Algorithm 3.2 in conjunction with either random or greedy partitioning. This elaborates the importance of gradual self-interference gain increments (i.e., step 4 of Algorithm 3.3) in finding a better RA compared to the direct application of Algorithm 3.2 with heuristic partitioning. In most cases there is no advantage to using high SINR approximation. These observations are very useful in practice since they illustrate that Algorithm 3.3 often works well when initialized with a reasonable starting point (e.g., BLSLA based initialization). In addition, we note that even with a very simple initialization, e.g., uniform initialization, Algorithm 3.3 yields substantial gains, especially at small and moderate SNR values (e.g., 0dB - 20dB).

Figure 3.12 shows the dependence of the average sum-rate [Figure 3.12(a)] and of the average network congestion [Figure 3.12(b)] on the numbers of channels C for Algorithm 3.3. We have considered the case SNR = 16dB and have considered a uniform initialization for Algorithm 3.3 where the initial power allocation $\mathbf{P}_0$ is chosen such that $[\mathbf{P}_0]_{l,c} = p_0^{\max}/(C.|\mathcal{O}_{\mathrm{tran}(l)}|)$. For comparison, we also plot the results for Algorithm 3.2 with random and greedy partitioning of nodes $\mathcal{N}$. The results are consistent with our previous observations in Figure 3.8, i.e., as the number of channels increases better performance in both the average sum-rate and the average network congestion is achieved. These benefits are again obtained by opportunistically exploiting the available multichannel diversity in the network via the proposed algorithms. Moreover, the results suggest that using Algorithm 3.3 in the RA can increase the gains very significantly, compared to the RA based on simple extensions to Algorithm 3.2, which runs with either random or greedy partitioning of nodes. For example, the relative gains in the average sum-rate are above 23% [Figure 3.12(a)] and the relative gains in the average network congestion are above 4.7% [Figure 3.12(b)] over the range of interest, $C = 1$ to $C = 8$.

3.4.3   Effect of self interference cancelation

For the considered network setups in this section, the channel power gains between nodes are given by (3.44) and the SNR operating point is given by (3.45). For illustration purposes we consider a single-channel case (i.e., $C = 1$). Moreover,
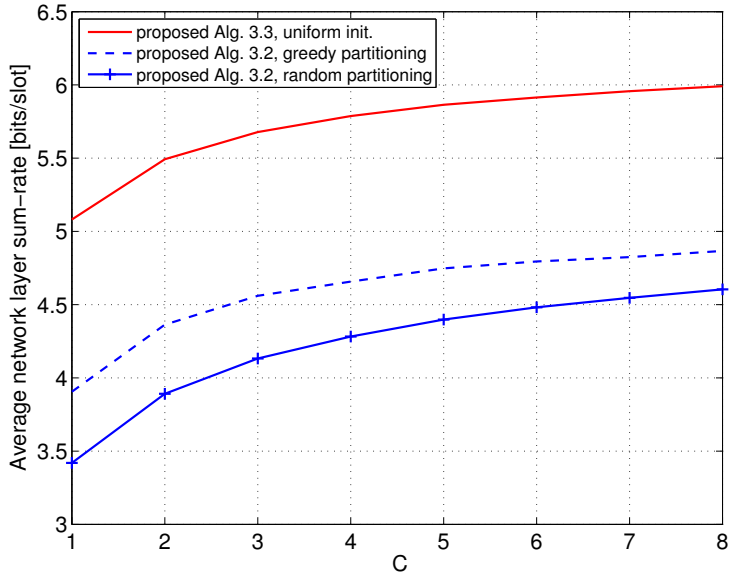
(a) Average network layer sum-rate $\sum_{n=1}^{9} \sum_{s \in \mathcal{S}_n} \bar{x}_n^s$
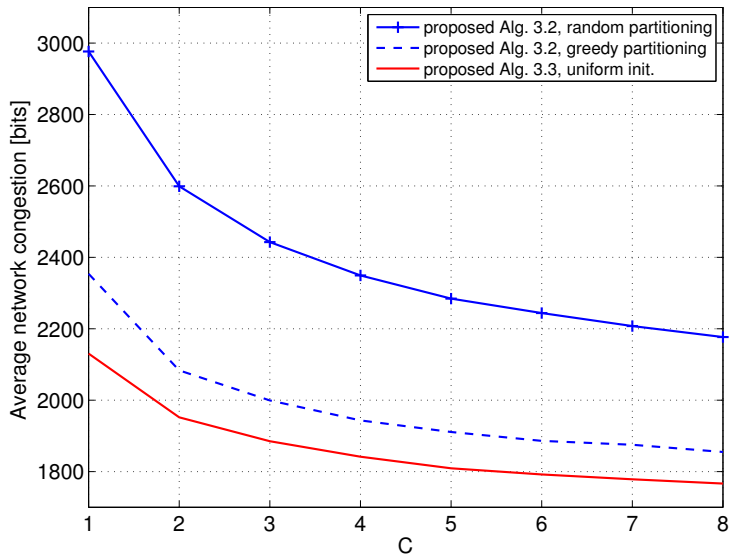
(b) Average network congestion $\sum_{n=1}^{9} \sum_{s=1}^{3} \bar{q}_n^s$

Fig 3.12. Dependence of the average sum-rate and the average network congestion on the number of channels C; $\mathrm{SNR} = 16$dB, [112] © 2011, IEEE.
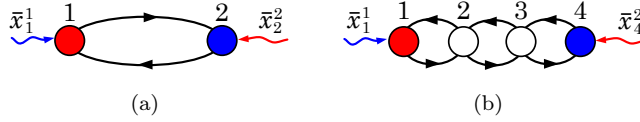
Fig 3.13. (a) Two node wireless network with $N = 2$ nodes, $L = 2$ links, and $S = 2$ commodities. Different commodities are represented by different color; (b) Tandem wireless network with $N = 4$ nodes and $S = 2$ commodities. Different commodities are represented by different color.

we set $D_0/d_0 = \sqrt{10}$ and $\eta = 4$. In all the simulations in this subsection, Algorithm 3.3 with BLSLA based initialization is considered.

Inspired by the Gaussian two-way channel [180], we first consider a simple two node wireless network as shown in Figure 3.13(a). There are two commodities, the first one arrives at node 1, and is intended for node 2; the second commodity arrives at node 2, and is intended for node 1. As explained in [180], a Gaussian two-way channel is equivalent to two independent Gaussian channels where perfect self interference cancelation is realized (i.e., $g = 0$). As a result, the sum-capacity of the symmetric Gaussian two-way channel becomes twice the capacity of either of the equivalent Gaussian channels. The considered two node network allows us to illustrate similar behavior in terms of the network layer average sum-rate.

Figure 3.14 shows the dependence of the average sum-rate [Figure 3.14(a)] and of the average network congestion [Figure 3.14(b)] on the self-interference gain g. We consider three link SNR values, $5, 16$, and 30dB, which correspond to low, medium, and high data rate systems respectively. The results show that the average sum-rate with perfect self interference cancelation (i.e., $g = 0$) is increased by a factor of 2 and the average network congestion reduced significantly, as compared to no self interference cancelation (i.e., $g = 1$); see Figure 3.14(a). Similar gains are achieved in terms of average network congestion as well; see Figure 3.14(b). The results also reveal that, even with an imperfect self interference cancelation technique, we can achieve the performance limits guaranteed by perfect self interference cancelation. For example, a decrease of the self-interference gain up to a value $g = 10^{-4}$ is enough to double the average sum-rate for link SNR $= 5$dB.

Let us now consider a tandem wireless network, as shown in Figure 3.13(b). There are two commodities, the first one arrives at node 1, and is intended for

126
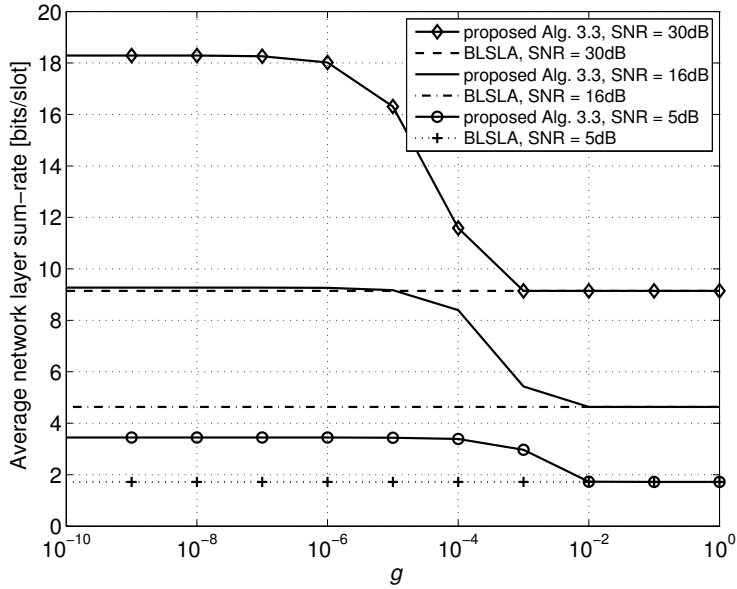
(a) Average network layer sum-rate $\sum_{s=1}^{2} \bar{x}_s^s$

(b) Average network congestion $\sum_{s=1}^{2} \bar{q}_s^s$

Fig 3.14. Dependence of the average sum-rate and of the average network congestion on the self-interference gain g in the case of the two node wireless network.

127
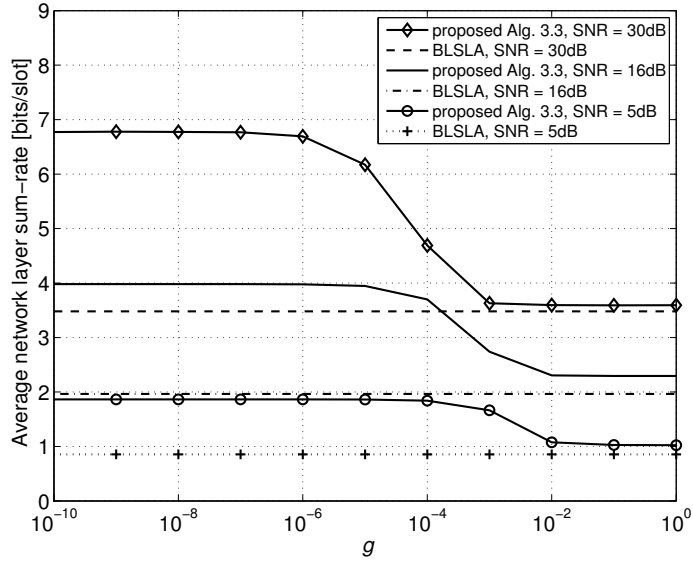
(a) Average network layer sum-rate $(\bar{x}_1^1 + \bar{x}_4^2)$

(b) Average network congestion $\sum_{n=1}^{4} \sum_{s=1}^{2} \bar{q}_n^s$

Fig 3.15. Dependence of the average sum-rate and the average network congestion on the self-interference gain g in the case of the four node tandem wireless network.

node 4; the second commodity arrives at node 4, and is intended for node 1. Thus we have $\mathcal{S}_1 = \{1\}, \mathcal{S}_4 = \{2\}$, and $\mathcal{S}_n = \emptyset$ for all $n \in \{2, 3\}$.

Figure 3.15 shows the dependence of the average sum-rate and of the average network congestion on the self-interference gain g for SNR values $5, 16$, and 30dB. The behavior is very similar to the case of Figure 3.14. For example, in the case SNR = 5dB, the results show that by decreasing the self-interference gain from $g = 10^{-1}$ to $g = 10^{-4}$ the average sum-rate is increased by a factor of around 1.82 [see Figure 3.15(a)] and the average network congestion has reduced significantly as well [see Figure 3.15(b)].

Let us next consider a fully connected multihop, multicommodity wireless network as shown in Figure 3.10. Figure 3.16 shows the dependence of the average sum-rate and of the average network congestion on the self-interference gain g for SNR values $5, 16$, and 30dB. Let us first consider the case of a low SNR value, i.e., SNR = 5dB. The results show that by decreasing the self-interference gain from $g = 10^{-1}$ to $g = 10^{-4}$ the average sum-rate is increased by a factor of about 1.22 [see Figure 3.16(a)]. From Figure 3.16(b), we see reductions in the the average network congestion as well. The network performance remains the same as in the case of perfect self interference cancelation for all values of $g < 10^{-4}$. In this region the network performance is limited by the interference between distinct nodes, and no further improvement is possible by only increasing the accuracy of the self interference cancelation. On the other hand, no gain in the network performance is achieved by using an imperfect self interference cancelation technique which leads to $g > 10^{-1}$. In this region the RA solution provided by Algorithm 3.3 is always *admissible* (i.e., no node transmits and receives simultaneously).

In each considered network setup (i.e., Figure 3.13(a),3.13(b), and 3.10) a similar behavior of the results holds for medium and high SNR values as well (i.e., SNR = 16 and 30dB). Moreover, as we change SNR from low values to high values, the accuracy level required by the self interference cancelation becomes more stringent. For example, in the case of the fully connected multihop, multicommodity wireless network in Figure 3.10, if the SNR operating point is changed from 5 to 30dB, then the accuracy level required by the self interference cancelation should be improved from $g = 10^{-1}$ to $g = 10^{-3}$ to start gaining in network layer performances. This is intuitively expected since, the larger the SNR operating point, the larger the power levels of the nodes, and therefore the
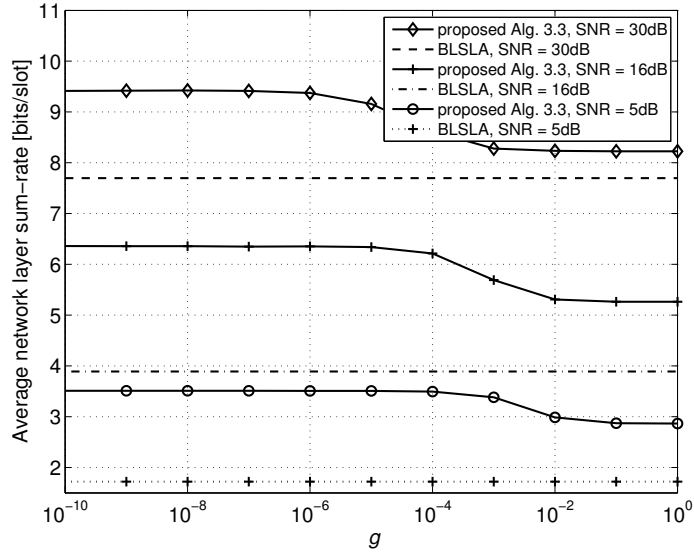
(a) Average network layer sum-rate $\sum_{n=1}^{9} \sum_{s \in \mathcal{S}_n} \bar{x}_n^s$
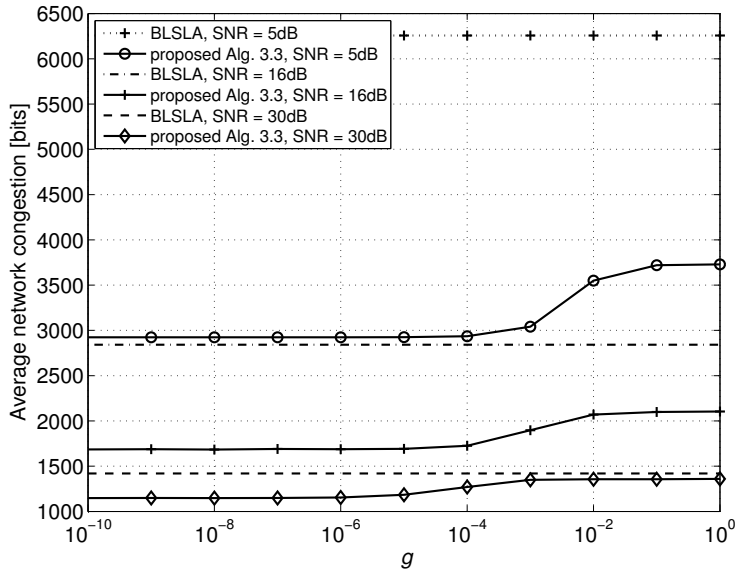
(b) Average network congestion $\sum_{n=1}^{9} \sum_{s=1}^{3} \bar{q}_n^s$

Fig 3.16. Dependence of the average sum-rate and of the average network congestion on the self-interference gain g in the case of the nine node multihop wireless network.

130

higher the accuracy level required by the self interference cancelation techniques to remove the increased transmit power at nodes.

Note that the relative gains due to self interference cancellation in the considered fully connected multihop network (Figure 3.10) are smaller compared to the relative gains experienced in the tandem wireless network [Figure 3.13(b)]. This behavior is intuitively explained by looking in to the network topology. When the self interference is significantly canceled, the resultant interference at the receiver node of any link in the case of the tandem multihop wireless network in Figure 3.13(b) is smaller on average to that of the multihop wireless network in Figure 3.10; note that any receiver node of the fully connected multihop network has many adjacent interfering nodes. Thus, with zero self interference, links in the tandem network can operate at larger rates, and therefore yields larger relative gains.

Finally, we show by an example, how to apply rate-equivalence (3.28) to find the required value of self-interference gain g in order to preserve network layer performances if the distance between nodes is scaled. Let us construct a new network by scaling the distances between the nodes of the original network (see Figure 3.10) by a factor of $\theta = \sqrt{10}$ and the maximum node transmission power $p_0^{\max}$ by a factor of $\theta^\eta = 100$ (note that $\eta = 4$). We refer to this new network as the *scaled network*. To illustrate the idea let us consider the case SNR = 5dB in Figure 3.11 and focus to the point $g = 10^{-4}$ for which the average sum-rate is 3.5 [bits/slot]. The value of g at this point can be considered as the minimum required accuracy level of self interference cancelation to achieve an average sum-rate of 3.5 bits/slot in the original network. Now we ask what is the required self-interference gain g_{new} that would result in the same average sum-rate value (i.e., 3.5 bits/slot) in the scaled network. From (3.28) it follows that the required accuracy level of self interference cancellation should be improved at least to a level of $g_{\mathrm{new}} = g/\theta^\eta = 10^{-4}/100 = 10^{-6}$.

3.4.4   NUM for networks with multiuser detection at receivers

The network in Figure 3.10 is considered. The assumptions and the simulation parameters are exactly the same as in Section 3.4.2.
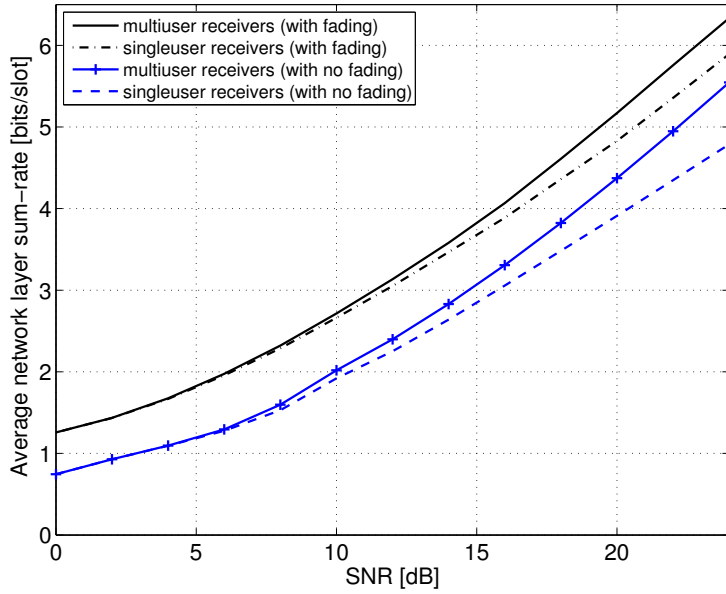
(a) Average network layer sum-rate $\sum_{n=1}^{9} \sum_{s \in \mathcal{S}_n} \bar{x}_n^s$
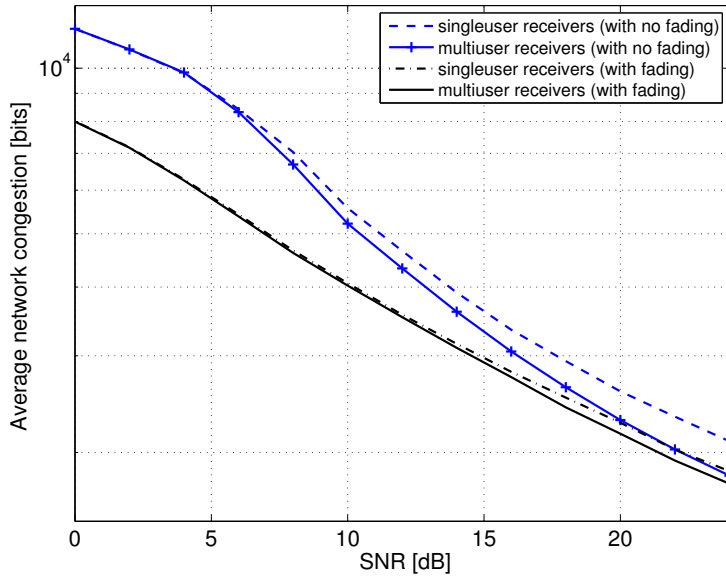
(b) Average network congestion $\sum_{n=1}^{9} \sum_{s=1}^{3} \bar{q}_n^s$

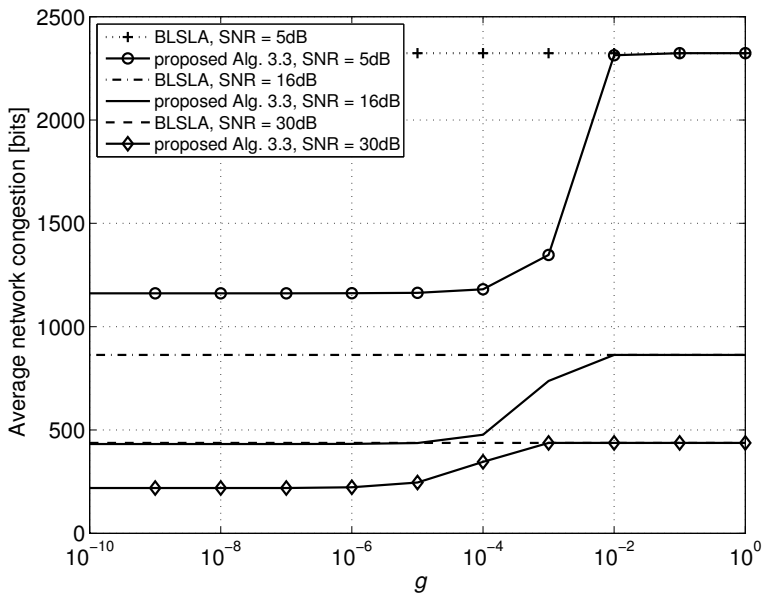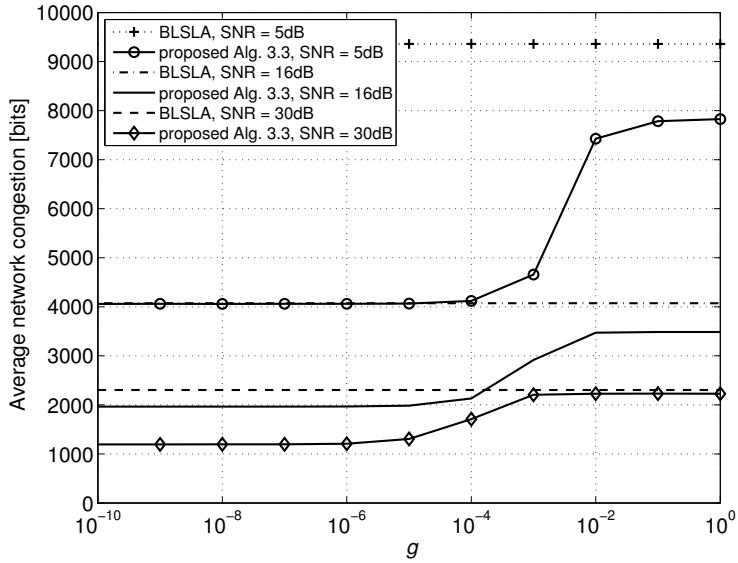**Fig 3.17. Dependence of the average sum-rate and the average network congestion on the SNR;** $C = 1$**, [112] © 2011, IEEE.**

Figure 3.17 shows the dependence of the average sum-rate [Figure 3.17(a)] and of the average network congestion [Figure 3.17(b)] on the SNR for RA where only one node is allowed to transmit in each slot and receivers perform multiuser detection. For illustration purposes we consider the single-channel case (i.e., $C = 1$). We also show the results for a nonfading case [i.e., by having $c_{ijc}(t) = 1$ in (3.24)] for comparison. Here we can make several observations. Fading can significantly improve the overall performance in average sum-rate and average network congestion. This observation has an analogy with multiuser diversity in downlink fading channels [9, Sec. 6.6]. Intuition suggests that when there are many links which fade independently, at any time slot there is a high probability that the resulting rate and power allocation yields a better schedule (see [17, Sec. 4.7]) compared to the nonfading case. There are significant advantages to having multiuser detection, especially for high SNR values. At low SNR, gains are marginal. Thus, multiuser detectors have a practical advantage over singleuser detectors, especially in a high SINR regime. For example, in a fading environment, at SNR = 24dB we obtain around a 7.5% increase in the average sum-rate and a 5% decrease in the average network congestion. In a nonfading environment multiuser detectors offer around a 16% increase in the average sum-rate and a 13.5% decrease in the average network congestion.

3.5   Summary and discussion

We have considered the power and rate control problem in a wireless network in conjunction with the next-hop routing / scheduling and the flow control problem. Thus, although our focus lies on the so-called resource allocation problem that is confined to the physical / medium access control layers, its formulation captures interactions with the higher-layers in a manner similar to the one employed in [44]. The result is a cross layer formulation. The problem, unfortunately, is NP-hard, and therefore there are no polynomial time algorithms to solve it. Our contribution has been to consider first a general access operation, but with a relatively simple form of receiver structure (bank of match filters), and then to limit the access operation to a single node at a time (either transmitting or receiving) but allowing for increased multiuser detector complexity at the receiver.

In the first case, we offer a new optimization methodology based on homo-

topy methods and complementary geometric programming solution methods. Numerical results showed that the our proposed algorithms perform close to exponentially complex optimal solution methods. In addition, they are in fact fast and are capable of handling large-scale problems.

The proposed method was also used to evaluate the gains achievable at the network layer when the network nodes employ self interference cancelation techniques with different degrees of accuracy. Numerical results have shown that the self interference cancelation requires a certain level of accuracy to obtain quantifiable gains at the network layer. The gains saturate after a certain cancelation accuracy. The level of accuracy required by the self interference cancelation techniques depends on many factors, such as the distances between the network nodes and the operating power levels of the network nodes. For the considered network setups, the numerical results showed that a self interference reduction in the range $20 - 60$dB leads to significant gains at the network layer. We emphasize that this level of accuracy is practically achievable, e.g., the recent proposals [170, 171, 187, 188] provide cost effective mechanisms for an up to 55dB reduction in the self-interference coefficient. These observations are indeed important in the context of certain future cellular systems. For example, as the trend in cellular systems is to increase the data rates of users by introducing very small cells (e.g., femtocells), the difference between transmitted and received power might not be very large. Therefore, self interference cancellation strategies can be of great benefit to such systems. Numerical results further show that the topology of the network has a substantial influence on performance gains. For example, in the case of tandem multihop wireless networks, the benefits due to self interference cancellation are more pronounced when compared to those of a multihop network, in which the nodes are located in a square grid.

In the second case, we obtain a complete solution and illustrate numerically the performance gain due to multiuser detector capability. The main benefit here is the simplicity of the proposed solution methods. As a result, these simple access protocols can be potentially useful in practical applications with more advanced communication systems.

134

4 WSRMax for downlink OFDMA systems

In this chapter, we propose a low-complexity method for WSRMax in OFDMA downlink systems. The method is based on the primal decomposition technique [36, 189]. By using numerical simulations, its performance is compared to the Lagrangian relaxation based algorithm [21], as well as to the optimal exhaustive search algorithm. Numerical results show that the proposed algorithm converges very fast compared to the Lagrangian relaxation based methods [21]. Although, the optimality of the final value can not be guaranteed due to the nonconvexity of the problem, the simulations show that rate region achieved by the proposed algorithm exactly matches the one obtained using the optimal exhaustive search algorithm.

4.1 System model and problem formulation

Consider a single antenna OFDMA downlink transmission with J users and C subcarriers, as shown in Figure 4.1. The signal received by user j in subcarrier c can be expressed as

$$y_{jc} = a_{jc}x_{jc}\sqrt{q_{jc}} + z_{jc}, \ j = 1, \ldots, J, \ c \in \mathcal{C}_j \ , \tag{4.1}$$

where j is the user index, c is the subcarrier index, $\mathcal{C}_j$ denotes the set of subcarriers allocated to user j, x_{jc} is the transmitted signal, q_{jc} is the power allocated, a_{jc} is channel gain, and z_{jc} is the received noise. We assume that a_{jc} is time invariant and its value is available at the base station. The noise samples are assumed to be independent and identically distributed as $z_{jc} \sim \mathcal{CN}(0, \sigma^2)$. Let b_{jc} denote the channel SNR of jth user in subcarrier c, i.e., $b_{jc} = |a_{jc}|^2/\sigma^2$ and β_j denote the weight associated with the rate of user j. The WSRMax problem subject to a sum-power constraint $p_0^{\max}$ can be formulated as [21]

$$
\begin{aligned}
\text{maximize} \quad & \sum_{j=1}^{J} \sum_{c \in \mathcal{C}_j} \beta_j \log(1 + q_{jc}b_{jc}) \\
\text{subject to} \quad & \sum_{j=1}^{J} \sum_{c \in \mathcal{C}_j} q_{jc} = p_0^{\max} \\
& \mathcal{C}_j \cap \mathcal{C}_l = \emptyset, \ \forall j \neq l \\
& q_{jc} \geq 0, \ j = 1, \ldots, J, \ c = 1, \ldots, C \ ,
\end{aligned}
\tag{4.2}
$$

where the variables are $(q_{jc})_{j=1,\ldots,J,c=1,\ldots,C}$ and $\mathcal{C}_j$ for all $j = 1, \ldots, J$.
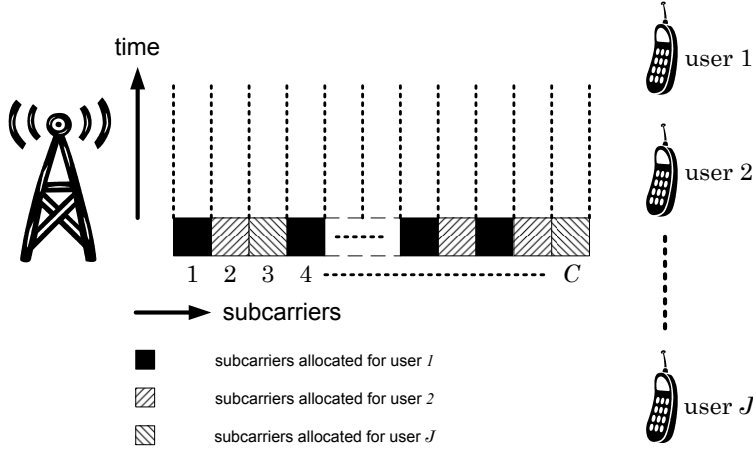
Fig 4.1. A downlink OFDMA system with J users and C subcarriers.

It is also useful to introduce a virtual system where each subcarrier can be used by all users in the same time. This results in a general OFDMA downlink channel where the signal received by user j in subcarrier c is given by

$$\bar{y}_{jc} = a_{jc} x_{jc} \sqrt{q_{jc}} + a_{jc} \sum_{i \neq j}^{J} x_{ic} \sqrt{q_{ic}} + z_{jc} \ . \tag{4.3}$$

The second term in the RHS represents the interference from other users. Assuming independent channel coding across users at the transmitter and independent decoding at receivers, the WSRMax problem for the virtual system can be formulated as

$$
\begin{aligned}
\text{maximize} \quad & \sum_{j=1}^{J} \sum_{c=1}^{C} \beta_j \log \left(1 + \frac{q_{jc}}{\sum_{i=1, i \neq j}^{J} q_{ic} + b_{jc}^{-1}} \right) \\
\text{subject to} \quad & \sum_{j=1}^{J} \sum_{c=1}^{C} q_{jc} = p_0^{\max} \\
& q_{jc} \geq 0, \ j = 1, \dots, J, \ c = 1, \dots, C \ ,
\end{aligned}
\tag{4.4}
$$

where the variable is $(q_{jc})_{j=1,\dots,J,c=1,\dots,C}$. The constraints associated with orthogonal subcarrier allocations in problem (4.2) have been dropped out and the interference among users allocated to the same subcarrier is reflected in the objective.

Here we can make several observations. First, any solution of problem (4.4) is such that the second constraint in problem (4.2) is automatically satisfied, for reasons that will be explained at the beginning of Section 4.2.2. In other words,

any solution of problem (4.4) is feasible for problem (4.2). Moreover, at any of these solutions the objective function of problem (4.4) will be exactly the same as the objective function of problem (4.2). Based on these observations it can be concluded that any solution of the auxiliary problem (4.4) is a solution for the original problem (4.2) as well.

The original problem (4.2) is combinatorial and it requires exponential complexity to find a global optimum. Although problem (4.4) is still nonconvex, it is noncombinatorial. Thus, in the following, we focus on solving problem (4.4) instead of solving the original problem (4.2). A similar approach has been used in [67] to solve the (nonweighted) sum-rate maximization problem, i.e., for the particular case $\beta_k = 1$, $k = 1, \ldots, J$. However, the methods proposed there do not apply to the general case of arbitrary weights, since the properties exploited in solving the sum-rate maximization problem are ruined when there are arbitrary weights. Due to the nonconvexity of problem (4.4) finding the global optimum is intractable. Thus, a successive approximation method inspired from primal decomposition techniques is presented in Section 4.2.

4.2 Algorithm derivation

4.2.1 Primal decomposition

To reveal the complicating constraints [189] we introduce C new variables $\bar{q}_c = \sum_{j=1}^{J} q_{jc}$, $c = 1, \ldots, C$, and reformulate problem (4.4) as follows:

$$
\begin{aligned}
&\text{maximize} &&\sum_{j=1}^{J} \sum_{c=1}^{C} \beta_j \log\left(1 + \frac{q_{jc}}{\bar{q}_c - q_{jc} + b_{jc}^{-1}}\right) \\
&\text{subject to} &&\sum_{c=1}^{C} \bar{q}_c = p_0^{\max} \\
& &&\sum_{j=1}^{J} q_{jc} = \bar{q}_c, \ c = 1, \ldots, C \\
& &&q_{jc} \geq 0, \ j = 1, \ldots, J, \ c = 1, \ldots, C \ ,
\end{aligned}
\tag{4.5}
$$

where the variables are $(q_{jc})_{j=1,\ldots,J,c=1,\ldots,C}$ and $(\bar{q}_c)_{c=1,\ldots,C}$. Note that $\bar{q}_c$ represents the total power allocated to subcarrier c. By treating $\bar{q}_m$ as complicating variables, problem (4.5) can be decomposed [36, 189] into a master problem and C subproblems, one subproblem for each subcarrier $c = 1, \ldots, C$. For a given

subcarrier c, the subproblem is given by

$$
\begin{aligned}
\text{maximize} \quad & \sum_{j=1}^{J} \beta_j \log \left(1 + \frac{q_{jc}}{\bar{q}_c - q_{jc} + b_{jc}^{-1}} \right) \\
\text{subject to} \quad & \sum_{j=1}^{J} q_{jc} = \bar{q}_c \\
& q_{jc} \geq 0, \ j = 1, \dots, J \ ,
\end{aligned}
\tag{4.6}
$$

where the variable is $(q_{jc})_{j=1,\dots,J}$. The master problem can be expressed as

$$
\begin{aligned}
\text{maximize} \quad & \sum_{c=1}^{C} f_c^{\star}(\bar{q}_c) \\
\text{subject to} \quad & \sum_{c=1}^{C} \bar{q}_c = p_0^{\max} \\
& \bar{q}_c \geq 0, \ c = 1, \dots, C \ ,
\end{aligned}
\tag{4.7}
$$

where the variable is $(\bar{q}_c)_{c=1,\dots,C}$ and $f_c^{\star}(\bar{q}_c)$ represents the optimal value of subproblem (4.6) for fixed $\bar{q}_c$.

4.2.2 Subproblem

In this section we focus on subproblem (4.6) and provide the solution in closed form. Let us first denote the feasible set of problem (4.6) by $\mathcal{P}$. Note that problem (4.6) is not a convex optimization problem since we have to *maximize* a convex function. However, its objective function is convex with respect to optimization variable $(q_{1c}, \dots, q_{Jc})$, its feasible set is a nonempty convex polyhedral set (i.e., a simplex [3]), and its objective is bounded above on $\mathcal{P}$. Thus, by following the approach of [67, Sec. III], from [190, Cor. 32.3.4] [17] it follows that the solutions of problems (4.6) must be achieved at one of the vertices of the polyhedral set $\mathcal{P}$. Consequently, the solutions of the cth subproblem can be expressed as

$$
(q_{1c}^{\star}, \dots, q_{Jc}^{\star}) = \bar{q}_c \mathbf{e}_{l_c} \ ,
\tag{4.8}
$$

where l_c represents the index of the user allocated to the cth subcarrier, i.e.,

$$
l_c = \arg \max_{j} \left(1 + \bar{q}_c b_{jc} \right)^{\beta_j} \ .
\tag{4.9}
$$

Solution (4.8) confirms that, even though in subproblem (4.6) all users are allowed to use all subcarriers, the optimal power allocation consists of allocating only one user per subcarrier. This guarantees that solution (4.8) is feasible for the original problem (4.2).

---

[17]If a convex function f is bounded above on a convex set $\mathcal{X} \subseteq \operatorname{dom} f$, then the maximum of f relative to $\mathcal{X}$ is attained at one of the finitely many extreme points of $\mathcal{X}$.

4.2.3 Master problem

In this section we first show that the master problem (4.7) is nonconvex, which means that standard convex optimization techniques (e.g., subgradient based methods) cannot be directly applied to solve problem (4.7). Therefore, for updating the complicating variables $\bar{q}_c$, we propose a method which uses a convex approximation of the master problem. In particular, we use a local convex approximation (a lower bound function) instead of the true objective of the master problem.

We start by substituting (4.8) and (4.9) in the objective of (4.6). Thus, $f_c^\star(\bar{q}_c)$ can be expressed as

$$f_c^\star(\bar{q}_c) = \beta_{l_c} \log\left(1 + \bar{q}_c b_{l_c c}\right) \tag{4.10}$$

$$= \max_j \beta_j \log\left(1 + \bar{q}_c b_{jc}\right) \ . \tag{4.11}$$

We note that the index l_c depends on $\bar{q}_c$ according to (4.9) and the function $f_c^\star(\bar{q}_c)$ is the pointwise maximum of a set of concave functions. Therefore $f_c^\star(\bar{q}_c)$ is not a concave function with respect to $\bar{q}_c$ in general [3]. This prevents direct application of standard convex optimization techniques, such as subgradient based methods for updating the complicating variables.

We next provide a method for updating the complicating variables by using successive convex lower bound approximations of the objective of the master problem. Let $l_c^{(i)}$, $c = 1, \ldots, C$ denote user to subcarrier allocations obtained by solving C subproblems of the form (4.6) for a given subcarrier power allocation $\bar{q}_c^{(i)}$, $c = 1, \ldots, C$. Here the superscript (i) represents an iteration index. Then, the objective of the master problem is approximated by the following lower bound:

$$\sum_{c=1}^{C} \beta_{l_c^{(i)}} \log\left(1 + \bar{q}_c b_{l_c^{(i)} c}\right) \leq \sum_{c=1}^{C} f_c^\star(\bar{q}_c), \ (\bar{q}_c)_{c=1,\ldots,C} \in \bar{\mathcal{P}} \ , \tag{4.12}$$

where $\bar{\mathcal{P}}$ denotes the feasible set of the master problem (4.7). Thus, the approximated master problem is as follows:

$$\begin{array}{ll}
\text{maximize} & \sum_{c=1}^{C} \beta_{l_c^{(i)}} \log\left(1 + \bar{q}_c b_{l_c^{(i)} c}\right) \\
\text{subject to} & \sum_{c=1}^{C} \bar{q}_c \leq p_0^{\max} \\
& \bar{q}_c \geq 0, \ c = 1, \ldots, C \ ,
\end{array} \tag{4.13}$$

where the variable is $(\bar{q}_c)_{c=1,\ldots,C}$. Note that in problem (4.13), the objective function of problem (4.7) is replaced by using the lower bound function given in (4.12). In addition, the sum-power constraint $\sum_{c=1}^{C} \bar{q}_c = p_0^{\max}$ of problem (4.7) is relaxed as $\sum_{c=1}^{C} \bar{q}_c \leq p_0^{\max}$. Of course, at the optimal solution of problem (4.13), the sum-power constraint must be active since the objective function of (4.13) is increasing in each $\bar{q}_c$.

The solution of the approximated master problem (4.13) can be found by the multilevel waterfilling algorithm [20]. In particular, the solution of problem (4.13) is given by the following multilevel waterfilling expression [20]:

$$\bar{q}_c^{\star} = \left(M \beta_{l_c^{(i)}} - b_{l_c^{(i)} c}^{-1} \right)^{+}, \; c = 1, \ldots, C \;, \tag{4.14}$$

where M is chosen such that the power constraint is satisfied with equality, i.e.,

$$M = \frac{p_0^{\max} + \sum_{\{c | \bar{q}_c^{\star} > 0\}} b_{l_c^{(i)} c}^{-1}}{\sum_{\{c | \bar{q}_c^{\star} > 0\}} \beta_{l_c^{(i)}}} \;. \tag{4.15}$$

The resulting solution is used as the subcarrier power allocation for the next iteration $i + 1$.

4.2.4 Algorithm for WSRMax in OFDMA downlink

The proposed iterative algorithm can be summarized as follows:

Algorithm 4.1. *Primal decomposition based method for OFDMA weighted sum-rate maximization*

1. Initialization; set $i = 1$. Given a feasible point $\left(\bar{q}_c^{(i)} \right)_{c=1,\ldots,C}$ for problem (4.13).
2. Let $\left(\bar{q}_c = \bar{q}_c^{(i)} \right)_{c=1,\ldots,C}$ and find $\left(l_c^{(i)} \right)_{c=1,\ldots,C}$ by using (4.9).
3. Solve the approximated the master problem (4.13). Denote the solution by $(\bar{q}_c^{\star})_{c=1,\ldots,C}$ and let $\left(\bar{q}_c^{(i+1)} = \bar{q}_c^{\star} \right)_{c=1,\ldots,C}$.
4. Stopping criterion; if the stopping criterion is satisfied STOP, otherwise set $i = i + 1$ and go to step 2.

Convergence behavior and exit criterion

In this section, we first investigate the monotonicity of the proposed algorithm. Then we provide a specific exit criterion, which certifies the convergence of the algorithm to a fixed power and subcarrier allocation. A simple graphical illustration is provided as well.

The following theorem establishes the monotonic behavior of the proposed algorithm.

Theorem 4.1. *For any iteration $i \geq 1$,*

$$\sum_{m=1}^{C} f_c^\star \left(\bar{q}_c^{(i)} \right) \leq \sum_{c=1}^{C} f_c^\star \left(\bar{q}_c^{(i+1)} \right) \ , \tag{4.16}$$

i.e., the proposed Algorithm 4.1 is an ascent algorithm.

Proof. From Algorithm 4.1 step 2, (4.10), and (4.11), it follows that the solution of problem (4.6) after the ith iteration of Algorithm 4.1 is given by

$$\begin{aligned}
f_c^\star \left(\bar{q}_c^{(i)} \right) &= \max_j \beta_j \log \left(1 + \bar{q}_c^{(i)} b_{jc} \right) \tag{4.17} \\
&= \beta_{l_c^{(i)}} \log \left(1 + \bar{q}_c^{(i)} b_{l_c^{(i)} c} \right) \ .
\end{aligned}$$

Now we can write the following chain of relations:

$$\begin{aligned}
\sum_{c=1}^{C} f_c^\star \left(\bar{q}_c^{(i)} \right) &\leq \sum_{c=1}^{C} \beta_{l_c^{(i)}} \log \left(1 + \bar{q}_c^{(i+1)} b_{l_c^{(i)} c} \right) \tag{4.18} \\
&\leq \sum_{c=1}^{C} \max_j \beta_j \log \left(1 + \bar{q}_c^{(i+1)} b_{jc} \right) \tag{4.19} \\
&= \sum_{c=1}^{C} f_c^\star \left(\bar{q}_c^{(i+1)} \right) \ , \tag{4.20}
\end{aligned}$$

where the first inequality follows from the third step of Algorithm 4.1, the second one follows trivially from the maximization over the users, and the equality follows from (4.17). □

The exit criterion for a such an ascent algorithm is typically chosen heuristically, e.g., the objective value increment between two successive iterations is below a certain predefined threshold. However, for the proposed algorithm we

are able to find an exit criterion, which certifies that the algorithm converges to a fixed power and subcarrier allocation and further improvement is not possible. This is established by the following theorem.

Theorem 4.2. *If at iteration $n + 1$ $(n \geq 1)$ we have $l_c^{(n)} = l_c^{(n+1)}$ for all $c = 1, \ldots, C$, then the following holds:*

1. *$l_c^{(i)} = l_c^{(n)}$ for all $i \geq n$ and $c = 1, \ldots, C$.*
2. *$\bar{q}_c^{(i)} = \bar{q}_c^{(n+1)}$ for all $i \geq n + 1$ and $m = 1, \ldots, C$.*
3. *$\sum_{c=1}^{C} f_c^{\star}\left(\bar{q}_c^{(i)}\right) - \sum_{c=1}^{C} f_c^{\star}\left(\bar{q}_c^{(n+1)}\right) = 0$, for all $i \geq n + 1$,*

i.e., the algorithm converges to a fixed power and subcarrier allocation.

Proof. Since b_{jc}s are continuous random variables, the probability of having multiple solutions for (4.9) is zero; for any $\bar{q}_c > 0$, (4.9) has multiple solutions if and only if $(1 + \bar{q}_c b_{jc})^{\beta_j} = (1 + \bar{q}_c b_{ic})^{\beta_i}$ for some $i \neq j$. Thus, in the following we assume that l_c given by (4.9) is unique

Now we can write the following:

$$l_c^{(n)} = l_c^{(n+1)} \text{ for all } c = 1, \ldots, C \Rightarrow \left(\bar{q}_c^{(n+1)} = \bar{q}_c^{(n+2)}\right)_{c=1,\ldots,C} \tag{4.21}$$

$$\Rightarrow l_c^{(n+1)} = l_c^{(n+2)} \text{ for all } c = 1, \ldots, C , \tag{4.22}$$

where (4.21) follows by noting that the objective function of (4.13) is strictly concave, and therefore it has a unique solution [3]. The second implication [i.e., (4.22)] follows because (4.9) has a unique solution. Thus, item 1) of the theorem follows directly from (4.22) and by induction. Furthermore, item 2) follows from item 1), (4.21), and by induction. Finally, item 3) follows trivially from item 2). $\square$

Thus, the exit criterion checks if the subcarrier allocation between two successive iterations remains unchanged. Such a point is a local optimum (possible global) in the sense that the objective can not be increased by changing the power allocation or subcarrier allocation only.

As a specific example, consider the simple OFDMA system with two subcarriers (i.e., $c = 1, 2$). By performing the variable transformations $\bar{q}_1 = (1 - t)p_0^{\max}$ and $\bar{q}_2 = tp_0^{\max}$, where t is a scalar such that $t \in [0, 1]$, we can express the variation of $\sum_{c=1}^{2} f_c^{\star}(\bar{q}_c)$ on $\bar{\mathcal{P}}$ as,

$$h(t) = \max_j \beta_j \log\left(1 + (1 - t)p_0^{\max} b_{j1}\right) + \max_j \beta_j \log\left(1 + tp_0^{\max} b_{j2}\right) ,$$
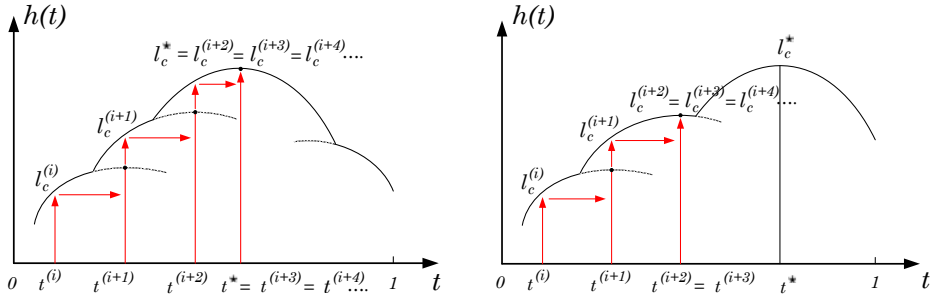
(a) Convergence to the global optimal solution.

(b) Convergence to a local optimal solution.

Fig 4.2. Convergence of Algorithm 4.1.

which is plotted in Figure 4.2. According to Figure 4.2(a) the global optimum is achieved at the iteration $(i + 3)$. Achieving global optimality is not always possible because quasiconcavity [3] of $h(t)$ cannot be guaranteed with random channel SNR, b_{jc}. Consequently, Algorithm 4.1 can converge to a local optimal solution as shown in Figure 4.2(b).

Complexity analysis

In this section we analyze and compare the computational complexity of the proposed Algorithm 4.1 to the Lagrangian relaxation based algorithm [21], as well as to the optimal exhaustive search algorithm. With J users and C subcarriers, altogether we have J^C user-subcarrier combinations. Therefore, finding optimal subcarrier and power assignment requires J^C searches. Combined with multilevel waterfilling at each instance of the subcarrier assignment, $O(CJ^C)$ operations are required to find the solution. The algorithm proposed in [21] for the weighted sum-rate maximization problem requires $O(CJ)$ operations to obtain a suboptimal solution. Our proposed Algorithm 4.1, described in Section 4.2.4, requires $O(CJ)$ operations in step 2 and $O(C \log C)$ operations in step 3, i.e., for ordering. In practice, it is reasonable to assume that $J \gg \log C$ for the following reasons. The number of users simultaneously serviced by the system can be very large. For example, in a Wi-Max system C can be up to 2048 [191]. However, the value of $\log C$ will not become very large (in a WiMax system $\log C = 11$ at most). Therefore, the complexity of Algorithm 4.1 can be approximated by $O(CJ)$.

143

Particularization to the sum-rate maximization

The problem of the sum-rate maximization (i.e., $\beta_j = 1$ for all $j = 1, \ldots, J$) in downlink OFDMA systems is solved in [67, Sec. III]. The solution method is exactly equivalent to only one iteration of Algorithm 4.1. Unlike the general weighted sum-rate maximization, in which user weights β_js are different, in the case of sum-rate maximization (i.e., $\beta_j = 1$ for all $j = 1, \ldots, J$) the index l_c *will not depend* on $\bar{q}_c$ according to (4.9). Thus, by using (4.10) and (4.11) the function $f_c^\star(\bar{q}_c)$ can be found as $f_c^\star(\bar{q}_c) = \log\left(1 + \bar{q}_c b_{l_c c}\right) = \log\left(1 + \bar{q}_c \cdot \max_i b_{ic}\right)$, which is concave with respect to $\bar{q}_c$ (recall that the function $f_c^\star(\bar{q}_c)$ is not concave with respect to $\bar{q}_c$ when the user weights β_js are different). As a result, the inequality given in (4.12) holds with equality and solving problem (4.7) gives the optimal subcarrier power allocation [67, Sec. III].

4.3   Numerical examples

The performance of the Algorithm 4.1 (Section 4.2.4) is compared to the dual decomposition based algorithm proposed in [21], denoted as Seong-WSRMax, as well as to the optimal algorithm based on exhaustive search. The Seong-WSRMax algorithm uses a bisection search method to update the dual variable λ [21, Sec. IV]. For initializing the bisection search interval $[\lambda_{\min}, \lambda_{\max}]$, we exploit the fact that the subgradient of the dual function can be analytically computed. Since the dual function is convex, the sign of its subgradients change as we pass through the minimum point of the dual function [21, eq. (11)]. Therefore, we use a grid search (with step size 1) to identify the interval, in which the subgradient of the dual function changes its sign, and it is used as the initial bisection search interval. Thus, the interval $[\lambda_{\min}, \lambda_{\max}]$ is guaranteed to contain the optimal value of the the dual function and the width of the initial interval is one, i.e., $(\lambda_{\max} - \lambda_{\min}) = 1$. Our proposed Algorithm 4.1 is initialized by allocating equal power to all subcarriers unless otherwise stated, i.e., $\left(\bar{q}_c^{(1)} = p_0^{\max}/C\right)_{c=1,\ldots,C}$.

We start by comparing the convergence behavior of Algorithm 4.1 and the Seong-WSRMax algorithm. For a fair comparison, we define the following metric:

$$\Delta D_{\mathrm{WSR}} = \mathrm{E}\left\{ \frac{|\, C_{\mathrm{opt}} - \hat{C}_{\mathrm{subopt}} \,|}{C_{\mathrm{opt}}} \right\}, \tag{4.23}$$

which is referred to as the average normalized WSR deviation, where C_{opt} is

the optimal WSR value obtained by exhaustive search, $\hat{C}_{\mathrm{subopt}}$ is the estimated objective value from either Algorithm 4.1 or the Seong-WSRMax algorithm, and expectation $\mathrm{E}\{\cdot\}$ is taken with respect to channel realization. An OFDMA system with $C = 8$ subcarriers and a uniform power delay profile with 4 channel taps is considered. We define the SNR operating point as

$$\mathrm{SNR} = \frac{p_0^{\max}}{\sigma^2 C} \ .$$

Figure 4.3 shows the convergence behavior of the considered algorithms with $\mathrm{SNR} = 10$dB for $J = 2$ and $J = 4$ users. The weights of the users are $(1, 2)$ for $J = 2$ and $(1, 2, 1, 2)$ for $J = 4$ [18]. The floor of the curves is due to the suboptimality of the algorithms. The results show that Algorithm 4.1 converges faster than the Seong-WSRMax algorithm and provides smaller average normalized WSR deviations. Specifically, for both cases $J = 2$ and $J = 4$, Algorithm 4.1 requires only 3 iterations on average to achieve an average normalized WSR deviation of 10^{-4} whilst the Seong-WSRMax algorithm requires around 15 iterations to reach the same accuracy level. It is intuitively obvious that the number of iterations required by Algorithm 4.1 is sensitive to the nature of the surface of the objective function $\sum_{c=1}^{C} f_c^\star(\bar{q}_c)$ of problem (4.7); see, e.g., Figure 4.2. In general, it is hard to quantify the number of iterations before convergence (or to quantify any bounds on the number of iterations) due to the nonconvexity of problem (4.5). However, the numerical results suggest that Algorithm 4.1 often converges fast in practice. It should be emphasized that the number of iterations required in the initialization of the Seong-WSRMax algorithm (i.e., the number of iterations required to find the initial bisection search interval) is not considered when drawing the curves. In particular, for the initialization process, the Seong-WSRMax algorithm requires a number of steps [each step has complexity of $O(JC)$] and our proposed Algorithm 4.1 requires none. Moreover, it is hard to find good initialization methods for the Seong-WSRMax algorithm (i.e., initialization for the bisection search method) requiring a compromise between the number of steps required in the initialization and the width of the initial searching interval ($\lambda_{\max} - \lambda_{\min}$). Consequently, additional precautions are required, and therefore in practical implementations Algorithm 4.1 is more favorable compared to the Seong-WSRMax algorithm.

---

[18] Numerical experience suggests that the algorithm behavior is insensitive to the weights.
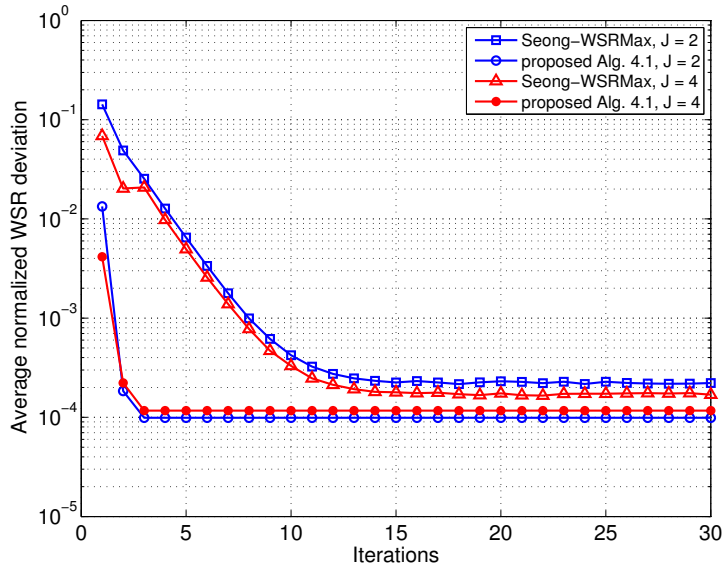
Fig 4.3. Average normalized WSR deviation vs. Iterations for $J = 2$ and 4 users, $C = 8$ subcarriers and $\mathrm{SNR} = 10$dB.
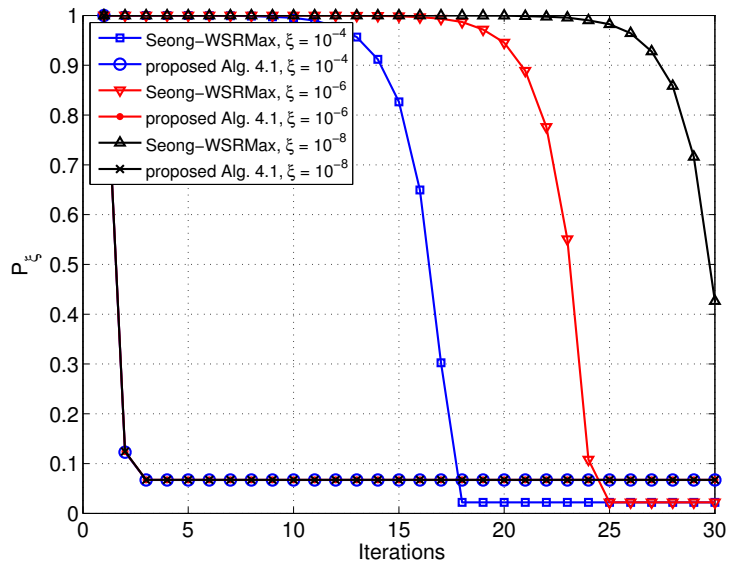
Fig 4.4. Probability of missing global optimum vs. Iterations for $J = 2$ users, $C = 8$ subcarriers and $\mathrm{SNR} = 10$dB.

Let us now consider another metric P_ξ, the probability of missing the global optimum, to compare the behavior of Algorithm 4.1 and the Seong-WSRMax algorithms. In particular, P_ξ is defined as

$$P_\xi = \text{Prob}\left\{ \mid C_{\text{opt}} - \hat{C}_{\text{subopt}} \mid > \xi \right\} , \qquad (4.24)$$

where ξ is a small number, which quantifies the maximum admissible deviation between C_{opt} and $\hat{C}_{\text{subopt}}$. It is considered that the global optimum is missed if $\hat{C}_{\text{subopt}}$ is more than ξ away from C_{opt}.

Figure 4.4 uses the same simulation setup as that in Figure 4.3 and depict the variation of probability of missing the global optimum, P_ξ with the number of iterations. The floor of probability P_ξ is again due to the suboptimality of both algorithms. The influence of ξ on P_ξ is totally indistinguishable in the case of Algorithm 4.1. This behavior shows that Algorithm 4.1 can achieve very close to optimal solutions within a very small number of iterations and then it remains there. The results further show that the P_ξ evaluated using the Seong-WSRMax algorithm is highly dependent on ξ. In particular, the smaller the deviations in the $\hat{C}_{\text{subopt}}$ from the optimal C_{opt}, the larger the number of iterations required by the Seong-WSRMax algorithm to reach the expected target value P_ξ. Therefore, independent from the ξ, Algorithm 4.1 allows us to find a suboptimal solution within a small number of iterations at the expense of a slight increase in P_ξ. These observations are very useful in practice since they carry significant information from the system design point of view. For example, consider a design requirement $P_{10^{-4}} \leq 0.3$. Here, the Seong-WSRMax algorithm requires 18 iterations. If we tighten the design requirement to $P_{10^{-6}} \leq 0.3$, then the number of iterations required by the Seong-WSRMax increases to 24. In contrast Algorithm 4.1 always requires just one iteration.

Figure 4.5 shows the achievable rate region computed by using all considered algorithms. Note that the standard way to characterize the boundary points in the 2-user rate region is by solving problem (4.4) for $\beta_1 = \alpha$ and $\beta_2 = 1 - \alpha$, where $\alpha \in [0, 1]$ [185]. The same simulation setup as in [21] was used, i.e., $C = 8$, $J = 2$, $p_0^{\max} = 16$ and the channel SNR vectors for users 1 and 2 are $10(1^2, 2^2, \ldots, C^2)$ and $10(C^2, (C-1)^2, \ldots, 1^2)$ respectively. Although the computational complexity of the proposed algorithm is much smaller compared to that of the optimal exhaustive search based method, Figure 4.5 indicates that the rate region obtained by Algorithm 4.1 almost coincides with the optimal rate

Fig 4.5. Data Rate Region with 2 **Users and** 8 **subcarriers.**
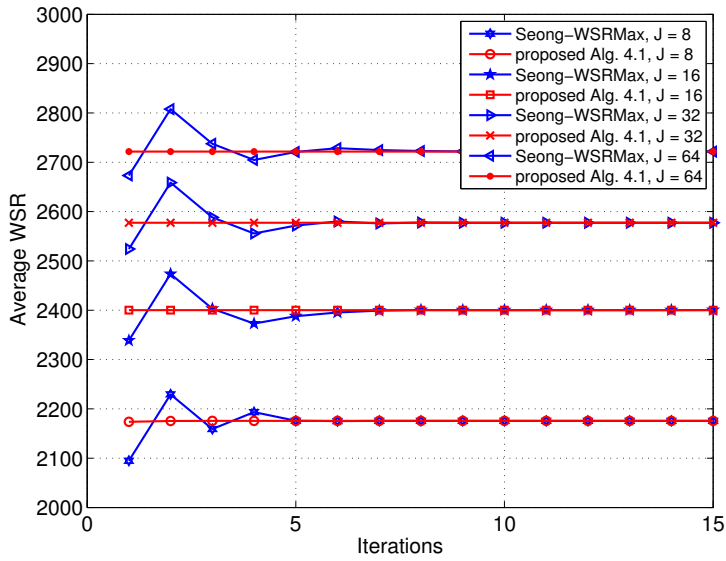
Fig 4.6. Evolution of expected WSR, $J = 8, 16, 32, 64$ **users,** $C = 256$ **subcarriers, and** $\mathrm{SNR} = 10$**dB.**

region. This behavior is expected since the average normalized WSR deviation (4.23) is in the order of 10^{-4} as shown in Figure 4.3.

Let us next compare the behavior of Algorithm 4.1 and the Seong-WSRMax algorithm with large numbers of subcarriers and users. Since, for large numbers of users and subcarriers the complexity of evaluating C_{opt} is prohibitively high, the metrics defined in (4.23) and (4.24) are not used. Instead, the behavior of Algorithm 4.1 is compared with that of the Seong-WSRMax algorithm.

In Figure 4.6, the evolution of the average WSR provided by Algorithm 4.1 is compared to the average WSR from the Seong-WSRMax algorithm, where the averaging is performed with respect to channel realization. An OFDMA system with $C = 256$ subcarriers, a uniform power delay profile with 128 channel taps and $J = 8, 16, 32, 64$ users is considered. The weights of the users are taken from the sequence $\{1, 2, 1, 2, \ldots, 1, 2\}$, e.g., when $J = 8$, weights are $(1,2,1,2,1,2,1,2)$. The SNR is assumed to be 10dB. The results show that even for a large number of carriers, Algorithm 4.1 converges fast compared to the Seong-WSRMax algorithm, independent of the number of users.

It is important to note that several instances of Algorithm 4.1 can be carried out independently, in parallel by starting them at several initialization points and by keeping track of the best solution found so far. To do this, in addition to running the Algorithm 4.1 with uniform initialization $\left(\bar{q}_c^{(i)} = p_0^{\mathrm{max}}/C\right)_{c=1,\ldots,C}$, n_{rand} instances of Algorithm 4.1 with arbitrarily chosen initialization points are carried out. Specifically, in step 1 of each parallel instance, we set $\left(\bar{q}_c^{(i)} = up_0^{\mathrm{max}}/C\right)_{c=1,\ldots,C}$, where u is a random number which is uniformly distributed between 0 and 1, i.e., $u \sim \mathcal{U}(0,1)$. Moreover, at the end of every iteration i of parallel algorithms, we keep track of the best point found so far.

Figure 4.7 shows the convergence behavior of the parallel version of Algorithm 4.1 and the Seong-WSRMax algorithm with SNR = 10dB for $n_{\mathrm{rand}} = 0, 1, 5$ and 10. The weights of the users are $(1, 2)$. The floor of the curves is due to the suboptimality of the algorithms. The results show that the parallel version of Algorithm 4.1 converges faster than the Seong-WSRMax algorithm and provides smaller values of average normalized WSR deviations. Specifically, for any value of n_{rand} parallel implementation requires only 3 iterations on average to achieve a constant level of average normalized WSR deviations, while the Seong-WSRMax algorithm requires around 15 iterations to reach a constant level. This remarkable reduction of the average normalized WSR deviation is
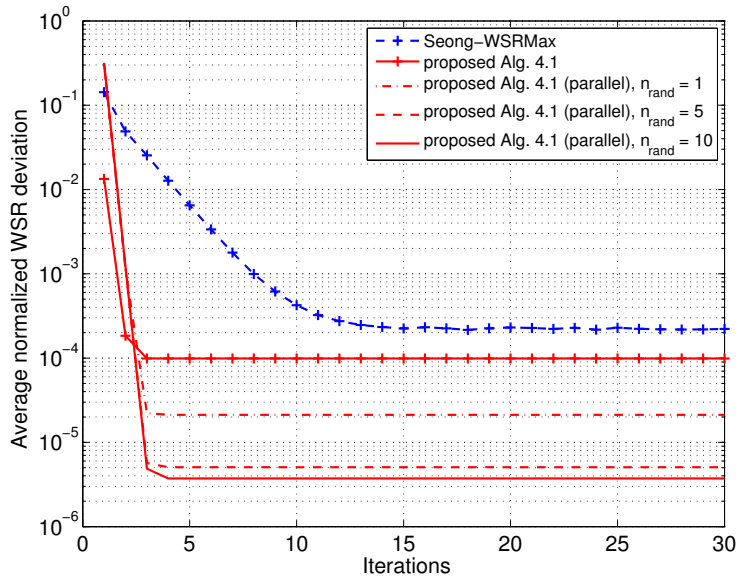
Fig 4.7. Average normalized WSR deviation vs. Iterations for $J = 2$ **users,** $C = 8$
subcarriers and $\mathrm{SNR} = 10$**dB.**
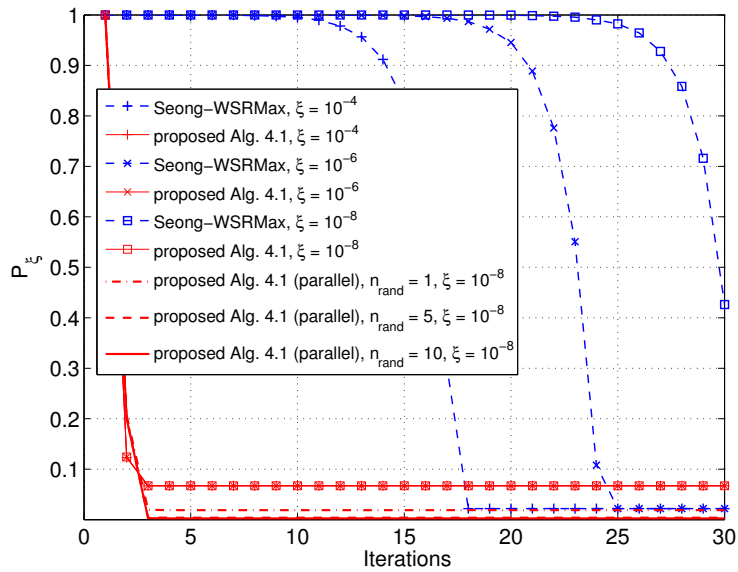
Fig 4.8. Probability of missing global optimum vs. Iterations for $J = 2$ **users,** $C = 8$
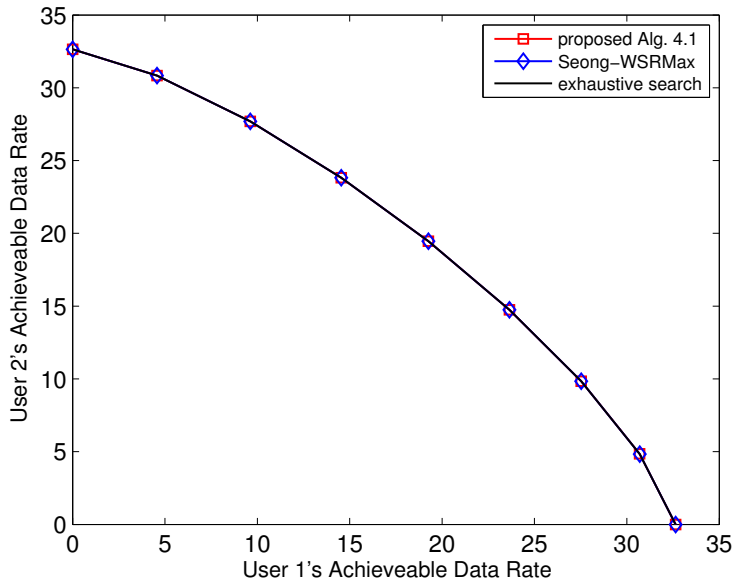**subcarriers and** $\mathrm{SNR} = 10$**dB.**

obtained when n_{rand} is changed from 0 to 1, 5, and 10. As n_{rand} increased the incremental gains that can be obtained in terms of average normalized WSR deviations are reduced.

Figure 4.8 uses the same simulation setup as that in Figure 4.7 and depicts the variation of the probability of missing the global optimum P_ξ with the number of iterations. The floor of probability P_ξ is again due to the suboptimality of both algorithms. Results again show that in the case of Algorithm 4.1, the influence of ξ on P_ξ is indistinguishable and the algorithm can achieve close to optimal solutions within a small number of iterations and then it remains there. Results further show that changing n_{rand} from 0 to 1, 5, and 10 yields significant improvements in P_ξ. As the number of parallel initialization points (i.e., n_{rand}) is increased, P_ξ become almost zero.

4.4   Summary and discussion

A joint subcarrier and power allocation algorithm, which is inspired by primal decomposition techniques, has been proposed for maximizing the WSR in multiuser OFDMA downlink systems. Although the original problem is nonconvex, the proposed algorithm finds fast a suboptimal, but still close to optimal, solution (i.e., more than 90% of the time). Performance can be further improved by using parallel implementations of the proposed algorithm; several instances of the algorithm are carried out independently, in parallel, by starting them at different initialization points. Unlike the dual decomposition based algorithms (e.g., [21]), our proposed method requires no additional precautions in the initialization. The algorithm's convergence to a suboptimal solution is possible within a very small number of iterations. Although the proposed primal decomposition based solution method does not rely on zero duality gap for proving optimality in the case of a large number of subcarriers, our computational experience with larger numbers of subcarriers suggests that the proposed algorithm is just as capable of finding the same solution as the dual decomposition based algorithm [21] (which is asymptotically optimal when the number of carriers grows to ∞), even with very few iterations.

5   A distributed approach for WSRMax in cellular networks with multiple antennas

This chapter provides an alternative distributed algorithm for the WSRMax problem in a multicell MISO downlink system, which does not rely on ZF beamforming or high SINR approximation. Our proposed method is based on primal decomposition methods and subgradient methods [189, 192]. Specifically, we first apply primal decomposition techniques [189] to split the problem into many subproblems and a master problem. The subproblems can be carried out independently, in parallel (one for each BS). Each BS optimizes locally the decision variables associated with its assigned users, i.e., transmit beamforming directions and transmit powers of each beam. Here we adopt an ascent algorithm originally proposed in [15, Sec. 4.3], which is based on second-order cone programming (SOCP) [193] and GP. The master problem resolves the out-of-cell interference levels, which are the complicating variables of the original WSRMax problem. Here we adopt a sequential convex approximation strategy [182] together with a subgradient method [192], which is carried out via BS coordination. We compare the proposed algorithm with the optimal branch and bound method proposed in [88] and the centralized one proposed in [15, Sec. 4.3]. The numerical results show that significant gains can be achieved by only a small amount of BS coordination. Of course, the global optimality of the solution cannot be guaranteed due to the nonconvexity of the original WSRMax problem.

5.1   System model and problem formulation

A multicell MISO downlink system, with N BSs, each equipped with T transmit antennas, is considered. The set of all BSs is denoted by $\mathcal{N}$ and we label them with the integer values $n = 1, \ldots, N$. The *transmission region* of each BS is modeled as a disc with radius R_{BS} centered at the location of the BS. A single data stream is transmitted for each user. We denote the set of all data streams in the system by $\mathcal{L}$ and label them with the integer values $l = 1, \ldots, L$. The transmitter node (i.e., the BS) of the lth data stream is denoted by $\mathrm{tran}(l)$ and the receiver node of the lth data stream is denoted by $\mathrm{rec}(l)$. We have

$\mathcal{L} = \cup_{n \in \mathcal{N}} \mathcal{L}(n)$, where $\mathcal{L}(n)$ denotes the set of data streams transmitted by the nth BS. Note that the users of the data streams transmitted by each BS are necessarily located inside the transmission region of the BS (see Figure 5.1).

The antenna signal vector transmitted by the nth BS is given by

$$\mathbf{x}_n = \sum_{l \in \mathcal{L}(n)} \sqrt{p_l} s_l \mathbf{v}_l \ , \tag{5.1}$$

where $p_l \in \mathbb{R}_+$, $s_l \in \mathbb{C}$, and $\mathbf{v}_l \in \mathbb{C}^T$ represent the power, the information symbol, and the transmit beamformer associated with the lth data stream. We assume that s_l and $\mathbf{v}_l$ are normalized such that $\mathrm{E}|s_l|^2 = 1$ and $\|\mathbf{v}_l\|_2 = 1$. Moreover, we assume independent data streams, i.e., $\mathrm{E}\{s_l s_j^*\} = 0$ for all $l, j \in \mathcal{L}$, where $l \neq j$.

The signal received at $\mathrm{rec}(l)$ is given by

$$y_l = \mathbf{h}_{ll}^{\mathrm{H}} \sqrt{p_l} s_l \mathbf{v}_l + \sum_{j \in \mathcal{L}(\mathrm{tran}(l)), j \neq l} \mathbf{h}_{jl}^{\mathrm{H}} \sqrt{p_j} s_j \mathbf{v}_j + \sum_{j \in \mathcal{L} \setminus \mathcal{L}(\mathrm{tran}(l))} \mathbf{h}_{jl}^{\mathrm{H}} \sqrt{p_j} s_j \mathbf{v}_j + z_l \tag{5.2}$$

$$= \mathbf{h}_{ll}^{\mathrm{H}} \sqrt{p_l} s_l \mathbf{v}_l + \sum_{j \in \mathcal{L}(\mathrm{tran}(l)), j \neq l} \mathbf{h}_{jl}^{\mathrm{H}} \sqrt{p_j} s_j \mathbf{v}_j + \sum_{i \in \mathcal{N} \setminus \{\mathrm{tran}(l)\}} \sum_{j \in \mathcal{L}(i)} \mathbf{h}_{jl}^{\mathrm{H}} \sqrt{p_j} s_j \mathbf{v}_j + z_l \ , \tag{5.3}$$

where $\mathbf{h}_{jl}^{\mathrm{H}} \in \mathbb{C}^{1 \times T}$ is the channel matrix between $\mathrm{tran}(j)$ and $\mathrm{rec}(l)$, and z_l is circular symmetric complex Gaussian noise with variance σ_l^2. Note that the second term in (5.3) represents the intra-cell interference and the third term represents the out-of-cell interference. The received SINR of the lth data stream is given by

$$\gamma_l = \frac{p_l |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_l|^2}{\sigma_l^2 + \sum_{j \in \mathcal{L}(\mathrm{tran}(l)), j \neq l} p_j |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_j|^2 + \sum_{i \in \mathcal{N} \setminus \{\mathrm{tran}(l)\}} w_{il}} \ , \tag{5.4}$$

where $w_{il} = \sum_{j \in \mathcal{L}(i)} p_j |\mathbf{h}_{jl}^{\mathrm{H}} \mathbf{v}_j|^2$, which represents the out-of-cell interference power from the ith BS to $\mathrm{rec}(l)$.

The out-of-cell interference term in (5.4) (i.e., $\sum_{i \in \mathcal{N} \setminus \{\mathrm{tran}(l)\}} w_{il}$) prevents resource allocation (or RA) on an intra-cell basis and demands centralized RA methods. To facilitate potential distributed algorithms for RA, we make the following assumption: transmissions from the ith BS *do interfere* with the lth data stream transmitted by BS $n \neq i$, if the distance between the ith BS and $\mathrm{rec}(l)$ is smaller than the threshold R_{int} [19]. The disc with the radius R_{int} centered at the location of any BS is referred to as the *interference region* of the

---

[19]Similar assumptions are made in [194] in the context of arbitrary wireless networks.
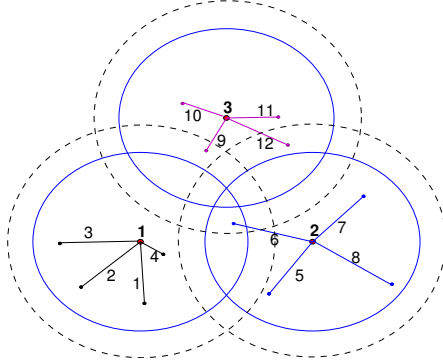
Fig 5.1. Multicell network, $\mathcal{N} = \{1, 2, 3\}$**,** $\mathcal{L} = \{1, \ldots, 12\}$**,** $\mathcal{L}(1) = \{1, \ldots, 4\}$**,** $\mathcal{L}(2) = \{5, \ldots, 8\}$**,** $\mathcal{L}(3) = \{9, \ldots, 12\}$**. The area inside the solid-lined circles around BS** $1, 2$**, and** 3 **represent the associated transmission regions of each BS and the area inside the dash-lined circles around the BSs represent the associated interference regions of each BS.**

BS (see Figure 5.1). Thus, if the ith BS is located at a distance larger than R_{int} to $\text{rec}(l)$, then the associated w_{il} components are set to zero [20]. Based on the assumption above, we can express γ_l as

$$\gamma_l = \frac{p_l |\mathbf{h}_{ll}^{\text{H}} \mathbf{v}_l|^2}{\sigma_l^2 + \sum_{j \in \mathcal{L}(\text{tran}(l)), j \neq l} p_j |\mathbf{h}_{ll}^{\text{H}} \mathbf{v}_j|^2 + \sum_{i \in \mathcal{N}_{\text{int}}(l)} w_{il}} , \qquad (5.5)$$

where $\mathcal{N}_{\text{int}}(l) \subseteq \mathcal{N} \setminus \{\text{tran}(l)\}$ is the set of out-of-cell interfering BSs that are located at a distance less than R_{int} to $\text{rec}(l)$. For example, in Figure 5.1 we have $\mathcal{N}_{\text{int}}(9) = \{1\}$, $\mathcal{N}_{\text{int}}(12) = \{2\}$, $\mathcal{N}_{\text{int}}(6) = \{1, 3\}$, and $\mathcal{N}_{\text{int}}(l) = \emptyset$ for all $l \in \mathcal{L} \setminus \{6, 9, 12\}$. Finally, it is useful to define the set $\mathcal{L}_{\text{int}}$ of data streams that are subject to out-of-cell interference, i.e., $\mathcal{L}_{\text{int}} = \{l \mid l \in \mathcal{L}, \mathcal{N}_{\text{int}}(l) \neq \emptyset\}$. For example, in Figure 5.1 we have $\mathcal{L}_{\text{int}} = \{6, 9, 12\}$.

Let β_l be an arbitrary positive weight associated with the lth data stream. We consider the case where all receivers are using *singleuser detection* (i.e., a receiver decodes its intended signal by treating all other interfering signals as noise). Assuming that the power allocation is subject to a maximum power constraint $\sum_{l \in \mathcal{L}(n)} p_l ||\mathbf{v}_l||_2 \leq p_n^{\max}$ for each BS $n \in \mathcal{N}$, the problem of WSRMax can be expressed as

---

[20]Setting certain w_{il}s to zero can be captured as a change in the statistical characteristics of noise z_l at $\text{rec}(l)$. However, those issues are extraneous to the main focus of this chapter.

$$\text{maximize} \quad \sum_{n\in\mathcal{N}}\sum_{l\in\mathcal{L}(n)}\beta_l\ln\left(1+\frac{p_l|\mathbf{h}_{ll}^{\mathrm{H}}\mathbf{v}_l|^2}{\sigma_l^2+\sum\limits_{j\in\mathcal{L}(\mathrm{tran}(l)),j\neq l}p_j|\mathbf{h}_{ll}^{\mathrm{H}}\mathbf{v}_j|^2+\sum\limits_{i\in\mathcal{N}_{\mathrm{int}}(l)}w_{il}}\right)$$

$$\text{subject to} \quad w_{il}=\sum_{j\in\mathcal{L}(i)}p_j|\mathbf{h}_{jl}^{\mathrm{H}}\mathbf{v}_j|^2,\ l\in\mathcal{L}_{\mathrm{int}},\ i\in\mathcal{N}_{\mathrm{int}}(l)$$
$$\sum_{l\in\mathcal{L}(n)}p_l||\mathbf{v}_l||_2^2\le p_n^{\mathrm{max}},\ n\in\mathcal{N}$$
$$||\mathbf{v}_l||_2=1,\quad p_l\ge 0,\ l\in\mathcal{L}\ ,$$

$$(5.6)$$

where the variables are $(p_l,\mathbf{v}_l)_{\in\mathcal{L}}$ and $(w_{il})_{\in\mathcal{L}_{\mathrm{int}},i\in\mathcal{N}_{\mathrm{int}}(l)}$. Here we assume natural logarithm $\ln(\cdot)$ instead of $\log(\cdot)$ without loss of generality to simplify the presentation.

5.2 Distributed algorithm derivation

In this section we derive a distributed algorithm for the WSRMax problem (5.6). The proposed algorithm is based on primal decomposition methods and subgradient methods. By treating out-of-cell interference powers $(w_{il})_{l\in\mathcal{L}_{\mathrm{int}},i\in\mathcal{N}_{\mathrm{int}}(l)}$ as complicating variables, we first break problem (5.6) into N subproblems (one for each BS) and a master problem. In the case of the subproblems, the associated variables are optimized by using SOCP and GP based approaches. In the case of the master problem, we adopt a subgradient method to update the complicating variables $(w_{il})_{l\in\mathcal{L}_{\mathrm{int}},i\in\mathcal{N}_{\mathrm{int}}(l)}$ via subproblem coordination.

5.2.1 Primal decomposition

We start by first reformulating problem (5.6) as

$$\text{minimize} \quad -\sum_{n\in\mathcal{N}}\sum_{l\in\mathcal{L}(n)}\beta_l\ln\left(1+\frac{p_l|\mathbf{h}_{ll}^{\mathrm{H}}\mathbf{v}_l|^2}{\sigma_l^2+\sum\limits_{j\in\mathcal{L}(\mathrm{tran}(l)),j\neq l}p_j|\mathbf{h}_{ll}^{\mathrm{H}}\mathbf{v}_j|^2+\sum\limits_{i\in\mathcal{N}_{\mathrm{int}}(l)}w_{il}}\right)$$

$$\text{subject to} \quad w_{il}\ge\sum_{j\in\mathcal{L}(i)}p_j|\mathbf{h}_{jl}^{\mathrm{H}}\mathbf{v}_j|^2,\ l\in\mathcal{L}_{\mathrm{int}},\ i\in\mathcal{N}_{\mathrm{int}}(l)$$
$$\sum_{l\in\mathcal{L}(n)}p_l||\mathbf{v}_l||_2^2\le p_n^{\mathrm{max}},\ n\in\mathcal{N}$$
$$||\mathbf{v}_l||_2=1,\quad p_l\ge 0,\ l\in\mathcal{L}\ ,$$

$$(5.7)$$

where the variables are $(p_l,\mathbf{v}_l)_{\in\mathcal{L}}$ and $(w_{il})_{\in\mathcal{L}_{\mathrm{int}},i\in\mathcal{N}_{\mathrm{int}}(l)}$. Problem (5.6) and (5.7)

are equivalent, since 1) function $\ln(\cdot)$ is increasing and 2) the objective function of problem (5.7) is increasing in w_{il}, and therefore the first set of constraints holds with equality at the optimal point.

Let $\mathcal{L}_{\text{int}}(n)$ denote the set of data streams for which base station n acts as an out-of-cell interferer. In particular, $\mathcal{L}_{\text{int}}(n) = \{l | l \in \mathcal{L}_{\text{int}}, n \in \mathcal{N}_{\text{int}}(l)\}$. By noting that the sets $\{(l, i) | l \in \mathcal{L}_{\text{int}}, i \in \mathcal{N}_{\text{int}}(l)\}$ and $\{(l, n) | n \in \mathcal{N}, l \in \mathcal{L}_{\text{int}}(n)\}$ are identical, we can rewrite the first inequality constraint of problem (5.7) as

$$w_{nl} \geq \textstyle\sum_{j \in \mathcal{L}(n)} p_j |\mathbf{h}_{jl}^{\text{H}} \mathbf{v}_j|^2, \ n \in \mathcal{N}, \ l \in \mathcal{L}_{\text{int}}(n) \ . \tag{5.8}$$

Now, we treat w_{nl} as complicating variables and use primal decomposition techniques to split problem (5.7) into N subproblems (one for each BS) and a master problem. The nth subproblem maximizes the weighted sum rate of nth BS by regarding w_{nl} fixed, i.e.,

$$
\begin{aligned}
\text{minimize} \quad & -\textstyle\sum_{l \in \mathcal{L}(n)} \beta_l \ln \left(1 + \frac{p_l |\mathbf{h}_{ll}^{\text{H}} \mathbf{v}_l|^2}{\sigma_l^2 + \sum\limits_{j \in \mathcal{L}(n), j \neq l} p_j |\mathbf{h}_{ll}^{\text{H}} \mathbf{v}_j|^2 + \sum\limits_{i \in \mathcal{N}_{\text{int}}(l)} w_{il}} \right) \\
\text{subject to} \quad & w_{nl} \geq \textstyle\sum_{j \in \mathcal{L}(n)} p_j |\mathbf{h}_{jl}^{\text{H}} \mathbf{v}_j|^2, \ l \in \mathcal{L}_{\text{int}}(n) \\
& \textstyle\sum_{l \in \mathcal{L}(n)} p_l \|\mathbf{v}_l\|_2^2 \leq p_n^{\max} \\
& \|\mathbf{v}_l\|_2 = 1, \quad p_l \geq 0, \ l \in \mathcal{L}(n) \ ,
\end{aligned}
$$
$$\tag{5.9}$$

with the variable $(p_l, \mathbf{v}_l)_{l \in \mathcal{L}(n)}$, while the master problem updates the complicating variables $(w_{nl})_{n \in \mathcal{N}, l \in \mathcal{L}_{\text{int}}(n)}$ to maximize the overall weighed sum rate [i.e., to maximize the objective of problem (5.7)]. To express the master problem compactly, let us denote the vector $(w_{nl})_{n \in \mathcal{N}, l \in \mathcal{L}_{\text{int}}(n)}$ of out-of-cell interference components by $\mathbf{w}$ and the optimal value of problem (5.9) by $f_n(\mathbf{w})$. The master problem is given by

$$
\begin{aligned}
\text{minimize} \quad & \textstyle\sum_{n \in \mathcal{N}} f_n(\mathbf{w}) \\
\text{subject to} \quad & \mathbf{w} \succeq \mathbf{0} \ ,
\end{aligned}
\tag{5.10}
$$

where the variable is $\mathbf{w}$.

5.2.2 Subproblem: BS optimization

Each subproblem (5.9) should be solved to compute the objective value of the master problem (5.10). However, note that problem (5.9) is NP-hard [23], and

therefore we have to rely on approximations. In this section, we adopt the ascent algorithm originally proposed in [15, Section 4.3] to find a suboptimal solution to problem (5.9).

Let us first introduce auxiliary variables γ_l for all $l \in \mathcal{L}(n)$ and reformulate problem (5.9) equivalently as

$$
\begin{aligned}
\text{minimize} \quad & -\sum_{l \in \mathcal{L}(n)} \beta_l \ln(1 + \gamma_l) \\
\text{subject to} \quad & \gamma_l \leq \frac{p_l |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_l|^2}{\sigma_l^2 + \sum_{j \in \mathcal{L}(n), j \neq l} p_j |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_j|^2 + \sum_{i \in \mathcal{N}_{\mathrm{int}}(l)} w_{il}}, \;\; l \in \mathcal{L}(n) \\
& w_{nl} \geq \sum_{j \in \mathcal{L}(n)} p_j |\mathbf{h}_{jl}^{\mathrm{H}} \mathbf{v}_j|^2, \;\; l \in \mathcal{L}_{\mathrm{int}}(n) \\
& \sum_{l \in \mathcal{L}(n)} p_l \|\mathbf{v}_l\|_2^2 \leq p_n^{\max} \\
& \|\mathbf{v}_l\|_2 = 1, \quad p_l \geq 0, \; l \in \mathcal{L}(n) \, ,
\end{aligned}
\tag{5.11}
$$

where the variable is $(p_l, \gamma_l, \mathbf{v}_l)_{\in \mathcal{L}(n)}$. The equivalence of problem (5.9) and (5.11) follows since the objective function of problem (5.11) is decreasing in γ_l, and therefore the first set of constraints holds with equality at the optimal point.

Particularized to our problem (5.11), the method proposed in [15, Section 4.3] is adopted to obtain the following algorithm, which yields a suboptimal solution for problem (5.11) [see Figure 5.2(a)].

Algorithm 5.1. *Finding a suboptimal solution for BS optimization (5.11)*

1. Initialization; given an initial beamformer configuration $\left(\mathbf{v}_l^{(0)}\right)_{\in \mathcal{L}(n)}$ and an initial power allocation $\left(p_l^{(0)}\right)_{l \in \mathcal{L}(n)}$. Set iteration index $i = 0$.
2. Compute $\hat{\gamma}_l$ for all $l \in \mathcal{L}(n)$ as follows:

$$
\hat{\gamma}_l = \frac{p_l^{(i)} |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_l^{(i)}|^2}{\sigma_l^2 + \sum_{j \in \mathcal{L}(n), j \neq l} p_j^{(i)} |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_j^{(i)}|^2 + \sum_{m \in \mathcal{N}_{\mathrm{int}}(l)} w_{ml}} \, .
$$

3. By fixing $\mathbf{v}_l = \mathbf{v}_l^{(i)}$ for all $l \in \mathcal{L}(n)$, solve the following problem for the

158

variable $(p_l, \gamma_l)_{l \in \mathcal{L}(n)}$:

$$
\begin{aligned}
\text{minimize} \quad & \prod_{l \in \mathcal{L}(n)} \left(\hat{\gamma}_l^{-\hat{\gamma}_l/(1+\hat{\gamma}_l)} (1 + \hat{\gamma}_l) \right)^{-\beta_l} \prod_{l \in \mathcal{L}(n)} \gamma_l^{-\beta_l \frac{\hat{\gamma}_l}{1+\hat{\gamma}_l}} \\
\text{subject to} \quad & \gamma_l \leq \frac{p_l |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_l|^2}{\sigma_l^2 + \sum\limits_{j \in \mathcal{L}(n), j \neq l} p_j |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_j|^2 + \sum\limits_{m \in \mathcal{N}_{\mathrm{int}}(l)} w_{ml}}, \ l \in \mathcal{L}(n) \\
& w_{nl} \geq \sum_{j \in \mathcal{L}(n)} p_j |\mathbf{h}_{jl}^{\mathrm{H}} \mathbf{v}_j|^2, \ l \in \mathcal{L}_{\mathrm{int}}(n) \\
& \sum_{l \in \mathcal{L}(n)} p_l \leq p_n^{\max} \\
& p_l \geq 0, \ l \in \mathcal{L}(n) .
\end{aligned}
$$

(5.12)

Denote the solution by $(p_l^\star, \gamma_l^\star)_{l \in \mathcal{L}(n)}$.

4. Stopping criterion; if the stopping criterion is satisfied STOP by returning the suboptimal solution $(\tilde{p}_l, \tilde{\gamma}_l, \tilde{\mathbf{v}}_l)_{l \in \mathcal{L}(n)}$, where $\tilde{p}_l = p_l^\star$, $\tilde{\gamma}_l = \gamma_l^\star$, and $\tilde{\mathbf{v}}_l = \mathbf{v}_l^{(i)}$. Otherwise, update the achieved SINR values $\gamma_l^{\mathrm{tmp}} = \gamma_l^\star$ for all $l \in \mathcal{L}(n)$.

5. By fixing $\gamma_l = \gamma_l^{\mathrm{tmp}}$ for all $l \in \mathcal{L}(n)$, solve the following problem for the variables $(p_l, \mathbf{v}_l)_{l \in \mathcal{L}(n)}$ and t:

$$
\begin{aligned}
\text{minimize} \quad & t \\
\text{subject to} \quad & \gamma_l \leq \frac{p_l |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_l|^2}{\sigma_l^2 + \sum\limits_{j \in \mathcal{L}(n), j \neq l} p_j |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_j|^2 + \sum\limits_{m \in \mathcal{N}_{\mathrm{int}}(l)} w_{ml}}, \ l \in \mathcal{L}(n) \\
& t^2 w_{nl} \geq \sum_{j \in \mathcal{L}(n)} p_j |\mathbf{h}_{jl}^{\mathrm{H}} \mathbf{v}_j|^2, \ l \in \mathcal{L}_{\mathrm{int}}(n) \\
& \sum_{l \in \mathcal{L}(n)} p_l ||\mathbf{v}_l||_2^2 \leq t^2 p_n^{\max} \\
& ||\mathbf{v}_l||_2 = 1, \ p_l \geq 0, \ l \in \mathcal{L}(n) .
\end{aligned}
$$

(5.13)

Denote the solution by $(p_l^\star, \mathbf{v}_l^\star)_{l \in \mathcal{L}(n)}$ and $t^\star$. Update $p_l^{(i+1)} = p_l^\star/(t^\star)^2$ and $\mathbf{v}_l^{(i+1)} = \mathbf{v}_l^\star$ for all $l \in \mathcal{L}(n)$. Set $i = i + 1$ and go to step 2.

The block diagram shown in Figure 5.2(a) summarizes Algorithm 5.1. The first step initializes the algorithm. The second step computes a feasible $(\gamma_l)_{l \in \mathcal{L}(n)}$ for problem (5.11), which is given by $(\hat{\gamma}_l)_{l \in \mathcal{L}(n)}$. The point $(\hat{\gamma}_l)_{l \in \mathcal{L}(n)}$ is used in step 3 to obtain an approximated variant of problem (5.11). In particular, the objective of (5.11) is replaced by using an upper bound function [21] and the

---

[21] From Lemma 3.1 we can readily obtain $-\sum_{l \in \mathcal{L}(n)} \beta_l \ln(1 + \gamma_l) = \ln \prod_{l \in \mathcal{L}(n)} (1 + \gamma_l)^{-\beta_l} \leq$ $\ln \prod_{l \in \mathcal{L}(n)} \left(\hat{\gamma}_l^{-\hat{\gamma}_l/(1+\hat{\gamma}_l)} (1 + \hat{\gamma}_l) \right)^{-\beta_l} \prod_{l \in \mathcal{L}(n)} \gamma_l^{-\beta_l \frac{\hat{\gamma}_l}{1+\hat{\gamma}_l}}$ for all $(\gamma_l > 0)_{l \in \mathcal{L}(n)}$, where $\hat{\gamma}_l$s are arbitrary positive numbers. In the objective function of problem (5.12), the natural logarithm 'ln' has been safely dropped since $\ln(\cdot)$ is an increasing function.

(a) Algorithm 5.1   (b) Algorithm 5.2
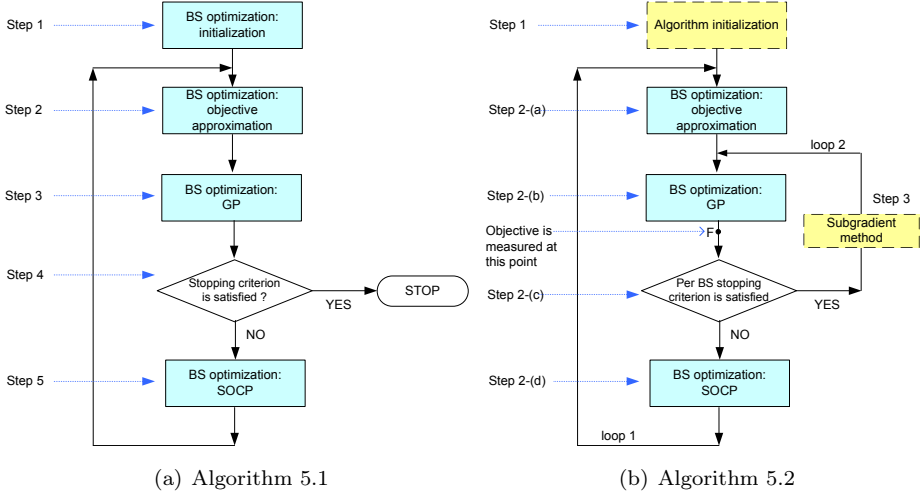
Fig 5.2. Block diagrams of the proposed algorithms.

optimization is carried out with respect to a subset of variables $(p_l, \gamma_l)_{l \in \mathcal{L}(n)}$ by considering the others, i.e., $(\mathbf{v}_l)_{l \in \mathcal{L}(n)}$ fixed. Note that problem (5.12) can easily be posed as a GP. Step 4 checks a stopping criterion. The algorithm terminates if the stopping criterion is satisfied. Otherwise, step 5 is carried out, where the beamformer directions $(\mathbf{v}_l)_{l \in \mathcal{L}(n)}$, as well as the power allocated to beamformers $(p_l)_{l \in \mathcal{L}(n)}$ are optimized to compute a *power margin* $t^\star$ at the BS such that $(\gamma_l)_{l \in \mathcal{L}(n)}$ is preserved. Indeed, problem (5.13) can be equivalently formulated as a SOCP (see [15, Section 4.3]), i.e.,

$$
\begin{aligned}
\text{minimize} \quad & t \\
\text{subject to} \quad & \begin{bmatrix} \sqrt{1 + \frac{1}{\gamma_l}}\,\mathbf{h}_{ll}^{\mathrm{H}}\mathbf{m}_l \\ \mathbf{M}_n^{\mathrm{H}}\mathbf{h}_{ll} \\ \sqrt{\sigma_l^2 + \sum_{i \in \mathcal{N}_{\mathrm{int}}(l)} w_{il}} \end{bmatrix} \succeq_{\mathrm{SOC}} 0, \ l \in \mathcal{L}(n) \\
& \begin{bmatrix} t\sqrt{w_{nl}} \\ \mathbf{M}_n^{\mathrm{H}}\mathbf{h}_{l_n l} \end{bmatrix} \succeq_{\mathrm{SOC}} 0, \ l \in \mathcal{L}_{\mathrm{int}}(n) \\
& \begin{bmatrix} t\sqrt{p_n^{\max}} \\ \mathbf{vec}(\mathbf{M}_n) \end{bmatrix} \succeq_{\mathrm{SOC}} 0,
\end{aligned}
\tag{5.14}
$$

where the variables are $\mathbf{M}_n = [\mathbf{m}_l]_{l \in \mathcal{L}(n)}$ and t, the notation $\succeq_{\mathrm{SOC}}$ denotes the generalized inequality with respect to the second-order cone [3, Section 2.2.3], and l_n is any arbitrary data stream such that $l_n \in \mathcal{L}(n)$. The solution of

problem (5.14) $[\mathbf{m}_l^\star]_{l \in \mathcal{L}(n)}$ is used to recover the solution of problem (5.13) as follows: $\mathbf{v}_l^\star = \mathbf{m}_l^\star / ||\mathbf{m}_l^\star||_2$, $p_l^\star = ||\mathbf{m}_l^\star||_2^2$.

Algorithm 5.1 is a descent algorithm as discussed in [15, Section 4.3]. Note that we have considered $(-\text{ WSR})$ in problem (5.11), and therefore the algorithm is descending. Of course, the value of WSR is ascending. We refer the reader to [15] for more details since the main focus of this chapter is to find a distributed algorithm for the original problem (5.6) [or problem (5.7)].

5.2.3 Master problem

Recall that computing the objective value of the master problem (5.10) requires the solution of each subproblem (5.9) that is NP-hard [23]. Moreover, even if we were able to solve the subproblems, we *cannot* directly apply standard subgradient methods to solve the master problem (5.10) since it is *not convex*. To handle these difficulties, we develop a method that solves successive approximated variants of the original master problem (5.10). Each approximated problem can be transformed into a convex problem by a change of variables. Thus, a subgradient method can be carried out to solve the resulting convex problem. It is increasingly important to note that the approximations and variable transformations mentioned above are such that we can always rely on the results of BS optimizations (see Section 5.2.2) to compute a subgradient for the subgradient method. This allows the coordination of the BS optimizations to find an approximate solution for the master problem (5.10), which resolves the out-of-cell interference.

We start by approximating the objective function of problem (5.10) from an upper bound function, which in turn is used to obtain an approximation of the master problem. We refer to the resulting approximation as *the approximated master* problem. Next, we derive an equivalent convex form of the approximated master problem, followed by the subgradient methods to solve it.

Derivation of an upper bound function for the master objective

To simplify the presentation, we first define two sets $\mathcal{H}$ and $\check{\mathcal{H}}$, where $\mathcal{H}$ is the feasible set of problem (5.11) and $\check{\mathcal{H}}$ is a set such that $\check{\mathcal{H}} \subset \mathcal{H}$. In particular, $\mathcal{H}$

is given by

$$
\mathcal{H} = \left\{ (p_l, \gamma_l, \mathbf{v}_l)_{l \in \mathcal{L}(n)} \left|
\begin{array}{l}
\gamma_l \leq \dfrac{p_l |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_l|^2}{\sigma_l^2 + \sum\limits_{j \in \mathcal{L}(n), j \neq l} p_j |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_j|^2 + \sum\limits_{i \in \mathcal{N}_{\mathrm{int}}(l)} w_{il}}, \ l \in \mathcal{L}(n) \\
w_{nl} \geq \sum_{j \in \mathcal{L}(n)} p_j |\mathbf{h}_{jl}^{\mathrm{H}} \mathbf{v}_j|^2, \ l \in \mathcal{L}_{\mathrm{int}}(n) \\
\sum_{l \in \mathcal{L}(n)} p_l \|\mathbf{v}_l\|_2^2 \leq p_n^{\max} \\
\|\mathbf{v}_l\|_2 = 1, \quad p_l \geq 0, \ l \in \mathcal{L}(n)
\end{array}
\right. \right\}.
$$
(5.15)

By fixing $\mathbf{v}_l = \check{\mathbf{v}}_l$ in (5.15), we obtain the subset $\check{\mathcal{H}}$ of $\mathcal{H}$. More specifically,

$$
\check{\mathcal{H}} = \left\{ (p_l, \gamma_l)_{l \in \mathcal{L}(n)} \left|
\begin{array}{l}
\gamma_l \leq \dfrac{p_l |\mathbf{h}_{ll}^{\mathrm{H}} \check{\mathbf{v}}_l|^2}{\sigma_l^2 + \sum\limits_{j \in \mathcal{L}(n), j \neq l} p_j |\mathbf{h}_{ll}^{\mathrm{H}} \check{\mathbf{v}}_j|^2 + \sum\limits_{i \in \mathcal{N}_{\mathrm{int}}(l)} w_{il}}, \ l \in \mathcal{L}(n) \\
w_{nl} \geq \sum_{j \in \mathcal{L}(n)} p_j |\mathbf{h}_{jl}^{\mathrm{H}} \check{\mathbf{v}}_j|^2, \ l \in \mathcal{L}_{\mathrm{int}}(n) \\
\sum_{l \in \mathcal{L}(n)} p_l \leq p_n^{\max} \\
p_l \geq 0, \ l \in \mathcal{L}(n)
\end{array}
\right. \right\}.
$$
(5.16)

Now we can write the following relations:

$$
f_n(\mathbf{w}) = \inf_{(p_l, \gamma_l, \mathbf{v}_l)_{l \in \mathcal{L}(n)} \in \mathcal{H}} - \sum_{l \in \mathcal{L}(n)} \beta_l \ln(1 + \gamma_l) \tag{5.17}
$$

$$
\leq \inf_{(p_l, \gamma_l)_{l \in \mathcal{L}(n)} \in \check{\mathcal{H}}} - \sum_{l \in \mathcal{L}(n)} \beta_l \ln(1 + \gamma_l) \tag{5.18}
$$

$$
= \inf_{(p_l, \gamma_l)_{l \in \mathcal{L}(n)} \in \check{\mathcal{H}}} \ln \left(\prod_{l \in \mathcal{L}(n)} (1 + \gamma_l)^{-\beta_l} \right) \tag{5.19}
$$

$$
\leq \inf_{(p_l, \gamma_l)_{l \in \mathcal{L}(n)} \in \check{\mathcal{H}}} \ln \left(\prod_{l \in \mathcal{L}(n)} \left(\check{\gamma}_l^{-\frac{\check{\gamma}_l}{1 + \check{\gamma}_l}} (1 + \check{\gamma}_l) \, \gamma_l^{-\frac{\check{\gamma}_l}{1 + \check{\gamma}_l}} \right)^{-\beta_l} \right) \tag{5.20}
$$

$$
= \ln \left(\underbrace{\inf_{(p_l, \gamma_l)_{l \in \mathcal{L}(n)} \in \check{\mathcal{H}}} \prod_{l \in \mathcal{L}(n)} \left(\check{\gamma}_l^{-\frac{\check{\gamma}_l}{1 + \check{\gamma}_l}} (1 + \check{\gamma}_l) \, \gamma_l^{-\frac{\check{\gamma}_l}{1 + \check{\gamma}_l}} \right)^{-\beta_l}}_{\check{f}_n(\mathbf{w})} \right) \tag{5.21}
$$

$$
= \ln \left(\check{f}_n(\mathbf{w}) \right). \tag{5.22}
$$

The first equality (5.17) follows from the definition of $f_n(\mathbf{z})$ and the equivalence of problem (5.9) and (5.11), (5.18) follows since $\check{\mathcal{H}} \subset \mathcal{H}$, (5.19) follows trivially by using the properties of $\ln(\cdot)$ function, (5.20) follows since $1 + \gamma_l \geq \check{\gamma}_l^{-\frac{\check{\gamma}_l}{1 + \check{\gamma}_l}} (1 + \check{\gamma}_l) \gamma_l^{-\frac{\check{\gamma}_l}{1 + \check{\gamma}_l}}$, where $\check{\gamma}_l$ is an arbitrary positive number (see Lemma 3.1), (5.21) follows since $\ln(\cdot)$ is a nondecreasing function, and $\check{f}_n(\mathbf{w})$ is the optimal value of

the following problem:

$$
\begin{aligned}
&\text{minimize} && \prod_{l\in\mathcal{L}(n)} \left(\check{\gamma}_l^{\,-\check{\gamma}_l/(1+\check{\gamma}_l)}(1+\check{\gamma}_l)\right)^{-\beta_l} \prod_{l\in\mathcal{L}(n)} \gamma_l^{\,-\beta_l \frac{\check{\gamma}_l}{1+\check{\gamma}_l}} \\
&\text{subject to} && (p_l,\gamma_l)_{l\in\mathcal{L}(n)} \in \check{\mathcal{H}} \ ,
\end{aligned}
\tag{5.23}
$$

where the variable is $(p_l,\gamma_l)_{l\in\mathcal{L}(n)}$. Note that problem (5.23) is equivalent to problem (5.12) with $(\mathbf{v}_l = \check{\mathbf{v}}_l)_{l\in\mathcal{L}(n)}$ and $(\hat{\gamma}_l = \check{\gamma}_l)_{l\in\mathcal{L}(n)}$.

From (5.17)-(5.22) we have $f_n(\mathbf{w}) \leq \ln\big(\check{f}_n(\mathbf{w})\big)$, which holds for all $n \in \mathcal{N}$. Thus we have

$$
\sum_{n\in\mathcal{N}} f_n(\mathbf{w}) \leq \sum_{n\in\mathcal{N}} \ln\big(\check{f}_n(\mathbf{w})\big) \ ,
\tag{5.24}
$$

which gives an upper bound on the objective function of (5.10).

The approximated master problem and its convex form

The approximated master problem is obtained by replacing the objective function of the original master problem (5.10) by the upper bound function given in (5.24) and by replacing the constraint $\mathbf{w} \succeq \mathbf{0}$ with $\mathbf{w} \succ \mathbf{0}$, i.e.,

$$
\begin{aligned}
&\text{minimize} && \sum_{n\in\mathcal{N}} \ln\big(\check{f}_n(\mathbf{w})\big) \\
&\text{subject to} && \mathbf{w} \succ \mathbf{0} \ ,
\end{aligned}
\tag{5.25}
$$

where the variable is $\mathbf{w}$. It is worth noting that the approximated master problem (5.25) still can allow ZF solutions (if any) by enforcing appropriate elements of $\mathbf{w}$ to be arbitrarily close to zero.

Next, we note for later reference the problem obtained from (5.25) by using the logarithmic change of variables $\bar{w}_{il} = \ln w_{il}$ (so $w_{il} = e^{\bar{w}_{il}}$), i.e.,

$$
\text{minimize} \quad \sum_{n\in\mathcal{N}} \ln\big(\check{f}_n(e^{\bar{\mathbf{w}}})\big) \ ,
\tag{5.26}
$$

with the variable $\bar{\mathbf{w}} = \{\bar{w}_{il}\}_{l\in\mathcal{L}_{\text{int}}, i\in\mathcal{N}_{\text{int}}(l)}$. Here we use the notation $e^{\mathbf{y}}$, where $\mathbf{y}$ is a vector, to mean componentwise exponentiation: $[e^{\mathbf{y}}]_k = e^{y_k}$. In the sequel, we show that problem (5.26) is indeed a convex reformulation of the approximated master problem (5.25).

To do this, we must show that the objective function of problem (5.26), i.e., $\sum_{n\in\mathcal{N}} \ln\big(\check{f}_n(e^{\bar{\mathbf{w}}})\big)$ is convex in $\bar{\mathbf{w}}$. The convexity of $\sum_{n\in\mathcal{N}} \ln\big(\check{f}_n(e^{\bar{\mathbf{w}}})\big)$ is readily verified by considering the convex form of GP (5.23). To see this, we first

restate GP (5.23) by explicitly characterizing its constrains, i.e.,

$$
\begin{aligned}
\text{minimize} \quad & \prod_{l\in\mathcal{L}(n)} \left(\check{\gamma}_l^{-\check{\gamma}_l/(1+\check{\gamma}_l)}(1+\check{\gamma}_l)\right)^{-\beta_l} \prod_{l\in\mathcal{L}(n)} \gamma_l^{-\beta_l \frac{\check{\gamma}_l}{1+\check{\gamma}_l}} \\
\text{subject to} \quad & \gamma_l \leq \frac{p_l|\mathbf{h}_{ll}^{\mathrm{H}}\check{\mathbf{v}}_l|^2}{\sigma_l^2 + \sum\limits_{j\in\mathcal{L}(n),j\neq l} p_j|\mathbf{h}_{ll}^{\mathrm{H}}\check{\mathbf{v}}_j|^2 + \sum\limits_{i\in\mathcal{N}_{\mathrm{int}}(l)} w_{il}}, \ l \in \mathcal{L}(n) \setminus \mathcal{L}_{\mathrm{local}}(n) \\
& \gamma_l \leq \frac{p_l|\mathbf{h}_{ll}^{\mathrm{H}}\check{\mathbf{v}}_l|^2}{\sigma_l^2 + \sum\limits_{j\in\mathcal{L}(n),j\neq l} p_j|\mathbf{h}_{ll}^{\mathrm{H}}\check{\mathbf{v}}_j|^2}, \quad l \in \mathcal{L}_{\mathrm{local}}(n) \\
& w_{nl} \geq \sum_{j\in\mathcal{L}(n)} p_j|\mathbf{h}_{jl}^{\mathrm{H}}\check{\mathbf{v}}_j|^2, \ l \in \mathcal{L}_{\mathrm{int}}(n) \\
& \sum_{l\in\mathcal{L}(n)} p_l \leq p_n^{\mathrm{max}} \\
& p_l \geq 0, \ l \in \mathcal{L}(n) ,
\end{aligned}
\tag{5.27}
$$

where the variable is $(p_l, \gamma_l)_{l\in\mathcal{L}(n)}$ and $\mathcal{L}_{\mathrm{local}}(n)$ is the subset of data streams transmitted by the nth BS, which are not interfered with by any out-of-cell interference, i.e., $\mathcal{L}_{\mathrm{local}}(n) = \{l \mid l \in \mathcal{L}(n), \mathcal{N}_{\mathrm{int}}(l) = \emptyset\}$. Let us next obtain the geometric program (5.27) in convex form by using the logarithmic change of variables $\bar{p}_l = \ln p_l$, $\bar{\gamma}_l = \ln \gamma_l$, and a logarithmic transformation of the objective and constraint functions as follows:

$$
\begin{aligned}
\text{minimize} \quad & \sum_{l\in\mathcal{L}(n)} \frac{\beta_l \check{\gamma}_l}{1+\check{\gamma}_l} \bar{\gamma}_l + \ln\left(\prod_{l\in\mathcal{L}(n)} \left(\check{\gamma}_l^{-\frac{\check{\gamma}_l}{1+\check{\gamma}_l}}(1+\check{\gamma}_l)\right)^{-\beta_l}\right) \\
\text{subject to} \quad & \ln\left(g_{ll}^{-1}e^{\bar{\gamma}_l} - \bar{p}_l\left(\sigma_l^2 + \sum_{j\in\mathcal{L}(n),j\neq l} g_{jl}e^{\bar{p}_j} + \sum_{i\in\mathcal{N}_{\mathrm{int}}(l)} w_{il}\right)\right) \leq 0, \\
& \hspace{6cm} l \in \mathcal{L}(n) \setminus \mathcal{L}_{\mathrm{local}}(n) \\
& \ln\left(g_{ll}^{-1}e^{\bar{\gamma}_l} - \bar{p}_l\left(\sigma_l^2 + \sum_{j\in\mathcal{L}(n),j\neq l} g_{jl}e^{\bar{p}_j}\right)\right) \leq 0, \ l \in \mathcal{L}_{\mathrm{local}}(n) \\
& \ln\left(\sum_{j\in\mathcal{L}(n)} g_{jl}w_{nl}^{-1}e^{\bar{p}_j}\right) \leq 0, \quad l \in \mathcal{L}_{\mathrm{int}}(n) \\
& \ln\left(\sum_{l\in\mathcal{L}(n)} (p_n^{\mathrm{max}})^{-1}e^{\bar{p}_l}\right) \leq 0 ,
\end{aligned}
\tag{5.28}
$$

where the variable is $(\bar{p}_l, \bar{\gamma}_l)_{l\in\mathcal{L}(n)}$ and $g_{jl} = |\mathbf{h}_{jl}^{\mathrm{H}}\check{\mathbf{v}}_j|^2$. Since the optimal value of GP (5.27) is $\check{f}_n(\mathbf{w})$, the optimal value of the GP in its convex form (i.e., problem (5.28) above) is given by $\ln(\check{f}_n(\mathbf{w}))$. Moreover, by setting $\mathbf{w} = e^{\bar{\mathbf{w}}}$ in problem (5.28), we note the following: 1) the constraint functions of problem (5.28) become jointly convex in $(\bar{p}_l, \bar{\gamma}_l)_{l\in\mathcal{L}(n)}$ and $\bar{\mathbf{w}}$, 2) the optimal value of problem (5.28) becomes $\ln\left(\check{f}_n(e^{\bar{\mathbf{w}}})\right)$. Therefore, by applying [36, Lem. 1], we conclude that $\ln\left(\check{f}_n(e^{\bar{\mathbf{w}}})\right)$ is convex in $\bar{\mathbf{w}}$. Consequently, the objective function

164

of problem (5.26), i.e., $\sum_{n\in\mathcal{N}}\ln\big(\check{f}_n\big(e^{\bar{\mathbf{W}}}\big)\big)$ becomes convex in $\bar{\mathbf{w}}$ as well.

Subgradient method to solve the convex form of the approximated master problem

In this subsection, we derive the subgradient method for solving problem (5.26). To do this, we need to find a subgradient of $\sum_{n\in\mathcal{N}}\ln\big(\check{f}_n\big(e^{\bar{\mathbf{W}}}\big)\big)$ at $\bar{\mathbf{w}}$.

A subgradient of $\sum_{n\in\mathcal{N}}\ln\big(\check{f}_n\big(e^{\bar{\mathbf{W}}}\big)\big)$ at $\bar{\mathbf{w}}$ is obtained as follows. For each $n\in\mathcal{N}$, solve problem (5.28) by fixing $\mathbf{w}=e^{\bar{\mathbf{W}}}$ and return a subgradient of $\ln\big(\check{f}_n\big(e^{\bar{\mathbf{W}}}\big)\big)$ at $\bar{\mathbf{w}}$. For a given n, let $(d_{il}^n(\bar{\mathbf{w}}))_{l\in\mathcal{L}_{\mathrm{int}},i\in\mathcal{N}_{\mathrm{int}}(l)}$ denote the subgradient of $\ln\big(\check{f}_n\big(e^{\bar{\mathbf{W}}}\big)\big)$ at $\bar{\mathbf{w}}$. Thus, a subgradient of $\sum_{n\in\mathcal{N}}\ln\big(\check{f}_n\big(e^{\bar{\mathbf{W}}}\big)\big)$ at $\bar{\mathbf{w}}$ is simply given by $\sum_{n\in\mathcal{N}}(d_{il}^n(\bar{\mathbf{w}}))_{l\in\mathcal{L}_{\mathrm{int}},i\in\mathcal{N}_{\mathrm{int}}(l)}$. The subgradient method for problem (5.26) is given by [192]

$$\bar{w}_{il}^{(j+1)}=\bar{w}_{il}^{(j)}-\theta^{(j)}\sum_{n\in\mathcal{N}}d_{il}^n\big(\bar{\mathbf{w}}^{(j)}\big)\ ,\ l\in\mathcal{L}_{\mathrm{int}},\ i\in\mathcal{N}_{\mathrm{int}}(l)\ , \qquad (5.29)$$

where j is the current iteration index of the subgradient method and $\theta^{(j)}\in\mathbb{R}_+$ is a step size. Let us next see the computation of $d_{il}^n(\bar{\mathbf{w}})$.

By applying [36, Lem. 1], $d_{il}^n(\bar{\mathbf{w}})$ for all $l\in\mathcal{L}_{\mathrm{int}},i\in\mathcal{N}_{\mathrm{int}}(l)$ is given by

$$d_{il}^n(\bar{\mathbf{w}})=\begin{cases}\dfrac{\lambda_l^\star\big(e^{\bar{\mathbf{W}}}\big)e^{\bar{w}_{il}}}{\sigma_l^2+\sum_{j\in\mathcal{L}(n),j\neq l}g_{jl}e^{\bar{p}_j^\star\big(e^{\bar{\mathbf{W}}}\big)}+\sum_{m\in\mathcal{N}_{\mathrm{int}}(l)}e^{\bar{w}_{ml}}} & \begin{array}{l}l\in\mathcal{L}(n)\setminus\mathcal{L}_{\mathrm{local}}(n),\\ i\in\mathcal{N}_{\mathrm{int}}(l)\end{array}\\[3mm] -\mu_l^\star\big(e^{\bar{\mathbf{W}}}\big) & l\in\mathcal{L}_{\mathrm{int}}(n),\ i=n\\[1mm] 0 & \text{otherwise}\ ,\end{cases}$$
$$(5.30)$$

where $(\lambda_l^\star(\mathbf{w}))_{l\in\mathcal{L}(n)\setminus\mathcal{L}_{\mathrm{local}}(n)}$ denotes the optimal Lagrange multipliers associated with the first set of constraints of problem (5.28), $(\mu_l^\star(\mathbf{w}))_{l\in\mathcal{L}_{\mathrm{int}}(n)}$ denotes the optimal Lagrange multipliers associated with the third set of constraints of (5.28), and $(\bar{p}_l^\star(\mathbf{w}),\bar{\gamma}_l^\star(\mathbf{w}))_{l\in\mathcal{L}(n)}$ denotes the optimal solution of problem (5.28). Since problem (5.28) and (5.27) are equivalent, we can further simplify (5.30) by considering the optimal sensitivity values (see [5, Section 3.3]) and the solution of problem (5.27) for $\mathbf{w}=e^{\bar{\mathbf{W}}}$, i.e.,

$$d_{il}^n(\bar{\mathbf{w}})=\begin{cases}\dfrac{\lambda_l^\star(\mathbf{w})w_{il}}{\sigma_l^2+\sum_{j\in\mathcal{L}(n),j\neq l}g_{jl}p_j^\star(\mathbf{w})+\sum_{m\in\mathcal{N}_{\mathrm{int}}(l)}\ w_{ml}} & \begin{array}{l}l\in\mathcal{L}(n)\setminus\mathcal{L}_{\mathrm{local}}(n),\\ i\in\mathcal{N}_{\mathrm{int}}(l)\end{array}\\[3mm] -\mu_l^\star(\mathbf{w}) & l\in\mathcal{L}_{\mathrm{int}}(n),\ i=n\\[1mm] 0 & \text{otherwise}\ ,\end{cases}$$
$$(5.31)$$

where $(\lambda_l^\star(\mathbf{z}))_{l \in \mathcal{L}(n) \setminus \mathcal{L}_{\mathrm{local}}(n)}$ is given by the optimal sensitivity values associated with the first set of constraints of problem (5.27), $(\mu_l^\star(\mathbf{z}))_{l \in \mathcal{L}_{\mathrm{int}}(n)}$ is given by the optimal sensitivity values associated with the third set of constraints of (5.27), and $(p_l^\star(\mathbf{z}), \gamma_l^\star(\mathbf{z}))_{l \in \mathcal{L}(n)}$ is the optimal solution of problem (5.27).
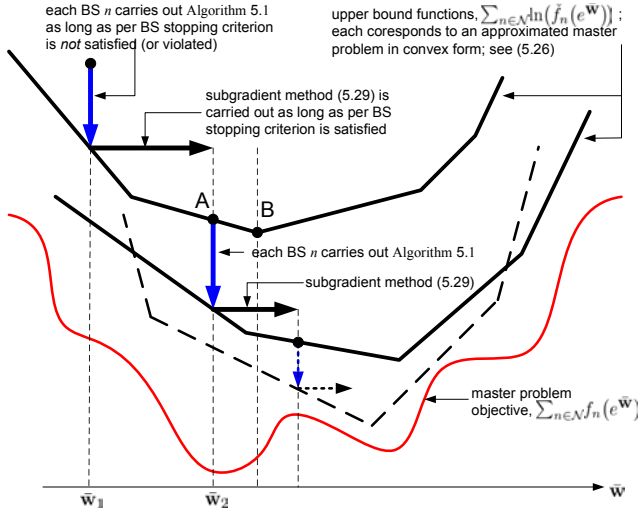
Finally, we recall that problem (5.23) [and therefore (5.27)] is equivalent to problem (5.12) with $(\mathbf{v}_l = \check{\mathbf{v}}_l)_{l \in \mathcal{L}(n)}$ and $(\hat{\gamma}_l = \check{\gamma}_l)_{l \in \mathcal{L}(n)}$. This equivalence is very important in deriving our distributed algorithm for WSRMax. In particular, the subproblem solution method, i.e., Algorithm 5.1 for all $n \in \mathcal{N}$ can be coordinated to find the subgradient $\sum_{n \in \mathcal{N}} (d_{il}^n(\bar{\mathbf{w}}))_{l \in \mathcal{L}_{\mathrm{int}}, i \in \mathcal{N}_{\mathrm{int}}(l)}$, and therefore to solve the convex form of the approximated master problem (5.26), as we will see in the following section.

5.2.4   Distributed algorithm for WSRMax

In this section we blend the solution methods 1) Algorithm 5.1, which finds a suboptimal solution to subproblem (5.9) (see Section 5.2.2) and 2) the subgradient method, which solves an approximation of the master problem (5.10) (see Section 5.2.3). The result is an algorithm, which solves a series of approximated variants of the original master problem (5.10) via a subgradient method. Subgradients for the subgradient method are computed by coordinating the subproblems or the BS optimizations. Note that each BS does *not* need to reveal the entirety of its own subproblem during the BS coordination; only a little communication is needed, and therefore the protocol between BSs can be very light or thin. The main skeleton of the proposed distributed algorithm is depicted in Figure 5.2(b). Note that the equivalence between problem (5.23) and problem (5.12) allows a smooth integration of the subgradient method (5.29) and Algorithm 5.1 in an iterative manner, as shown in Figure 5.2(b).

A graphical interpretation of the behavior of the proposed algorithm by using the objective functions of the master problem (5.10) and the approximated master problems of the form (5.25) are depicted in Figure 5.3. First, each BS n carries out Algorithm 5.1 in parallel for a globally agreed out-of-cell interference $\mathbf{w}$ until a stopping criterion is satisfied (see Figure 5.3). Once the stopping criterion is satisfied, each BS has just solved a GP of the form (5.27); see step 2-(b) of Figure 5.2(b). Then the results of the BS optimizations are collected via BS coordination to carry out the subgradient method (5.29) for some approximated

each BS *n* carries out Algorithm 5.1 as long as per BS stopping criterion is *not* satisfied (or violated)

upper bound functions, $\sum_{n \in \mathcal{N}} \ln(\check{f}_n(e^{\bar{\mathbf{w}}}))$; each coresponds to an approximated master problem in convex form; see (5.26)

subgradient method (5.29) is carried out as long as per BS stopping criterion is satisfied

A B

each BS *n* carries out Algorithm 5.1

subgradient method (5.29)

master problem objective, $\sum_{n \in \mathcal{N}} f_n(e^{\bar{\mathbf{w}}})$

$\bar{\mathbf{w}}_1$ $\bar{\mathbf{w}}_2$ $\bar{\mathbf{w}}$

Fig 5.3. The behavior of Algorithm 5.2; the objective function of problem (5.10) and (5.25) are shown in the domain of $\bar{\mathbf{w}}$.

master problem of the form (5.26). The subgradient method (5.29) is carried out as long as the per BS stopping criterion is satisfied. This allows updating the out-of-cell interference to a new value; see the change of $\bar{\mathbf{w}}$ from $\bar{\mathbf{w}}_1$ to $\bar{\mathbf{w}}_2$ in Figure 5.3. Once the per BS stopping criterion is violated, each BS n can again carry out Algorithm 5.1 in parallel for the new out-of-cell interference values of $\mathbf{w}$ until the stopping criterion is satisfied. The process continues in an iterative manner (see Figure 5.3). The detailed algorithm is as follows (see Figure 5.2(b) for a concise block diagram).

Algorithm 5.2. *Distributed algorithm for WSRMax*

1. Initialization; given the globally agreed initial out-of-cell interference $\mathbf{w}$, an initial beamformer configuration $\left(\mathbf{v}_l^{(0)}\right)_{\in \mathcal{L}(n)}$, and an initial power allocation $\left(p_l^{(0)}\right)_{\in \mathcal{L}(n)}$. Set BS iteration index $i_n = 0$ for all $n \in \mathcal{N}$.
2. For $n = 1$ to N

   (a) Compute

   $$\hat{\gamma}_l = \frac{p_l^{(i_n)} |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_l^{(i_n)}|^2}{\sigma_l^2 + \sum\limits_{j \in \mathcal{L}(n), j \neq l} p_j^{(i_n)} |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_j^{(i_n)}|^2 + \sum\limits_{i \in \mathcal{N}_{\mathrm{int}}(l)} w_{il}}, \; l \in \mathcal{L}(n) \; .$$

167

(b) By fixing $\mathbf{v}_l = \mathbf{v}_l^{(i_n)}$ for all $l \in \mathcal{L}(n)$, solve the following problem for the variable $(p_l, \gamma_l)_{l \in \mathcal{L}(n)}$:

$$\text{minimize} \quad \prod_{l \in \mathcal{L}(n)} \left(\hat{\gamma}_l^{-\hat{\gamma}_l/(1+\hat{\gamma}_l)}(1+\hat{\gamma}_l)\right)^{-\beta_l} \prod_{l \in \mathcal{L}(n)} \gamma_l^{-\beta_l \frac{\hat{\gamma}_l}{1+\hat{\gamma}_l}}$$

$$\text{subject to} \quad \gamma_l \leq \frac{p_l |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_l|^2}{\sigma_l^2 + \sum\limits_{j \in \mathcal{L}(n), j \neq l} p_j |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_j|^2 + \sum\limits_{i \in \mathcal{N}_{\mathrm{int}}(l)} w_{il}}, \ l \in \mathcal{L}(n) \setminus \mathcal{L}_{\mathrm{local}}(n)$$

$$\gamma_l \leq \frac{p_l |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_l|^2}{\sigma_l^2 + \sum\limits_{j \in \mathcal{L}(n), j \neq l} p_j |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_j|^2}, \ l \in \mathcal{L}_{\mathrm{local}}(n)$$

$$w_{nl} \geq \sum_{j \in \mathcal{L}(n)} p_j |\mathbf{h}_{jl}^{\mathrm{H}} \mathbf{v}_j|^2, \ l \in \mathcal{L}_{\mathrm{int}}(n)$$

$$\sum_{l \in \mathcal{L}(n)} p_l \leq p_n^{\mathrm{max}}$$

$$p_l \geq 0, \ l \in \mathcal{L}(n) .$$

(5.32)

Denote the solution by $(p_l^{\star}(\mathbf{w}), \gamma_l^{\star}(\mathbf{w}))_{l \in \mathcal{L}(n)}$.

(c) Per BS stopping criterion; if the stopping criterion is satisfied set $(\bar{w}_{il} = \ln w_{il})_{l \in \mathcal{L}_{\mathrm{int}}, i \in \mathcal{N}_{\mathrm{int}}(l)}$, return $d_{il}^n(\bar{\mathbf{w}})$ by using $(\lambda_l^{\star}(\mathbf{w}))_{l \in \mathcal{L}(n) \setminus \mathcal{L}_{\mathrm{local}}(n)}$ and $(\mu_l^{\star}(\mathbf{w}))_{l \in \mathcal{L}_{\mathrm{int}}(n)}$, which correspond to the optimal sensitivity values of the first and third constraints of problem (5.32), respectively, and go to step 3. Otherwise, update achieved SINR values $\gamma_l^{\mathrm{tmp}} = \gamma_l^{\star}(\mathbf{w})$ for all $l \in \mathcal{L}(n)$ and set the subgradient iteration index $j = 0$.

(d) By fixing $\gamma_l = \gamma_l^{\mathrm{tmp}}$ for all $l \in \mathcal{L}(n)$, solve the following problem for the variables $(p_l, \mathbf{v}_l)_{l \in \mathcal{L}(n)}$ and t:

$$\text{minimize} \quad t$$

$$\text{subject to} \quad \gamma_l \leq \frac{p_l |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_l|^2}{\sigma_l^2 + \sum\limits_{j \in \mathcal{L}(n), j \neq l} p_j |\mathbf{h}_{ll}^{\mathrm{H}} \mathbf{v}_j|^2 + \sum\limits_{i \in \mathcal{N}_{\mathrm{int}}(l)} w_{il}}, \ l \in \mathcal{L}(n)$$

$$t^2 w_{nl} \geq \sum_{j \in \mathcal{L}(n)} p_j |\mathbf{h}_{jl}^{\mathrm{H}} \mathbf{v}_j|^2, \ l \in \mathcal{L}_{\mathrm{int}}(n)$$

$$\sum_{l \in \mathcal{L}(n)} p_l \|\mathbf{v}_l\|_2^2 \leq t^2 p_n^{\mathrm{max}}$$

$$\|\mathbf{v}_l\|_2 = 1, \ p_l \geq 0, \ l \in \mathcal{L}(n) .$$

(5.33)

Denote the solution by $(p_l^{\star}, \mathbf{v}_l^{\star})_{l \in \mathcal{L}(n)}$ and $t_n^{\star}$. Update $p_l^{(i_n+1)} = p_l^{\star}/(t_n^{\star})^2$ and $\mathbf{v}_l^{(i_n+1)} = \mathbf{v}_l^{\star}$ for all $l \in \mathcal{L}(n)$. Set $i_n = i_n + 1$ and go to step 2-(a).

3. Set $\bar{\mathbf{w}}^{(j)} = \bar{\mathbf{w}}$. Carry out (5.29) to yield $(\bar{w}_{il}^{(j+1)})_{l \in \mathcal{L}_{\mathrm{int}}, i \in \mathcal{N}_{\mathrm{int}}(l)}$ and update the out-of-cell interference by setting $\mathbf{w} = (e^{\bar{w}_{il}^{(j+1)}})_{l \in \mathcal{L}_{\mathrm{int}}, i \in \mathcal{N}_{\mathrm{int}}(l)}$. Increment subgradient iteration index j, i.e., set $j = j + 1$ and go to step 2-(b).
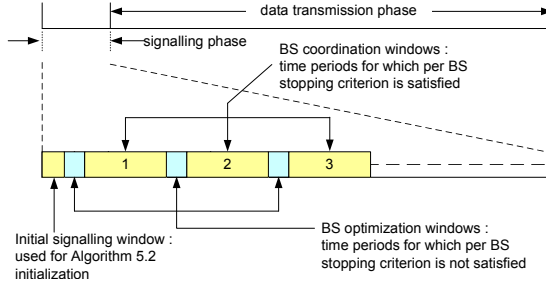
Fig 5.4. An example signalling frame structure.

The first step initializes Algorithm 5.2. Steps 2-(a) to 2-(d) represent the per BS optimizations that can be carried out in a parallelized or decentralized fashion by each BS as long as the per BS stopping criterion is violated. Note that steps 2-(a) to 2-(d) of Algorithm 5.2 are analogous to steps 2 to 5 of Algorithm 5.1; see Figure 5.2. In step 2-(c), once each BS has realized that the per BS stopping criterion has been satisfied, BS coordination is initiated. For example each BS n ($\in \mathcal{N}$) computes its own $d_{il}^n(\bar{\mathbf{w}})$ and those parameters are exchanged between BSs to carry out step 3. Step 3 is, of course, the subgradient method (5.29), which in turn updates the global out-of-cell interference variable $\mathbf{w}$. For clarity, the solid-lined blocks in Figure 5.2(b) represent the steps that are carried out in fully decentralized fashion and the dash-lined blocks represent the steps that are carried out via BS coordination.

Figure 5.4 suggests a signalling strategy between BSs so that Algorithm 5.2 can be carried out in a decentralized fashion with backhaul message exchanges. In particular, each BS's transmissions are synchronized and the data transmission phase of each BS is preceded by a signalling phase, in which the rate/power allocation of each BS is determined via WSRMax; see Figure 5.4. The signalling phase consists of three types of time slots called *initial signalling window*, *BS optimization window*, and *BS coordination window*. The initial signalling window is used for step 1 of Algorithm 5.2, i.e., the initialization step. The latter two (i.e., BS optimization window and BS coordination window) are repeated until the data transmission phase is reached, as shown in Figure 5.4. We define the BS coordination windows to be the time periods for which the per BS stopping criterion [see step 2-(c)] of Algorithm 5.2 is satisfied. Therefore, during any BS coordination window, step 2(c), step 3, and step 2(b) of Algorithm 5.2 is carried out repeatedly; see loop 2 of Figure 5.2(b). The BS optimization windows

are defined to be the time periods for which the per BS stopping criterion is violated or *not* satisfied. Therefore, during BS optimization windows, step 2(c), step 2-(d), step 2-(a), and step 2(b) of Algorithm 5.2 are carried out repeatedly; see loop 1 of Figure 5.2(b). Typically, we may assume that the time period of any BS optimization window is significantly smaller compared to the time period of any BS coordination window because of the following reasons: 1) the extensive computational power available at BSs enabling the BS optimization to be carried out very fast, 2) BS coordination requires backhaul message exchanges between BSs, which in turn demand stringent time requirements.

5.3   Numerical examples

In this section we run our proposed Algorithm 5.2 (Section 5.2.4) in multiuser multicell environments and the benefits due to different degrees of BS coordination are numerically evaluated. As benchmarks, we consider the centralized suboptimal algorithm proposed in [15, Section 4.3] and the optimal branch and bound method proposed in [88].

Since the main goal of this section is to numerically evaluate the performance of Algorithm 5.2, for simplicity, in our simulations we consider a slightly modified signalling strategy compared to the one described at the end of Section 5.2.4; see Figure 5.4. In particular, we assume that during any BS optimization window shown in Figure 5.4, each BS carries out loop 1 [see Figure 5.2(b)] of Algorithm 5.2 repeatedly only for a fixed $J_{\mathrm{BS-opt}}$ number of iterations (instead of for a fixed time period). Moreover, during any BS coordination window, each BS coordinates to carry out loop 2 of Algorithm 5.2 repeatedly for a fixed J_{subgrad} number of iterations (instead of for a fixed time period). Note that in either loop 1 or in loop 2 of Algorithm 5.2, step 2-(b) or the 'BS optimization: GP' is always carried out; see Figure 5.2(b). Therefore, the overall iterations of loop 1 and loop 2 in Algorithm 5.2 during the signalling phase can be determined by the number of times that step 2-(b) is carried out. We refer to the combined loop 1 and loop 2 iterations in Algorithm 5.2 as *GP iterations* for simplicity.

We consider an exponential path loss model, where the channel gains between

BSs and users are given by

$$\mathbf{h}_{ij} = \sqrt{\left(\frac{d_{ij}}{d_0}\right)^{-\eta}} \; \mathbf{c}_{ij} \; , \tag{5.34}$$

where d_{ij} is the distance from the transmitter of the ith data stream to the receiver of jth data stream, d_0 is the *far field reference distance* [176], η is the path loss exponent, and $\mathbf{c}_{ij} \in \mathbb{C}^T$ such that $\mathbf{c}_{ij} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ (i.e., frequency-flat fading with uncorrelated antennas). The first term of (5.34) represents the path loss factor and the second term models Rayleigh small-scale fading. An arbitrarily generated set $\acute{\mathcal{C}}$ of fading coefficients where $\acute{\mathcal{C}} = \{\mathbf{c}_{ij} \mid i, j \in \mathcal{L}\}$ is referred to as a *single fading realization*. The variance of the noise is considered equal for all data streams, i.e., $\sigma_l^2 = N_0$ for all $l \in \mathcal{L}$ and the maximum power constraint is assumed the same for all nodes, i.e., $p_n^{\max} = p_0^{\max}$ for all $n \in \mathcal{N}$. We define the SNR operating point at a distance d [distance units] as

$$\mathrm{SNR}(d) = \begin{cases} \dfrac{p_0^{\max}}{N_0} & d \leq d_0 \\ \dfrac{p_0^{\max}}{N_0} \left(\dfrac{d}{d_0}\right)^{-\eta} & \text{otherwise} . \end{cases} \tag{5.35}$$

In all our simulations we set $d_0 = 1$, $\eta = 4$, $p_0^{\max}/N_0 = 45$ dB, $\mathrm{SNR}(R_{\mathrm{int}}) = 0$ dB, where R_{int} is the radius of the interference region of each BS, and $\mathrm{SNR}(R_{\mathrm{BS}}) = 8$ dB, where R_{BS} is the radius of the transmission region of each BS.

In our simulations two multicell multiuser wireless cellular networks, as shown in Figure 5.5 are considered. In the case of the first network [i.e., Figure 5.5(a)], there are $N = 2$ BSs with $T = 4$ antennas at each one. The BSs are located such that the distance between the two BSs is $D_{\mathrm{BS}} = 1.5 \times R_{\mathrm{BS}}$. In the case of the second network [i.e., Figure 5.5(b)], there are $N = 3$ BSs with $T = 4$ antennas at each one. Moreover, the BSs are located such that they form an equilateral triangle and the distance between any two BSs is $D_{\mathrm{BS}} = 1.5 \times R_{\mathrm{BS}}$. There are 4 users per each BS located inside the transmission region of the BS. The locations of users associated with BSs are arbitrarily chosen, as shown in Figure 5.5. A single data stream is transmitted for each user.

To see the behavior of Algorithm 5.2, we first consider a nonfading case and run the algorithm in both networks shown in Figure 5.5. Figure 5.6 shows the objective value of problem (5.6) computed at point 'F' of Algorithm 5.2 [see Figure 5.2(b)] in the case of the two considered networks for arbitrary generated
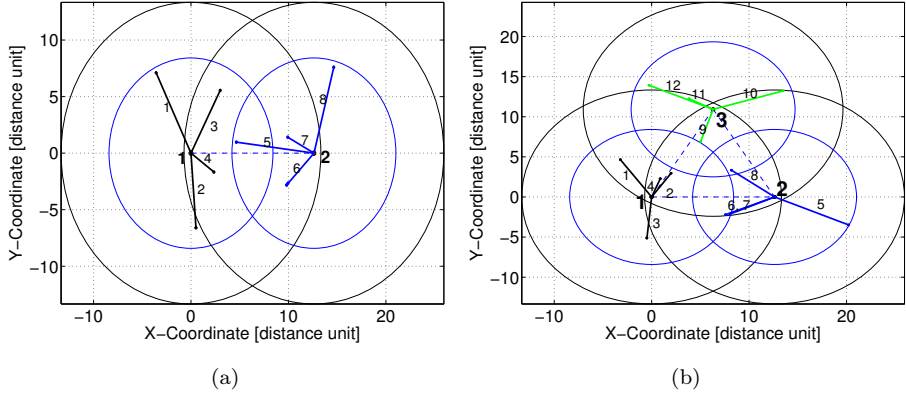
Fig 5.5. (a) Multicell network 1, $\mathcal{N} = \{1, 2\}$**,** $\mathcal{L} = \{1, \ldots, 8\}$**,** $\mathcal{L}(1) = \{1, \ldots, 4\}$**,** $\mathcal{L}(2) = \{5, \ldots, 8\}$**,** $\mathcal{L}_{\mathrm{int}} = \{3, \ldots, 7\}$**; (b) Multicell network 2,** $\mathcal{N} = \{1, 2, 3\}$**,** $\mathcal{L} = \{1, \ldots, 12\}$**,** $\mathcal{L}(1) = \{1, \ldots, 4\}$**,** $\mathcal{L}(2) = \{5, \ldots, 8\}$**,** $\mathcal{L}(3) = \{9, \ldots, 12\}$**,** $\mathcal{L}_{\mathrm{int}} = \{1, 2, 4, 6, \ldots, 11\}$**.**

fading realizations. Note that the X-axes of Figure 5.6 represent GP iterations, i.e., combined loop 1 and loop 2 iterations in Algorithm 5.2. Plots are drawn for the cases of $J_{\mathrm{BS-opt}} = 15$ and $J_{\mathrm{subgrad}} = 1, 5, 10, 20, 50$. Note that J_{subgrad} is a measure of the degree of BS coordination, e.g., $J_{\mathrm{subgrad}} = j$ means that the subgradient method is carried out j consecutive times during any BS coordination window. The weights β_l of each data stream are arbitrarily chosen from the interval $(0, 1]$. In step 1 of Algorithm 5.2, the components of the initial out-of-cell interference vector $\mathbf{w}$ are chosen on the order of noise variance N_0. Moreover, the initial beamformers $\left(\mathbf{v}_l^{(0)}\right)_{l \in \mathcal{L}(n)}$ are randomly generated and a uniform initial beamformer power allocation is chosen, i.e., $\left(p_l^{(0)} = p_0^{\max}/T\right)_{l \in \mathcal{L}(n)}$.

Let us first focus on the case of $J_{\mathrm{subgrad}} = 1$ in Figure 5.6(a) to describe the behavior of the algorithm. Each solid circle on the curve represents the subgradient iteration of Algorithm 5.2 carried out during BS coordination windows [see loop 2 in Figure 5.2(b)]. The solid line segment of the curve, starting at any circle represents BS optimizations performed in parallel during the BS optimization windows [see loop 1 in Figure 5.2(b)]. Note that the BS optimizations are always nondecreasing steps; nondecreasing because we have plotted the positive WSR value instead of the negative value. The flattening of the solid line segments means that BS optimizations cannot further improve the system objective. Note that the subgradient method is not an ascent algorithm [192] in general, and therefore it does not necessarily increase the

172
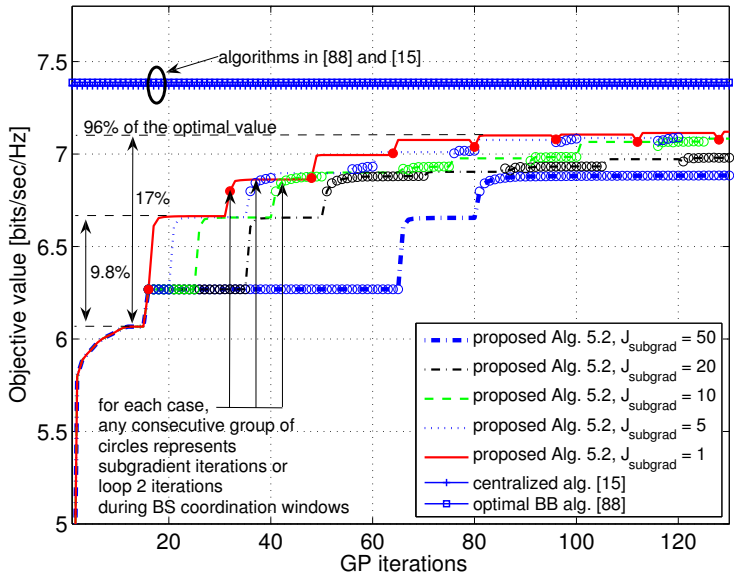
(a)
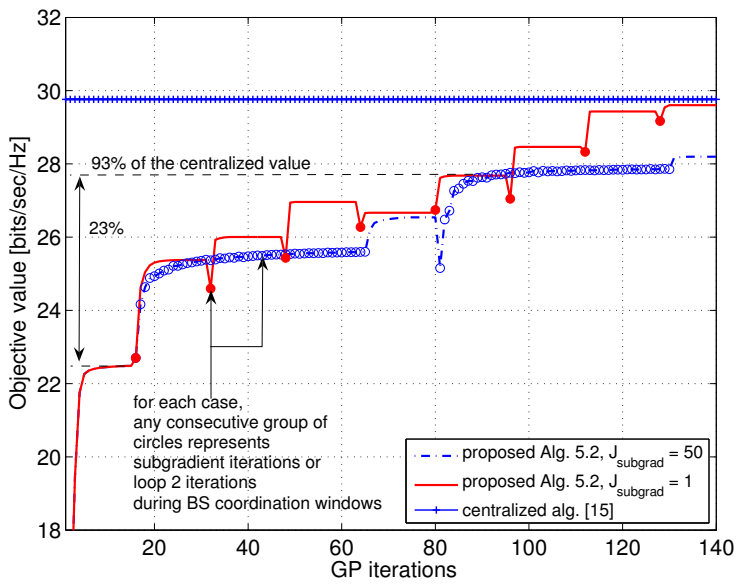
(b)

Fig 5.6. Objective value versus GP iteration: (a) Multicell network 1; (b) Multicell network 2.

objective value. Of course, if one needs to obtain a monotonic algorithm, it is achieved easily by keeping track of the best point found so far [192]. Results show that BS coordination can resolve the out-of-cell interference (i.e., $\mathbf{w}$) via the subgradient method, which increases substantially the system objective value. For example, before the very first subgradient iteration in BS coordination window 1, the algorithm achieves 6.066 bits/sec/Hz and before the subgradient iteration in BS coordination window 2, the algorithm achieves 6.6643 bits/sec/Hz. Therefore, after the very first subgradient iteration in BS coordination window 1, the overall system objective has been improved by more than 9.8% compared to the noncoordinating case [see Figure 5.6(a)]. A 17% gain is achieved after the subgradient iteration in the fifth BS coordination window [see Figure 5.6(a)]. As a reference, we consider the overall system objective value achieved by using the optimal branch and bound algorithm given in [88] and the centralized algorithm proposed in [15, Section 4.3]. Results show that our proposed algorithm achieves close to optimal solutions, even in a distributed manner, with a small amount of BS coordination. For example, five BS coordinations, i.e., five subgradient iterations yield 96% of the objective value obtained by the optimal branch and bound algorithm [88] and the centralized method given in [15, Section 4.3].

Figure 5.6(a) further shows that the value of J_{subgrad}, which parameterizes the degree of BS coordination, has a significant impact on the overall system objective value. It is interesting to note that a smaller number of *consecutive* subgradient iterations (e.g., $J_{\mathrm{subgrad}} = 1,\ 5$) during any BS coordination window can perform better compared to a larger number of consecutive subgradient iterations (e.g., $J_{\mathrm{subgrad}} = 20, 50$). This behavior is very important in practice to reduce significantly the backhaul message exchanges during any BS coordination window (see Figure 5.4). We can intuitively explain the behavior by considering the two points 'A' and 'B' in Figure 5.3. In particular, point 'A' corresponds to a smaller J_{subgrad}, where the (convex form) approximated master problem (5.26) is solved to a low accuracy. Point 'B' corresponds to a larger J_{subgrad}, where the (convex form) approximated master problem is solved to a high accuracy. Of course, point 'B' is better than point 'A' for the (convex form) approximated master problem. But the accuracy of the solution is absolutely irrelevant as far as the original master problem (5.10) is concerned; see the master objective depicted in Figure 5.3. The main concern is to have a suggestive point for the next approximation. Therefore, refining the approximation more often (which

174

corresponds to a smaller J_{subgrad}), rather than solving some approximated master problem to a high accuracy (i.e., a larger J_{subgrad}) is more beneficial.

Figure 5.6(b) shows the algorithm's behavior in the case of network 2 in Figure 5.5(b). As a reference, here we consider only the centralized algorithm [15, Section 4.3]. Note that the problem dimension in the case of network 2 is substantially larger, and therefore the computational complexity of evaluating the optimal value by using the branch and bound algorithm [88] is prohibitively high. Results show that the behavior of Algorithm 5.2 is almost the same as Figure 5.6(a). The network can yield substantial gains by carrying out a lower number of subgradient iterations during any BS coordination window, i.e., less backhaul message exchanges between BSs. For example, after the subgradient iteration in the fifth BS coordination window, the overall system objective value has been improved by 23% compared to the noncoordinating case [see Figure 5.6(b)]. Moreover, 93% of the objective value given by the centralized algorithm [15, Section 4.3] is achieved as well [see Figure 5.6(b)].

Next we present an example to show that our proposed distributed Algorithm 5.2 can sometimes outperform the centralized algorithm given in [15, Section 4.3]. Figure 5.7 shows the objective value of problem (5.6) computed at point 'F' of Algorithm 5.2 in the case of network 1 and network 2 for some other arbitrary generated fading realizations; see Figure 5.7(a) and 5.7(b), respectively. The algorithm parameters are the same as in Figure 5.6 except for the fading realizations. Note that Algorithm 5.2 outperforms the centralized one in the case of both networks. This behavior is intuitively expected since both algorithms, i.e., Algorithm 5.2 and the one given in [15, Section 4.3] are suboptimal methods to problem (5.6), and therefore optimality is not guaranteed.

In order to see the average behavior of the proposed algorithm, finally, we consider a fading case. We run Algorithm 5.2 for 500 fading realizations with $J_{\mathrm{subgrad}} = 1$. Recall that the algorithm parameter $J_{\mathrm{subgrad}} = 1$ means that during any BS coordination window only one subgradient iteration is carried out. The average objective value of problem (5.6) achieved at point 'F' of Algorithm 5.2 [see Figure 5.2(b)] is computed before each BS coordination window.

Figure 5.8 shows the dependence of the average objective value on the number of BS coordinations in the case of the considered network 1 and 2. Note that any BS coordination represents a single BS coordination window or a subgradient iteration since $J_{\mathrm{subgrad}} = 1$. Results show that the performance
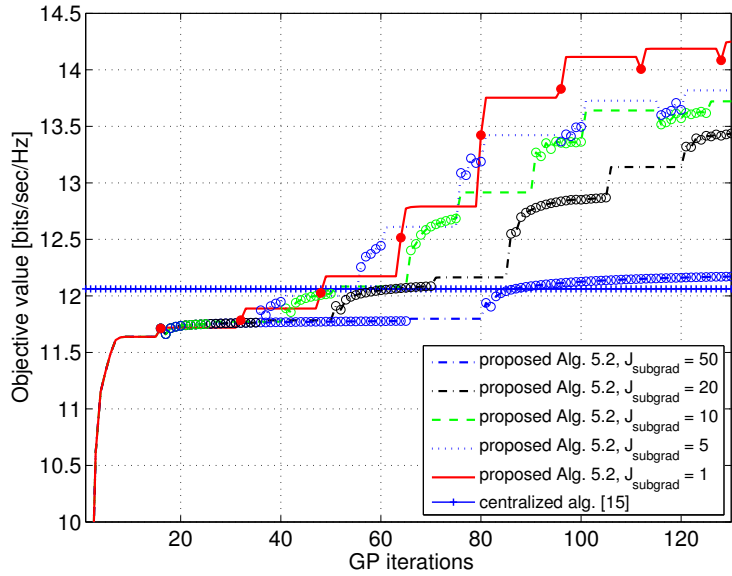
(a)
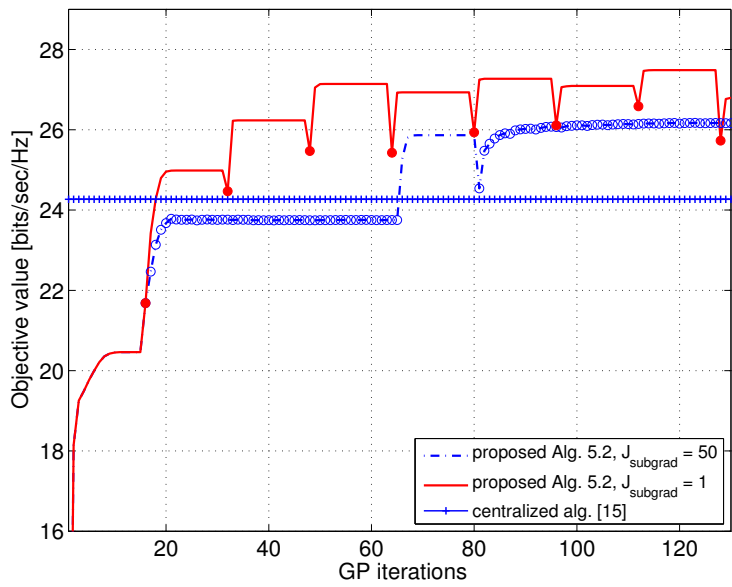
(b)

Fig 5.7. Objective value versus GP iteration: (a) Multicell network 1; (b) Multicell network 2.

Fig 5.8. Average objective value versus number of BS coordinations: (a) Multicell
network 1; (b) Multicell network 2.

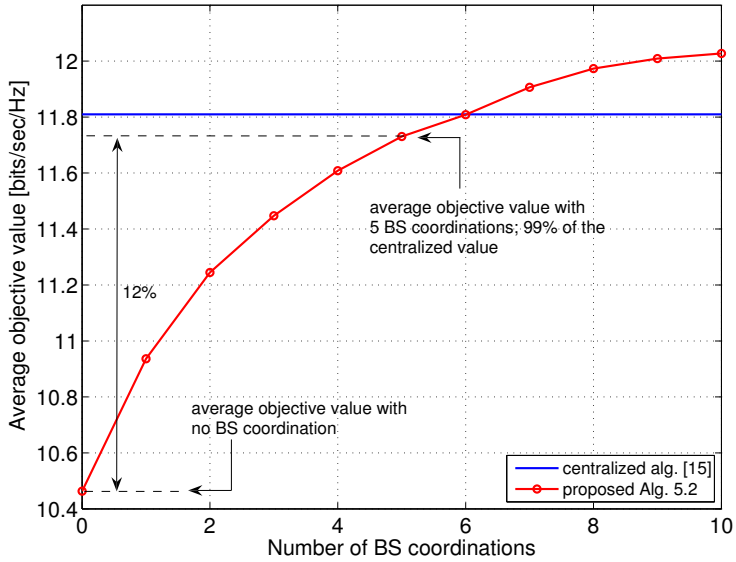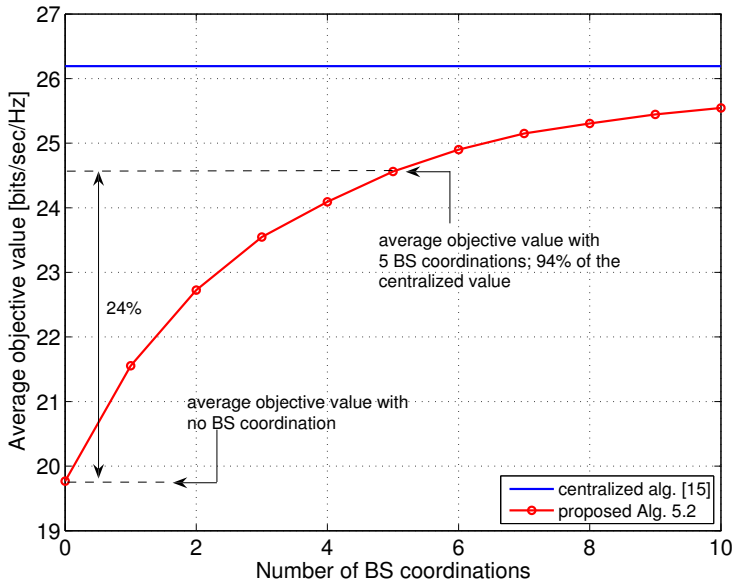gains of the proposed algorithm are very significant. For network 1, a more than 12% improvement in the average objective value is achieved within five BS coordinations when compared to the noncoordinating case [see Figure 5.8(a)]. Moreover, the achieved average objective value in five BS coordinations is around 99% of the value given by the centralized algorithm [15, Section 4.3]. For network 2, within five BS coordinations, a more than 24% improvement in the average objective value is achieved when compared to the noncoordinating case [see Figure 5.8(b)]. Moreover, the corresponding average objective value is more than 94% of the value given by the considered centralized algorithm. Here Algorithm 5.2 performs very closely to the centralized algorithm, but has not outperformed the centralized one, as in the case of network 2 [see Figure 5.8(a)]. This is intuitively expected for the following reasons. The larger the number of BSs, the larger the amount of out-of-cell interference coupling between BSs. Therefore, a centralized approach, of course, can resolve the interference coupling more reliably.

5.4 Summary and discussion

We considered the WSRMax problem in a multicell downlink system. Indeed, the problem is NP-hard in general. A distributed algorithm based on primal decomposition and subgradient methods has been derived to finds a suboptimal solution to the problem. In particular, the main problem was split into many subproblems (one for each BS) and a master problem. An ascent algorithm based on second-order cone programming and geometric programming were adopted in the case of subproblems, where the beamforming directions and the beamforming powers of each BS are optimized for fixed out-of-cell interference values in a fully decentralized fashion. The master problem resolves out-of-cell interference values, which are the complicating variables of the main problem. A sequential convex approximation strategy together with a subgradient method were blended to tackle the master problem. The subgradient method relies on subproblem coordinations (i.e., BS coordinations) to find a subgradient. Because of the subgradient iterations, the overall algorithm is not necessarily an ascent algorithm. Nevertheless, one can readily obtain an ascent algorithm by keeping track of the best point found during the iterations of our proposed method.

Numerical results have been presented to see the benefits due to different

degrees of BS coordination. Results showed that the proposed algorithm could significantly improve the overall system objective value with a small amount of BS coordination. In addition, we compared our proposed distributed algorithm with an optimal branch and bound based method. The considered simulation scenario indicated that our proposed algorithm can achieve WSR values, which are substantially close to the optimal value. Since the computational complexity of the optimal branch and bound based method was prohibitively expensive, we considered a suboptimal centralized algorithm designed for the same problem as another benchmark. The average behavior of the proposed algorithm has been numerically compared with that of the centralized variant. Results showed that for all considered network setups, more than 94% of the average objective given by the centralized algorithm has been achieved by our proposed algorithm within five BS coordinations.

6   Conclusions and future work

In this chapter we first summarize the conclusions of this thesis, highlighting our contributions and the main results. Finally, we highlight some future research directions.

6.1   Conclusions

Radio resource management algorithms for wireless communication networks by applying optimization techniques were investigated in this thesis. A greater emphasis was placed on the general WSRMax problem, which is NP-hard; it plays an important role in various problems of recent interest in wireless communication, including NUM, cross-layer design, link scheduling, and many others.

The first chapter was mainly intended to highlight the motivation for the research and to review earlier and parallel work. In Chapter 2, a global optimization approach based on the branch and bound technique was developed to solve the nonconvex WSRMax problem with an optimality certificate. Efficient analytic bounding methods were developed and their effect on the convergence of the BB algorithm was analyzed numerically. Though the convergence speed was dramatically increased by improving the lower bound, the benefits of improving the upper bounding methods are imperceptible. This suggests that a grater emphasis should be placed on exploring better lower bounding techniques. Unlike other branch and bound based algorithms for WSRMax, our method does not rely on the convertibility of the problem into a DC (difference of convex functions) problem. Therefore, the proposed method applies to a broader class of WSRMax problems (e.g., WSRMax in multicast wireless networks). More importantly, it can tackle *any* system performance metric that can be expressed as a Lipschitz continuous and increasing function of SINR values (not restricted to WSRMax).

The worst case complexity of the proposed BB algorithm increases exponentially with the problem size for a given accuracy. Even a small-scale problems (e.g., one with few tens of variables) can take a long time, and therefore such global methods are applicable when the computing time requirements are not

critical. This means that, in the case of wireless networks, where the computing time is often crucial, the proposed global optimization method is hardly applicable. However, it is indeed useful to provide performance benchmarks, e.g., for evaluating the performance loss encountered by any suboptimal method for WSRMax.

The link-interference model considered in Chapter 2 is very general; it can accommodate a wide variety of network topologies. Moreover, the considered link-interference model supports different node capabilities, including single- or multipacket transmission, single- or multipacket reception, and many others. Numerical examples of diverse application domains of WSRMax were presented to demonstrate our proposed BB algorithm.

Fast, suboptimal algorithms for the WSRMax problem in general multi-commodity, multichannel wireless networks were developed in Chapter 3. The proposed algorithms were carried out within a general cross-layer utility maximization framework and the quantitative impact of gains that can be achieved at the network layer in terms of end-to-end rates and network congestion was numerically evaluated.

First, a general access operation with a relatively simple form of receivers structure (a bank of match filters) was considered; a receiver decodes each of its intended signals by treating all other interfering signals as noise. The proposed algorithms were based on complementary geometric programming and homotopy methods. The algorithm based on homotopy methods handled the self-interference problem without combinatorial constraints to enforce simultaneous transmissions and receptions in the same frequency band; thus the combinatorial nature of the problem has been circumvented. Numerical results showed that our algorithms could exploit the multichannel diversity via dynamic power allocation across the available channels. It was interesting to see numerically that they performed closely to the exponentially complex optimal BB algorithm. The proposed methods also provide a mechanism, based on which the gains achievable at the network layer could be evaluated when the network nodes employ self interference cancelation techniques with different degrees of accuracy. Numerical results showed that quantifiable gains at the network layer have been achieved after a certain level of the self interference cancelation accuracy, e.g., a self interference reduction in the range $20 - 60$dB led to significant gains at the network layer.

182

The latter part of Chapter 3 considered a case where all receivers perform multiuser detection and the gains that can be achieved at the network layer were numerically evaluated. The proposed solution methods here were obtained by imposing additional constraints, such as that only one node can transmit to others at a time or that only one node can receive from others at a time. The main benefit of such constraints was the problem tractability and simplicity of the solution methods. Consequently, these simple access protocols can be potentially useful in practice with more advanced communication systems.

Note that all the algorithms purposed in Chapter 3 are reliant on the well known interior-point methods. Therefore, the proposed algorithms are fast, compared to the optimal BB algorithm proposed in Chapter 2 and they can be deployed in relatively large-scale problems (e.g., one with few hundreds of variables), which are not, of course, apparently handled by exponentially complex optimal algorithms. However, to facilitate the use of the algorithms in real-time applications further improvements in the efficiency are required. There are many important issues and methods in linear algebra that can be used to improve the efficiency of the proposed algorithms. For example, the structure of the problem can be heavily exploited. We may assume in a wireless network that, the problem variables (e.g., power and SINRs) are not fully coupled due to exponential path losses. Such assumptions can result desirable sparsity patterns in the involved problem data (e.g., sparse matrices), which in turn can greatly affect the efficiency of the proposed methods.

In Chapter 4, the WSRMax problem in OFDMA downlink systems was considered and a low-complexity suboptimal algorithm was developed. The proposed algorithm was based on primal decomposition techniques and it jointly optimizes the subcarrier and power allocation in the OFDMA downlink. The algorithm is fast compared to the optimal exhaustive search based algorithms, and numerical results showed that its performance results in close to optimal solution, e.g., more than 90% of the time. It was interesting to see numerically that the performance has been further increased by carrying out parallel instances of the algorithm (with a different initialization point for each one). Numerical results showed that the convergence of the proposed algorithm to a suboptimal solution was possible within a small number of iterations, independent of the number of subcarriers or users.

A distributed algorithm for WSRMax in the context of multiantenna multicell

downlink systems was proposed in Chapter 5. The proposed method was based on classical primal decomposition techniques and subgradient methods, where the main problem was split into many subproblems (one for each BS) and a master problem. A GP and a SOCP based ascent method was developed to tackle the subproblems, where the BSs' own decision variables, such as beamformers' directions and power allocation, are locally optimized in a fully distributed fashion. The BS optimizations, or the subproblems, are coordinated to find an approximate solution for the master problem, where the out-of-cell interference levels are resolved. In the case of the master problem, a sequential convex approximation strategy, together with a subgradient method, have been adopted. One obvious limitation of the proposed method is the use of single antenna receivers instead of multiantennas, which can also be considered as multiantenna receivers with fixed linear receivers. However, such limitations are desirable, especially when producing low-complexity receivers.

Numerical examples showed that the proposed algorithm could significantly improve the overall system objective value with a small amount of BS coordination. Comparisons were made numerically between the proposed distributed algorithm and the optimal BB algorithms. Simulation indicated that the proposed algorithm could perform substantially close to the optimal BB based algorithm. Finally, the average behavior of the proposed algorithm was numerically compared with that of a centralized algorithm. Results showed that, more than 94% of the average objective of the centralized algorithm could be achieved within five BS coordinations. These observations are indeed favorable in the context of large-scale practical communication systems; BS coordination can be carried out with a small amount of backhaul message exchanges to determine distributively the rate/power allocation of each BS before any data transmission phase.

6.2   Future work

Indeed, many problems of recent interest in wireless communication networks can be posed in the framework of mathematical optimization. One important direction of interest is radio resource management strategies for new cellular systems with very small cells, such as femtos. To leverage the successful deployment of such future cellular systems, efficient resource management solutions are increasingly important. For example, new sophisticated protocols are to be investigated

in order to enable the coordination between the entities macrocell-macrocell, macrocell-femtocell, and femtocell-femtocell such that overall system performance is improved. Due to the the emergence of large scale wireless networks, there can be hundreds or thousands of femtocells deployed in future cellular systems. Therefore, the scalability of the signalling protocols are central from a systems perspective so that they can be implemented without explosion. In other words, the protocols should be light or thin to reduce signalling overheads. Application of mathematical optimization tools as a basis for such protocol designs is still to be explored, e.g., classical optimization approaches such as mathematical decomposition methods (pimal/dual decomposition), subgradient methods, and more refined variants (with superior convergence properties), such as consensus optimization via the alternating direction method of multipliers.

Another important direction yet to be investigated is the deployment of self-interference cancellation electronics in future wireless networks. It seems that the trend in certain future cellular systems is to enhance the system coverage by introducing very small femtocells. Therefore, the difference between transmit and received signal power levels of the nodes inside femtocells are not necessarily high. These observations suggest that a moderate level of self-interference cancellation carried out at a node can scale down the self-interference power to an acceptable level so that its reception is uninterrupted. One other example in which such self-interference cancellation mechanisms can potentially be useful is with relay nodes. To facilitate such simultaneous transmissions and receptions sophisticated transceivers that support realtime optimization and signal processing capabilities are needed of course.

There are still many unsolved, but interesting problems in the context of MIMO wireless networks. For example, the joint optimization of beamforming patterns, transmit powers, and link activations to maximize a weighted sum of achievable link rates in MIMO networks is still an open problem. Of course, there are many local optimization methods, even though the optimality is not guaranteed. But the problem is not yet solved globally. Therefore, it is very important to seek global optimization methods, which act as performance benchmarks for any local algorithm and provide a way of evaluating the performance loss encountered by those local methods. Usually, not all the problems we encounter in wireless communication networks are convex and exponentially complex (worst case) global optimization methods are needed to

185

find the solution. However, in practice, such global approaches are, of course, not useful; fast computation of at least a suboptimal solution is highly desirable. Therefore, seeking algorithms based on fast local optimization methods for nonconvex problems are increasingly important. Sequential convex programming, alternating convex optimization, and convex-concave procedure are some local optimization methods that can be adopted for nonconvex optimization problems. Distributed resource management, in the context of general MIMO wireless networks, is another important and challenging problem, which has not been investigated in literature. Due to the explosion of the size of wireless networks and the network nodes' limited knowledge of the overall problem data, distributed resource management strategies are increasingly important from both practical and theoretical perspectives.

References

1. Dahlman E, Parkvall S, Sköld J & Beming P (2007) 3G Evolution HSPA and LTE for Mobile Broadband. Academic Press is an imprint of Elsevier, San Diego, CA.
2. Lu WW (2002) Broadband Wireless Mobile: 3G and Beyond. John Wiley & Sons Ltd, West Sussex, England.
3. Boyd S & Vandenberghe L (2004) Convex Optimization. Cambridge University Press, Cambridge, UK.
4. Boyd S, Parikh N, Chu E, Peleato B & Eckstein J (2010) Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning 3(1): 1–122.
5. Boyd S, Kim SJ, Vandenberghe L & Hassibi A (2007) A tutorial on geometric programming. Optimization and Engineering 8(1): 67–127.
6. Boyd S (2007) Convex optimization I. [Online]. Available: `http://www.stanford.edu/class/ee364a/`.
7. Boyd S (2007) Convex optimization II. [Online]. Available: `http://www.stanford.edu/class/ee364b/`.
8. Boyd S & Vandenberghe L (2010) Additional exercises for convex optimization. [Online]. Available: `http://www.stanford.edu/~boyd/cvxbook/bv_cvxbook_extra_exercises.pdf`.
9. Tse D & Viswanath P (2005) Fundamentals of Wireless Communication. Cambridge University Press, Cambridge, UK.
10. Goldsmith A (2005) Wireless Communications. Cambridge University Press, New York, USA.
11. Starr T, Sorbara M, Cioffi J & Silverman PJ (2003) DSL Advances. Prentice-Hall, Upper Saddle River, NJ 07458.
12. Tassiulas L & Ephremides A (1992) Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. IEEE Trans. Automat. Contr. 37(12): 1936–1949.
13. Yu W (2002) Competition and cooperation in multi-user communication environments. Ph.D. thesis, Department of Electrical Engineering, Stanford University, CA, USA. [Online]. Available: `http://www.comm.utoronto.ca/~weiyu/thesis_2.pdf`.
14. Palomar DP (2003) A unified framework for communications through MIMO channels. Ph.D. thesis, Department of Signal Theory and Communications, Technical University of Catalonia, Barcelona, Spain.
15. Codreanu M (2007) Multidimensional Adaptive Radio Links for Broadband Communications. Ph.D. thesis, Centre for Wireless Communications, University of Oulu. Acta Universitatis Ouluensis, Oulu, Finland. [Online]. Available: `http://herkules.oulu.fi/isbn9789514286223`.
16. Neely MJ (2003) Dynamic power allocation and routing for satellite and wireless networks with time varying channels. Ph.D. thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA.

17. Georgiadis L, Neely MJ & Tassiulas L (2006) Resource allocation and cross-layer control in wireless networks. Found. Trends Net. 1(1): 1–144.

18. Neely MJ (2010) Stochastic Network Optimization with Application to Communication and Queueing Systems, volume 7 of *Synthesis Lectures on Communication Networks*. Morgan & Claypool, San Rafael, CA.

19. Palomar DP & Chiang M (2006) A tutorial on decomposition methods for network utility maximization. IEEE J. Select. Areas Commun. 24(8): 1439–1451.

20. Hoo LMC, Halder B, Tellado J & Cioffi M (2004) Multi user transmit optimization for multiuser broadcast channels: Asymptotic FDMA capacity region and algorithms. IEEE Trans. Commun. 52(6): 922–930.

21. Seong K, Mohseni M & Cioffi M (2006) Optimal resource allocation for OFDMA downlink systems. In: Proc. IEEE Int. Symp. Inform. Theory, pp. 1394–1398. Seattle, USA.

22. Luo ZQ & Yu W (2006) An introduction to convex optimization for communications and signal processing. IEEE J. Select. Areas Commun. 24(8): 1426–1438.

23. Luo Z & Zhang S (2008) Dynamic spectrum management: Complexity and duality. IEEE J. Select. Areas Commun. 2(1): 57–73.

24. Horst R, Pardolos P & Thoai N (2000) Introdiction to Global Optimization, volume 48. Kluwer Academic Publishers, Dordrecht, Boston, London, second edition.

25. Audet C, Hansen P & Savard G (2005) Essays and Surveys in Global Optimization. Springer Science + Business Media, Inc, 233 Spring street, NY 10013, USA.

26. Nelson R & Kleinrock L (1985) Spatial TDMA: Collision-free multihop channel access protocol. IEEE Trans. Commun. 33(9): 934–944.

27. Kelly FP, Maulloo A & Tan D (1998) Rate control for communication networks: Shadow prices, proportional fairness, and stability. J. Op. Research Soc. 49: 237–252.

28. Kelly FP (1997) Charging and rate control for elastic traffic. European Trans. Telecommun. 8: 33–37.

29. Chiang M, Low SH, Calderbank AR & Doyle JC (2006) Layering as optimization decomposition: Framework and examples. In: Proc. IEEE Inform. Theory Workshop, pp. 52–56. Punta del Este, Uruguay.

30. Chiang M, Low SH, Calderbank AR & Doyle JC (2006) Layering as optimization decomposition: Current status and open issues. In: Proc. IEEE Int. Conf. Inf. Sci. and Sys., pp. 355–362.

31. Chiang M, Low SH, Calderbank AR & Doyle JC (2006) Layering as optimization decomposition: Questions and answers. In: Proc. IEEE Military Commun. Conf., pp. 1–10. Washington, DC, USA.

32. Stolyar AL (2005) Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm. Que. Sys. 50(4): 401–457.

33. Chiang M & Bell J (2004) Balancing supply and demand of bandwidth in wireless cellular networks: Utility maximization over powers and rates. In: Proc. IEEE INFOCOM, volume 4, pp. 2800–2811. Hong Kong.

34. Chiang M (2005) Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control. IEEE J. Select. Areas Commun. 23(1): 104–116.

35. Johansson B, Soldati P & Johansson M (2006) Mathematical decomposition techniques for distributed cross-layer optimization of data networks. IEEE J. Select. Areas Commun. 24(8): 1535–1547.

36. Palomar DP & Chiang M (2007) Alternative distributed algorithms for network utility maximization: Framework and applications. IEEE Trans. Automat. Contr. 52(12): 2254–2269.

37. Chiang M, Low SH, Calderbank AR & Doyle JC (2007) Layering as optimization decomposition: A mathematical theory of network architectures. Proceedings of the IEEE 95(1): 255–312.

38. Lin X & Shroff NB (2004) Joint rate control and scheduling in multihop wireless networks. Tech. rep., Purdue University. [Online]. Available: `http://cobweb.ecn.purdue.edu/~linx/papers.html`.

39. Lin X & Shroff NB (2004) Joint rate control and scheduling in multihop wireless networks. In: Proc. IEEE Int. Conf. Dec. and Cont., volume 5, pp. 1484–1489. Atlantis, Paradise Island, Bahamas.

40. Lin X & Shroff NB (2005) The impact of imperfect scheduling on cross-layer rate control in wireless networks. In: Proc. IEEE INFOCOM, volume 3, pp. 1804–1814. Miami, USA.

41. Neely MJ, Modiano E & Rohrs CE (2005) Dynamic power allocation and routing for time varying wireless networks. IEEE J. Select. Areas Commun. 23(1): 89–103.

42. Neely MJ (2006) Super-fast delay tradeoffs for utility optimal fair scheduling in wireless networks. IEEE J. Select. Areas Commun. 24(8): 1–12.

43. Lin X, Shroff NB & Srikant R (2006) A tutorial on cross-layer optimization in wireless networks. IEEE J. Select. Areas Commun. 24(8): 1452–1463.

44. Neely MJ, Modiano E & Li C (2008) Fairness and optimal stochastic control for heterogeneous networks. IEEE/ACM Trans. Networking 16(2): 396–409.

45. Yi Y & Chiang M (2008) Stochastic network utility maximization: A tribute to Kelly's paper published in this journal a decade ago. European Trans. Telecommun. 19(4): 421–442.

46. Eryilmaz A & Srikant R (2006) Joint congestion control, routing and MAC for stability and fairness in wireless networks. IEEE J. Select. Areas Commun. 24(8): 1514–1524.

47. Eryilmaz A & Srikant R (2007) Fair resource allocation in wireless networks using queue-length-based scheduling and congestion control. IEEE/ACM Trans. Networking 15(6): 1333–1344.

48. Pantelidou A & Ephremides A (2009) A cross-layer view of wireless multicasting under uncertainty. In: Proc. IEEE Inform. Theory Workshop, pp. 110–114. Volos, Greece.

49. Tassiulas L & Ephremides A (1993) Dynamic server allocation to parallel queues with randomly varying connectivity. IEEE Trans. Inform. Theory 39(2): 466–478.

50. Eryilmaz A, Srikant R & Perkins JR (2005) Stable scheduling policies for fading wireless channels. IEEE/ACM Trans. Networking 13(2): 411–424.

51. Wu X & Srikant R (2005) Regulated maximal matching: A distributed scheduling algorithm for multi-hop wireless networkswith node-exclusive spectrum sharing. In: Proc. IEEE Conf. on Decision and Cont., and the Europ. Cont. Conf., pp. 5342–5347. Seville, Spain.

52. Neely MJ, Modiano E & Rohrs CE (2003) Power allocation and routing in multibeam satellites with time-varying channels. IEEE/ACM Trans. Networking 11(1): 138–152.

53. Sharma G, Mazumdar RR & Shroff NB (2006) On the complexity of scheduling in wireless networks. In: Proc. ACM int. conf. on Mobile comp. and net., pp. 227–238. Los Angeles, CA, USA.

54. Papadimitriou CH & Steiglitz K (1982) Combinatorial Optimization: Algorithms and Complexity. Prentice-Hall, Englewood Cliffs New Jersey.

55. Cendrillon R, Yu W, Moonen M, Verlinden J & Bostoen T (2006) Optimal multiuser spectrum balancing for digital subscriber lines. IEEE Trans. Commun. 54(5): 922–933.

56. Xu Y, Panigrahi S & Le-Ngoc T (2006) A concave minimization approach to dynamic spectrum management for digital subscriber lines. In: Proc. IEEE Int. Conf. Commun., pp. 84–89. Istanbul, Turkey.

57. Xu Y, Le-Ngoc T & Panigrahi S (2008) Global concave minimization for optimal spectrum balancing in multi-user DSL networks. IEEE Trans. Signal Processing 56(7): 2875–2885.

58. Eriksson K (2010). Dynamic resource allocation in wireless networks. Examensarbete utfört i Kommunikationssystem vid Tekniska högskolan i Linköping. [Online]. Available: `http://liu.diva-portal.org/smash/get/diva2:326290/FULLTEXT01`.

59. Al-Shatri H & Weber T (2010) Optimizing power allocation in interference channels using D.C. programming. In: Proc. Workshop on Resource Alloc. in Wireless Net., pp. 367–373. Avignon, France.

60. Tsiaflakis P, Vangorp J, Moonen M & Verlinden J (2007) A low complexity optimal spectrum balancing algorithm for digital subscriber lines. Els. Sig. Processing 87(7).

61. Tsiaflakis P, Tan W, Yi Y, Chiang M & Moonen M (2008) Optimality certificate of dynamic spectrum management in multi-carrier interference channels. In: Proc. IEEE Int. Symp. Inform. Theory, pp. 1298–1302. Toronto, Canada.

62. Qian L, Zhang YJA & Huang J (2009) MAPEL: Achieving global optimality for a non-convex wireless power control problem. IEEE Trans. Wireless Commun. 8(3): 1553–1563.

63. Cendrillon R, Huang J, Chiang M & Moonen M (2007) Autonomous spectrum balancing for digital subscriber lines. IEEE Trans. Signal Processing 55(8): 4241–4257.

64. Tsiaflakis P, Diehl M & Moonen M (2008) Distributed spectrum management algorithms for multiuser DSL networks. IEEE Trans. Signal Processing 56(10): 4825–4843.

65. Lui R & Yu W (2005) Low-complexity near-optimal spectrum balancing for digital subscriber lines. In: Proc. IEEE Int. Conf. Commun., pp. 1947–1951. Seoul, Korea.

66. Papandriopoulos J & Evans S (2006) Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks. In: Proc. IEEE Int. Conf. Commun., pp. 3270–3275. Istanbul, Turkey.

67. Li G & Liu H (2005) On the optimality of the OFDMA networks. IEEE Commun.

Lett. 9(5): 438–440.

68. Jang J & Lee KB (2003) Transmit power adaptation for multiuser OFDM systems. IEEE J. Select. Areas Commun. 21(2): 171–178.

69. Yu W & Cioffi M (2002) FDMA capacity of gaussian multiple-access channels with ISI. IEEE Trans. Commun. 50(1): 102–111.

70. Bae C & Cho D (2007) Fairness-aware adaptive resource allocation scheme in multihop OFDMA systems. IEEE Commun. Lett. 11(2): 134–136.

71. Shen Z, Andrews JG & Evans BL (2005) Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints. IEEE Trans. Wireless Commun. 4(6): 2726–2737.

72. Codreanu M, Tölli A, Juntti M & Latva-aho M (2007) Joint design of Tx-Rx beamformers in MIMO downlink channel. IEEE Trans. Signal Processing 55(9): 4639–4655.

73. Tölli A, Codreanu M & Juntti M (2008) Cooperative MIMO-OFDM cellular system with soft handover between distributed base station antennas. IEEE Trans. Wireless Commun. 7(4): 1428–1440.

74. Shi S, Schubert M & Boche H (2007) Weighted sum-rate optimization for multiuser MIMO systems. In: Proc. Conf. Inform. Sciences Syst. (CISS), pp. 425–430. Baltimore, MD, USA.

75. Agarwal R & Cioffi JM (2008) Beamforming design for the MIMO downlink for maximizing weighted sum-rate. In: Proc. IEEE Int. Symp. Inform. Theory and its Applications, pp. 1–6. Auckland, New Zealand.

76. Zhang L, Xin Y, Liang YC & Poor HV (2009) Cognitive multiple access channels: optimal power allocation for weighted sum rate maximization. IEEE Trans. Commun. 57(9): 2754–2762.

77. Guthy C, Utschick W, Hunger R & Joham M (2010) Efficient weighted sum rate maximization with linear precoding. IEEE Trans. Signal Processing 58(4): 2284–2297.

78. Christensen SS, Agarwal R, Carvalho E & Cioffi J (2008) Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design. IEEE Trans. Wireless Commun. 7(12): 4792–4799.

79. Vucic N, Shi S & Schubert M (2010) DC programming approach for resource allocation in wireless networks. In: Proc. Workshop on Resource Alloc. in Wireless Net., pp. 360–366. Avignon, France.

80. Stojnic M, Vikalo H & Hassibi B (2006) Rate maximization in multi-antenna broadcast channels with linear preprocessing. IEEE Trans. Wireless Commun. 5(9): 2338 –2342.

81. Hanly SV & Tse DNC (1998) Multiaccess fading channels–Part II: Delay-limited capacities. IEEE Trans. Inform. Theory 44(7): 2816–2831.

82. Boyd S (2007) Branch-and-bound methods. [Online]. Available: `http://www.stanford.edu/class/ee364b/lectures/bb_slides.pdf`.

83. Weeraddana PC, Codreanu M, Latva-aho M & Ephremides A (2011) Weighted sum-rate maximization for a set of interfering links via branch and bound. IEEE Trans. Signal Processing 59(8): 3977–3996.

84. Joshi S, Weeraddana PC, Codreanu M & Latva-aho M (2011) Weighted sum-rate maximization for MISO downlink cellular networks via branch and bound. IEEE

Trans. Signal Processing, submitted .

85. Weeraddana PC, Codreanu M, Latva-aho M & Ephremides A (2011) Optimal maxweight scheduling in a multihop wireless network via branch and bound. In: Proc. IEEE Int. Symp. Inform. Theory, pp. 2787–2791. Saint Petersburg, Russia.

86. Weeraddana PC, Codreanu M, Latva-aho M & Ephremides A (2010) Weighted sum-rate maximization for a set of interfering links via branch and bound. In: Proc. Annual Asilomar Conf. Signals, Syst., Comp. Pacific Grove, CA, USA.

87. Codreanu M, Weeraddana PC, Latva-Aho M & Ephremides A (2011) Weighted sum-rate maximization in singlecast and multicast wireless networks - global optimum via branch and bound. In: Proc. IEEE Int. Symp. Pers., Indoor, Mobile Radio Commun., to appear. Toronto, Canada.

88. Joshi S, Weeraddana PC, Codreanu M & Latva-Aho M (2011) Weighted sum-rate maximization for MISO downlink cellular networks via branch and bound. In: Proc. Annual Asilomar Conf. Signals, Syst., Comp., to appear. Pacific Grove, CA, USA. [Online]. Available: `http://www.ee.oulu.fi/~chathu/Asilomar_Joshi_2011.pdf`.

89. Phuong NTH & Tuy H (2003) A unified monotonic approach to generalized linear fractional programming pp. 229–259. Kluwer Academic Publishers.

90. Win MZ & Scholtz RA (2000) Ultra-wide bandwidth time-hopping spread-spectrum impulse radio for wireless multiple-access communications. IEEE Trans. Commun. 48(4): 679–691.

91. Radunovic B & Boudec JL (2004) Optimal power control, scheduling, and routing in UWB networks. IEEE J. Select. Areas Commun. 22(7): 1252–1270.

92. Souilmi Y, Knopp R & Caire G (2003) Coding strategies for UWB interference-limited peer-to-peer networks. In: Proc. Int. Symp. on Modelling and Opt. in Mobile, Ad-hoc and Wireless Networks. INRIA Sophia-Antipolis, France.

93. Radunovic B & Boudec JL (2005) Power control is not required for wireless networks in the linear regime. In: Proc. IEEE Int. Symp. World of Wireless, Mobile and Mult. Networks, volume 5, pp. 417–427. Taormina, Giardini Naxos, Italy.

94. Cruz RL & Santhanam AV (2003) Optimal routing, link scheduling and power control in multi-hop wireless networks. In: Proc. IEEE INFOCOM, volume 1, pp. 702–711. San Diego, CA.

95. Oneill DC, Julian D & Boyd S (2004) Optimal routes and flows in congestion constrained ad-hoc networks. In: Proc. IEEE Veh. Technol. Conf., pp. 702–711. Los Angeles, CA.

96. Julian D, Chiang M, OŠNeil D, & Boyd S (2002) Qos and fairness constrained convex optimization of resource allocation for wireless cellular and ad-hoc networks. In: Proc. IEEE INFOCOM, volume 2, pp. 477–486. New York, USA.

97. Avriel M & Williams AC (1970) Complementary geometric programming. SIAM J. Appl. Math. 19(1): 125–141.

98. Chiang M (2005) Geometric programming for communication systems 2(1-2): 1–154.

99. Chiang M, Tan CW, Palomar DP, O'Neill D & Julian D (2007) Power control by geometric programming. IEEE Trans. Wireless Commun. 6(7): 2640–2651.

100. Matskani E, Sidiropoulos ND & Tassiulas L (2011) Convex approximation

algorithms for back-pressure power control of wireless multi-hop networks. In: Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, pp. 3032–3035. Prague, Czech Republic.

101. Weeraddana PC, Codreanu M & Latva-aho M (2009) Cross-layer resource allocation for wireless networks via signomial programming. In: Proc. IEEE Global Telecommun. Conf., pp. 1–6. Honolulu, Hawaii, USA.

102. Codreanu M, Weeraddana PC & Latva-aho M (2009) Cross-layer utility maximization subject to stability constraints for multi-channel wireless networks. In: Proc. Annual Asilomar Conf. Signals, Syst., Comp., pp. 776–780. Pacific Grove, CA, USA.

103. Weeraddana PC, Codreanu M & Latva-aho M (2009) On the advantages of using multiuser receivers in wireless ad-hoc networks. In: Proc. IEEE Veh. Technol. Conf., pp. 1–6. Anchorage, Alaska, USA.

104. Weeraddana PC, Codreanu M, Latva-aho M & Ephremides A (2010) On the effect of self-interference cancelation in multihop wireless networks. EURASIP J. Wireless Comm. and Netw. 2010, Article ID 513952, 10 pages. Doi:10.1155/2010/513952.

105. Weeraddana PC, Codreanu M, Latva-aho M & Ephremides A (2010) The benefits from simultaneous transmission and reception in wireless networks. In: Proc. IEEE Inform. Theory Workshop, pp. 1–5. Dublin, Ireland.

106. Hajek B & Sasaki G (1988) Link scheduling in polynomial time. IEEE Trans. Inform. Theory 34(5): 910–917.

107. Borbash SA & Ephremides A (2006) The feasibility of matchings in a wireless network. IEEE Trans. Inform. Theory 52(6): 2749 – 2755.

108. ElBatt T & Ephremides A (2004) Joint scheduling and power control for wireless ad hoc networks. IEEE Trans. Wireless Commun. 3(1): 74 – 85.

109. Borbash SA & Ephremides A (2006) Wireless link scheduling with power control and SINR constraints. IEEE Trans. Inform. Theory 52(11): 5106–5111.

110. Bui L, Eryilmaz A, Srikant R & Wu X (2008) Asynchronous congestion control in multi-hop wireless networks with maximal matching-based scheduling. IEEE/ACM Trans. Networking 16(4): 826–839.

111. Kim TS, Yang Y, Hou JC & Krishnamurthy SV (2009) Joint resource allocation and admission control in wireless mesh networks. In: Proc. Int. Symp. on Modelling and Opt. in Mobile, Ad-hoc and Wireless Networks, pp. 1–10. Seoul, Korea.

112. Weeraddana PC, Codreanu M, Latva-aho M & Ephremides A (2011) Resource allocation for cross-layer utility maximization in wireless networks. IEEE Trans. Veh. Technol. 60(6): 2790–2809.

113. Weeraddana PC, Codreanu M, Latva-aho M & Ephremides A (2010) Resource allocation for cross-layer utility maximization in multi-hop wireless networks in the presence of self interference. In: Proc. Int. Symp. on Modelling and Opt. in Mobile, Ad-hoc and Wireless Networks, pp. 70–75. Avignon, France.

114. Allgower EL & Georg K (2003) Introduction to Numerical Continuation Methods. SIAM, Philadelphia, PA.

115. Sesia S, Toufik I & Baker M (2009) LTE - the UMTS long term evolution: From Theory to Practice. A John Wiley & Sons publication, UK.

116. Ali SH, Lee K & Leung VCM (2007) Dynamic resource allocation in OFDMA

wireless metropolitan area networks [Radio Resource Management and Protocol Engineering for IEEE 802.16]. IEEE Trans. Wireless Commun. 14(1): 6–13.

117. Kwon T, Lee H, Choi S, Kim J, Cho D, Cho S, Yun S, Park W & Kim K (2005) Design and implementation of a simulator based on a cross-layer protocol between MAC and PHY layers in a WiBro compatible.IEEE 802.16e OFDMA system. IEEE Commun. Mag. 43(12): 136–146.

118. Huang C, Juan H, Lin M & Chang C (2007) Radio resource management of heterogeneous services in mobile WiMAX systems [Radio Resource Management and Protocol Engineering for IEEE 802.16]. IEEE Trans. Wireless Commun. 14(1): 20–26.

119. Rhee W & Cioffi M (2000) Increase in capacity of multiuser OFDM system using dynamic subchannel allocation. In: Proc. IEEE Veh. Technol. Conf., volume 2, pp. 1085–1089. Tokyo, Japan.

120. Mao Z & Wang X (2006) Branch-and-bound approach to OFDMA radio resource allocation. In: Proc. IEEE Veh. Technol. Conf., pp. 1–5. Montréal, Canada.

121. Wong CY, Cheng RS, Letaief KB & Murch RD (1999) Multiuser OFDM with adaptive subcarrier, bit, and power allocation. IEEE J. Select. Areas Commun. 17(10): 1747–1758.

122. Zhang G (2004) Subcarrier and bit allocation for real-time services in multiuser OFDM systems. In: Proc. IEEE Int. Conf. Commun., volume 5, pp. 2985– 298. Paris, France.

123. Kivanc D, Li G & Liu H (2003) Computationally efficient bandwidth allocation and power control for OFDMA. IEEE Trans. Wireless Commun. 2(6): 1150Ũ1158.

124. Ho WWL & Liang YC (2009) Optimal resource allocation for multiuser MIMO-OFDM systems with user rate constraints. IEEE Trans. Veh. Technol. 58(3): 1190–1203.

125. Ergen M, Coleri S & Varaiya P (2003) QoS aware adaptive resource allocation techniques for fair schesuling in OFDMA based broadband wireless access systems. IEEE Trans. Broadcast. 49(4): 362–370.

126. Li G & Liu H (2006) Downlink radio resource allocation for multi-cell OFDMA system. IEEE Trans. Wireless Commun. 5(12): 3451–3459.

127. Li G & Liu H (2006) Resource allocation for OFDMA relay networks with fairness constraints. IEEE J. Select. Areas Commun. 24(11): 2061–2069.

128. Kim I, Park I & Lee YH (2006) Use of linear programming for dynamic subcarrier and bit allocation in multiuser OFDM. IEEE Trans. Veh. Technol. 55(4): 1195–1207.

129. Weeraddana PC, Codreanu M, Wei L & Latva-aho M (2010) Primal decomposition-based method for weighted sum-rate maximization in downlink OFDMA systems. EURASIP J. Wireless Comm. and Netw. 2010, Article ID 324780, 9 pages. Doi:10.1155/2010/324780.

130. Weeraddana PC, Wei L, Codreanu M & Latva-aho M (2008) Weighted sum-rate maximization for downlink OFDMA systems. In: Proc. Annual Asilomar Conf. Signals, Syst., Comp., pp. 990–994. Pacific Grove, CA, USA.

131. Weeraddana PC, Wei L, Codreanu M & Latva-aho M (2008) Adaptive subcarrier and power allocation for OFDMA systems. In: Proc. of the IFIP Wireless Days Conf., pp. 1–5. Dubai, UAE.

132. Weeraddana PC, Codreanu M & Latva-aho M (2010) An efficient close to optimal radio resource allocation mechanism towards LTE downlink transmission. In: Proc. URSI/IEEE Convention Radio Science, pp. 71–74. Oulu, Finland.

133. Weeraddana PC, Codreanu M, Wei L & Latva-aho M (2008) Low complexity adaptive subcarrier and power allocation scheme for downlink OFDMA systems. In: Proc. Int. Symp. Wireless Pers. Multimedia Commun. Lapland, Finland. CD-Rom.

134. Nedić A & Ozdaglar A (2011) Cooperative distributed multi-agent optimization. Convex Optimization in Signal Processing and Communications, (D. P. Palomar and Y. C. Eldar, eds.) , Cambridge University Press. pp. 340–386.

135. Yang B & Johansson M (2011) Distributed optimization and games: A tutorial overview. Networked Control Systems, Lecture Notes in Control and Information Sciences, Springer-Verlag London. 406: 109–148.

136. Papandriopoulos J & Evans JS (2009) SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks. IEEE Trans. Inform. Theory 55(8): 3711–3724.

137. Park SH, Park H & Lee I (2010) Distributed beamforming techniques for weighted sum-rate maximization in MISO interference channels. IEEE Commun. Lett. 14(12): 1131–1133.

138. Björnson E, Bengtsson M & Ottersten B (2010) Optimality properties and low-complexity solutions to coordinated multicell transmission. In: Proc. IEEE Global Telecommun. Conf., pp. 1–6. Miami, Florida, USA.

139. Björnson E, Jaldén N, Bengtsson M & Ottersten B (2010) Optimality properties, distributed strategies, and measurement-based evaluation of coordinated multicell OFDMA transmission. IEEE Trans. Signal Processing, submitted .

140. Venturino L, Prasad N & Wang X (2010) Coordinated linear beamforming in downlink multi-cell wireless networks. IEEE Trans. Wireless Commun. 9(4): 1451–1461.

141. Liu YF, Dai YH & Luo ZQ (2011) Coordinated beamforming for MISO interference channel: Complexity analysis and efficient algorithms. IEEE Trans. Signal Processing 59(3): 1142–1157.

142. Weeraddana PC, Codreanu M, Latva-aho M & Ephremides A (2011) Multicell downlink weighted sum-rate maximization: A distributed approach. IEEE Trans. Signal Processing, submitted .

143. Weeraddana PC, Codreanu M & Latva-aho M (2011) Multicell downlink weighted sum-rate maximization: A distributed approach. In: Proc. Annual Asilomar Conf. Signals, Syst., Comp., to appear. Pacific Grove, CA, USA.

144. Lee BO, Je HW, Sohn I, Shin OS & Lee KB (2008) Interference-aware decentralized precoding for multicell MIMO TDD systems. In: Proc. IEEE Global Telecommun. Conf., pp. 1–5. New Orleans, LA, USA.

145. Zhang R & Cui S (2010) Cooperative interference management with MISO beamforming. IEEE Trans. Signal Processing 58(10): 5450–5458.

146. Larsson EG & Jorswieck EA (2008) Competition versus cooperation on the MISO interference channel. IEEE J. Select. Areas Commun. 26(7): 1059–1069.

147. Jorswieck EA & Larsson EG (2008) The MISO interference channel from a game-theoretic perspective: A combination of selfishness and altruism achieves

pareto optimality. In: Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, pp. 5364–5367. Las Vegas, Nevada, USA.

148. Shi C, Berry RA & Honig ML (2008) Distributed interference pricing with MISO channels. In: Proc. Annual Allerton Conf. Commun., Cont., Computing, pp. 539–546. Urbana-Champaign, USA.

149. Shi C, Berry RA & Honig ML (2009) Monotonic convergence of distributed interference pricing in wireless networks. In: Proc. IEEE Int. Symp. Inform. Theory, pp. 1619–1623. Seoul, Korea.

150. Shi C, Berry RA & Honig ML (2009) Local interference pricing for distributed beamforming in MIMO networks. In: Proc. IEEE Military Commun. Conf., pp. 1–6. Boston, MA.

151. Ho ZKM & Gesbert D (2010) Balancing egoism and altruism on the MIMO interference channel. IEEE J. Select. Areas Commun., submitted [Online]. Available: `http://arxiv.org/PS_cache/arxiv/pdf/0910/0910.1688v1.pdf`.

152. Björnson E, Zakhour R, Gesbert D & Ottersten B (2010) Cooperative multicell precoding: Rate region characterization and distributed strategies with instantaneous and statistical CSI. IEEE Trans. Signal Processing 58(8): 4298–4310.

153. Zakhour R & Gesbert D (2009) Coordination on the MISO interference channel using the virtual SINR framework. In: Proc. ITG Workshop Smart Antennas. Berlin, Germany.

154. Zakhour R & Gesbert D (2010) Distributed multicell-MISO precoding using the layered virtual SINR framework. IEEE Trans. Wireless Commun. 9(8): 2444–2448.

155. Zakhour R, Ho Z & Gesbert D (2009) Distributed beamforming coordination in multicell MIMO channels. In: Proc. IEEE Veh. Technol. Conf., pp. 1–5. Barcelona, Spain.

156. Tölli A, Pennanen H & Komulainen P (2011) Decentralized minimum power multi-cell beamforming with limited backhaul signaling. IEEE Trans. Wireless Commun. 10(2): 570–580.

157. Song B, Cruz RL & Rao BD (2004) A simple joint beamforming and power control algorithm for multi-user MIMO wireless networks. In: Proc. IEEE Veh. Technol. Conf., pp. 247–251. Los Angeles, CA.

158. Yu HDW (2010) Coordinated beamforming for the multicell multi-antenna wireless system. IEEE Trans. Wireless Commun. 9(5): 1748–1759.

159. Jorswieck EA, Larsson EG & Danev D (2008) Complete characterization of the pareto boundary for the MISO interference channel. IEEE Trans. Signal Processing 56(10): 5292–5296.

160. Matskani E, Sidiropoulos ND, Luo ZQ & Tassiulas L (2008) Convex approximation techniques for joint multiuser downlink beamforming and admission control. IEEE Trans. Wireless Commun. 7(7): 2682–2693.

161. Matskani E, Sidiropoulos ND, Luo ZQ & Tassiulas L (2007) A second-order cone deflation approach to joint multiuser downlink beamforming and admission control. In: Proc. IEEE Works. on Sign. Proc. Adv. in Wireless. Commun., pp. 1–5. Helsinki, Finland.

162. Matskani E, Sidiropoulos ND, Luo ZQ & Tassiulas L (2007) Joint multiuser downlink beamforming and admission control: A semidefinite relaxation approach. In: Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, volume 3, pp.

585–588. Honolulu, Hawaii, USA.

163. Cao M, Wang X, Kim S & Madihian M (2007) Multi-hop wireless backhaul networks: A cross-layer design paradigm. IEEE J. Select. Areas Commun. 25(4): 738–748.

164. Kodialam M & Nandagopal T (2005) Characterizing achievable rates in multi-hop wireless mesh networks with orthogonal channels. IEEE/ACM Trans. Networking 13(4): 868–880.

165. Shi Y & Hou YT (2008) A distributed optimization algorithm for multi-hop cognitive radio networks. In: Proc. IEEE INFOCOM, pp. 1292 –1300. Phoenix, AZ, USA.

166. Cao M, Raghunathan V, Hanly S, Sharma V & Kumar PR (2007) Power control and transmission scheduling for network utility maximization in wireless networks. In: Proc. IEEE Int. Conf. Dec. and Cont., pp. 5215–5221. New Orleans, LA, USA.

167. Barreto DE & Chiu SS (2007) Decomposition methods for cross-layer optimization in wireless networks. In: Proc. IEEE Wireless Commun. and Networking Conf., pp. 270–275. Kawloon, Hong Kong.

168. Suzuki H, Itoh K, Ebin Y & Sato M (1999) A booster configuration with adaptive reduction of transmitter-receiver antenna coupling for pager systems. In: Proc. IEEE Veh. Technol. Conf., volume 3, pp. 1516–1520. Amsterdam, Netherlands.

169. Halperin D, Anderson T & Wetherall D (2008) Taking the sting out of carrier sense: Interference cancellation for wireless LANs. In: Proc. ACM int. conf. on Mobile comp. and net., pp. 339–350. San Francisco, California, USA.

170. Radunović B, Gunawardena D, Proutiere A, Singh N, Balan V & Key P (2009) Efficiency and fairness in distributed wireless networks through self-interference cancellation and scheduling. Tech. Rep. MSR-TR-2009-27, Microsoft Research. [Online]. Available: `http://research.microsoft.com/apps/pubs/default.aspx?id=79933`.

171. Radunović B, Gunawardena D, Key P, Proutiere A, Singh N, Balan V & Dejean G (2009) Rethinking indoor wireless: Low power, low frequency, full-duplex. Tech. Rep. MSR-TR-2009-148, Microsoft Research. [Online]. Available: `http://research.microsoft.com/apps/pubs/default.aspx?id=104950`.

172. Balakrishnan V, Boyd S & Balemi S (1991) Branch and bound algorithm for computing the minimum stability degree of parameter-dependent linear systems. Int. J. Robust and Nonlinear Con. 1(4): 295–317.

173. Song B, Cruz RL & Rao BD (2005) Network duality and its application to multi-user MIMO wireless networks with SINR constraints. In: Proc. IEEE Int. Conf. Commun., volume 4, pp. 2684–2689. Seoul, Korea.

174. Horn R & Johnson C (1994) Topics in Matrix Analysis. Cambridge University Press, Cambridge, UK.

175. Boyd S (2006) GGPLAB: A simple matlab toolbox for geometric programming. [Online]. Available: `http://www.stanford.edu/~boyd/ggplab/`.

176. Kumar A, Manjunath D & Kuri J (2008) Wireless Networking. ELSEVIER Inc., Burlington, MA , USA.

177. Abou-Faycal IC, Trott MD & Shamai S (2001) The capacity of discrete-time memoryless Rayleigh-fading channels. IEEE Trans. Inform. Theory 47(4): 1290 –1301.

178. Weingarten H, Steinberg Y & Shamai S (2006) The capacity region of the Gaussian multiple-input multiple-output broadcast channel. IEEE Trans. Inform. Theory 52(9): 3936 – 3964.

179. Tse DNC & Hanly SV (1998) Multiaccess fading channels. I. polymatroid structure, optimal resource allocation and throughput capacities. IEEE Trans. Inform. Theory 44(7): 2796–2815.

180. Cover TM & Thomas JA (2006) Elements of Information Theory. John Wiley, New York, USA, second edition.

181. Marks BR & Wright GP (1978) A general inner approximation algorithm for nonconvex mathematical programs. Oper. Research 26(4): 681–683.

182. Boyd S (2007) Sequential convex programming. [Online]. Available: `http://www.stanford.edu/class/ee364b/lectures/seq_slides.pdf`.

183. Kojima M, Megiddo N & Noma T (1989) Homotopy continuation methods for nonlinear complementarity problems. Math. of Op. Research 16: 754–774.

184. Nesterov Y & Nemirovsky A (1994) Interior-Point Polynomial Algorithms in Convex Programming. SIAM.

185. Tse DN (1997) Optimal power allocation over parallel gaussian broadcast channels. Unpublished, a short summary published in Proc. of Int. Symp. Inform. Theory, Ulm, Germany .

186. Pantelidou A & Ephremides A (2010) A cross-layer view of optimal scheduling. IEEE Trans. Inform. Theory 56(11): 5568–5580.

187. Radunović B, Gunawardena D, Key P, Proutiere A, Singh N, Balan V & Dejean G (2010) Rethinking indoor wireless mesh design: Low power, low frequency, full-duplex. In: Proc. IEEE Works. on Wireless Mesh Networks, pp. 1–6. Boston, USA.

188. Choi JI, Jain M, Srinivasan K, Levis P & Katti S (2010) Achieving single channel, full duplex wireless communication. In: Proc. ACM int. conf. on Mobile comp. and net., pp. 1–12. Chicago, Illinois, USA.

189. Boyd S (2007) Primal and dual decomposition. [Online]. Available: `http://www.stanford.edu/class/ee364b/lectures/decomposition_slides.pdf`.

190. Rockafellar RT (1970) Convex Analysis. Prinston, N.J., Princeton University Press.

191. Nuaymi L (2007) WiMax Technology for Broadband Wireless Access. John Wiley & Sons publication, UK.

192. Boyd S (2007) Subgradient methods. [Online]. Available: `http://www.stanford.edu/class/ee364b/lectures/subgrad_method_slides.pdf`.

193. Lobo MS, Vandenberghe L, Boyd S & Lebret H (1998) Applications of second-order cone programming. Linear Algebra and Applications 284: 193–228.

194. Gupta P & Kumar PR (2000) The capacity of wireless networks. IEEE Trans. Inform. Theory 46(2).

195. Zander J (1992) Performance of optimum transmitter power control in cellular radio systems. IEEE Trans. Veh. Technol. 41(1): 57–62.

196. Foschini G & ZMiljanic (1993) A simple distributed autonomus power control algorithm and its covergence. IEEE Trans. Veh. Technol. 42(4): 641–646.

197. Yates RD & Huang C (1995) Integrated power control and base station assignment. IEEE Trans. Veh. Technol. 44(3): 638 –644.

198. Hanly SV (1995) An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity. IEEE J. Select. Areas Commun. 13(7): 1332 –1340.

199. Seneta E (2006) Non-Negative Matrices and Markov Chains. Springer Science + Business Media, Inc, 233 Spring street, NY 10013, USA. Revised printing.

200. Yu W, Rhee W, Boyd S & Cioffi J (2004) Iterative water-filling for gaussian vector multiple-access channels. IEEE Trans. Inform. Theory 50(1): 145–152.

201. Jafar SA (2003) Fundamental capacity limits of multiple antenna wireless systems. Ph.d. dissertation, Stanford university, CA, USA.

Appendix 1 : Proof of Theorem 2.2

Theorem 2.2 shows certain similarities to the classical feasibility conditions derived in [195–198]. These conditions were derived based on the Perron-Frobenius theory [199] by assuming the primitiveness of $\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}$. We give a slightly more general proof based on the theory of M-matrices [174, p. 112], which circumvent the technical condition of $\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}$ being primitive. Thus they hold for any nonnegative matrix $\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}$.

To prove the first statement we show that $\rho\left(\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}\right) < 1$ is necessary for $\boldsymbol{\gamma} \in \mathcal{G}$. Recall that (2.42) can be expressed as $\mathbf{A}(\boldsymbol{\gamma})\mathbf{p} \geq \mathbf{b}(\boldsymbol{\gamma})$. Thus, we can write the following necessary (but not sufficient) condition for $\boldsymbol{\gamma} \in \mathcal{G}$:

$$\boldsymbol{\gamma} \in \mathcal{G} \Rightarrow \exists\, \mathbf{p} \geq \mathbf{0} \text{ such that } \mathbf{A}(\boldsymbol{\gamma})\mathbf{p} \geq \mathbf{b}(\boldsymbol{\gamma}) . \tag{1.0.1}$$

The condition above is easily derived by ignoring the second set of inequalities (i.e., the power constraints) in the description of $\mathcal{G}$ in (2.6). Strict positivity of $\boldsymbol{\gamma}$ implies that $\mathbf{b}(\boldsymbol{\gamma}) > \mathbf{0}$ and $\mathbf{p} > \mathbf{0}$. This observation together with the condition (1.0.1) yield the following necessary conditions for $\boldsymbol{\gamma} \in \mathcal{G}$:

$$\boldsymbol{\gamma} \in \mathcal{G} \Rightarrow \exists\, \mathbf{p} > \mathbf{0} \text{ such that } \mathbf{A}(\boldsymbol{\gamma})\mathbf{p} > \mathbf{0} . \tag{1.0.2}$$

Finally, [174, Th. 2.5.3, items 12 and 2] states that $\exists\, \mathbf{p} > \mathbf{0}$ such that $\mathbf{A}(\boldsymbol{\gamma})\mathbf{p} > \mathbf{0}$ if and only if $\rho\left(\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}\right) < 1$. Consequently, we can rewrite (1.0.2) equivalently as $\boldsymbol{\gamma} \in \mathcal{G} \Rightarrow \rho\left(\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}\right) < 1$ which, by the contraposition, is equivalent to $\rho\left(\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}\right) \geq 1 \Rightarrow \boldsymbol{\gamma} \notin \mathcal{G}$.

The second part follows directly from the description of $\mathcal{G}$ in (2.6), where the SINR constraints (2.42) are satisfied with equality, i.e., $\mathbf{A}(\boldsymbol{\gamma})\mathbf{p} = \mathbf{b}(\boldsymbol{\gamma})$. Note that since the nonnegative matrix $\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}$ has the spectral radius smaller than one, i.e., $\rho\left(\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}\right) < 1$, the matrix $\mathbf{A}(\boldsymbol{\gamma}) = \mathbf{I} - \mathbf{B}(\boldsymbol{\gamma})\mathbf{G}$ is invertible and its inverse has nonnegative entries, i.e., $\mathbf{A}^{-1}(\boldsymbol{\gamma}) \geq \mathbf{0}$ [174, Th. 2.5.3, items 2 and 17]. Thus $\mathbf{p} = \mathbf{A}^{-1}(\boldsymbol{\gamma})\mathbf{b}(\boldsymbol{\gamma}) \geq \mathbf{0}$.

We prove the third part by showing that $\mathbf{p}^{\star} = \mathbf{A}^{-1}(\boldsymbol{\gamma})\mathbf{b}(\boldsymbol{\gamma})$ is the minimum power vector [22] (with respect to generalized inequality $\preceq_{\mathbb{R}_+^L}$) which satisfies the

---

[22] A point $\mathbf{p} \in \mathcal{S}$ is the minimum element of set $\mathcal{S}$ with respect to generalized inequality $\preceq_{\mathbb{R}_+^L}$ if and only if $\mathcal{S} \subset \mathbf{p} + \mathbb{R}_+^L$ [3, Sec. 2.4.2].

SINR constraints in (2.42), i.e., $\mathbf{p}^\star$ is the unique solution of the following vector optimization problem: [23]

$$
\begin{aligned}
&\text{minimize (w.r.t. } \mathbb{R}_+^L) \quad \mathbf{p} \\
&\text{subject to} \qquad\qquad \mathbf{A}(\boldsymbol{\gamma})\mathbf{p} \geq \mathbf{b}(\boldsymbol{\gamma}) \ ,
\end{aligned}
\tag{1.0.3}
$$

where the variable is $\mathbf{p}$. Since $\mathbf{p}^\star$ is the minimum power vector that achieves SINR values $\boldsymbol{\gamma}$, if it violates any power constraint then any other power vector $\mathbf{p}$ that achieves $\boldsymbol{\gamma}$ must also violate those power constraints, because $\mathbf{p}^\star \leq \mathbf{p}$.

A standard technique for solving vector optimization problems is scalarization [3, Sec. 4.7.4]. We choose an arbitrary $\boldsymbol{\lambda} > \mathbf{0}$ and solve the following scalar optimization:

$$
\begin{aligned}
&\text{minimize} \quad \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{p} \\
&\text{subject to} \quad \mathbf{A}(\boldsymbol{\gamma})\mathbf{p} \geq \mathbf{b}(\boldsymbol{\gamma}) \ ,
\end{aligned}
\tag{1.0.4}
$$

where the variable is $\mathbf{p}$. Let us make the change of variable $\mathbf{y} = \mathbf{A}(\boldsymbol{\gamma})\mathbf{p}$ and rewrite problem (1.0.4) as

$$
\begin{aligned}
&\text{minimize} \quad \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{A}(\boldsymbol{\gamma})^{-1}\mathbf{y} \\
&\text{subject to} \quad \mathbf{y} \geq \mathbf{b}(\boldsymbol{\gamma}) \ ,
\end{aligned}
\tag{1.0.5}
$$

where the variable is $\mathbf{y}$. Recall that $\mathbf{A}^{-1}(\boldsymbol{\gamma}) \geq \mathbf{0}$ (since $\rho\left(\mathbf{B}(\boldsymbol{\gamma})\mathbf{G}\right) < 1$), and therefore the gradient of the objective has positive entries, i.e., $(\mathbf{A}(\boldsymbol{\gamma})^{-1})^{\mathrm{T}}\boldsymbol{\lambda} \geq \mathbf{0}$. Thus, the optimal solution does not depend on $\boldsymbol{\lambda}$ and it is given by $\mathbf{y}^\star = \mathbf{b}(\boldsymbol{\gamma})$. This, in turn, implies that the optimal solution of problem (1.0.4) [and, implicitly of problem (1.0.3)] is given by $\mathbf{p}^\star = \mathbf{A}^{-1}(\boldsymbol{\gamma})\mathbf{b}(\boldsymbol{\gamma})$.

---

[23] For a more detailed discussion of vector optimization, see [3, Sec. 4.7].

Appendix 2 : Compute γ_{ImpCGP} via complementary geometric programming (CGP)

We show in this appendix how to compute efficiently γ_{ImpCGP} via CGP [97], when $f_0(\gamma) = \sum_{l \in \mathcal{L}} -\beta_l \log(1 + \gamma_l)$. Note that this is the only place where the exact expression of the rate function (2.1) has been explicitly taken into account. In the derivation of all other bounds only the monotonicity property has been used. We start by equivalently reformulating problem (2.60) as

$$
\begin{aligned}
\text{minimize} \quad & \prod_{l \in \mathcal{L}} \left(1 + \gamma_l \right)^{-\beta_l} \\
\text{subject to} \quad & \gamma_{l,\min} \leq \gamma_l \leq \gamma_{l,\max}, \ l \in \mathcal{L} \\
& \gamma_l \leq \frac{g_{ll} p_l}{\sigma^2 + \sum_{j \neq l} g_{jl} p_j}, \ l \in \mathcal{L} \\
& \sum_{l \in \mathcal{O}(n)} p_l \leq p_n^{\max}, \ n \in \mathcal{T} \\
& p_l \geq 0, \ l \in \mathcal{L} \ ,
\end{aligned}
\qquad (2.0.1)
$$

where the variables are $(p_l)_{l \in \mathcal{L}}$ and $(\gamma_l)_{l \in \mathcal{L}}$. The equivalence between problem (2.60) and problem (2.0.1) follows from the monotonically increasing property of $\log(\cdot)$ function and the explicit description of the constraints. To obtain a sub-optimal solution, we adopt Algorithm 3.2 (see Section 3.2.2) in a straightforward manner as follows:

Algorithm 2.0.1. *CGP based algorithm for finding γ_{ImpCGP}*

1. Given tolerance $\varepsilon > 0$. Let $\hat{\gamma} = \bar{\mathbf{a}}_{l^\star}$.
2. Solve the following GP:
$$
\begin{aligned}
\text{minimize} \quad & \prod_{l \in \mathcal{L}} \gamma_l^{-\beta_l \frac{\hat{\gamma}_l}{1+\hat{\gamma}_l}} \\
\text{subject to} \quad & \gamma_{l,\min} \leq \gamma_l \leq \gamma_{l,\max}, \ l \in \mathcal{L} \\
& \gamma_l \leq \frac{g_{ll} p_l}{\sigma^2 + \sum_{j \neq l} g_{jl} p_j}, \ l \in \mathcal{L} \\
& \sum_{l \in \mathcal{O}(n)} p_l \leq p_n^{\max}, \ n \in \mathcal{T} \ ,
\end{aligned}
\qquad (2.0.2)
$$
   with the variables $(p_l)_{l \in \mathcal{L}}$ and $(\gamma_l)_{l \in \mathcal{L}}$. Denote the solution by $(p_l^\star)_{l \in \mathcal{L}}$ and $(\gamma_l^\star)_{l \in \mathcal{L}}$.
3. If $\max_{l \in \mathcal{L}} |\gamma_l^\star - \hat{\gamma}_l| > \varepsilon$ set $(\hat{\gamma}_l = \gamma_l^\star)_{l \in \mathcal{L}}$ and go to step 2. Otherwise set $\gamma_{\mathrm{ImpCGP}} = \hat{\gamma}$ and STOP.

Appendix 3 : A direct monomial approximation

In this appendix we derive a monomial approximation for the objective function of problem (3.11), which results in the same successive approximation steps as in Algorithm 3.2. The derivation is given in [15, Lem. 4.2.2] and outlined here for completeness. Let us first prove the following lemma.

Lemma 3.0.1. *Let $m(\boldsymbol{\gamma}) = d \prod_{c \in \mathcal{C}} \prod_{l \in \mathcal{L}} \gamma_{lc}^{a_{lc}}$ be a monomial function [5] used to approximate the objective function (3.11), i.e., $f(\boldsymbol{\gamma}) = \prod_{c \in \mathcal{C}} \prod_{l \in \mathcal{L}} (1 + \gamma_{lc})^{-\beta_l}$, near an arbitrary point $\hat{\boldsymbol{\gamma}} \in \mathbb{R}_+^{L \times C}$. The parameters d and a_{lc} of the best monomial local approximation are given by*

$$a_{lc} = -\beta_l \hat{\gamma}_{lc} (1 + \hat{\gamma}_{lc})^{-1}, \quad d = f(\hat{\boldsymbol{\gamma}}) \prod_{c \in \mathcal{C}} \prod_{l \in \mathcal{L}} \hat{\gamma}_{lc}^{-a_{lc}} , \qquad (3.0.1)$$

where $\hat{\gamma}_{lc} = [\hat{\boldsymbol{\gamma}}]_{lc}$.

Proof. The monomial function m is the best local approximation of f near the point $\hat{\boldsymbol{\gamma}}$ if [5],

$$m(\hat{\boldsymbol{\gamma}}) = f(\hat{\boldsymbol{\gamma}}), \quad \nabla m(\hat{\boldsymbol{\gamma}}) = \nabla f(\hat{\boldsymbol{\gamma}}) . \qquad (3.0.2)$$

By replacing the expressions of m and f in (3.0.2) we obtain the following system of equations:

$$\begin{cases} d \prod_{c \in \mathcal{C}} \prod_{l \in \mathcal{L}} \hat{\gamma}_{lc}^{a_{lc}} = f(\hat{\boldsymbol{\gamma}}) \\ a_{lc} \hat{\gamma}_{lc}^{-1} d \prod_{c \in \mathcal{C}} \prod_{l \in \mathcal{L}} \hat{\gamma}_{lc}^{a_{lc}} = -\dfrac{\beta_l f(\hat{\boldsymbol{\gamma}})}{(1 + \hat{\gamma}_{lc})}, \ c \in \mathcal{C}, \ l \in \mathcal{L} , \end{cases}$$

which has the solution given by (3.0.1). $\qquad \square$

By using the local approximation given by Lemma 3.0.1 in the objective function of problem (3.11), and ignoring the multiplicative constant d which does not affect the problem solution, we obtain the identical successive approximation steps as in Algorithm 3.2.

Appendix 4 : Extension to multichannel SIC

In this appendix we present the multichannel extension of the material presented in Section 3.3. The assumptions remain the same as in Section 3.3, i.e., at every node $n \in \mathcal{N}$ the transmitter performs superposition coding over its outgoing links $\mathcal{O}(n)$ independently in each channel $c \in \mathcal{C}$ and every receiving node $n \in \mathcal{N}$ performs successive interference cancelation to decode the signals of incoming links $l \in \mathcal{I}(n)$ in each channel $c \in \mathcal{C}$. In every channel $c \in \mathcal{C}$, the SIC receiver at every node $n \in \mathcal{N}$ has to decode and cancel out the signals of all its incoming links $\mathcal{I}(n)$ and any subset of the remaining links in its complement set $\mathcal{L} \setminus \mathcal{I}(n)$ to obtain the largest set of achievable rates. Let us denote the set of links which are decoded at the node n, associated with each channel $c \in \mathcal{C}$ by $\mathcal{D}_c(n)$. Here the set $\mathcal{D}_c(n) = \mathcal{I}(n) \cup \mathcal{U}_c(n)$ for some $\mathcal{U}_c(n) \subseteq \mathcal{L} \setminus \mathcal{I}(n)$. Furthermore, let $\mathcal{R}_c^{\text{SIC}}(\mathcal{D}_c(1), \ldots, \mathcal{D}_c(N), p_{1c}^{\max}, \ldots, p_{Nc}^{\max})$ denote the achievable rate region associated with channel $c \in \mathcal{C}$ for the given $\mathcal{D}_c(1), \ldots, \mathcal{D}_c(N)$ and maximum node transmission power $p_{1c}^{\max}, \ldots, p_{Nc}^{\max}$, where $p_{nc}^{\max}$ is the maximum transmission power allocated to channel $c \in \mathcal{C}$ at node $n \in \mathcal{N}$. By taking the union of all possible combinations of sets $\mathcal{D}_c(1), \ldots, \mathcal{D}_c(N)$, the achievable rate region associated with channel $c \in \mathcal{C}$ for a given maximum node transmission power $p_{1c}^{\max}, \ldots, p_{Nc}^{\max}$ can be expressed as

$$\mathcal{R}_c^{\text{SIC}}(p_{1c}^{\max}, \ldots, p_{Nc}^{\max}) = \bigcup_{\mathcal{D}(1), \ldots, \mathcal{D}(N) | \forall n \in \mathcal{N} \ \exists \mathcal{U}(n) \subseteq \mathcal{L} \setminus \mathcal{I}(n) \ \text{s.t.} \ \mathcal{D}(n) = \mathcal{I}(n) \cup \mathcal{U}(n)} \mathcal{R}_c^{\text{SIC}}(\mathcal{D}(1), \ldots, \mathcal{D}(N), p_{1c}^{\max}, \ldots, p_{Nc}^{\max}) . \quad (4.0.1)$$

Let $\boldsymbol{\pi}_{nc} = \big(\pi_{nc}(1), \ldots, \pi_{nc}(|\mathcal{D}_c(n)|)\big)$ represent an arbitrary permutations of the links in $\mathcal{D}_c(n)$ which describes the decoding and cancelation order at node n in channel c. The rate region $\mathcal{R}_c^{\text{SIC}}(\mathcal{D}_c(1), \ldots, \mathcal{D}_c(N), p_{1c}^{\max}, \ldots, p_{Nc}^{\max})$ is obtained by considering all possible combinations of decoding orders for all nodes, i.e., all possible $\prod_{n \in \mathcal{N}}(|\mathcal{D}_c(n)| \, !)$ combinations $\boldsymbol{\pi}_c \triangleq \boldsymbol{\pi}_{1c} \times \boldsymbol{\pi}_{2c} \times \ldots \times \boldsymbol{\pi}_{Nc}$. Thus, the achievable rate region associated with channel $c \in \mathcal{C}$ for given $\mathcal{D}_c(1), \ldots, \mathcal{D}_c(N)$

and maximum node transmission power, $p_{1c}^{\max}, \ldots, p_{Nc}^{\max}$ can be expressed as

$$
\mathcal{R}_c^{\mathrm{SIC}}(\mathcal{D}_c(1), \ldots, \mathcal{D}_c(N), p_{1c}^{\max}, \ldots, p_{Nc}^{\max})
$$

$$
= \bigcup_{\pi_c} \left\{ (r_1, \ldots, r_L) \middle| \begin{array}{l} r_{\pi_{nc}(l)} \leq \log\left(1 + \dfrac{G_{\pi_{nc}(l)nc}(t)\, p_{\pi_{nc}(l)c}}{\sigma^2 + \sum_{j>l} G_{\pi_{nc}(j)nc}(t)\, p_{\pi_{nc}(j)c}}\right), \\[2mm] \forall (n,l) \text{ s.t. } n \in \mathcal{N},\ l \in \{1, \ldots, |\mathcal{D}_c(n)|\} \\[2mm] \sum_{l \in \mathcal{O}(n)} p_{lc} \leq p_{nc}^{\max},\ n \in \mathcal{N} \\[2mm] p_{lc} \geq 0,\ l \in \mathcal{L} \end{array} \right\},
$$

where G_{lnc}, $l \in \mathcal{L}$, $n \in \mathcal{N}$, $c \in \mathcal{C}$ represents the power gain from the transmitter of link l to the receiver at *node n* in channel c and p_{lc} represents the power allocated for link $l's$ signal in channel c. Note that we assume equal channel bandwidths for all $c \in \mathcal{C}$. By having superposition coding at the transmitters and SIC at the receivers, the achievable rate region for the interference channel can be expressed as

$$
\mathcal{R}^{\mathrm{SIC}}(p_1^{\max}, \ldots, p_N^{\max})
$$

$$
= \left\{ (r_1, \ldots, r_L) \middle| \begin{array}{l} (r_1, \ldots, r_L) \in \sum_{c \in \mathcal{C}} \mathcal{R}_c^{\mathrm{SIC}}(p_{1c}^{\max}, \ldots, p_{Nc}^{\max}) \\[2mm] \sum_{c \in \mathcal{C}} p_{nc}^{\max} \leq p_n^{\max},\ n \in \mathcal{N} \\[2mm] p_{nc}^{\max} \geq 0,\ n \in \mathcal{N},\ c \in \mathcal{C} \end{array} \right\}.
$$

The RA subproblem at the third step of Dynamic Cross-Layer Control Algorithm 3.1 is shown in (3.31). Finding the solution of this problem is extremely difficult, as we have already mentioned in the single-channel case, i.e., $C = 1$. However, by limiting the access protocol so that only one node can transmit in all its outgoing links in each slot, the problem can be identified as weighted sum-rate maximization over the capacity region of parallel Gaussian broadcast channels [185]. When only one node can receive from all its incoming links in each slot, the problem can be cast as weighted sum-rate maximization over the capacity region of Gaussian vector multiple access channel [200],[201, Sec. 6].

Appendix 5 : The barrier method

In this appendix we outline the basic steps involved in solving problem (3.33) using the barrier method [3, Sec. 11.3.1]. For the sake of notational simplicity let us define $\bar{\sigma}_{\rho_n(k)} = \sigma^2/g_{\rho_n(k)\rho_n(k)}$ for $k = 1, \ldots, |\mathcal{O}(n)|$ and $\bar{\sigma}_{\rho_n(|\mathcal{O}(n)|+1)} = 0$. Furthermore, let $\mathbf{u}_i$ be the ith column of the upper triangular matrix $\mathbf{U} \in \mathbb{R}_+^{|\mathcal{O}(n)| \times |\mathcal{O}(n)|}$ with all nonzero entries being equal to 1.

By denoting the feasible set of rate allocation vector $\mathbf{r}_n = (r_{\rho_n(1)} \ldots r_{\rho_n(|\mathcal{O}(n)|)})$ [185, Sec. 3], problem (3.33) can be equivalently expressed as

$$
\begin{aligned}
\text{maximize} \quad & \sum_{l \in \mathcal{O}(n)} \beta_l r_l \\
\text{subject to} \quad & u_i(\mathbf{r}_n) \leq 0, \ i = 1, \ldots, |\mathcal{O}(n)| + 1 ,
\end{aligned}
\tag{5.0.1}
$$

where the variable is $\mathbf{r}_n$. The function $u_i(\mathbf{r}_n)$ can be compactly expressed as

$$
u_i(\mathbf{r}_n) = \begin{cases} -\mathbf{e}_i^{\mathrm{T}} \mathbf{r}_n & 1 \leq i \leq |\mathcal{O}(n)| \\ \displaystyle\sum_{j=1}^{|\mathcal{O}(n)|} b_n(j) e^{\mathbf{u}_j^{\mathrm{T}} \mathbf{r}_n} - \bar{\sigma}_{\rho_n(1)} - p_n^{\max} & i = |\mathcal{O}(n)| + 1 , \end{cases}
$$

where $b_n(j) = \bar{\sigma}_{\rho_n(j)} - \bar{\sigma}_{\rho_n(j+1)}$. Problem (5.0.1) is a convex optimization problem [3], and therefore can be solved efficiently. It is worth noting that, given any feasible $\mathbf{r}_n$ the corresponding power variables $p_{\rho_n(k)}, \ k = 1, \ldots, |\mathcal{O}(n)|$ are given by [185]

$$
p_{\rho_n(k)} = \left(e^{r_{\rho_n(k)}} - 1 \right) \sum_{i \geq k} b_n(i) e^{\sum_{k < j \leq i} r_{\rho_n(j)}} .
$$

The barrier method [3, Sec. 11.3.1] can be used to solve problem (5.0.1). The gradient and the Hessian of the function $u_i(\mathbf{r}_n)$ are given by

$$
\nabla u_i(\mathbf{r}_n) = \begin{cases} -\mathbf{e}_i & 1 \leq i \leq |\mathcal{O}(n)| \\ \sum_{j=1}^{|\mathcal{O}(n)|} b_n(j) e^{\mathbf{u}_j^{\mathrm{T}} \mathbf{r}_n} \mathbf{u}_j & i = |\mathcal{O}(n)| + 1 \end{cases}
$$

and

$$
\nabla^2 u_i(\mathbf{r}_n) = \begin{cases} \mathbf{0} & 1 \leq i \leq |\mathcal{O}(n)| \\ \sum_{j=1}^{|\mathcal{O}(n)|} b_n(j) e^{\mathbf{u}_j^{\mathrm{T}} \mathbf{r}_n} \mathbf{u}_j \mathbf{u}_j^{\mathrm{T}} & i = |\mathcal{O}(n)| + 1 . \end{cases}
$$

The expressions above are used to evaluate the gradient and the Hessian of the logarithmic barrier function [3, Sec. 11.2.1].

381. Muhos, Matti (2011) Early stages of technology intensive companies

382. Laitinen, Ossi (2011) Utilisation of tube flow fractionation in fibre and particle analysis

383. Lasanen, Kimmo (2011) Integrated analogue CMOS circuits and structures for heart rate detectors and other low-voltage, low-power applications

384. Herrala, Maila (2011) Governance of infrastructure networks : development avenues for the Finnish water and sewage sector

385. Kortelainen, Jukka (2011) EEG-based depth of anesthesia measurement : separating the effects of propofol and remifentanil

386. Turunen, Helka (2011) CO_2-balance in the athmosphere and CO_2-utilisation : an engineering approach

387. Juha, Karjalainen (2011) Broadband single carrier multi-antenna communications with frequency domain turbo equalization

388. Martin, David Charles (2011) Selected heat conduction problems in thermomechanical treatment of steel

389. Nissinen, Jan (2011) Integrated CMOS circuits for laser radar transceivers

390. Nissinen, Ilkka (2011) CMOS time-to-digital converter structures for the integrated receiver of a pulsed time-of-flight laser rangefinder

391. Kassinen, Otso (2011) Efficient middleware and resource management in mobile peer-to-peer systems

392. Avellan, Kari (2011) Limit state design for strengthening foundations of historic buildings using pretested drilled spiral piles with special reference to St. John's Church in Tartu

393. Khatri, Narendar Kumar (2011) Optimisation of recombinant protein production in *Pichia pastoris* : Single-chain antibody fragment model protein

394. Paavola, Marko (2011) An efficient entropy estimation approach

395. Khan, Zaheer (2011) Coordination and adaptation techniques for efficient resource utilization in cognitive radio networks

396. Koskela, Timo (2011) Community-centric mobile peer-to-peer services: performance evaluation and user studies

397. Karsikas, Mari (2011) New methods for vectorcardiographic signal processing

SERIES EDITORS

A

SCIENTIAE RERUM NATURALIUM

Senior Assistant Jorma Arhippainen

B

HUMANIORA

Lecturer Santeri Palviainen

C

TECHNICA

Professor Hannu Heusala

D

MEDICA

Professor Olli Vuolteenaho

E

SCIENTIAE RERUM SOCIALIUM

Senior Researcher Eila Estola

F

SCRIPTA ACADEMICA

Director Sinikka Eskelinen

G

OECONOMICA

Professor Jari Juga

EDITOR IN CHIEF

Professor Olli Vuolteenaho

PUBLICATIONS EDITOR

Publications Editor Kirsti Nurkkala

UNIVERSITY of OULU
OULUN YLIOPISTO