*Les Cahiers du GERAD* ISSN: 0711–2440

**Reformulation and Decomposition for Traffic Routing in Optical Networks**

B. Vignac, F. Vanderbeck,
B. Jaumard

G–2010–12

February 2010

# Reformulation and Decomposition Approaches for Traffic Routing in Optical Networks


**Benoit Vignac**

*Institut de Mathématiques de Bordeaux*
*Université Sciences et Technologies - Bordeaux I*
*& CIRRELT, Université de Montréal*
vignac@crt.umontreal.ca

**François Vanderbeck**

*Institut de Mathématiques de Bordeaux*
*Université Sciences et Technologies - Bordeaux I*
fv@math.u-bordeaux1.fr

**Brigitte Jaumard**

*GERAD & CIISE*
*Concordia University*
bjaumard@ciise.concordia.ca



February 2010

*Les Cahiers du GERAD*

G–2010–12

**Abstract**

We consider a multi-layer network design model arising from a real-life telecommunication application where traffic routing decisions imply the installation of expensive nodal equipment. Customer requests come in the form of bandwidth reservations for a given origin destination pair. Bandwidth demands are expressed as a multiple of nominal granularities. Each request must be single-path routed. Grooming several requests on the same wavelength and multiplexing wavelengths in the same optical stream allow the packing of more traffic. However, each addition or withdrawal of a request from a wavelength requires optical to electrical conversion for which a specific portal equipment is needed. The objective is to minimize the number of such equipment. We deal with backbone optical networks, therefore with networks with a moderate number of nodes (14 to 20) but thousands of requests. Further difficulties arise from the symmetries in wavelength assignment and traffic loading. Traditional multi-commodity network flow approaches are not suited for this problem. Four alternative models relying on Dantzig-Wolfe and/or Benders' decomposition are introduced and compared. The formulations are strengthened using symmetry breaking restrictions, variable domain reduction, zero-one decomposition of integer variables, and cutting planes. The resulting dual bounds are compared to the values of primal solutions obtained through hierarchical optimization and rounding procedures. For realistic size instances, our best approaches provide solutions with optimality gap of approximately 5% on average in around 2 hours of computing time.

# Introduction

To accommodate the increase of traffic in telecommunication networks, today's optical networks have huge capacity (tens of Tb/s) thanks to new technologies. The wavelength bandwidth utilization is increased by packing several requests on the same wavelength, a technique called *traffic grooming* [18]. Moreover, several streams can be multiplexed on an optical signal, each of them supported by a different wavelength, a technique called *wavelength division multiplexing* (WDM). However, packing multiple requests together on the same optical stream still requires to convert the signal in the electrical domain at each traffic aggregation or disaggregation, at an origin or a destination node or at a switch. These so-called opto-electronic or O/E/O *conversions* require the installation of expensive optical ports. Hence, traffic grooming and routing decisions along with wavelength assignments must be optimized to reduce opto-electronic system installation cost, while satisfying quality of service requirements (like limiting end-to-end delay). This optimization problem is known as the *grooming, routing and wavelength assignment* (GRWA) problem.

The telecommunication backbone network that we consider is defined by a *physical network* with given nodes and edges. We assume that (*i*) each edge is made of two *optical fiber links* with opposite signal transportation directions, (*ii*) each link has a fiber capacity made of a uniform number of wavelengths and (*iii*) each wavelength has the same transport capacity. Traffic demands take the form of bandwidth reservations. Each request is defined by its origin and destination and a bandwidth requirement that is selected from a discrete set of standard granularities (larger granularities are multiple of smaller ones). Because of transmission protocols (ATM, MPLS...), a request must be single-path routed. Its optical route, or *lightpath*, is defined by a sequence of *optical hops*, each of which is defined by a path in the physical network along which the signal remains into the optical domain with no electrical conversion at intermediate nodes (that are optically *bypassed*).

Hence, traffic routing can be viewed as defining a lightpath for each request in a *logical network* whose nodes are those of the physical network and whose arcs represent optical hops, each of which is associated with a physical path. The logical network is a multi-digraph containing as many arcs between two nodes as the number of physical paths between them. A sequence of arcs of the logical network defines a lightpath (with associated physical paths in the underlying physical network for each optical hop). In the sequel, we sometimes model traffic in the *aggregate logical network*, where different optical hops with the same end-nodes are represented by a single aggregate optical hop. This aggregate network holds a single arc between two nodes if there exists at least a physical path between them along which one can establish an optical hop. Hence, we call it the *connectivity network*. A path is the connectivity network is called a *virtual route* as the associated physical routing is not defined.

The bandwidth capacity of an arc of the connectivity network is the sum of the capacties of the logical network arcs between these nodes. Thanks to wavelength division multiplexing, each optical hop (i.e., each arc of the logical network) can carry a separate stream for each wavelength. However, the feasibility of a wavelength assignment must be checked in the underlying physical network. Each physical fiber link can carry at most one stream per wavelength. Using *grooming*, multiple requests can share the same optical hop on the same wavelength provided their cumulative bandwidth requirement does not exceed wavelength capacity. For the O/E/O conversion, an *optical port* must be installed at each end-nodes of the optical hops. The overall optical port installation cost is therefore measured by twice the number of optical hops that are used.

When routing a request, one must make sure that the end-to-end delay remains reasonable. The conversion delay at O/E/O nodes plus the fiber link transmission delay must satisfy quality of service (QoS) criteria, especially for real time applications like voice or video-conference ones [29]. Furthermore, one should also account for the fact that, if an optical hop is too long, the stream must be regenerated using repeaters that also imply O/E/O conversions. As we only deal with backbone networks, we cannot control the accumulated delay into the access and the metropolitan networks. Hence, we can only attempt to limit delay in the backbone network; we use two business rules: (*i*) we limit the O/E/O conversion delay by restricting lightpaths to involve at most 2 optical hops (i.e., at most one O/E/O at an intermediate node); (*ii*) the physical path used by a request must be one of the three elementary shortest paths that exists in the network between its origin and destination nodes.

In summary, the specific restrictions assumed in this study are:

**Assumption 1 (single-path routing)** *Each request must be single-path routed (no "bifurcation" is allowed). This assumption makes the problem harder because one must follow individual flows for each origin-destination and granularity by defining separate "commodities" for which an integer flow solution is required. One cannot aggregate these commodities or relax the problem to continuous flow.*

**Assumption 2** *(***Divisibility of request granularities)** *Each request takes the form of a bandwidth reservation that takes value in a discrete set of standard granularities; each possible granularity is a multiple of smaller granularities and a divider of the wavelength transport capacity. In our study, request granularities are selected in* $\{1, 3, 12, 48\}$*; they are measured in* OC *(*1 OC = *51,84 Mb/s); the wavelength capacity is* $U = 192$ OC*. Because of this property, the bin packing problem underlying the bandwidth capacity check is trivial to solve.*

**Assumption 3** *(***2-hop routing)** *The number of optical hops on a lightpath is bounded by two in order to limit the* O/E/O *delay. This restriction has limited impact on the port installation cost. Indeed, we observed that imposing single-hop routing results in a significant increase in port installation costs, while allowing 3 or more optical hops results in only very marginal decrease in optical port installation costs and may increase the delay [24].*

**Assumption 4** *(***Physical path length)** *The overall length of the physical route of each request is at most the length of the third elementary shortest path between its source and destination. In practice, for each origin destination pair* $(s, d)$*, we restrict our attention to the three shortest paths as physical support to the lightpath for routing* $(s, d)$*-requests.*

In addition, some models developed herein rely on the following restrictive assumptions.

**Assumption 5** *(***Restrictive grooming configurations)** *We further limit the solution space by restricting our attention to simple grooming scenarios (as detailed later and illustrated in Figure 2) such as a single-hop, the fusion of two single-hop streams, or their bifurcation, etc.*

**Assumption 6** *(***Wavelength continuity)** *We sometimes consider a further restriction, requiring wavelength continuity on each lightpath. If the signal is assigned to a given wavelength on an optical hop, it must be re-sent on the same wavelength on the next optical hop.*

Many variants of the GRWA problem have been studied in the literature that differ by their objective function and constraints; a classification can be found in [5]. Maximization of the throughput was studied in [26, 27, 28] under a restricted optical port resource. However, in view of the large available capacity, minimizing the network cost, for which the optical port cost is a major component, is a more meaningful objective as studied in [8, 12, 15]. When every request has the same granularity, as assumed in [15] and [28], a grooming ratio is defined as the number of requests that can be groomed on the same optical hop and Assumption 1 is naturally satisfied [11]. However, in real-life applications, bandwidth requirements follow a discrete value distribution as modeled in [27]. To the best of our knowledge, no study has yet enforced a maximum number of optical hops for each request (except in [25] for IP over WDM networks with the objective of maximizing the throughput). However, multi-hop routing without any restriction may lead to unacceptable end-to-end delays. To obtain an easier GRWA problem, some decisions are sometimes fixed a priori. When the set of optical hops is given in advance [12], it removes the issue of defining a physical routing for each optical hop. When the number of wavelengths/links is assumed sufficiently large [8, 15], the connectivity topology can always be implemented on the physical topology and the wavelength assignment becomes trivial. Finally, some studies do not assume single-path routing and therefore deal with continuous flow models (see, e.g, [16]).

The GRWA problem is proved NP-hard in [27] (for its simplest variants). Greedy heuristics have been proposed in [26] and [27] (where an oracle is used that does shortest path routing in an auxiliary graph), and in [15] (that assumes a given set of optical hops and exploits a multi-commodity flow dual formulation). A tabu search, a genetic algorithm and a multi-objective evolutionary algorithm have also been considered respectively in [14] and [21]; but these publications offer no lower bounds to evaluate the heuristic solutions quality. A hierarchical optimization procedure is used in [12], where the GRWA problem is decomposed into two parts, GR and WA, that are solved sequentially. To obtain primal solutions of the GR problem, a multi-commodity flow formulation is solved, where optical hop design variables are not restricted to take integer values (linear relaxation); then, the fractional solution is rounded up.

Exact solution approaches have mostly been based on multi-commodity flow formulations [12, 15, 27]. However, the size of the resulting mathematical program is too large for medium to large instances. Indeed, the GRWA model implies many more commodities than standard telecommunication routing problems. The single-path routing assumption imposes integer flows and the definition of a separate commodity for each origin-destination pair and for each granularity; this leads to a hard integer capacitated network design problem on the logical network where arcs represent optical hops. Moreover, multiple wavelengths implicitly duplicate this support graph in as many layers as the number of available wavelengths. Capacitated network design problems are already very hard to solve when the number of commodities is lower than 100 on non complete graphs with less than 30 nodes as shown in [3, 4, 6, 7]. On instances with 435 commodities, the authors of [6] report an average gap of around 30 % after one hour of computation using CPLEX [13] and an extra hour of post-processing. Most of the GRWA instances are on networks with 10 to 30 nodes but involve hundreds of requests. The number of wavelengths is often greater than 10. This leads to instance sizes that are beyond what can be solved using exact capacitated network design approaches. The instances considered in this study imply backbone networks with 14 and 21 nodes, thousand of requests, real-life granularity distribution (see [19]), and tens of wavelengths.

In this study, we compare four exact optimization based approaches, two models rely on Benders' decomposition, one model on Dantzig-Wolfe decomposition, and the fourth on a combination of these two approaches. When using Benders' decomposition, the first stage decisions are either the grooming and virtual routing, or the grooming and physical routing. When using Dantzig-Wolfe decomposition, the subproblems are either associated with feasible traffic loading for the restricted set of grooming configurations (under Assumption 5), or traffic loading for a single wavelength (under Assumption 6). Dual bounds are obtained by solving the LP relaxation of Benders' (resp. Dantzig-Wolfe) master program. However, in Benders' approaches, we only solve a relaxed master where no Benders' cuts are generated (the sub-problem is an integer linear program). Thus, in practice, we carry on a simple hierarchical optimization with no feedback loop. Benders' approaches give rise to compact formulations that can be handled directly by CPLEX. Dual bounds can then be improved through a branch-and-bound method and integer solutions to the master are derived through CPLEX build-in heuristics. It remains to solve the second stage problems to recover a primal solution if feasible. For Dantzig-Wolfe approaches, we develop our own column generation procedure to solve the master LP and obtain primal solutions using a rounding heuristic.

The rest of the paper is organized as follows. In Section 1, we formally describe the GRWA problem, provide an initial formulation, review the impact of our assumptions, and explain the symmetries. In Sections 2 and 3, we present the Dantzig-Wolfe and Benders' decomposition approaches. Section 4 outlines ways to reduce symmetries, to restrict the domains of the variables, to derive a zero-one reformulation that enables tighter relations between the variables and to strengthen the formulations with cutting planes. Then, in Sections 5 and 6, we summarize the algorithms used for each of the four solution approaches and compare the numerical results.

# 1 Description of the GRWA Problem

Let graph $G = (V, A)$ represents the physical network with $n = |V|$ nodes and $m = |A|$ arcs. Each arc $a \in A$ is associated with a directional optical fiber link of the physical network, whose length is denoted by $l_a$. Each optical fiber link can carry up to $W$ wavelengths. Let $\Lambda = \{1, \ldots, W\}$. Each wavelength has a transport capacity $U = 192$ OC. A request $r \in R$ is defined by a triplet $r = (s, d, t)$ where $(s, d) \in V^2$ denotes its origin and destination nodes, respectively, and $t \in T = \{1, 3, 12, 48\}$ is the granularity of the bandwidth reservation. Let $D_{sdt} \in \mathbb{N}$ be the number of $(s,d,t)$-requests and $D_{sd} = \sum_{t \in T} t\, D_{sdt}$ be the aggregate traffic demand between $s$ and $d$ expressed in OC. Let $K$ be the set of distinct $(s,d,t)$ requests, each of which is a *request-commodity* that is defined by a quadruplet $k = (s,d,t,D)$, where $s$ and $d$ represent origin and destination nodes of the traffic, $t$ the bandwidth granularity, and $D = D_{sdt} > 0$ the $(s,d,t)$ demand. Notation $s_k$ (resp. $d_k$, $t_k$, and $D_k$) stands for the source node of *request-commodity* $k$ (resp. destination node, granularity, and demand). $K_{sd}$ (resp. $K_{sdt}$) denotes the set of request-commodities with source $s$ and destination $d$ (and granularity $t$). Let $L^{sd}_{\max}$ be the maximum length of the physical paths of requests that go from $s$ to $d$. Given Assumption 4, $L^{sd}_{\max}$ is the length of the third shortest $(s,d)$-path.

The GRWA problem can be viewed as a 2-layer multi-commodity capacitated network design problem. The first layer models grooming and virtual routing in the connectivity network, where there is one flow type for each *request-commodity* $k \in K$. The second layer models the design of the connectivity network in the physical network and the wavelength assignment. To properly model wavelength clash and path length constraints, one must indeed associate a path in the physical network to each optical hop and assign a specific wavelength to this physical path. We formulate the problem in terms of variables $x$ representing flows on lightpaths and design variables $y$ defining the optical hops that are put in place and their associated physical path and wavelength assignment. Let $\mathcal{P}_{sd}$ be the set of feasible paths from $s$ to $d$. The associated polyhedron is

$$\mathcal{P}_{sd} = \Bigg\{ z \in \{0,1\}^m : \sum_{(u,v) \in A} l_{uv}\, z_{uv} \le L^{sd}_{\max} \tag{1}$$

$$\sum_{v \in V:(s,v) \in A} z_{sv} = 1 = \sum_{u \in V:(u,d) \in A} z_{ud}; \sum_{u \in V:(u,v) \in A} z_{uv} - \sum_{u \in V:(v,u) \in A} z_{vu} = 0 \quad \forall v \in V \setminus \{s,d\} \Bigg\}, \tag{2}$$

where variable $z_{uv} = 1$ if arc $a = (u,v)$ is in the path. Given Assumption 4, $\mathcal{P}_{sd}$ is restricted to the three shortest paths from $s$ to $d$. We assume that the network does not contain two paths of the same length. Let $\mathcal{P} = \cup_{sd} \mathcal{P}_{sd}$; note that $|\mathcal{P}| = 3\,n(n-1)$.

Let $\mathcal{L}_{sd}$ be the set of feasible lightpaths from $s$ to $d$. Its polyhedral description is:

$$\mathcal{L}_{sd} = \Bigg\{ (y,w) \in \{0,1\}^{|\mathcal{P}| \times (n-2)} : \sum_{p \in \mathcal{P}} \sum_{a \in A} l_a\, z^p_a\, y_p \le L^{sd}_{\max} \tag{3}$$

$$\sum_{v \in V \setminus \{s\}} \sum_{p \in \mathcal{P}_{sv}} y_p = 1 = \sum_{v \in V \setminus \{d\}} \sum_{p \in \mathcal{P}_{vd}} y_p; \sum_{p \in \mathcal{P}_{sv}} y_p = w_v = \sum_{p \in \mathcal{P}_{vd}} y_p \quad \forall v \in V \setminus \{s,d\} \tag{4}$$

$$\sum_{v \in V \setminus \{s,d\}} w_v \le 1 \Bigg\}, \tag{5}$$

where variable $y_p = 1$ if path $p$ forms an optical hop of the lightpath; $w_v = 1$ if $v$ is an intermediate node of the lightpath where a O/E/O conversion takes place; $z^p_a$ are not variables but fixed arc indicators that define path $p$: $z^p_a = 1$ if path $p \in \mathcal{P}$ uses arc $a \in A$. Let $\mathcal{L} = \cup_{sd} \mathcal{L}_{sd}$; note that $|\mathcal{L}| \le |\mathcal{P}|(n-2)$ (each path $p \in \mathcal{P}$ can yield a lightpath for each of its intermediate nodes).

The GRWA can be formulated using variables:
$x^k_\ell$ = the amount of request-commodity $k$ that is routed along lightpath $\ell \in \mathcal{L}_{s_k d_k}$;
$y^\lambda_p = 1$ if an optical hop is routed along path $p \in \mathcal{P}$ and assigned with wavelength $\lambda$.

Then, the problem takes the form:

$$\min \sum_{p \in \mathcal{P}} \sum_{\lambda \in \Lambda} y^\lambda_p \tag{6}$$

$$\sum_{\ell \in \mathcal{L}_{s_k d_k}} x^k_\ell \quad = \quad D_k \quad \forall k \in K \tag{7}$$

$$\sum_{p\in\mathcal{P}} z_a^p\, y_p^\lambda \quad \leq \quad 1 \quad \forall \lambda \in \Lambda,\, a \in A \tag{8}$$

$$\sum_{k\in K}\sum_{\ell\in\mathcal{L}_{s_k d_k}} t_k\, y_p^\ell\, x_\ell^k \quad \leq \quad U(\sum_{\lambda\in\Lambda} y_p^\lambda) \quad \forall p \in \mathcal{P} \tag{9}$$

$$x_\ell^k \quad \in \quad I\!N \quad \forall k \in K,\, l \in \mathcal{L}_{s_k d_k} \tag{10}$$

$$y_p^\lambda \quad \in \quad \{0,1\} \quad \forall \lambda \in \Lambda, p \in \mathcal{P}, \tag{11}$$

where $y_p^\ell$ is not a variable but the indicator of a solution of (3–5) defining lightpath $\ell \in \mathcal{L}$ and $z_a^p$ describes the path $p$ of the associated solution of (1–2).

The *shortest path routing* constraints are modeled by (3). The *Single-path routing* are modeled by (4) and (10). The *2-hop restrictions* are modeled by (5). The objective (6) is to minimize the number of optical hops that are used (equivalent to half the number of installed O/E/O ports). Constraints (7) model *demand satisfaction*. Constraints (8), together with the restriction of the wavelength index to the set $\Lambda$, model *Link capacity and Wavelength clash* constraints. Each link can carry at most $W$ wavelengths and hence, it can be part of at most one optical hop for each wavelength. The bandwidth of the stream that is carried on an optical hop for a given wavelength is bounded by $U$. These *wavelength capacity* constraints are modeled by (9) although they seem to model only a surrogate relaxation. One would a priori need to measure individual traffic assignment for each physical link and each wavelength stream. However, wavelength clash constraints (8) guarantee that optical hops are arc disjoint for a given wavelength. Hence, it is enough to check capacity for each optical hop and wavelength pair. Furthermore, because of the traffic granularity and divisibility (Assumption 2), one can aggregate the traffic load for all wavelengths and yet correctly enforce wavelength bandwidth capacity:

**Proposition 1** *Under Assumption 2, a solution satisfying the surrogate capacity constraints (9) can be decomposed into wavelength assigned flows that satisfy bandwidth capacity constraints.*

**Proof:** Consider the traffic through a given optical hop $p$. Decomposing this traffic per wavelength $\lambda$ while obeying individual knapsack constraints amounts to solving a bin packing feasibility problem, with $y_p = \sum_{\lambda\in\Lambda} y_p^\lambda \geq \left\lceil \frac{\sum_{k\in K}\sum_{\ell\in\mathcal{L}_{s_k d_k}} t_k y_p^\ell x_\ell^k}{U} \right\rceil$ available bins of capacity $U$ (the latter inequality being implied by constraints (9)). Under Assumption 2, this bin packing problem can be solved using a trivial first fit decreasing greedy procedure (see [20]). It consists in $(i)$ sorting the $x_\ell^k$ traffic bundles, for $k \in K$ and $\ell$ such that $y_p^\ell = 1$, in the decreasing order of their granularities; $(ii)$ assigning traffic bundles in that order to the $y_p$ bins each of which is indexed by a $\lambda$. All the bins are filled at full capacity $U$, but the last one whose load is $\sum_{k\in K}\sum_{\ell\in\mathcal{L}_{s_k d_k}} t_k y_p^\ell x_\ell^k - \left\lfloor \frac{\sum_{k\in K}\sum_{\ell\in\mathcal{L}_{s_k d_k}} t_k y_p^\ell x_\ell^k}{U} \right\rfloor U$. ∎

The flow disaggregation argument used in the above proof can be extended to a flow redistribution argument that shall be used throughout the paper:

**Observation 1 (bandwidth re-assignment)** *Given the traffic granularity set and their divisibility outlined in Assumption 2, the total bandwidth reservation made on an optical hop $p$, noted $x_p$, can be aggregated and re-partitioned between the $y_p = \lceil \frac{x_p}{U} \rceil$ wavelengths in any fashion that meets the wavelength capacity $U$, with no consequences on the rest of the solution nor its cost.*

Hence, the are many symmetries in the feasible solution set, not only among possible wavelength assignments, but also due to possible granularity exchanges. Our reformulations shall aim at reducing these symmetries. For now, note that the wavelength index on the $y$ variables are required to model wavelength clash constraints.

Let us review Assumptions 1 to 6 and discuss their impact on the problem formulation. If we relax Assumption 1, the $x$ variables can take continuous values. We can omit constraints (10). If we do not make Assumption 2, the surrogate capacity constraints (9) are no more enough to enforce bandwidth capacity; we have to set a capacity constraint for each wavelength and optical hop. Then, one would need to specify the wavelength used on each optical hop that is part of a lightpath. This implies a dramatic increase in

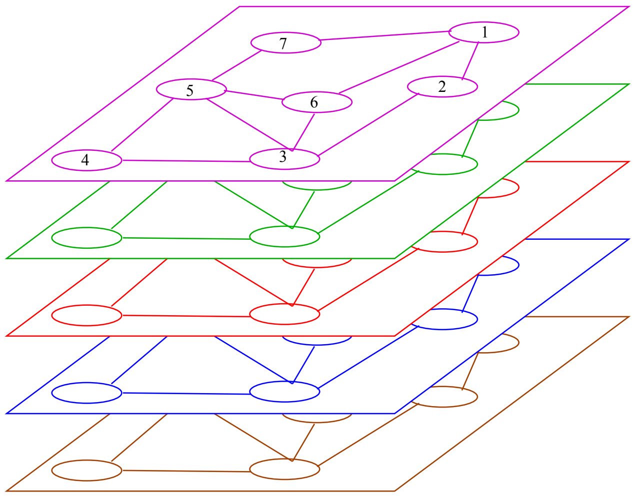


Figure 1: Stack of independent wavelength traffic routing

the number of variables. If Assumption 3 is relaxed, one would have to modify the definition of the routed lightpath to allow more than 2 hops. To enforce Assumption 6, one would need to use variables $x_\ell^{k\lambda}$ to represent the amount of request-commodity $k$ routed on $\ell$ and assigned to wavelength $\lambda$; furthermore one need to disaggregate capacity constraints (9). The resulting formulation would be:

$$\min\{\sum_{p\in\mathcal{P}}\sum_{\lambda\in\Lambda} y_p^\lambda : \sum_{\lambda,\ell\in\mathcal{L}_{s_kd_k}} x_\ell^{k\lambda} = D_k \quad \forall k\in K;\ \sum_{p\in\mathcal{P}} z_a^p\, y_p^\lambda \le 1 \quad \forall\lambda\in\Lambda,\ a\in A; \tag{12}$$

$$\sum_k \sum_{\ell\in\mathcal{L}_{s_kd_k}} t_k\, y_p^\ell\, x_\ell^{k\lambda} \le U y_p^\lambda \quad \forall\lambda\in\Lambda,\ p\in\mathcal{P};\ x_\ell^{k\lambda}\in I\!N \quad \forall\lambda, k\in K,\ l\in\mathcal{L}_{s_kd_k};\ y_p^\lambda\in\{0,1\} \quad \forall\lambda, p\in\mathcal{P}\}. \tag{13}$$

Then, a solution to GRWA can be decomposed into $W$ independent traffic routings, each of them using its own wavelength (this is illustrated in Figure 1).

Relaxing the path length restriction (Assumption 4) significantly eases the problem. One can then use a formulation in terms of arc flow $x$ in the connectivity network and design variables $y$ defining aggregate optical hops, while physical paths are modeled using arc flow variables $z$. Let $x_{ij}^k$ be the amount of *request-commodity* $k=(s,d,t,D)$ that is routed on aggregate optical hop $(i,j)\in V^2$ (because of the 2-hop constraints, these variables are only defined for $i=s_k$ or $j=d_k$). Let $y_{ij}$ be the number of aggregate optical hops $(i,j)\in V^2$ that are used and must be physically established. Let $z_{uv}^{ij\lambda}=1$ if a stream is assigned to wavelength $\lambda$ on aggregate optical hop $(i,j)\in V^2$ and it uses a route passing through physical arc $(u,v)\in A$. The resulting model (when Assumption 4 is relaxed) takes the form:

$$\min \sum_{(i,j)\in V^2} y_{ij} \tag{14}$$

$$\sum_{j\in V\setminus\{s_k\}} x_{s_kj}^k = \sum_{i\in V\setminus\{d_k\}} x_{id_k}^k = D_k \quad \forall k\in K \tag{15}$$

$$x_{s_ki}^k - x_{id_k}^k = 0 \quad \forall k\in K, i\in V \tag{16}$$

$$\sum_{k\in K} t_k\, x_{ij}^k \le U\, y_{ij} \quad \forall (i,j)\in V^2 \tag{17}$$

$$\sum_{v\in V\setminus\{j\}}\sum_{\lambda\in\Lambda} z_{vj}^{ij\lambda} = \sum_{v\in V\setminus\{i\}}\sum_{\lambda\in\Lambda} z_{iv}^{ij\lambda} = y_{ij} \quad \forall (i,j)\in V^2 \tag{18}$$

$$\sum_{u\in V\setminus\{v\}} z_{vu}^{ij\lambda} - \sum_{u\in V\setminus\{v\}} z_{uv}^{ij\lambda} = 0 \quad \forall (i,j)\in V^2,\ \lambda\in\Lambda,\ v\in V\setminus\{i,j\} \tag{19}$$

$$\sum_{(i,j)\in V^2} z_{uv}^{ij\lambda} \leq 1 \quad \forall \lambda \in \Lambda,\ (u,v) \in A \tag{20}$$

$$x_{ij}^k \in I\!N \quad \forall k \in K, (i,j) \in V^2 : i = s_k \text{ or } j = d_k \tag{21}$$

$$y_{ij} \in I\!N \quad \forall (i,j) \in V^2 \tag{22}$$

$$z_{uv}^{ij\lambda} \in \{0,1\} \quad \forall (i,j) \in V^2,\ \lambda \in \Lambda,\ (u,v) \in A. \tag{23}$$

Another useful relaxation of formulation (6–11) comes from the relaxation of the wavelength clash constraints (8). Then, the wavelength index on the $y$ variables is no longer required and the optical hops with the same end-nodes can be aggregated. Let $y_{ij}$ be the number of optical hops established between nodes $i$ and $j$. It is defined for each arc $(i,j)$ of the *connectivity network*. Lightpaths with the same end-nodes and O/E/O node can also be aggregated. The resulting aggregate lightpaths set is denoted by $\tilde{\mathcal{L}}$. The relaxed problem can be formulated in terms of virtual routing in the connectivity network and aggregate optical hop selection. We shall refer to this relaxation as the **virtual routing formulation**. It takes the form:

$$\min \sum_{(i,j)\in V^2} y_{ij} \tag{24}$$

$$[\text{VR}] \quad \sum_{\ell\in\tilde{\mathcal{L}}_{s_k d_k}} x_\ell^k = D_k \quad \forall k \in K \tag{25}$$

$$\sum_{k\in K} \sum_{\ell\in\tilde{\mathcal{L}}_{s_k d_k}} t_k\, y_{ij}^\ell\, x_\ell^k \leq U\, y_{ij} \quad \forall (i,j) \in V^2 \tag{26}$$

$$x_\ell^k \in I\!N \quad \forall k \in K,\ l \in \tilde{\mathcal{L}}_{s_k d_k} \tag{27}$$

$$y_{ij} \in I\!N \quad \forall (i,j) \in V^2. \tag{28}$$

The LP relaxation of this formulation provides the same dual bound as the LP relaxation of formulation (6–11). Moreover, these bounds are not better than the trivial combinatorial bound on the number of required optical hops given wavelength capacity: $\lceil \sum_k t_k D_k / U \rceil$.

**Proposition 2** *The dual bounds obtained from the linear relaxations of (6–11) and (24–28) both give the trivial dual bound* $\sum_{k\in K} t_k D_k / U$.

**Proof:** An optimal solution to the LP relaxation of (24–28) is obtained by routing each request on a single-hop aggregate lightpath. Indeed, setting $y_{ij} = \sum_{k\in K_{ij}} t_k D_k / U$, $(i,j) \in V^2$ yields a feasible solution. Any other feasible LP solution can only cost more. Actually, if a request was routed over a two-hop aggregate lightpath, port installation cost would be incurred at the intermediate node. As formulation (24–28) is a relaxation of formulation (6–11), its LP solution $\sum_{k\in K} t_k D_k / U$ is also a lower bound for formulation (6–11). Thus, to obtain the bound result for formulation (6–11), it is enough to exhibit an LP solution reaching that bound. In fact, we show that we can recover the "single-hop solution" from any feasible LP solution $(\bar{y}, \bar{x})$ to (6–11). The idea is to keep the same physical routing while modifying the virtual routing so that each request is routed on a single-hop lightpath. Assume, w.l.o.g., that $\bar{y}_p = \frac{\sum_{\ell,k} y_p^\ell\, t_k\, \bar{x}_\ell^k}{U}$. We build a "single-hop solution" $(\tilde{x}, \tilde{y})$ as follows. For each $(s,d) \in V$ and $p \in \mathcal{P}_{sd}$, let $\bar{x}_p^k$ be the total $k$ demand that is physically routed on $(s,d)$-path $p$ (some of this traffic might be two-hop traffic). Then, set $\tilde{x}_\ell^k = \bar{x}_p^k$ for $\ell \in \mathcal{L}_{sd}$ such that $y_p^\ell = 1$, and set $\tilde{y}_p = \frac{\sum_{k\in K_{sd}} t_k\, \bar{x}_p^k}{U}$, so as to satisfy wavelength capacity constraints, while $\tilde{y}_p^\lambda = \frac{\tilde{y}_p}{W}$. Observe that $\sum_{p\in\mathcal{P}} z_a^p \tilde{y}_p = \sum_{p\in\mathcal{P}} z_a^p \bar{y}_p$ $\forall a \in A$, since we have not changed the physical path traffic assignment. Hence, the wavelength clash constraints remain satisfied by solution $\tilde{y}$. Moreover, the demands remain satisfied and the total cost can only decrease since some intermediate electrical conversions have been removed. ∎

# 2 Dantzig-Wolfe Reformulations

Here, we consider two possible Dantzig-Wolfe decomposition approaches to the GRWA problem (also known as price or constraint decomposition [23]) under restrictive Assumptions 5 or 6. Note that formulation (6–11) expressed in terms of flows on lightpaths can itself be viewed as a reformulation resulting from a Dantzig-Wolfe decomposition. A compact "original" formulation in terms of flows on connectivity and physical arcs generalizing (14–23) could be derived and applying Dantzig-Wolfe reformulation to this original formulation would give (6–11). However, formulation (6–11) can be handled directly (it does not require the use of a column generation approach) because the path length constraint significantly restrict the number of feasible physical paths and lightpaths which can therefore be enumerated.

## 2.1 Grooming Pattern Formulation

Under Assumption 5, the grooming is restricted to simple **optical hop configurations** that are listed below and illustrated in Figure 2.

**Single-Hop configuration:** This is the simplest case, it consists in an optical hop $p \in \mathcal{P}_{sd}$, where the traffic is only accepted on the origin-destination pair $(s, d)$.

**Two-hop configuration:** It is composed of two optical hops $p_1 \in \mathcal{P}_{si}$ and $p_2 \in \mathcal{P}_{id}$, $(s, i)$ with distinct nodes $s$, $i$ and $d$. The grooming can only involve traffic on origin-destination pairs $(s, d)$, $(s, i)$ and $(i, d)$. The generation of two-hop configurations is done so as to guarantee that the path length constraint is satisfied. We select one of the three shortest $(s, d)$-paths and consider its internal nodes as potential intermediate node $i$, while checking if the corresponding physical paths $p_1$ and $p_2$ satisfy path length constraints.

**Three-hop splitting configuration:** It is composed of three optical hops $p_1 \in \mathcal{P}_{si}$, $p_2 \in \mathcal{P}_{id_1}$ and $p_3 \in \mathcal{P}_{id_2}$ with distinct nodes $s$, $i$, $d_1$ and $d_2$. The traffic can be groomed on origin-destination pairs $(s, d_1)$, $(s, d_2)$, $(s, i)$ $(i, d_1)$, and $(i, d_2)$. To guarantee that the path length constraints are satisfied, we generate a three-hop splitting configuration for a given pair of paths $(p_1 \in \mathcal{P}_{sd_1}, p_2 \in \mathcal{P}_{sd_2})$ that happens to split at an intermediate node $i$. We then verify that resulting induced sub-paths satisfy path length constraints.

**Three-hop merging configuration:** It is the symmetric of the previous case where the support is made of three optical hops $p_1 \in \mathcal{P}_{s_1 i}$, $p_2 \in \mathcal{P}_{s_2 i}$ and $p_3 \in \mathcal{P}_{id}$ with distinct nodes $s_1$, $s_2$, $i$ and $d$. The only traffic is on origin-destination pairs $(s_1, d)$, $(s_2, d)$, $(s_1, i)$ $(s_2, i)$, and $(i, d)$. The generation of three-hop merging configuration is similar to the generation of three-hop splitting configurations.

**Three-hop interlaced-lightpaths configuration:** It is composed of three optical hops $p_1 \in \mathcal{P}_{s_1 s_2}$, $p_2 \in \mathcal{P}_{s_2 d_1}$ and $p_3 \in \mathcal{P}_{d_1 d_2}$ with distinct nodes $s_1$, $s_2$, $d_1$ and $d_2$. The only traffic is on origin-destination pairs $(s_1, d_1)$, $(s_2, d_2)$, $(s_1, s_2)$ $(s_2, d_1)$, and $(d_1, d_2)$. The generation of three-hop interlaced-lightpaths configuration is similar to the generation of three-hop splitting configurations.

An optical hop configuration defines the way in which lightpaths can share their optical hops and hence the cost of installing O/E/O converter ports at nodes. It fixes the routing pattern but not the traffic load. A **grooming pattern** is defined for a fixed optical hop configuration by fixing the amount of traffic for each origin-destination pair concerned carried out by the optical hop configuration, in such a way that wavelength capacity $U$ is not exceeded. A global solution can then be expressed in terms of a grooming pattern selection whose total traffic meets the demand and for which there exists a feasible wavelength assignment avoiding clashes (we enforce the latter by selecting optical hop configurations that are arc disjoint). This reformulation arises from a relaxation of constraints (7–8) in formulation (6–11). Then, the solution of the remaining problem decomposes into independently routed traffic on grooming patterns. The optical hop configuration determines the way in which different traffic may interact with each other, but they are no interactions among grooming patterns. We have further restricted the solution space by limiting the number of interaction modes to the 5 optical hop configuration types presented in Section 2.1.

Let $\mathcal{O}$ denote the restricted set of optical hop configurations that has been pre-generated by enumeration from the shortest path lists. Each optical hop configuration $o \in \mathcal{O}$ is defined by an optical hop indicator vector $y^o$ with $y^o_p = 1$ if optical hop $p$ is used in the optical hop configuration $o$. Let $\mathcal{G}(o)$ denote the set of feasible **grooming patterns** build from optical hop configuration $o \in \mathcal{O}$ and $\mathcal{G} = \cup_{o \in \mathcal{O}} \mathcal{G}(o)$. Each grooming pattern $g \in \mathcal{G}$ is defined by a traffic flow and an optical hop indicator vector $(x^g, y^g)$ where $x^g_k$ gives the number of demands $k \in K$ that are routed over $g$ (note that due to the optical hop restrictions, very few components are positive), and $y^g_p = 1$ if optical hop $p$ is used in the optical hop configuration underlying

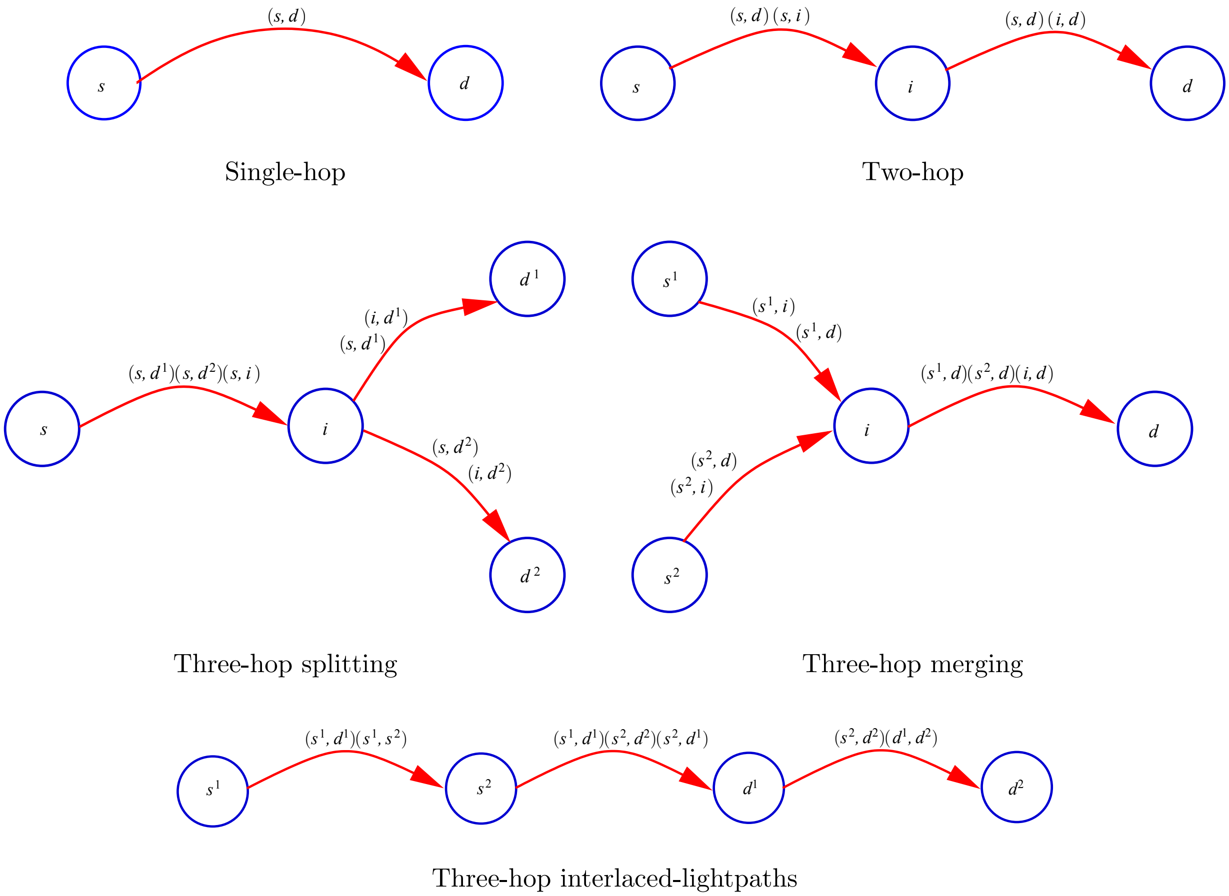


Figure 2: Optical hop configurations that define the supports of *grooming patterns*

grooming pattern $g$. For a fixed optical hop configuration $o \in \mathcal{O}$, $\mathcal{G}(o)$ admits a polyhedral description. For instance, for a two-hop configuration between nodes $s$ and $d$ with an O/E/O at node $i$,

$$\mathcal{G}(sid) = \{x \in I\!N^{3|T|} : \sum_{k \in K_{sd}} t_k \, x_k + \sum_{k \in K_{si}} t_k \, x_k \leq U; \sum_{k \in K_{sd}} t_k \, x_k + \sum_{k \in K_{id}} t_k \, x_k \leq U\}, \tag{29}$$

saying that the traffic must share the wavelength capacity $U$ on each optical hop (the full capacity, $U$, can be used because grooming patterns will be assigned a wavelength in such as way that the corresponding signal does not share links with any other signal). We further restrict traffic load to carry at least some traffic of each type, for otherwise the grooming pattern could be decomposed into simpler grooming patterns with possibly lower cost, and we do not load more traffic that the demand. In the example of the polyhedral description for the two-hop configuration $(s, i, d)$, we add the constraints: $\sum_{k \in K_{sd}} x_k^g \geq 1$, $\sum_{k \in K_{si}} x_k^g + \sum_{k \in K_{id}} x_k^g \geq 1$ and $x_k^g \leq D_k \quad \forall k \in K_{sd} \cup K_{si} \cup K_{id}$.

The reformulation of the GRWA problem in terms of variables $\mu_g$, that counts the number of times the grooming pattern $g \in \mathcal{G}$ is used, is as follows:

$$\min \sum_{g \in \mathcal{G}} \sum_{p \in \mathcal{P}} y_p^g \, \mu_g \tag{30}$$

$$\sum_{g \in \mathcal{G}} x_k^g \, \mu_g = D_k \quad \forall k \in K \tag{31}$$

$$\sum_{g \in \mathcal{G}} y_p^g \, \mu_g = \sum_{\lambda \in \Lambda} y_p^\lambda \quad \forall p \in \mathcal{P} \tag{32}$$

$$\sum_{p \in \mathcal{P}} z_a^p \, y_p^\lambda \leq 1 \quad \forall \lambda \in \Lambda, \, a \in A \tag{33}$$

$$y_p^\lambda \in \{0,1\} \quad \forall p \in \mathcal{P},\ \lambda \in \Lambda \tag{34}$$

$$\mu_g \in I\!N \quad \forall g \in \mathcal{G}. \tag{35}$$

The LP relaxation of this formulation can be solved by column generation. The pricing problem takes the form:

$$\min_{o\in\mathcal{O}}\{\sum_{p\in\mathcal{P}}(1-\rho_p)\, y_p^o - \max_{g\in\mathcal{G}(o)}\{\sum_{k\in K}\pi_k\, x_k^g\}\}, \tag{36}$$

where $(\pi, \rho)$ are the dual variables associated with constraints (31) and (32) respectively. For a fixed optical hop configuration, $o \in \mathcal{O}$, the pricing problem reduces to a loading problem subject to knapsack constraints. This problem can be solved in pseudo-polynomial time under the divisibility Assumption 2 (as presented in Section 4). The dual bound given by the LP relaxation is in theory better than that of formulation (6–11) because the subproblem captures the knapsack capacity constraints. However, the dual bound is subject to restrictive Assumption 5. The clash constraints remain in the master program and require the use of wavelength indexing on the $y$'s, leading to symmetry in the representation of the solutions. Moreover, enforcing integrality is not easy as it requires to fix both grooming pattern selection and optical hop variables. Hence, we shall not use this reformulation directly.

## 2.2 Wavelength Routing Configuration Formulation

Under the wavelength continuity Assumption 6, the routing on a given wavelength is independent from the routing on other wavelengths. Then, after relaxing the demand covering constraints in formulation (12–13), the problem decomposes into a subproblem for each wavelength whose solution defines the traffic carried by this wavelength and the associated optical hops that are used (as illustrated in Figure 1). The subproblem solutions shall be called *wavelength routing configurations*. If we further make the grooming restriction Assumption 5, the solution to this subproblem can itself be decomposed into arc disjoint grooming patterns. This leads to a nested decomposition approach.

Let $\mathcal{C}$ be the set of feasible wavelength routing configurations. Each configuration $c \in \mathcal{C}$ is defined by a grooming pattern indicator vector $\mu^c$, with $\mu_g^c = 1$ if grooming pattern $g$ is used. The polyhedral description of $\mathcal{C}$ is:

$$\mathcal{C} = \{\mu \in \{0,1\}^{|\mathcal{G}|} : \sum_{g\in\mathcal{G}}\sum_{p\in\mathcal{P}} z_a^p\, y_p^g\, \mu_g \le 1 \quad \forall a \in A;\ \sum_{g\in\mathcal{G}} x_k^g\, \mu_g \le D_k \quad \forall k \in K\}$$

where we make sure that the total traffic load does not exceed the demand for each commodity $k$. Alternatively, a wavelength routing configuration can be defined by the traffic load and the number of used optical hops: $(x^c, y^c) = (\sum_{g\in\mathcal{G}} x^g\, \mu_g^c, \sum_{g\in\mathcal{G}} y_p^g\, \mu_g^c)$. Then, the GRWA problem can be reformulated as:

$$\min \sum_{c\in\mathcal{C}}\sum_{p\in\mathcal{P}} y_p^c\, \nu_c \tag{37}$$

$$[\mathrm{WRC}] \quad \sum_{c\in\mathcal{C}} x_k^c\, \nu_c = D_k \quad \forall k \in K \tag{38}$$

$$\sum_{c\in\mathcal{C}} \nu_c \le W \tag{39}$$

$$\nu_c \in I\!N \quad \forall c \in \mathcal{C}, \tag{40}$$

where variable $\nu_c = 1$ if configuration $c \in \mathcal{C}$ is used. This **Wavelength Routing Configuration** reformulation eliminates the symmetry in the wavelength assignment. The LP relaxation of this formulation can be solved using a nested column generation approach where the pricing problem is itself solved by column generation. Wavelength routing configurations are generated by solving a pricing problem of the form:

$$\min\{\sum_{g\in\mathcal{G}}(\sum_{p\in\mathcal{P}} y_p^g - \sum_{k\in K}\pi_k\, x_k^g)\mu_g : \mu \in \mathcal{C}\} - \sigma, \tag{41}$$

where $(\pi, \sigma)$ are the dual variables associated with constraints (38) and (39), respectively. When solving problem (41) by column generation, the pricing sub-problem is:

$$\min_{o\in\mathcal{O}}\{\sum_{p} y_p^o - \max_{g\in\mathcal{G}(o)}\{\sum_{k\in K}\pi_k\, x_k^g\}\}. \tag{42}$$

The LP relaxation bound is in theory better than that given by grooming pattern formulation (30–35) as more constraints are included in the definition of $\mathcal{C}$ than $\mathcal{G}$. But this dual bound is only valid under two restrictive assumptions: Assumptions 5 and 6.

Note that other decompositions than the ones presented before are possible. For instance, after relaxing constraints (7) and (9), the problem decomposes into one topology design problem for each wavelength. The subproblem consists in finding paths that can be assigned to the same wavelength (it is a maximal weight stable set problem). Although the subproblem is not trivial, the master LP relaxation bound is typically the same as that of Proposition 2, because the wavelength clash constraints are fairly easy to satisfy given the large number of wavelengths that are available.

# 3 Benders' Decomposition and Hierarchical Optimization

Another form of decomposition is Benders' (also known as resource or variable decomposition [23]). One adopt an hierarchical approach fixing first the "important" decision variables that set the resource levels for the second stage problem.

## 3.1 Grooming and Physical Routing First, Wavelength Assignment Second

In formulation (6–11), when one fixes first the traffic flow decisions, $x_l^k$, and the aggregate decisions of establishing hops, $y_p = \sum_\lambda y_p^\lambda$, the GRWA problem is reduced to a wavelength assignment feasibility problem. The traditional Benders' reformulation approach consists in projecting (6–11) onto the space of the important variables. The so-called Benders' master program takes the form:

$$\min \sum_{p \in \mathcal{P}} y_p + \phi(y) \tag{43}$$

$$[\text{GPR}] \quad \sum_{\ell \in \mathcal{L}_{s_k d_k}} x_\ell^k \; = \; D_k \quad \forall k \in K \tag{44}$$

$$\sum_{k \in K} \sum_{\ell \in \mathcal{L}_{s_k d_k}} y_p^\ell \, t_k \, x_\ell^k \; \leq \; U \, y_p \quad \forall p \in \mathcal{P} \tag{45}$$

$$x_\ell^k \; \in \; I\!N \quad \forall k \in K, \, \ell \in \mathcal{L}_{s_k d_k} \tag{46}$$

$$y_p \; \in \; I\!N \quad \forall p \in \mathcal{P}, \tag{47}$$

where $\phi(y) = \infty$ if the optical hop establishment decisions $y$ cannot be associated with a wavelength assignment that satisfies clash constraints and zero otherwise. This master problem captures the **Grooming and Physical Routing (GPR)** decisions of the GRWA problem. Note that, having dropped the wavelength indexing, the master does not suffer the symmetry in wavelength assignment that was present in (6–11). To improve the model, one can add further necessary (but not sufficient) conditions for the existence of a feasible wavelength assignment:

$$\sum_{p \in \mathcal{P}} z_a^p \, y_p \leq W \quad \forall a \in A \tag{48}$$

that states that on each link there is at most $W$ optical hops. Given a feasible fixed master integer solution $(\bar{x}, \bar{y})$ to (43–47), the feasibility check entails solving a *wavelength assignment* (WA) sub-problem:

$$[\text{WA}] \quad \sum_{\lambda \in \Lambda} y_p^\lambda \; = \; \bar{y}_p \quad \forall p \in \mathcal{P} \tag{49}$$

$$\sum_{p \in \mathcal{P}} z_a^p \, y_p^\lambda \; \leq \; 1 \quad \forall \lambda \in \Lambda, \, a \in A \tag{50}$$

$$y_p^\lambda \; \in \; \{0, 1\} \quad \forall p \in \mathcal{P}, \, \lambda \in \Lambda. \tag{51}$$

This second stage problem captures the wavelength assignment decisions (WA) of the GRWA problem. Observe that the WA problem is not trivial (it does not reduce to an application of the flow decomposition theorem) because it involves multiple commodities (one for each hop $p$) that are linked by constraints (50).

Benders' approach entails a polyhedral characterization of the set $\{y : \phi(y) = 0\}$. The master is iteratively augmented with feasibility cuts when its solution $\bar{y}$ does not lead to a feasible second stage problem. When the second stage problem is an LP, Farkas Lemma can be invoked to obtain a cut in the $y$ variables from the dual solution to the feasibility subproblem when the latter is infeasible [23]. However, there is no such generic procedure when the second stage problem is an integer program as in our case (for examples of ad-hoc Benders' approaches with integer subproblems, see [10]). In this application, the sub-problem is either feasible (and optimality of the master is reached), or infeasible (and master feasibility cuts should be generated). Note that solving the LP relaxation of the sub-problem is of no use because it is always feasible (this can be proved using the same argument than in the proof of Proposition 2) and thus would never return feasibility cuts. In [12], a sufficient property on the set of selected optical hops for the feasibility of the WA problem has been presented. They might be used to generate Benders' feasibility cuts but we did not investigate on this issue. Hence, we only consider a relaxation where no feasibility cuts are added to the master. In other words, we ignore the term $\phi(y)$. Solving this relaxed problem provides a valid dual bound. When a primal solution to the first stage problem is found, we check whether the associate WA subproblem is feasible, otherwise we discard it. Thus, the Benders' approach considered here reduces to a hierarchical optimization procedure.

Benders' decomposition can also be applied to formulation (30–35), leading to the master problem:

$$\min \sum_{g\in\mathcal{G}} \sum_{p\in\mathcal{P}} y_p^g \mu_g + \phi(\mu) \tag{52}$$

$$[\text{BGP}] \quad \sum_{g\in\mathcal{G}} x_k^g\,\mu_g = D_k \quad \forall k \in K \tag{53}$$

$$\mu_g \in I\!N \quad \forall g \in \mathcal{G}, \tag{54}$$

that we call **Benders' Grooming Pattern** formulation. Let $\phi(\mu) = \infty$ if the decisions of establishing optical hops implied by solution $\mu$ cannot be associated with a wavelength assignment that satisfies clash constraints, and zero otherwise. For a fixed first stage solution $\bar{\mu}$, the second stage problem is again [WA] given in (49–51), where $\bar{y}_p = \sum_{g\in\mathcal{G}} y_p^g \bar{\mu}_g$.

## 3.2 Grooming and Virtual Routing First, Path and Wavelength Assignment Second

Note that, if we aggregate optical hops with the same end-nodes in formulation (43–47) without the term $\phi(y)$, we obtain formulation (24–28). Then, the second stage problem differs from [WA] given in (49–51). Because the physical routing of the optical hops is not fixed in the first stage problem, the second stage problem also involves finding a feasible physical path assignment for the aggregate lightpaths. A primal solution of the first stage problem (24–28) is often infeasible for the second stage problem. Instead, we make use of a relaxed two-step procedure to attempt to recover a feasible primal solution to the original problem. We start by solving the **Grooming and Path Assignment** problem:

$$\min \sum_{p\in\mathcal{P}} y_p \tag{55}$$

$$[\text{GPA}] \quad \sum_{p\in\mathcal{P}_{ij}} y_p \geq \bar{y}_{ij} \quad \forall (i,j) \in V^2 \tag{56}$$

$$\sum_{\ell\in\mathcal{L}_{s_k d_k}} x_\ell^k = D_k \quad \forall k \in K \tag{57}$$

$$\sum_{k\in K} \sum_{\ell\in\mathcal{L}_{s_k d_k}} y_p^\ell\, t_k\, x_\ell^k \leq U\, y_p \quad \forall p \in \mathcal{P} \tag{58}$$

$$x_\ell^k \in I\!N \quad \forall k \in K,\ \ell \in \mathcal{L}_{s_k d_k} \tag{59}$$

$$y_p \in I\!N \quad \forall p \in \mathcal{P}. \tag{60}$$

Constraints (56) set a lower bound on the number of optical hops that are established between each pair of nodes while other constraints are the same as in [GPR]. Note that we allow a modification of the routing of the flows to get more flexibility. Then, if problem (55–60) is feasible, we solve the feasibility check WA problem (49–51) in a second step.

In formulation [BGP] given in (52–54), one can also aggregate the optical hop configuration that are logically equivalent (same end nodes and O/E/O node) leading to the set $\tilde{\mathcal{O}}$ of aggregate optical hop configurations and associated set $\tilde{\mathcal{G}}$ of aggregate grooming patterns. This relaxation leads to formulation:

$$\min \sum_{g \in \tilde{\mathcal{G}}} \sum_{p \in \mathcal{P}} y_p^g \mu_g + \phi(\mu) \tag{61}$$

$$\text{[BGPA]} \quad \sum_{g \in \tilde{\mathcal{G}}} x_k^g \, \mu_g = D_k \quad \forall k \in K \tag{62}$$

$$\mu_g \in I\!N \quad \forall g \in \tilde{\mathcal{G}}. \tag{63}$$

Given a fixed master integer solution $\bar{\mu}$ to (61–63), $\phi(\bar{\mu}) = 0$ if there exists associated physically routed grooming configurations and feasible wavelength assignment. Let $\bar{\iota}_o = \sum_{g \in \tilde{\mathcal{G}}(0)} \bar{\mu}_g$ be the number of times the aggregate optical hop configuration $o \in \tilde{\mathcal{O}}$ is used. The feasibility check sub-problem takes the form of a *grooming pattern and wavelength assignment* (GPWA) problem:

$$\text{[GPWA]} \quad \sum_{\lambda \in \Lambda} y_p^\lambda = \sum_{o \in \mathcal{O}} y_p^o \iota_o \quad \forall p \in \mathcal{P} \tag{64}$$

$$\sum_{o' \in \mathcal{O} : o' \equiv o} \iota_{o'} = \bar{\iota}_o \quad \forall o \in \tilde{\mathcal{O}} \tag{65}$$

$$\sum_{p \in \mathcal{P}} z_a^p \, y_p^\lambda \leq 1 \quad \forall \lambda \in \Lambda, \, a \in A \tag{66}$$

$$\iota_o \in I\!N \quad \forall o \in \tilde{\mathcal{O}} \tag{67}$$

$$y_p^\lambda \in \{0,1\} \quad \forall p \in \mathcal{P}, \, \lambda \in \Lambda. \tag{68}$$

where $o' \equiv o$ means that optical hop configuration, $o' \in \mathcal{O}$, is virtually equivalent to aggregate optical hop configuration, $o \in \tilde{\mathcal{O}}$. Constraints (64) states that the number of wavelengths reserved for physical path $p$ depends on the selected optical hop configurations. Constraints (65) enforce a selection of optical hop configurations that reproduce the connectivity selection.

# 4 Implementation Features

The above models share some similarities. In this section, we present some formulation strengthening through partial reformulation, variable domain reduction, an extended 0-1 formulation, and valid inequalities that can be useful for several models. They can result in better LP bounds, eliminate the symmetries resulting from bandwidth re-assignment as outlined in Observation 1, bring stability in the solution process, or help our primal heuristics. We also detail the pricing procedures for the column generation approaches.

## 4.1 Enhanced Demand Covering Constraints

Due to the divisibility of the granularities, they can exist many ways to cover the overall demand on a given pair of end-nodes. For instance, an aggregate demand of OC $-$ 48 can be covered by either a OC $-$ 48 bandwidth reservation or four reservations of OC $-$ 12, etc. More formally, given Assumption 2, for each $t \in T$, $t' \in T$, such that $t \geq t'$, we have $\frac{t}{t'} \in I\!N$. Hence, if we have reserved $t$ units of capacity on a lightpath for a demand $k \in K_{sdt}$, this capacity may be used to either route demand $k$ or, equivalently, $\frac{t}{t'}$ demands $k' \in K_{sdt'}$. These symmetric representations of the same solution and the resulting instability in the solution process can be avoided by reformulating demand covering constraints.

The idea is to build granularity exchanges in the formulation. One simply needs to give another meaning to $x_\ell^k$. It represents the number of bandwidths streams of capacity $t_k$ that are reserved for $(s_k, d_k)$-traffic on lightpath $\ell \in \mathcal{L}$. This capacity reservation can be used to cover any $(s_k, d_k)$-traffic with granularity $t_k$

or lower. Then, the demand constraints can be aggregated. The capacity reservation for $(s_k, d_k)$-traffic of granularity $t_k$ or higher must be sufficient to cover the associated demands, but not greater than the total $(s_k, d_k)$-demand converted in granularity $t_k$. Thus, demand covering constraints (7) can be replaced by:

$$\underbrace{\left\lfloor \sum_{k' \in K_{s_k d_k}} \frac{t_{k'}}{t_k} D_{k'} \right\rfloor}_{\mathrm{LHS}_k} \geq \sum_{\ell \in \mathcal{L}_{s_k d_k}} \sum_{\substack{k' \in K_{s_k d_k}: \\ t_{k'} \geq t_k}} \frac{t_{k'}}{t_k} x_\ell^{k'} \geq \underbrace{\sum_{\substack{k' \in K_{s_k d_k}: \\ t_{k'} \geq t_k}} \frac{t_{k'}}{t_k} D_{k'}}_{\mathrm{RHS}_k} \quad \forall k \in K. \tag{69}$$

Along the lines of the bandwidth re-assignment Observation 1, we note that

**Observation 2** *Under Assumption 2, any solution to (6, 69, 8–11) can be transformed into a solution of formulation (6–11) by disaggregating granularity reservations that exceed their associated demands in order to cover lower granularity demands.*

## 4.2 Domain Reduction for Bandwidth Reservations

For bandwidth reservation variables $x_\ell^k$ and their aggregate value $x^k = \sum_\ell x_\ell^k$, we define both lower and upper bounds, denoted $\underline{b}_k^\ell$ and $\overline{b}_k^\ell$, $\underline{b}_k$ and $\overline{b}_k$, respectively, with

$$\underline{b}_k^\ell \leq \underline{b}_k \quad \text{and} \quad \overline{b}_k^\ell \leq \overline{b}_k \quad \forall k \in K,\ \ell \in \mathcal{L}_{s_k d_k}. \tag{70}$$

Moreover,

$$\underline{b}_k \geq \sum_{\ell \in \mathcal{L}_{s_k d_k}} \underline{b}_k^\ell \text{ and } \overline{b}_k \leq \sum_{\ell \in \mathcal{L}_{s_k d_k}} \overline{b}_k^\ell \quad \forall k \in K. \tag{71}$$

At the outset, the lower bounds are set to 0 and the upper bounds to $\infty$. Strengthening one of these bounds (as those presented below) may imply strengthening the others through relations (70–71). We also define a domain of discrete values with the same index convention. The default domain for $x_\ell^k$ is $Q_k^\ell = \{\underline{b}_k^\ell, \ldots, \overline{b}_k^\ell\}$ and similarly for $Q_k$.

Tightening these bounds allows us to strengthen the formulation and to eliminate some symmetries resulting from the bandwidth re-assignment property of Observation 1.

**Observation 3 (selected representative of the symmetry class of bandwidth reservations on an optical hop)** *On a given optical hop, we assume, w.l.o.g., that at least $y_p - 1$ wavelengths are used at their full capacity $U$. Moreover, we define the capacity reservation in the largest possible granularities.*

Hence, we may forbid a capacity reservation of $x_\ell^k \geq \frac{t_k^+}{t_k}$ units of $t_k$, where $t_k^+$, defined for $t_k \in T$, is the successor of $t_k$ in the set $T \cup \{U\}$ sorted by increasing granularities: $t_k^+ = \min\{t \in T \cup \{U\} : t > t_k\}$. Indeed, we rather use the symmetric solution where a global capacity reservation $t_k^+$ is made instead of the $\frac{t_k^+}{t_k}$ reservations of capacity $t_k$.

An interesting special case arises when a physical path $p$ can only be used by a single $(s, d)$ commodity:

**Observation 4 (Bandwidth reservation on a dedicated single-hop lightpath)** *If $p \in \mathcal{P}_{sd}$ while for all $(i, j) \neq (s, d)$ and all $\ell \in \mathcal{L}_{ij}$, $y_p^\ell = 0$, then, we can restrict our attention to solutions where the total flow on optical hop $p$, $x_p$, satisfies $x_p \bmod U = 0$ or $x_p = D_{sd} \bmod U$.*

Another symmetry breaking restriction consists in bounding the capacity reservation level on multi-hop lightpaths:

**Observation 5 (non-degenerated multi-hops)** *The maximum $(s, d)$ reservation routed over a multi-hop lightpath can be restricted, w.l.o.g., to be strictly lower than $U$.*

Indeed, we can remove $U$ units of $(s,d)$ flow from a multi-hop lightpath and route it on a single-hop lightpath without increasing the cost of the solution.

Such considerations allow us to refine capacity reservation bounds. Let $t_{sd}^{\max}$ be the largest $t \in T$ for which $D_{sdt}$ is positive: $t_{sd}^{\max} = \max\{t \in T : D_{sdt} > 0\}$. Then, given $k \in K,\ \ell \in \mathcal{L}_{s_k d_k}$, we can derive valid upper bounds:

$$\bar{b}_k^\ell \leq \frac{t_k^+}{t_k} - 1 \text{ if } t_k < t_{s_k d_k}^{\max} \tag{72}$$

$$\bar{b}_k^\ell \leq \frac{U}{t_k} - 1 \text{ if } \ell \text{ is a multiple-hop lightpath} \tag{73}$$

$$\bar{b}_k = \left\lfloor \sum_{\substack{k' \in K_{s_k d_k}: \\ t_{k'} \leq t_k}} \frac{t_{k'}}{t_k} D_{k'} \right\rfloor . \tag{74}$$

Bound (72) specifies that, when $t_k < t_{s_k d_k}^{\max}$ the capacity reserved by $k$ must be strictly lower than $\frac{t_k^+}{t_k}$ following Observation 3. Bound (73) considers the case of a multi-hop lightpath where the capacity reservation cannot cover the wavelength capacity $U$ as stated in Observation 5. Bound (74) specifies that a capacity reservation cannot exceed the cumulative demand for that granularity and lower granularities. Regarding lower bounds, we can state:

$$\underline{b}_k = D_k, \text{ if } t_k = t_{s_k d_k}^{\max}; \tag{75}$$

$$\underline{b}_k = \max\left\{0, D_k - (\mathrm{LHS}_{k^+} - \mathrm{RHS}_{k^+})\frac{t_k^+}{t_k}\right\}, \text{ if } t_k < t_{s_k d_k}^{\max},\ k^+ = (s_k, d_k, t_k^+). \tag{76}$$

Aggregate bound (75) imposes that the number of capacity reservation of granularity $t_k$ is greater than the $(s_k, d_k, t_k)$-demand when it cannot be met with larger granularities. Aggregate bound (76) imposes that the minimum number of capacity reservations for granularity $t_k$ is equal to the $(s_k, d_k, t_k)$-demand minus the surplus of larger granularities expressed in $t_k$ units. We use the tightest of these upper and lower bounds and propagate them through relations (70–71).

Another tightening can come from preprocessing the demand vector. Indeed, observe that if, for $k \in K$ such that $t_k < t_k^{\max}$, we have $D_k > \bar{b}_k$, then some fraction of the demand $k$ must be covered using larger granularities. Hence, $\left\lceil \frac{t_k(D_k - \bar{b}_k)}{t_k^+} \right\rceil$ can be added to demand $(s_k, d_k, t_k^+)$ and $\frac{t_k^+}{t_k}\left\lceil \frac{t_k(D_k - \bar{b}_k)}{t_k^+} \right\rceil$ units of demand can be removed from demand $k$. We perform these preprocessing operations for each $(s,d) \in V^2$, starting with the smallest granularity. Table 1 gives an example of demand and the associated bounds, as well as left and right hand-side values of the aggregate demand constraints (69). Note that all these bounds are also valid when dealing with aggregate lightpath set $\tilde{L}$ used in virtual routing formulation (24–28). (73) are valid.

### 4.3 Zero-One Extended Formulation

A 0-1 extended formulation for (6–11) enables us to tighten the model with valid inequalities that have been proposed for capacitated network design problems in the binary case [3]. We shall use these tightening when

Table 1: An example of demand vector for different granularities and associated capacity reservation bounds

| $t_k$ | $D_k$ | $\bar{b}_k$ | $\underline{b}_k$ | single-hop lightpath $\bar{b}_k^\ell$ | multi-hop lightpath $\bar{b}_k^\ell$ | $\mathrm{RHS}_k$ | $\mathrm{LHS}_k$ |
|---|---|---|---|---|---|---|---|
| 48 | 5 | 7 | 5 | 7 | 3 | 5 | 7 |
| 12 | 3 | 8 | 0 | 3 | 3 | 23 | 28 |
| 3 | 6 | 21 | 0 | 3 | 3 | 98 | 113 |
| 1 | 47 | 47 | 2 | 2 | 2 | 341 | 341 |

dealing with model [VR] given in (24–28) and model [GPR] given in (43–47). The 0-1 reformulation arises from the unary decomposition of capacity reservation variables $x_\ell^k$. We implement the change of variables:

$$x_\ell^k = \sum_{q \in Q_k^\ell} q\, x_{\ell q}^k \text{ with } \sum_{q \in Q_k^\ell} x_{\ell q}^k = 1,\ x_{\ell q}^k \in \{0,1\}\ \forall q \in Q_k^\ell, \tag{77}$$

where $x_{\ell q}^k = 1$ if a capacity reservation of $q$ units of granularity $t_k$ is made over lightpath $\ell \in \mathcal{L}_{s_k d_k}$. Observe that it is sufficient to consider $q \in Q_k^\ell \setminus \{0\}$, as $x_\ell^k = 0$ can be achieved by setting all $x_{\ell q}^k = 0$. In the sequel, when dealing with 0-1 transformations, we shall consider that 0 has therefore been omitted from the value domain $\tilde{Q}_k^\ell = Q_k^\ell \setminus \{0\}$ and $\sum_{q \in \tilde{Q}_k^\ell} x_{\ell q}^k \leq 1$.

Note that for a single-hop aggregate lightpath $\ell \in \tilde{\mathcal{L}}$ and $t_k = t_{s_k d_k}^{\max}$, $\overline{b}_k^\ell$ can be quite large as it is not constrained by bounds (72) or (73). To avoid dealing with too many $x_{\ell q}^k$ variables in such a case, we introduce a refined variable decomposition. If on a single-hop lightpath $x_\ell^k = q$ with $q \bmod \frac{U}{t_k} = 0$, there are exactly $\frac{q\, t_k}{U}$ wavelengths for the single-hop lightpath that are fully loaded at capacity $U$ and are fully dedicated to this bandwidth reservation. The idea is to count these dedicated single-hop wavelength-assigned lightpaths apart and to define $x_\ell^{ak}$ as the residual reservation. Then, $x_\ell^k = x_\ell^{ak} + x_\ell^{bk}$ with

$$x_\ell^{ak} \in Q_k^{a\ell} = \{0, \ldots, \frac{U}{t_k} - 1\} \text{ and } x_\ell^{bk} \in Q_k^{b\ell} = \{0, \frac{U}{t_k}, 2\frac{U}{t_k}, \ldots, \left\lfloor \frac{\overline{b}_k^\ell\, t_k}{U} \right\rfloor \frac{U}{t_k}\}. \tag{78}$$

We then apply the classical unary decomposition (77) to $x_\ell^{ak}$ and $x_\ell^{bk}$ separately. Note that we can omit the subscript $a$ or $b$ and note $x_{q\ell}^k$ for $q \in Q_k^{a\ell} \cup Q_k^{b\ell}$ without any confusion as $Q_k^{a\ell} \cap Q_k^{b\ell} = \{0\}$, i.e., the 0-1 transformation is $x_\ell^k = \sum_{q \in Q_k^{a\ell} \cup Q_k^{b\ell}} q x_{q\ell}^k$, $\sum_{q \in Q_k^{a\ell} \cup Q_k^{b\ell}} q x_{q\ell}^k = 1$. With this refined decomposition, and given the divisibility Assumption 2, where $U$ and $T$ are fixed, the above defined unary decompositions are technically polynomial. Note that for a single-hop lightpath $\ell \in \mathcal{L}_{sd}$ such that its underlying optical hop $p \in \mathcal{P}_{sd}$ fits the case of Observation 4, $\tilde{Q}_k^{a\ell}$ is either empty or a singleton.

## 4.4 Valid Inequalities

We present valid inequalities of the literature that we have found useful for the GRWA problem. Their linearization and dynamic generation is specific to the model in which we use them (it shall be presented in Section 5 where we discuss the practical solution procedure for the four models that we selected).

The **cut set** inequalities are well-known valid inequalities for the network loading problem [2, 7, 16]. There is an exponential number of such inequalities, but we restrict our attention to a polynomial subset (that is the most helpful computationally) that sets lower bounds on the number of outgoing and incoming optical hops for each node:

$$\sum_{d \in V \setminus \{s\}} \sum_{p \in \mathcal{P}_{sd}} y_p \geq \left\lceil \frac{\sum_{d \in V \setminus \{s\}} D_{sd}}{U} \right\rceil \quad \forall s \in V\ , \quad \sum_{s \in V \setminus \{d\}} \sum_{p \in \mathcal{P}_{sd}} y_p \geq \left\lceil \frac{\sum_{s \in V \setminus \{d\}} D_{sd}}{U} \right\rceil \quad \forall d \in V. \tag{79}$$

Beyond the cut set inequalities, one can derive lower bounds on the number of wavelengths reserved over all lightpaths $\ell \in \mathcal{L}_{sd}$ that need to be setup for each $(s,d)$ pair. These **$(s,d)$-lightpath** cuts are expressed as non linear constraints for now:

$$\sum_{\ell \in \mathcal{L}_{sd}} \left\lceil \frac{\sum_{k \in K_{sd}} t_k\, x_\ell^k}{U} \right\rceil \geq \left\lceil \frac{D_{sd}}{U} \right\rceil \quad \forall (s,d) \in V^2. \tag{80}$$

The model can be refined by counting separately the number of wavelengths reserved for each granularity $t_k$. The **$k$-lightpath cuts** take the form:

$$\sum_{\ell \in \mathcal{L}_{sd}} \left\lceil \frac{t_k\, x_\ell^k}{U} \right\rceil \geq \left\lceil \frac{t_k\, b_k}{U} \right\rceil \quad \forall k \in K. \tag{81}$$

Note that for $k$ such that $t_k < t^{\max}_{s_k d_k}$ and $\underline{b}_k > 0$, $\left\lceil \frac{t_k\, \underline{b}_k}{U} \right\rceil = 1$.

We also consider cuts that enforce upper bounds on the number of lightpaths with high bandwidth reservation. Let

$$\overline{\alpha}_{sd} = \min\{u \in [1,U] : u \times (1 + \max\{\left\lfloor \frac{D_{sd}}{U} \right\rfloor, 1\}) > D_{sd}\} = \left\lfloor \frac{D_{sd}}{\max\{\left\lfloor \frac{D_{sd}}{U} \right\rfloor, 1\} + 1} \right\rfloor, \tag{82}$$

be the largest "average" $(s,d)$-bandwidth that could be reserved on a lightpath. Then, all lightpaths cannot have higher than average bandwidth reservation. Hence, we define the following **bandwidth reservation upper bound cuts**:

$$\sum_{\ell \in \mathcal{L}_{sd}} (\left\lfloor \frac{\sum_{k \in K_{sd}} t_k\, x^k_\ell}{U} \right\rfloor + \delta((\sum_{k \in K_{sd}} t_k\, x^k_\ell) \bmod U > \overline{\alpha}_{sd})) \leq \max\{\left\lfloor \frac{D_{sd}}{U} \right\rfloor, 1\} \quad \forall (s,d) \in V^2, \tag{83}$$

where $\sum_{\ell \in \mathcal{L}_{sd}} \left\lfloor \frac{\sum_{k \in K_{sd}} t_k\, x^k_\ell}{U} \right\rfloor$ counts the number of wavelengths at full capacity dedicated to $(s,d)$ traffic and $\delta(x_\ell \bmod U > \overline{\alpha}_{sd}) = 1$ if extra wavelength is used for lightpath $\ell$ to carry larger than average bandwidth.

The 0-1 transformation allows us to enforce tighter relations between flow (bandwidth reservation) and design variables. The so-called **strong linking** inequalities are standard in network design problems with binary variables. They state that if flow variable $x$ uses a link, the latter must be setup as measured by the associated design variable $y$. In our model, design variables, $y_p = \sum_{\lambda \in \Lambda} y^\lambda_p$, are not binary but general integer, hence strong linking inequalities take the form of GUB constraints:

$$\sum_{q \in \tilde{Q}^\ell_k} \left\lceil \frac{q\, t_k}{U} \right\rceil x^k_{\ell q} \leq y_p \quad \forall k \in K,\, p \in \mathcal{P},\, \ell \in \mathcal{L}_{s_k d_k} : y^\ell_p = 1. \tag{84}$$

One can also derive Gomory cuts and cover cuts from knapsack type constraints. For instance, from the flow bound constraints $\underline{b}_k \leq \sum_{\ell \in \mathcal{L}_{s_k d_k}} \sum_{q \in \tilde{Q}^\ell_k} q\, x^k_{\ell q} \leq \overline{b}_k$ we derive Gomory cuts, so-called **$k$-demand cuts**, of the form

$$\sum_{\ell \in \mathcal{L}_{s_k d_k}} \sum_{q \in \tilde{Q}^\ell_k} \left\lfloor \frac{q}{\epsilon} \right\rfloor x^k_{\ell q} \leq \left\lfloor \frac{\overline{b}_k}{\epsilon} \right\rfloor \quad \text{and} \quad \sum_{\ell \in \mathcal{L}_{s_k d_k}} \sum_{q \in \tilde{Q}^\ell_k} \left\lceil \frac{q}{\epsilon} \right\rceil x^k_{\ell q} \geq \left\lceil \frac{\underline{b}_k}{\epsilon} \right\rceil, \tag{85}$$

for all $k \in K$ and integer $\epsilon \in \{1, \ldots, \max_{\ell \in \mathcal{L}_{s_k d_k}} \overline{b}^\ell_k\}$. Similarly, we can derive Gomory cuts from the enhanced demand constraints (69). These **aggregated $k$-demand cuts** take the form

$$\sum_{\ell \in \mathcal{L}_{s_k d_k}} \sum_{\substack{k' \in K_{s_k d_k} : \\ t_{k'} \geq t_k}} \sum_{q \in \tilde{Q}^\ell_{k'}} \left\lfloor \frac{q\, t_{k'}}{\epsilon\, t_k} \right\rfloor x^q_{\ell q} \leq \left\lfloor \frac{\mathrm{LHS}_k}{\epsilon} \right\rfloor \text{ and } \sum_{\ell \in \mathcal{L}_{s_k d_k}} \sum_{\substack{k' \in K_{s_k d_k} : \\ t_{k'} \geq t_k}} \sum_{q \in \tilde{Q}^\ell_{k'}} \left\lceil \frac{q\, t_{k'}}{\epsilon\, t_k} \right\rceil x^q_{\ell q} \geq \left\lceil \frac{\mathrm{RHS}_k}{\epsilon} \right\rceil, \tag{86}$$

for all $k \in K : t_k \neq t^{\max}_{s_k d_k}$, $\overline{b}_k > D_k$ or $\overline{b}_k \geq \frac{t^+_k}{t_k}$ and all integer $\epsilon \in \{2, \ldots, \max_{\ell \in \mathcal{L}_{s_k d_k}} \overline{b}^\ell_k\}$ (the conditions on $k$ are there to avoid generating cuts identical to (85)). We also use **Lifted Knapsack Cover** and **$\epsilon$-split and $c$-strong** inequalities proposed by [3]. Details and further inequalities are presented in [24].

### 4.5 Optimizing Bandwidth Reservation on a Grooming Pattern

The pricing problem (36) or (42) relies on a core sub-problem consisting in making optimal bandwidth reservation for a fixed single-hop lightpath. Given the symmetry breaking restrictions outlined in Observation 3, there is a unique way to reserve a total bandwidth $b \leq U$ for a given $(s,d)$-commodity. We further restrict bandwidth reservation to ensure that the resulting grooming patterns define so-called *proper columns*, i.e., a pattern that do not carry more traffic than the aggregate demand bound defined in (69). Generating proper columns is known to strengthen the column generation formulation as shown in [22]. Thus, given a lightpath $\ell$ on a $(s,d)$ pair, a bandwidth reservation $b \leq U$ is feasible if the following polyhedron is not empty:

$$X_{sd}(b) = \{x_k \in I\!N^{|K_{sd}|} : \sum_{k \in K_{sd}} t_k\, x_k = b \tag{87}$$

$$\sum_{k' \in K_{sd}: t_{k'} \geq t_k} \frac{t_{k'}}{t_k}\, x_{k'} \leq \mathrm{LHS}_k \quad \forall k \in K_{sd} \tag{88}$$

$$x_k \leq \frac{t_k^+}{t_k} - 1 \quad \forall k \in K_{sd} : t_k \neq t_{sd}^{\max}\}. \tag{89}$$

If $X_{sd}(b)$ is not empty, it admits a unique solution. This feasibility problem can be solved in $O(|T|)$, after sorting the items in the decreasing order of their granularities, using a first fit decreasing procedure.

Then, given the reward vector $\pi$, computed from the dual variables associated with demand constraints (69), the function

$$\psi^{sd} : u \in \{0, \ldots, U\} \to \psi^{sd}(u) = \max\{\sum_{k \in K_{sd}} \pi_k\, x_k : 0 \leq b \leq u,\ X_{sd}(b) \neq \emptyset,\ x \in X_{sd}(b)\}, \tag{90}$$

can be computed in $O(U|T|)$. Now, for any optical hop configuration, $o \in \mathcal{O}$, one can determine the optimal traffic loading, $\psi_o = \max_{g \in \mathcal{G}(o)}\{\sum_{k \in K} \pi_k x_k^g\}$, using the solutions of the pre-computed function $\psi^{sd}$ as follows.

**For a $(s,d)$-single-hop configuration,** $\psi_o = \psi^{sd}(\min\{D_{sd}, U\})$ is computed in $O(1)$.

**For a $(s,i,d)$-two-hop configuration,** $\psi_o = \max_{1 \leq u \leq \min\{D_{sd},U\}}\{\psi^{sd}(u) + \psi^{si}(U-u) + \psi^{id}(U-u)\}$ is computed in $O(U)$.

**For a $(s_1,s_2,i,d)$-three-hop merging configuration,**

$$\psi_o = \max\{\psi^{s_1 d}(u_1) + \psi^{s_2 d}(u_2) + \psi^{id}(U - u_1 - u_2) + \psi^{s_1 i}(U - u_1) + \psi^{s_2 i}(U - u_2) :$$
$$1 \leq u_1 \leq \min\{D_{s_1 d}, U-1\}, 1 \leq u_2 \leq \min\{D_{s_2 d}, U-1\}, u_1 + u_2 \leq U\}$$

is computed in $O(U^2)$.

**For three-hop splitting configuration and three-hop interlaced-lightpaths,** the computation are equivalent to that of a three-hop merging configuration.

## 4.6 Generating Wavelength Configuration

The pricing problem of wavelength routing configuration that is given in (41) can be formulated in terms of grooming patterns. For each optical hop configuration, $o \in \mathcal{O}$, consider the optimal grooming pattern, $g_o^\star$, with minimum reduced cost, $\bar{c}_g = \sum_p y_p^g - \sum_{k \in K} \pi_k\, x_k^g$, only if $\bar{c}_g < 0$. Then, wavelength routing configuration pricing problem (41) can be written as:

$$\min\{\sum_{o \in \mathcal{O}} \bar{c}_{g_o^\star}\, \mu_{g_o^\star} : \sum_{o \in \mathcal{O}} \sum_{p \in \mathcal{P}} z_a^p\, y_p^o\, \mu_{g_y^\star} \leq 1 \quad \forall a \in A\}, \tag{91}$$

where the constraints formulate wavelength clashes on physical arcs. It is a maximum weight stable set problem in a conflict graph where nodes represent the optimal grooming patterns that are linked by a conflict edge if they share a physical arc. In our computational experiments it can be solved in reasonable time using a MIP solver.

In practice, we consider a tighter pricing problem that yield *proper columns*, i.e., columns that do not carry an aggregate traffic higher than the aggregate demand bound $\mathrm{LHS}_k$ defined in (69). If the solution of (91) does not satisfy the proper column bound constraints:

$$\sum_{k' \in K_{s_k d_k}: t_{k'} \geq t_k} \frac{t_{k'}}{t_k}\, x_{k'}^c \leq \mathrm{LHS}_k \quad \forall k \in K, \tag{92}$$

we have to consider alternative grooming patterns that do not correspond to the minimum possible reduced cost for a given optical hop configuration. We have developed two methods for solving the wavelength routing configuration pricing problem with traffic bounds (92): a greedy heuristic and an exact method making use of an in-situ column generation technique [17].

The greedy approach works as follows. At each iteration, we keep a list of the most negative reduced cost "*proper*" grooming pattern for each hop configuration (if any). Then, we select the optical hop configuration solution, $\arg\min_{o\in\mathcal{O}}\{\bar{c}_{g_o^\star}\}$, and we update the solution: (*i*) We mark the arcs used by the selected optical hop configuration, (*ii*) we update the demand vector according to the flow on the selected grooming pattern $g_o^\star$, and (*iii*) we update the list of optical hop configurations by regenerating the optimal "*proper*" grooming pattern for each $o \in \mathcal{O}$ for which the current optimum carries a traffic exceeding the modified bounds, or uses arcs that are marked. The greedy algorithm stops when all arcs are marked or there is no more optical hop configurations such that $\bar{c}_{g_o^\star} < 0$. When selecting the best optical hop configuration solution, $\arg\min_{o\in\mathcal{O}}\{\bar{c}_{g_o^\star}\}$, we first test if not selecting it frees up arcs that can then be used by a subset of other patterns leading to a better reduced cost; if so, we move on to the next optical hop configuration.

The in-situ column generation is a method to generate columns (grooming patterns in our case) directly in the master program (which is defined by the wavelength configuration pricing problem in our case) [17]. Note that, for a given $o \in \mathcal{O}$, an alternative to $g_o^\star$ must be considered only if $g_o^\star$ carries traffic that is involved in some violated proper column constraints (92). Let $\bar{K}$ be the set of demands for which constraints (92) are violated. Let $\tilde{\mathcal{O}} \subset \mathcal{O}$ be the optical configurations for which $g_o^\star$ is involved in some violated constraints (92), and $\bar{\mathcal{O}} = \mathcal{O} \setminus \tilde{\mathcal{O}}$. For each $o \in \tilde{\mathcal{O}}$, we define an indicator variable $\iota_o = 1$ if we chose to use a grooming pattern (defined by $x_k^o$) for this optical hop configuration. Then, the wavelength configuration pricing problem can be written as:

$$\min \sum_{o\in\bar{\mathcal{O}}} \bar{c}_{g_o^\star}\, \mu_{g_o^\star} + \sum_{o\in\tilde{\mathcal{O}}} \Big(\sum_{p} y_p^o \iota_o - \sum_{k\in K} \pi_k\, x_k^o\Big) \tag{93}$$

$$\sum_{o\in\bar{\mathcal{O}}} \sum_{p\in\mathcal{P}} z_a^p\, y_p^o\, \mu_{g_o^\star} + \sum_{o\in\tilde{\mathcal{O}}} \sum_{p\in\mathcal{P}} z_a^p\, y_p^o \iota_o \leq 1 \quad \forall a \in A \tag{94}$$

$$\sum_{o\in\tilde{\mathcal{O}}} \sum_{k'\in K: t_{k'} \geq t_k} \frac{t_{k'}}{t_k} x_k^o \leq \mathrm{LHS}_k \quad \forall k \in \bar{K} \tag{95}$$

$$\iota_o \in \{0,1\} \quad \forall o \in \tilde{\mathcal{O}} \tag{96}$$

$$(x_k^o, \iota_o) \in \mathcal{G}(o) \quad \forall k \in K,\ o \in \tilde{\mathcal{O}} \tag{97}$$

$$\mu_{g_o^\star} \in \{0,1\} \quad \forall o \in \bar{\mathcal{O}}. \tag{98}$$

The mathematical programming formulation involves replacing $(x_k^o, \iota_o) \in \mathcal{G}(o)$ by the polyhedral description of $\mathcal{G}(o)$ as given in a specific example in (29) where bound constraints are transformed into variable bounds implying setup variable $\iota_o$.

# 5 Selected Models and Algorithms

Four of the above models have been found tractable for computational purposes. We now describe briefly the overall algorithm used for each of them. We discuss how the elements of the previous section are used in combination and the selected implementation strategies.

## 5.1 A Cutting Plane Approach for the Virtual Routing Formulation

The Virtual Routing formulation (24–28) is denoted by [VR]. It models a relaxation of the GRWA problem, considering only the first stage of the Bender's approach of Section 3.2 that involves grooming and virtual routing decisions. Its LP optimum provides a valid dual bound, although it is the trivial bound of Proposition 2. We tackle (24–28) directly by branch-and-bound using CPLEX with default cut generation and primal heuristics. We also test the impact of the cut set constraints (79).

A variant denoted by [VR 0-1] is to consider the 0-1 form of (24–28), making use of the unary decomposition presented in Section 4.3 and strengthening it with the valid inequalities of Section 4.4. Observe that, following the variable split (78), when $x^{bk}_{\ell q} = 1$ for $q \in Q^{b\ell}_k$, we can count the $\frac{q\,t_k}{U}$ dedicated single-hops directly in the cost function and write the capacity constraints only in terms of the $x^{ak}_{\ell q}$ variables, $q \in Q^{a\ell}_k$. This in turn leads to a smaller domain for the $y_{ij}$. Hence, the valid inequalities presented next in Section 4.4, shall typically be more efficient (in particular strong linking inequalities [1] for a $0-1$ reformulation of the bifurcated capacitated network design problem using strong linking inequalities). The resulting formulation is:

$$\min \sum_{(i,j)\in V^2} y_{ij} + \sum_{k\in K} \sum_{\ell \in \tilde{\mathcal{L}}_{s_k d_k}} \sum_{q\in \tilde{Q}^{b\ell}_k} \frac{q\,t_k}{U}\, x^{bk}_{\ell q} \tag{99}$$

$$\text{[VR 0-1]} \quad \text{LHS}_k \geq \sum_{\substack{k'\in K_{s_k d_k}:\\ t_{k'}\geq t_k}} \sum_{\ell \in \tilde{\mathcal{L}}_{s_k d_k}} \sum_{q\in \tilde{Q}^{\ell}_{ak'}} \frac{t_{k'}}{t_k}\, q\, x^{ak'}_{\ell q} \geq \text{RHS}_k \quad \forall k \in K \tag{100}$$

$$\sum_{k\in K} \sum_{\ell \in \tilde{\mathcal{L}}_{s_k d_k}} \sum_{q\in \tilde{Q}^{a\ell}_k} y^{\ell}_{ij}\, t_k\, q\, x^{ak}_{\ell q} \leq U\, y_{ij} \quad \forall (i,j) \in V^2 \tag{101}$$

$$\sum_{q\in \tilde{Q}^{a\ell}_k} x^{ak}_{\ell q} \leq 1 \quad \forall k \in K,\ \ell \in \tilde{\mathcal{L}}_{s_k d_k} \tag{102}$$

$$\sum_{q\in \tilde{Q}^{b\ell}_k} x^{bk}_{\ell q} \leq 1 \quad \forall k \in K:\ t_k = t^{\max}_{s_k d_k},\ \ell \in \tilde{\mathcal{L}}_{s_k d_k} \tag{103}$$

$$x^k_{\ell q} \in \{0,1\}\ \forall k \in K, \ell \in \tilde{\mathcal{L}}_{s_k d_k}, q \in \tilde{Q}^{a\ell}_k \cup \tilde{Q}^{b\ell}_k \tag{104}$$

$$y_{ij} \in I\!N \quad \forall (i,j) \in V^2. \tag{105}$$

We use the cut set constraints (79), strong linking cuts (84), Gomory cuts (85) and (86), lifted cover cuts on constraints (100) and (101) and $\epsilon$-split $c$-strong cuts on constraints (101) with $1 \leq \epsilon \leq 4$. Cut set, strong linking and Gomory cuts can be enumerated at each iteration of the cutting plane method, whereas for lifted knapsack cover and $\epsilon$-split $c$-strong inequalities, we use the separation method presented in [3] and [9]. The cuts can sometimes bring better primal solutions as well. Note that each of the above presented cuts admits a linear expression in terms of the $x^k_{\ell q}$. For instance, the strong linking cuts takes the form:

$$\sum_{q\in \tilde{Q}^{a\ell}_k} x^k_{\ell q} \leq y_{ij} \quad \forall (i,j) \in V^2,\ k \in K,\ \ell \in \mathcal{L}_{s_k d_k}:\ y^{\ell}_{ij} = 1. \tag{106}$$

Primal solutions to [VR] or [VR 0-1] that are obtained with CPLEX, need to be check for feasibility using the two-step procedure described in Section 3.2. We first solve (55–60) and then (49–51). We observed that adding constraints to enforce request-commodity upper bounds (74) helps CPLEX primal heuristic to find better primal solutions.

## 5.2 A Column and Cut Generation Approach for the Wavelength Routing Configuration Formulation

The Wavelength Routing Configuration formulation (37–40) is denoted by [WRC]. Its LP relaxation is solved using a nested column generation approach which requires a large computing time. The enhanced demand covering constraints take the form:

$$\sum_{c\in \mathcal{C}} \sum_{\substack{k'\in K_{s_k d_k}:\\ t_{k'}\geq t_k}} \frac{t_{k'}}{t_k}\, x^c_{k'}\, \nu_c \geq \text{RHS}_k \quad \forall k \in K. \tag{107}$$

Aggregate bounds $\underline{b}_k$ and $\overline{b}_k$ and $\text{LHS}_k$ are not explicitly formulated in [WRC] because they induce difficulties for finding a feasible solution of the LP relaxation. However, omitting these bounds shall result in weaker dual bounds.

The master LP is initialized with a set of artificial columns that are later eliminated from the solution by increasing their cost if needed. The pricing problem is initially solved by the greedy heuristic. As the number of optical hop configurations can be large, we first restrict the set of optical hop configurations to the single-hop and two-hop configurations, then when there is no more negative reduced cost wavelength routing configurations, we add the splitting and merging and finally, we add the interlaced configurations. The exact pricing method is only applied when the heuristic fails to identify a negative reduced cost column; only then, in the latter iterations, the pricing problem solution value can be used to compute strong Lagrangian dual bounds on the master LP.

To improve the dual bound, we use cut set inequalities (79) that takes the form

$$\sum_{d \in V \setminus \{s\}} \sum_{c \in \mathcal{C}} \sum_{p \in \mathcal{P}_{sd}} y_p^c \, \nu_c \geq \left\lceil \frac{\sum_{d \in V \setminus \{s\}} D_{sd}}{U} \right\rceil \quad \forall s \in V$$

and

$$\sum_{s \in V \setminus \{d\}} \sum_{c \in \mathcal{C}} \sum_{p \in \mathcal{P}_{sd}} y_p^c \, \nu_c \geq \left\lceil \frac{\sum_{s \in V \setminus \{d\}} D_{sd}}{U} \right\rceil \quad \forall d \in V,$$

where indicator $y_p^c$ is one if configuration $c$ uses path $p$. We also use the $(s,d)$-lightpath inequalities (80): $\sum_{c \in \mathcal{C}} y_{sd}^c \, \nu_c \geq \left\lceil \frac{D_{sd}}{U} \right\rceil \quad \forall (s,d) \in V^2$, where indicator $y_{sd}^c$ is equal to the number of grooming patterns on $c$ that carry $(s,d)$ traffic, and $k$-lightpath cuts (81): $\sum_{c \in \mathcal{C}} y_k^c \, \nu_c \geq \left\lceil \frac{t_k \, \underline{b}_k}{U} \right\rceil \quad \forall k \in K : \underline{b}_k > 0$, where indicator $y_k^c$ is equal to the number of grooming patterns on $c$ that carry a bandwidth reservation of granularity $t_k$. Bandwidth reservation upper bound cuts (83) become: $\sum_{c \in \mathcal{C}} y_{sd\overline{\alpha}_{sd}}^c \, \nu_c \leq \max\{\left\lfloor \frac{D_{sd}}{U} \right\rfloor, 1\} \quad \forall (s,d) \in V^2$ where indicator $y_{sd\overline{\alpha}_{sd}}^c$ is equal to the number of grooming patterns on $c$ that carry more than $\alpha_{sd}$ unit of $(s,d)$ flow. The dual variables associated with these cuts either result in modified dual values $\pi_k$ or in fixed cost linked with the choice of optical hops in the pricing problem and sub-problems. Each time the master LP is solved to optimality we enumerate the valid inequalities and add the violated ones.

As restrictive Assumptions 5 and 6 are made, the resulting dual bound is not strictly valid for GRWA but can be taken as an estimation of the port installation cost. The primal bounds are valid. They are obtained using a rounding heuristic. At each iteration, we get the LP solution $\bar{\nu}$ and we first attempt to round down the master LP solution. For all wavelength routing configurations with $\bar{\nu}_c \geq 1$, we fix $\nu_c = \lfloor \bar{\nu}_c \rfloor$ in the current partial primal solution. If there are no candidate $c$ with $\bar{\nu}_c \geq 1$, we select a wavelength routing configuration to be rounded-up and fix $\nu_c = \lceil \bar{\nu}_c \rceil$ in the primal solution. We use a selection criterion based on the largest ratio of granted demand over cost among the tenth wavelength routing configurations with the largest $\bar{\nu}_c$. It helps to avoid selecting columns with low traffic. After fixing a partial solution, the residual master program is re-optimized by column generation but using the heuristic pricing solver only. The procedure is reiterated until all demands are covered or the residual master is infeasible.

## 5.3 A Cutting Plane Hierarchical Optimization Approach for the Grooming and Physical Routing Formulation

We consider a relaxation of the Grooming and Physical Routing formulation (43–48) denoted by [GPR] where we ignore the term $\phi(y)$, considering only the first stage decision of the Benders' approach of Section 3.1. Its LP relaxation is solved using CPLEX. The resulting dual bound is that of Proposition 2. As for the [VR] formulation (24–28), we use the 0-1 reformulation, denoted by [GPR 0-1], and cutting planes to improve the dual bound. We implemented the restrictions presented in Observation 4, and it slightly improves the dual bound. Primal solutions are obtained using CPLEX default primal heuristics. We check a posteriori the true feasibility of the proposed primal solution by solving the [WA] problem (49–51) using CPLEX.

## 5.4 A Column and Cut Generation Hierarchical Optimization Approach for the Grooming Pattern Formulation

Benders' reformulation (61–63) of the Grooming Pattern formulation is denoted by [BGPA]. We obtain dual bounds by solving the master LP using a column generation procedure, ignoring the term $\phi(y)$ (we consider only the grooming and virtual routing decisions corresponding to the first stage of the approach of Section 3.2). The master LP is initialized with a set of single-hop configuration grooming patterns that define a feasible integer solution. We strengthen the formulation using the upper bound on the aggregate traffic in the enhanced demand constraints: $\sum_{g \in \tilde{\mathcal{G}}} \sum_{k' \in K_{s_k d_k}: t_{k'} \geq t_k} \frac{t_{k'}}{t_k} x^g_{k'} \mu_g \leq \text{LHS}_k \ \forall k \in K$. The solution of the grooming pricing problem was discussed in Section 4.5. At each iteration of the column generation procedure, we add multiple columns to the master (all the grooming patterns with a negative reduced cost). The LP dual bound is not strictly valid for the GRWA problem as it is obtained under Assumption 5. To improve the dual bound, we use the same valid inequalities as for [WRC] formulation given in (37–40). However, in this case, we add all of them a priori to the initial master LP as their number is relatively small. Valid primal bounds are obtained using a rounding heuristic similar to the one used for [WRC]. To check if there exists a feasible wavelength assignment for the selected aggregate optical hop configurations of the master primal solution, we use CPLEX to solve the second stage [GPWA] problem defined in (64–68).

## 6 Numerical Results and Comparison

The four approaches have been tested and compared on realistic size data sets. We have generated four different instances for NSF network (14 nodes and 21 edges) and the EON network (20 nodes and 39 edges). Table 2 summarizes the characteristics of each instance and gives the total number of requests, the overall required bandwidth, the number of wavelengths per fiber link (it is the minimal number of wavelengths required by the trivial single-hop primal solution). We also include the trivial dual bound of Proposition 2 rounded-up, $tDB$, the traditional cut set dual bound, $csDB$:

$$csDB = \max \left\{ \sum_{s \in V} \left\lceil \frac{\sum_{d \in V \setminus \{s\}} D_{sd}}{U} \right\rceil, \sum_{d \in V} \left\lceil \frac{\sum_{s \in V \setminus \{d\}} D_{sd}}{U} \right\rceil \right\},$$

our best truly valid dual bound (in the absence of any restrictive assumption), $DB$, our best primal bounds over all four approaches, $PB$, and the single-hop routing solution primal bound, $shPB$. The number between brackets along side the best bounds refer to the model that releases this best truely valid bound: [1] denotes model [VR], [2] denotes model [VR 0-1] and [3] denotes model [GPR 0-1].

In Table 3 we compare the four approaches in terms of dual bounds at the root node, $rDB$, and dual bounds obtained at the root after adding cuts, $rcDB$, until no more cuts can be generated from our separation procedure (with no time limit). The last row provides the average computation time in seconds. A "*" indicates that the bound is only valid under restrictive Assumptions 5 or 6. In Table 4, we present the dual bounds that can be obtained through CPLEX default branching and automatic cut generations run with a time limit of 1 hour, $bDB$, and those obtained when running CPLEX default branch-and-cut with a time limit of 1 hour after adding all the cuts of Section 4.4 at the root node, $cbDB$. This can only be done for the direct formulations that do not require dynamic column generation. In Tables 5 and 6, we compare primal bounds and gaps to the best dual bound (in percent) obtained with our primal heuristics (either hierarchical optimization or rounding procedure). In Table 5 we do not use the cuts of Section 4.4, while the results of Table 6 are obtained making use of our cutting plane procedure. We give the average computational time

Table 2: Instances characteristics and associated dual and primal bounds

| Instance | $\sum_k D_k$ | $\sum_k t_k D_k$ | $W$ | tDB | csDB | DB | PB(gap) | shPB (gap) |
|---|---|---|---|---|---|---|---|---|
| NSF 1 | 1,667 | 34,007 | 19 | 178 | 185 | 198 [1] | 209 [1] (5.5) | 269 (35.8) |
| NSF 2 | 1,332 | 26,918 | 13 | 141 | 149 | 159 [1][2][3] | 170 [2] (6.9) | 182 (14.4) |
| NSF 3 | 1,949 | 40,064 | 25 | 209 | 216 | 227 [1][2] | 246 [1] (8.3) | 364 (60.3) |
| NSF 4 | 60,303 | 114,328 | 47 | 596 | 603 | 614 [1] | 624 [1] (1.6) | 685 (11.5) |
| EON 1 | 6,639 | 78,837 | 33 | 411 | 420 | 461 [3] | 484 [1] (4.9) | 605 (31.2) |
| EON 2 | 2,292 | 36,547 | 18 | 191 | 201 | 237 [2][3] | 258 [1] (8.8) | 380 (60.3) |
| EON 3 | 3,667 | 79,592 | 36 | 415 | 425 | 468 [3] | 493 [2] (5.3) | 741 (58.3) |
| EON 4 | 120,676 | 217,337 | 67 | 1,132 | 1,143 | 1,188 [3] | 1211 [1] (1.9) | 1,329 (11.8) |
| average | 24,816 | 78,454 | 32.2 | 409.1 | 417.7 | 444 | 461.8 (5.4) | 569.3 (35.5) |

Table 3: Comparing dual bound at the root node

| Instance | [VR] | | [VR 0-1] | | [WRC] | | [GPR 0-1] | | [BGPA] | |
|---|---|---|---|---|---|---|---|---|---|---|
| | rDB | rcDB | rDB | rcDB | rDB* | rcDB* | rDB | rcDB | rDB* | rcDB* |
| NSF1 | 178 | 185 | 178 | 196 | 192 | 199 | 178 | 195 | 194 | 199 |
| NSF2 | 141 | 149 | 141 | 159 | 160 | 161 | 141 | 159 | 161 | 161 |
| NSF3 | 209 | 216 | 210 | 225 | 209 | 234 | 210 | 221 | 215 | 236 |
| NSF4 | 596 | 603 | 597 | 610 | 596 | 604 | 596 | 604 | 597 | 604 |
| EON1 | 411 | 420 | 423 | 454 | 440 | 441 | 429 | 457 | 447 | 457 |
| EON2 | 191 | 201 | 205 | 236 | 258 | 261 | 205 | 236 | 261 | 262 |
| EON3 | 415 | 425 | 419 | 443 | 415 | 425 | 454 | 465 | 427 | 475 |
| EON4 | 1,132 | 1,143 | 1,144 | 1,173 | 1,132 | 1,143 | 1,161 | 1,177 | 1,144 | 1,152 |
| average | 409.13 | 417.75 | 414.63 | 437.00 | 425.25 | 433.50 | 421.75 | 439.25 | 430.75 | 443.25 |
| time | 0.02 | 0.05 | 0.16 | 1,845 | 35,828 | >100,000 | 0.31 | 80,312 | 441 | 5,573 |

Table 4: Comparing dual bounds obtained after running CPLEX default branch-and-cut for 1 hour

| Instance | [VR] | | [VR 0-1] | | [GPR 0-1] | |
|---|---|---|---|---|---|---|
| | bDB | cbDB | bDB | cbDB | bDB | cbDB |
| NSF1 | 198 | 198 | 196 | 197 | 190 | 196 |
| NSF2 | 159 | 159 | 158 | 159 | 151 | 159 |
| NSF3 | 227 | 227 | 224 | 227 | 216 | 224 |
| NSF4 | 614 | 614 | 610 | 613 | 599 | 607 |
| EON1 | 457 | 457 | 451 | 456 | 447 | 461 |
| EON2 | 236 | 236 | 234 | 237 | 229 | 237 |
| EON3 | 444 | 444 | 438 | 444 | 459 | 468 |
| EON4 | 1,179 | 1,180 | 1,170 | 1,177 | 1,169 | 1,188 |
| average | 439.25 | 439.38 | 435.13 | 438.75 | 432.5 | 442.5 |
| time | 3,614 | 3,612 | 3,623 | 1,845 | 3,656 | 83,923 |

Table 5: Comparing primal bounds $PB$ obtained with the primal heuristics of this paper without adding cutting planes before hand

| Instance | [VR] | [VR 0-1] | [WRC] | [GPR 0-1] | [BGPA] |
|---|---|---|---|---|---|
| NSF1 | 212 (7.1) | 212 (7.1) | 240 (21.2) | $\infty$ | 220 (11.1) |
| NSF2 | 173 (8.8) | 170 (6.9) | 190 (19.5) | $\infty$ | 175* (10.1) |
| NSF3 | 246 (8.3) | 254 (11.8) | 269 (18.5) | 265 (16.7) | 258 (13.6) |
| NSF4 | 627 (2.1) | $\infty$ | 643 (4.7) | $\infty$ | 642 (4.5) |
| EON1 | 484 (4.9) | $\infty$ | 526 (14.1) | $\infty$ | 512 (11.1) |
| EON2 | 258 (8.8) | 261 (10.1) | 309 (30.3) | 272 (14.7) | 292 (23.2) |
| EON3 | 494 (5.5) | 493 (5.3) | 557 (19.1) | 564 (20.5) | 528 (12.8) |
| EON4 | 1,230 (3.5) | $\infty$ | 1,239 (4.2) | $\infty$ | 1,236 (4.1) |
| average | 465.50 (6.1) | | 496.63 (16.4) | | 482.88 (11.3) |
| time | 10,571 | 5,447 | 57,954 | 3,282 | 6,128 |

Table 6: Comparing primal bound $PB$ obtained with the primal heuristics of this paper applied after running the cutting plane procedure of this paper

| Instance | [VR] | [VR 0-1] | [GPR 0-1] | [BGPA] |
|---|---|---|---|---|
| NSF1 | 209 (5.5) | 212 (7.1) | 223 (12.6) | 218 (10.1) |
| NSF2 | 172 (8.1) | 172 (8.1) | 171 (7.5) | 174 (9.4) |
| NSF3 | 246 (8.3) | 259 (14.1) | 274 (20.7) | 255 (12.3) |
| NSF4 | 624 (1.6) | $\infty$ | $\infty$ | 637 (3.7) |
| EON1 | 484 (4.9) | $\infty$ | $\infty$ | 501 (8.6) |
| EON2 | 258 (8.8) | 262 (10.5) | 291 (22.7) | 285 (20.2) |
| EON3 | 494 (5.5) | 753 (60.9) | 671 (43.3) | 524 (11.9) |
| EON4 | 1,211 (1.9) | $\infty$ | $\infty$ | 1,221 (2.7) |
| average | 462.25 (5.6) | | | 476.88 (9.9) |
| time | 8,477 | 6,973 | 83,923 | 36,090 |

on last row. For the feasibility subproblems, CPLEX is called with its default settings without a time limit. The analysis of the results can be summarized as follows:

1. With the Virtual Routing formulation (24–28), [VR], the LP dual bound is equal to the trivial bound. Cut set inequalities only allow to get the trivial cut set dual bound, but they help to find good primal solutions. During the one hour of CPLEX default branch-and-cut, the dual bounds are greatly improved, mostly by the built-in cutting plane method, whereas the improvement of the dual bound due to a pure branch-and-bound (without CPLEX built-in cut generation) is marginal. Adding the cut sets before calling on CPLEX MIP solver does make a slight difference in terms of the quality of dual bounds. The primal solutions obtained by the three stage procedure, with a time limit of one hour for each of the first 2 stages, are very good on average. The computational time is around 6000 (resp. 7000) seconds on average when solving the three stage procedure without (resp. with cut set) inequalities. The second stage alone takes less than 3600 seconds in average. The valid solution obtained in the second stage has often the same cost as the first stage incumbent.

   For [VR 0-1], we observe that the LP relaxation value is better than the trivial dual bound thanks to our bound preprocessing. The dual bounds obtained applying CPLEX default branch-and-cut to the original formulation (without the cuts of Section 4.4) are not better than for the integer [VR] version. However, adding the cuts of Section 4.4 greatly improves the dual bound at the root node at the expense of largely increasing the computation time (we did not set any time limit on the cutting plane procedure and we consider all the feasible inequalities of Section 4.4 when generating cuts). Note that the dual bound after our cutting plane method is better than the dual bound obtained after one hour of CPLEX branch-and-cut without using our cutting plane procedure. It shows the interest of the cuts of Section 4.4. Combining these cuts with those generated with the built-in cutting plane method of CPLEX yields extra improvements. The dual bounds obtained from the [VR 0-1] formulation are very close to those obtained from [VR]. Because of the large number of variables in [VR 0-1], CPLEX primal heuristics tend to be less effective (we are not always able to find integer solutions to the first stage problem [VR 0-1]). However, it gives two of the best primal bounds. For many instances, no primal solutions were found at the first stage and the second stage problem is not called. Hence, in Table 5, the average computational time is smaller for [VR 0-1] than for [VR]. However, when using all the families of valid inequalities of Section 4.4, the average time is multiplied by 4 (the cutting plane procedure takes as much as 12,000 seconds in average). This computational time could be reduced by letting CPLEX use its own lifted cover cut inequalities as their separation method should be better than our trivial separation method.
2. For the Wavelength Routing Configuration formulation (37–40), [WRC], the LP relaxation is very hard to solve to optimality. The computational times are very large. The resulting dual bound is not as good as that of formulation [BGPA], given in (52–54), because we do not use all the possible formulation strengthening such as the aggregate flow upper bound $\text{LHS}_k$. The cutting plane method was stopped prematurely after 100,000 seconds of computing time. The primal bounds have a poor quality because fixing wavelength configuration are very aggregate decisions that often end up with little flexibility at the end of the rounding procedure. (We do not report on combining the primal heuristic with cut generation, as the computing times are too large.)
3. For the Grooming and Physical Routing formulation (43–47) in its 0-1 form, [GPR 0-1], the LP relaxation bound is better than the trivial bound. This is due to the domain restriction of the traffic reservation variables on single-hop lightpaths. The dual bounds that are obtained at the end of the CPLEX branch-and-cut procedure (within one hour) are very good. When it is combined with our cutting plane procedure, it gives the best average dual bounds. However, with this formulation, very few primal bounds were obtained and, moreover, their costs are not as good as for the [VR] formulation. This can be explained by the fact that [GPR 0-1] has many more variables than [VR], which makes CPLEX heuristics less efficient. The computing times of the root dual bound with our cutting plane method are quite large: [GPR 0-1] involves more constraints and more induced cuts than [VR]. However, a close look to the results shows that the average time is greatly affected by the EON4 instance that takes more than 500,000 seconds. We have use all the families of valid inequalities of Section 4.4 and we perform separation exactly. Considering a restricted set of valid inequalities and developing heuristic separation could allow us to reduce the computational times (while we could expect to keep very good dual bounds).

4. For the Benders' reformulation (61–63) of the Grooming Pattern formulation, [BGPA], using cuts not only improves the dual bound but also helps in getting better primal bounds on average. Even though the dual bounds are subject to restrictive Assumptions, they are very close to the best dual bounds obtained by the formulation based on the original variables. The primal bounds are weaker than the ones obtained from [VR], but the procedure is robust as the primal solutions were always validated in the second stage feasibility check. The computing times are distributed as follows. Optimizing the LP relaxation takes around 400 seconds without cuts and 5500 seconds using all the cuts of Section 4.4; the primal bound computation takes in average 6000 seconds without cuts and 36000 when using all the cuts (re-optimizing the residual master LP using the cutting plane procedure takes up to 170 000 seconds for EON3). Even though this formulation is based on a restriction, the model is far from trivial to solve.

# Conclusion

The grooming, routing and wavelength assignment (GRWA) problem is quite challenging to solve given the large number of requests and the inherent symmetry in wavelength assignment and alternative traffic loading patterns. To derive dual bounds we have developed and compared (*i*) relaxations (considering only the first stage decisions of a Benders' approach), (*ii*) Dantzig-Wolfe reformulations, and (*iii*) an original hybridization of these two techniques. An important feature in our approaches is the symmetry breaking by way of restricting the solution set to solutions involving large granularity reservations and favoring single-hop routing. Further enhancements are obtained through a 0-1 decomposition with variable domain reduction (the refinement (78) is original and quite helpful computationally) and cut generation inspired from the literature. Primal bounds are obtained using (*i*) a hierarchical optimization procedure based on CPLEX build-in primal heuristics, (*ii*) a rounding procedure (equivalent to a depth first plunging into a branch-and-price tree), or (*iii*) a combination of these two. For both dual and primal bounds, we manage to exploit CPLEX capabilities rather than compete with it.

The best dual bounds are obtained with the hybrid Dantzig-Wolfe / Benders' approach (formulation [BGPA]) when comparing root node results (even though they rely on a restrictive assumption). Comparing what can be obtained with CPLEX default MIP solver, the best dual bounds, in average, are obtained from Benders' [GPR 0-1] formulation that takes advantage of the bound strengthening and the valid inequalities. The best primal bounds are derived from the [VR] formulation that shows the efficiency of the built-in CPLEX heuristics, while the hybrid Dantzig-Wolfe / Benders' approach provides good primal solutions and also shows its robustness. Zero-One extended formulation provides better dual bounds at the root node and more opportunity to derive cutting planes, but they demand (much) larger computing times. Moreover, primal heuristics are not as efficient in these larger variable spaces. Our pure column generation approach, using [WRC], is comparatively much more time consuming. Even though its LP value defines good dual bound, applying branching and cutting to a compact formulation can yield better bounds much faster.

Computing times could probably be much improved by developing more efficient separation procedure for the cuts of Section 4.4. Further developments would also include ad-hoc branching schemes for the column generation approach to examine whether branching could improve dual bounds significantly (although the root node computing time is already quite large).

# References

[1] F. Antonio and G. Bernard. 0-1 reformulations of the multicommodity capacitated network design problem. *Discrete Applied Mathematics*, 157(6):1229–1241, 2009.

[2] A. Atamtürk. On capacitated network design cut-set polyhedra. *Mathematical Programming*, 92:425–437, 2000.

[3] A. Atamtürk and D. Rajan. On splittable and unsplittable capacitated network design arc-set polyhedra. *Mathematical Programming*, 92:315–333, 2002.

[4] C. Barnhart, C.A. Hane, and P.H. Vance. Using branch-and-price-and-cut to solve origin-destination integer multicommodity flow problems. *Operations Research*, 48(2):318–326, 2000.

[5] R. S. Barr, M. S. Kingsley, and R. A. Patterson. *Handbook of optimization in telecommunications*, chapter Telecommunications network grooming, pages 837–862. Springer US, 2006.

[6] E. Bartolini and M. Aristide. Algorithms for the non-bifurcated network design problem. *Journal of Heuristics*, 15(3):259–281, 2009.

[7] B. Brockmullüller, O. Günlük, and L.A. Wolsey. Designing private line networks - polyhedral analysis and computation. Technical report, CORE discussion paper 9647, Université Catholique de Louvain, 1999.

[8] M. Brunato and R. Battiti. A multistart randomized greedy algorithm for traffic grooming on mesh logical topologies. In F. Neri A. Bianco, editor, *Next Generation Optical Network Design and Modelling*. Kluwer, 2003.

[9] Z. Gu, G. L. Nemhauser, and M. W. P. Savelsbergh. Lifted cover inequalities for 0-1 integer programs: Computation. *INFORMS Journal on Computing*, 10:427–437, 1998.

[10] M. Hamdouniand, G. Desaulniers, and F. Soumis. Parking buses in a depot using block patterns: A benders decomposition approach for minimizing type mismatches. *Computers and Operations Research*, 34(11):3362–3379, 2007.

[11] J.Q. Hu. Traffic grooming in WDM ring networks: A linear programming solution. *Journal of Optical Networking*, 1(11):397–408, November 2002.

[12] J.Q. Hu and B. Leida. Traffic grooming, routing and wavelength assignment in optical WDM mesh networks. In *IEEE Infocom 2004*, pages 495–501, March 2004.

[13] Ilog. *CPLEX 10.1 User's Manual*, 2006.

[14] B. Jaumard, Y. Solari, and A. Houle. Dimensioning WDM optical networks with minimum MSPP configuration. In *IASTED International Conference OCSN,Optical Communications Systems and Networks*, July 2004.

[15] V.R. Konda and T.Y. Chow. Algorithm for traffic grooming in optical networks to minimize the number of transceivers. In *IEEE Workshop Hight Performance Switching and Rooting*, pages 218–221, May 2001.

[16] A. M. C. A. Koster, S. Orlowski, C. Raack, G. Baier, and T. Engel. *Single-layer Cuts for Multi-Layer Network Design Problems*, chapter 1, pages 1–23. Springer, College Park, MD, U.S.A., 2008. Selected proceedings 9th INFORMS Telecommunications Conference.

[17] J. Lee. In situ column generation for a cutting-stock problem. *Computers and Operations Research*, 34(8):2345–2358, 2007.

[18] B. Mukherjee. *Optical WDM Network*. Springer, 2006.

[19] C. Ou, K. Zhu, H. Zang, J. Zang, H. Zhu, L. H. Sahasrabuddhe, and B. Mukherjee. Traffic grooming for survivable WDM networks: Dedicated protection. *Journal of Optical Networking*, 3(1):50–74, 2004.

[20] Y. Pochet and R. Weismantel. The sequential knapsack polytope. *SIAM Journal on Optimization*, 8:248–264, 1998.

[21] P. Prathombutr, J. Stach, and E.K. Park. An algorithm for traffic grooming in WDM mesh networks with multiple objectives. *Telecommunication Systems*, 28(3-4):369–386, March 2005.

[22] F. Vanderbeck and M.W.P. Savelsbergh. A generic view of Dantzig-Wolfe decomposition in mixed integer programming. *Operations Research Letters*, 34(3):296–306, 2006.

[23] F. Vanderbeck and L.A. Wolsey. Reformulation and decomposition of integer programs. Technical report, CORE DP 2009/16, Université Catholique de Louvain, 2009.

[24] B. Vignac. *Reformulation and Decomposition Based Methods for a Multi-Layer Capacitated Network Design Problem*. PhD thesis, Université de Montréal / Université Bordeaux 1, 2009.

[25] H. Yen and S. Lee. QoS aware traffic grooming and integrated routing on IP over WDM networks. *Photonic Network Communications*, 14(1):1–10, August 2007.

[26] H. Zhu, H. Zang, K. Zhu, and B. Mukherjee. A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks. *IEEE/ACM Transactions on Networking*, 11(2):285–299, April 2003.

[27] K. Zhu and B. Mukherjee. Traffic grooming in an optical WDM mesh network. *IEEE Journal on Selected Areas in Communications*, 20(1):122–133, January 2002.

[28] K. Zhu, H. Zang, and B. Mukherjee. Design of WDM mesh networks with sparse grooming capability. In *IEEE Globecom*, pages 2696–2700, November 2002.

[29] A. Zymolka. *Design of Survivable Optical Networks by Mathematical Optimization*. Ph.d. thesis, TU Berlin, 2007.