Reformulation and Decomposition Approaches for Traffic Routing in Optical Networks

Benoit Vignac, François Vanderbeck, Brigitte Jaumard

Reformulation and Decomposition Approaches for Traffic Routing in Optical Networks

B. Vignac, Université de Bordeaux and Université de Montréal, vignac@crt.umontreal.ca

F. Vanderbeck, Université de Bordeaux and INRIA, fv@math.u-bordeaux1.fr

B. Jaumard, Concordia University, bjaumard@cse.concordia.ca

Revised in November 2015

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 multiples 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 a more efficient use of network capacity. However, each addition or withdrawal of a request from a wavelength requires optical to electrical conversion and the use of cross-connect equipment with expensive ports of high densities. The objective is to minimize the number of required ports of the cross-connect 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. Instead, 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 discretization 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 two hours of computing time.

Introduction

To accommodate the increase of traffic in telecommunication networks, today's optical networks have huge capacity (tens of $\mathrm{Tb/s}$) thanks to new technologies. The wavelength bandwidth utilization is increased by packing several requests on the same wavelength, a technique called *traffic grooming* [26]. 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 delays). 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) wavelengths have 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 transport protocols (SONET, SDH...), 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 so-called segment, i.e., a subpath in the physical network along which the signal remains into the optical domain with no electrical conversion at intermediate nodes (that are optically *bypassed*). Thus, a lightpath establishes an optical end-to-end connection from a source node to a destination node. Note that the optical signal can be carried by a different wavelength on each of its intermediate optical hops.

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. Note that the logical network is a multi-digraph as there are as many arcs between two nodes as the number of possible physical paths between them, each of which being a potential support for an optical hop. In the sequel, we sometimes model traffic in the *aggregated logical network*, where different optical hops with the same end-nodes are represented by a single aggregated optical hop. This aggregated network holds a single arc between two nodes if there exists at least one physical path between them along which one can establish an optical hop. Hence, we call it the *connectivity network*.

The transport capacity of an arc of the connectivity network is the sum of the transport capacities of the logical network arcs between these nodes, and is equal to the transport capacity of the wavelength(s) supporting it. Using *grooming*, multiple requests can share the same optical hop on the same wavelength provided their cumulative bandwidth requirement does not exceed the wavelength transport capacity. For the O/E/O conversion, a port (i.e., the combination of optical transceivers and electronic terminal equipment needed to access a wavelength) must be installed at each end-node of the optical hops. The overall 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 services [42]. Furthermore, one should also account for the fact that, if an optical hop is too long, not only the transmission delays may impact the end-to-end delays [25], but the signal must be regenerated using, e.g., optical amplifiers. 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. Path length are measured as the sum of the physical link lengths. We assume that the network does not contain two paths of the same length.

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* $U = \text{OC-}192$*. Because of this divisibility 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 single-hop routing restrictions result in a significant increase in port installation costs. On the other hand, relaxing two-hop routing restrictions (allowing 3 or more optical hops) results in only very marginal decrease in optical port installation costs and may increase the delay [35].*

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 (to simplify the model formulation or to make the solution approach tractable – or both).

Assumption 5. *(***Restrictive grooming configurations***) The solution space can be further restricted to simple grooming scenarios (as detailed later and illustrated in Figure 1) such as a single-hop, the fusion of a two-hop with single hops, the bifurcation of two-hops, or their combination.*

Assumption 6. *(***Wavelength continuity***) Requiring wavelength continuity on each lightpath amounts to assuming that 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 [4]. Maximization of the throughput was studied in [39, 40, 41] 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 [7, 17, 20]. When every request has the same granularity, as assumed in [20] and [41], 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 [16]. However, in real-life applications, bandwidth requirements follow a discrete value distribution as modeled in [40]. To the best of our knowledge, no study has yet enforced a maximum number of optical hops for each request (except in [38] 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 [17], it removes the issue of defining a physical routing for each optical hop. When the number of wavelengths/links is assumed sufficiently large [7, 20], the connectivity topology can always be implemented on the physical topology and the wavelength assignment becomes trivial. Observe that in our model, wavelength assignment is more constrained because of the extra restrictions on the number of optical hops and on the path length (Assumptions 3 and 4). Finally, some studies do not assume single-path routing (as in Assumption 1) and therefore deal with easier continuous flow models (see, e.g, [21]).

The GRWA problem is proved NP-hard in [40] (for its simplest variants). Greedy heuristics have been proposed in [39] and [40] (where an oracle is used that does shortest path routing in an auxiliary graph), and in [20] (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 [19] and [30]; but these publications offer no lower bounds to evaluate the heuristic solutions quality. A hierarchical optimization procedure is used in [17], 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. Other hierarchical approaches have been more recently proposed in [8, 36, 37], but they are all heuristic and do not provide any guaranteed bounds. In addition, a heuristic column generation has been proposed in [32].

Exact solution approaches have mostly been based on multi-commodity flow formulations [17, 20, 40]. 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 [2, 3, 5, 6]. On instances with 435 commodities, the authors of [5] report an average gap of around 30 % after one hour of computation using CPLEX [18] and an extra hour of post-processing. The GRWA instances considered in this study assume backbone networks with 14 and 21 nodes, but involve thousands of requests, real-life granularity distribution (see [28]), and tens of wavelengths. This leads to instance sizes that are beyond what can be solved using exact capacitated network design approaches.

In this study, we analyse several possible formulations of the GRWA problem. We consider in particular four models: one of which relies on Dantzig-Wolfe decomposition, two models rely on Benders' decomposition, and the fourth on a combination of these two approaches. We develop and compare exact optimization based approaches for these models: a nested column generation approach, hierarchical optimization approaches with either two stages (grooming and physical routing first, wavelength assignment second) or three stages (grooming and virtual routing first, physical routing second, and wavelength assignment third), and a hybrid method combining hierarchical optimization and column generation. In column generation approaches, 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 Dantzig-Wolfe (resp. Benders') master programs. However, in Benders' approaches, we only solve a relaxed master where no Benders' feasibility cuts are generated (the sub-problem is an integer linear program). This simple hierarchical optimization with no feedback loop provides valid bounds. Benders' approaches give rise to compact formulations that can be handled directly by CPLEX. Dual bounds can then be improved through a branch-and-cut method and integer solutions to the master are derived through CPLEX built-in heuristics. Then, 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, 3, and 4, we present respectively the Dantzig-Wolfe and Benders' decomposition approaches and their combination. Section 5 outlines ways to reduce symmetries, to restrict the domains of the variables, to take advantage of zero-one discretization that enables to express tighter relations between the variables and to strengthen the formulations with cutting planes. Then, in Sections 6 and 7, we summarize the algorithms used for each of the four solution approaches and compare the numerical results. Our results set new benchmarks for the GRWA, both in terms of the size of the instances that are dealt with (up to 120,000 requests on 21 nodes) and by providing solution guarantees (comparing our best bounds allows one to estimate the optimality gap to 5% on average). We conclude with an analysis of the numerical tests (best results are obtained with the approaches that exploit best the commercial MIP-solver capabilities), and a summary of the advanced and innovative features experimented in our approaches (reformulation that avoid symmetries, nested decomposition, column generation applied to a relaxed Benders master program, domain reduction that exploits the characteristics of our application).

Table 1: Notations used for indices, variables and solution vectors

Indices:	
u, v, s, d	nodes in the physical network,
a	arc in the physical network,
k	request-commodity,
p	path in the physical network,
i, j	nodes in the virtual network,
ℓ	lightpath in the virtual network,
λ	wavelength.
Variable or Solution Vectors:	
$x_{\ell k}$	traffic of request-commodity k that is routed along lightpath ℓ,
$y_{p\lambda}$	is 1 if path p is used to build a hop that carries a signal on wavelength λ,
y_p	number of hops built using path p,
y_{ij}	number of hops built between nodes i and j of virtual network,
z_{uv}	is 1 if arc (u, v) is used in defining a path that shall support a hop.

1. Description of the GRWA Problem

Let graph $G = (V, A)$ represent 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, which are represented by the set $\Lambda = \{1, \ldots, W\}$. Each wavelength has a transport capacity $U = \text{OC-192}$. 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. In R, several requests can be identical. Let $D_{sdt} > 0$ represent the number of (s, d, t)-requests. Let K be the set of distinct (s, d, t) requests, each of which is a *request-commodity* that is defined by a quadruplet $k = (s, d, t, D = D_{sdt})$. 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). Let $D_{sd} = \sum_{t \in T} t\, D_{sdt}$ be the aggregated traffic demand between s and d expressed in OC. K_{sd} (resp. K_{sdt}) denotes the set of request-commodities with source s and destination d (resp. 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 (assigning physical paths to optical hops) and the wavelength assignment. To properly model arc capacity and path length constraints, one must indeed associate a path in the physical network with each optical hop and assign a specific wavelength to this physical path. To formulate the problem, we make use of mainly three variable classes that are summarised in Table 1, along with the index set, for further reference: x is a traffic vector whose components represent flows on lightpaths, y is a design vector whose components define the optical hops that are put in place in the virtual network in terms of the physical paths in the underlying network, along with their wavelength assignment, and z is a design vector whose components define the physical path in terms of arc in the physical network. For simplicity, we use the same vector notation x, y, z to designate a given solution to a subproblem. In that case the vector has an upper script associated with the subproblem. To classify the various formulations considered in the sequel (see Table 3 for a summary), we use a formulation name abbreviation starting with 'O' for the original formulation, 'D' for a Dantzig-Wolfe (D-W) reformulation, 'B' for a formulation deriving from Benders, 'P' for a pricing subproblem (in a D-W approach), and 'F' for a second stage feasibility subproblem (in a Benders' approach).

Let $\mathcal{P}_{sd}$ be the set of feasible paths from s to d. It can be described by an integer polyhedron:

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

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

where variable $z_{uv} = 1$ if arc (u,v) is in the path. Given Assumption 4, $\mathcal{P}_{sd}$ is restricted to the three shortest paths from s to d. Let $\mathcal{P} = \cup_{sd}\mathcal{P}_{sd}$; note that $|\mathcal{P}| \leq 3\,n(n-1)$.

Let $\mathcal{L}_{sd}$ be the set of feasible lightpaths from s to d. It can be described by an integer polyhedron:

$$\mathcal{L}_{sd} = \left\{(y,\delta) \in \{0,1\}^{|\mathcal{P}|\,(n-2)} : \sum_{p\in\mathcal{P}}\sum_{a\in A} l_a\, z_a^p\, y_p \leq L_{\max}^{sd}\right., \tag{3}$$

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

$$\left.\sum_{v\in V\setminus\{s,d\}} \delta_v \leq 1\right\}, \tag{5}$$

where variable $y_p = 1$ if path p forms an optical hop of the lightpath; $\delta_v = 1$ if v is an intermediate node of the lightpath where a O/E/O conversion takes place; z_a^p are here input data, i.e., indicators of the arcs that define path p: $z_a^p = 1$ if path $p \in \mathcal{P}$ uses arc $a \in A$. Let $\mathcal{L} = \cup_{sd}\mathcal{L}_{sd}$; note that $|\mathcal{L}| \leq |\mathcal{P}|(n-2)$ (each path $p \in \mathcal{P}$ can yield a lightpath for each of its intermediate nodes).

The GRWA can be formulated using variables:
$x_{\ell k}$ = the amount of request-commodity $k \in K$ that is routed along lightpath $\ell \in \mathcal{L}_{s_k d_k}$;
$y_{p\lambda} = 1$ if an optical hop is routed along path $p \in \mathcal{P}$ and assigned with wavelength $\lambda \in \Lambda$.
Then, the problem admits an **Original Compact** formulation which takes the form:

$$\min \sum_{p\in\mathcal{P}}\sum_{\lambda\in\Lambda} y_{p\lambda} \tag{6}$$

$$\text{[OC]} \quad \sum_{\ell\in\mathcal{L}_{s_k d_k}} x_{\ell k} = D_k \quad \forall k \in K \tag{7}$$

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

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

$$x_{\ell k} \in \mathbb{Z}_+ \quad \forall k \in K,\; l \in \mathcal{L}_{s_k d_k} \tag{10}$$

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

where y_p^ℓ and z_a^p are input data, i.e., solution indicator vectors as opposed to variables: y_p^ℓ is 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* is 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 Λ, model *link capacity and wavelength clash* constraints. Each link can carry at most W wavelengths and hence, it supports at most W optical hops. 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 each wavelength. 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 λ 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) and (11)). Under Assumption 2, this bin packing problem can be solved using a trivial first fit decreasing greedy procedure (see [29]). It consists in (i) sorting the $x_{\ell k}$ traffic bundles, for $k \in K$ and $\ell \in \mathcal{L}$ 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 λ. 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, there 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. Then, we can omit constraints (10). If we do not make Assumption 2, the surrogate capacity constraints (9) are not sufficient to enforce wavelength transport 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 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. 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. Assumption 5 shall take its meaning in Section 2.1. To enforce Assumption 6, one would need to use variables $x_{\ell k \lambda}$ to represent the amount of request-commodity k routed on lightpath ℓ and assigned to wavelength λ; furthermore one need to disaggregate capacity constraints (9). Then, as we shall see in Section 2.2, a solution to GRWA can be decomposed into W independent traffic routings, each of them using its own wavelength.

In the sequel, model [OC] shall be reformulated using constraint and/or variable decomposition techniques [34], i.e.,

- either one identifies a subset of constraints as defining a subsystem and one reformulates the problem using variables that represent a selection of the solution to that subsystem; this is the so-called Dantzig-Wolfe decomposition principle which can equivalently be presented as the result of a Lagrangian relaxation of the constraints that do not define the subsystem;
- or one identifies a subset of variables that define primary decisions and one formulates the problem in two stages: optimizing first on the primary decisions, second on the remaining variables; this is the so-called Benders' decomposition principle that can involve more than two stages (one of our reformulation involves 3 stages);
- or a combination of the two above.

2. Dantzig-Wolfe Reformulations

Here, we consider two possible Dantzig-Wolfe decomposition approaches to the GRWA problem under restrictive Assumptions 5 or 6. Note that the [OC] formulation given by (6-11) is expressed in terms of flows on lightpaths. It can itself be viewed as a reformulation resulting from a Dantzig-Wolfe decomposition applied to a formulation written in terms of flows on connectivity and physical arcs. However, the [OC] formulation (6-11) does not require the use of a column generation approach because it has a polynomial number of variables given Assumptions 3 and 4.

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 1.

Single-Hop configurations: 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 configurations: It is composed of two optical hops defined by their associated paths: $p_1 \in \mathcal{P}_{si}$ and $p_2 \in \mathcal{P}_{id}$, 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 configurations: It is composed of three optical hops defined by their associated paths: $p_1 \in \mathcal{P}_{si}$, $p_2 \in \mathcal{P}_{id_1}$ 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 configurations: It is the reverse of the previous case where the support is made of three optical hops defined by their associated paths: $p_1 \in \mathcal{P}_{s_1 i}$, $p_2 \in \mathcal{P}_{s_2 i}$ 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 configurations is similar to the generation of three-hop splitting configurations.

Three-hop interlaced-lightpaths configurations: It is composed of three optical hops defined by their associated paths: $p_1 \in \mathcal{P}_{s_1 s_2}$, $p_2 \in \mathcal{P}_{s_2 d_1}$ 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 configurations 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 hop carried out by the 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

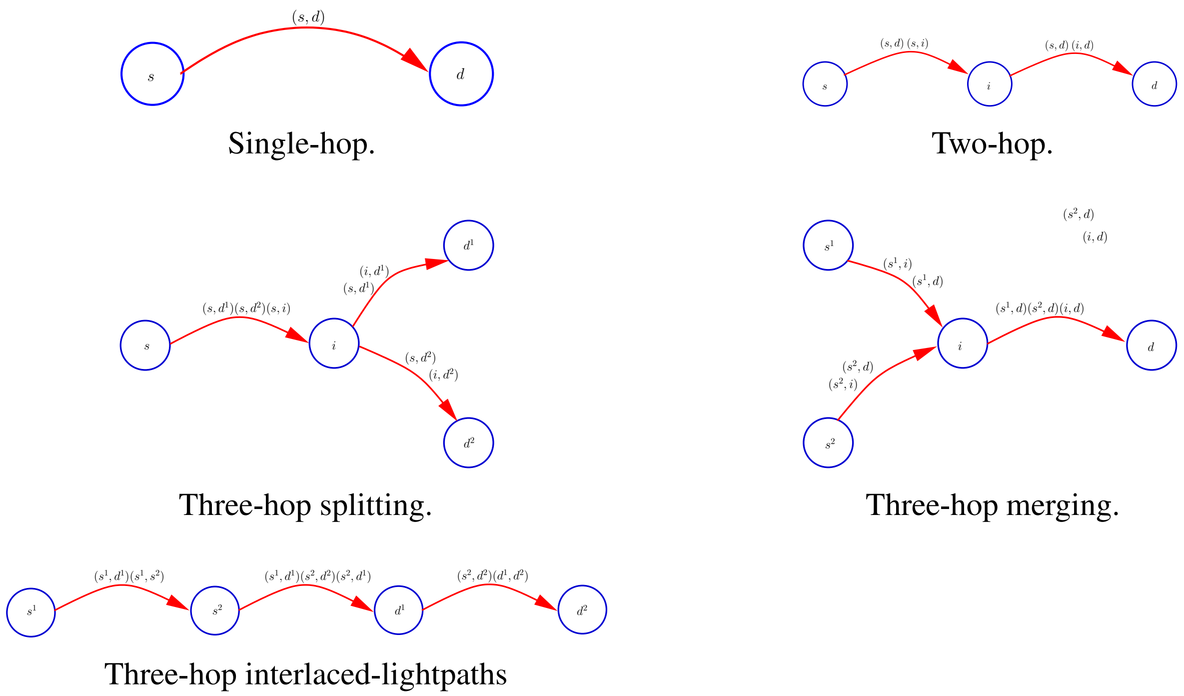

Figure 1: Optical hop configurations that define the supports of *grooming patterns*.

the latter by selecting optical hop configurations that are arc disjoint). The grooming pattern reformulation arises from considering [OC] constraints (9-10) as a subsystem on which we base a Dantzig-Wolfe decomposition. Equivalently this reformulation can be obtained by applying a Lagrangian relaxation of constraints (7-8): the solution of the remaining problem decomposes into independently routed traffic on grooming patterns. Note that the optical hop configuration determines the way in which different traffic may interact with each other, but there are no interactions among grooming patterns. We have further restricted the solution space by limiting the number of interaction modes to the above 5 optical hop configuration types illustrated in Figure 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** built from optical hop configurations $o \in \mathcal{O}$ 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 grooming pattern g. For a fixed optical hop configuration $o \in \mathcal{O}$, $\mathcal{G}(o)$ admits an integer polyhedral description. For instance, for a two-hop configuration between nodes s and d with an O/E/O at node i,

$$\mathcal{G}(s,i,d) = \{x \in {\mathbb{Z}_+}^{3|T|} : \sum_{k\in K_{sd}} t_k\, x_k + \sum_{k\in K_{si}} t_k\, x_k \le U; \sum_{k\in K_{sd}} t_k\, x_k + \sum_{k\in K_{id}} t_k\, x_k \le U\}, \tag{12}$$

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 a 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 than the demand. In the example of the polyhedral description for the two-hop configuration (s,i,d), we add the constraints: $\sum_{k\in K_{sd}} x_k^g \ge 1$, $\sum_{k\in K_{si}} x_k^g + \sum_{k\in K_{id}} x_k^g \ge 1$ and $x_k^g \le D_k \quad \forall k \in K_{sd} \cup K_{si} \cup K_{id}$.

The reformulation of the GRWA problem in terms of variables μ_g, whose value represents the number of times the grooming pattern $g \in \mathcal{G}$ is used, shall be named the **Dantzig-Wolfe Grooming Pattern** formulation:

$$\begin{aligned}
&\min \sum_{g\in\mathcal{G}} \sum_{p\in\mathcal{P}} y_p^g\, \mu_g && (13)\\
\text{[DGP]}\quad &\sum_{g\in\mathcal{G}} x_k^g\, \mu_g = D_k \quad \forall k \in K && (14)\\
&\sum_{g\in\mathcal{G}} y_p^g\, \mu_g = \sum_{\lambda\in\Lambda} y_{p\lambda} \quad \forall p \in \mathcal{P} && (15)\\
&\sum_{p\in\mathcal{P}} z_a^p\, y_{p\lambda} \le 1 \quad \forall \lambda \in \Lambda,\ a \in A && (16)\\
&y_{p\lambda} \in \{0,1\} \quad \forall p \in \mathcal{P},\ \lambda \in \Lambda && (17)\\
&\mu_g \in \mathbb{Z}_+ \quad \forall g \in \mathcal{G}. && (18)
\end{aligned}$$

The LP relaxation of this formulation can be solved by column generation. The problem **Pricing Grooming Configuration** takes the form:

$$\text{[PGC]} \equiv \min_{o\in\mathcal{O}} \{ \sum_{p\in\mathcal{P}} (1-\rho_p)\, y_p^o - \max_{g\in\mathcal{G}(o)} \{ \sum_{k\in K} \pi_k\, x_k \} \}, \tag{19}$$

where (π, ρ) are the dual variables associated with constraints (14) and (15) 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 5). The dual bound given by the LP relaxation is in theory better than that of the [OC] formulation given in (6-11) because the subproblem captures the knapsack capacity constraints. However, the dual bound is subject to the 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; but it will be the basic formulation to a Benders' decomposition approach (see Section 4).

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. Relaxing the demand covering constraints (7) in a Lagrangian fashion leads to a Dantzig-Wolfe decomposition based on subsystem (8)-(11), i.e., 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. The subproblem solutions shall be called *wavelength routing configurations*. If we further make the grooming restriction of 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 μ^c, with $\mu^c_g = 1$ if grooming pattern g is used. An integer polyhedral description of $\mathcal{C}$ is:

$$\mathcal{C} = \{\mu \in \{0,1\}^{|\mathcal{G}|} : \sum_{g \in \mathcal{G}} \sum_{p \in \mathcal{P}} z^p_a \, y^g_p \, \mu_g \leq 1 \quad \forall a \in A; \sum_{g \in \mathcal{G}} x^g_k \, \mu_g \leq D_k \quad \forall k \in K\}$$

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: $(x^c, y^c) = (\sum_{g \in \mathcal{G}} x^g \, \mu^c_g, \sum_{g \in \mathcal{G}} y^g_p \, \mu^c_g)$. Then, the GRWA problem can be modeled by the **Dantzig-Wolfe Wavelength Routing Configuration** formulation:

$$\min \sum_{c \in \mathcal{C}} \sum_{p \in \mathcal{P}} y^c_p \, \nu_c \tag{20}$$

$$\text{[DWRC]} \quad \sum_{c \in \mathcal{C}} x^c_k \, \nu_c = D_k \quad \forall k \in K \tag{21}$$

$$\sum_{c \in \mathcal{C}} \nu_c \leq W \tag{22}$$

$$\nu_c \in \mathbb{Z}_+ \quad \forall c \in \mathcal{C}, \tag{23}$$

where variable $\nu_c = 1$ if configuration $c \in \mathcal{C}$ is used and W is the number of available wavelengths. This [DWRC] formulation 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. Hence, the problem of **Pricing Wavelength Configuration** takes the form:

$$\text{[PWC]} \equiv \min\{\sum_{g \in \mathcal{G}} (\sum_{p \in \mathcal{P}} y^g_p - \sum_{k \in K} \pi_k \, x^g_k) \mu_g : \mu \in \mathcal{C}\} - \sigma, \tag{24}$$

where (π, σ) are the dual variables associated with constraints (21) and (22), respectively. When solving problem (24) by column generation, the pricing sub-problem is:

$$\min_{o \in \mathcal{O}} \{ \sum_{p} y_p^o - \max_{g \in \mathcal{G}(o)} \{ \sum_{k \in K} \pi_k \, x_k^g \} \}. \tag{25}$$

The LP relaxation bound is in theory better than that of the grooming pattern formulation (13-18) because more constraints are included in the definition of $\mathcal{C}$ than that of $\mathcal{G}$. But this dual bound is only valid under two restrictive assumptions: Assumptions 5 and 6.

3. Benders' Decomposition and Hierarchical Optimization

Another form of decomposition is Benders' (also known as resource or variable decomposition [34]). One adopts a 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 the [OC] formulation given in (6-11), when one fixes first the traffic flow decisions, $x_{\ell k}$, 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 the [OC] formulation given in (6-11) onto the space of the important variables. Then, the so-called Benders' master program captures the grooming and physical routing decisions of the GRWA problem. Hence, the program is called **Benders' Grooming and Physical Routing (BGPR)** formulation:

$$\min \sum_{p \in \mathcal{P}} y_p + \phi(y) \tag{26}$$

$$\text{[BGPR]} \quad \sum_{\ell \in \mathcal{L}_{s_k d_k}} x_{\ell k} = D_k \quad \forall k \in K \tag{27}$$

$$\sum_{k \in K} \sum_{\ell \in \mathcal{L}_{s_k d_k}} t_k \, y_p^\ell \, x_{\ell k} \leq U \, y_p \quad \forall p \in \mathcal{P} \tag{28}$$

$$x_{\ell k} \in \mathbb{Z}_+ \quad \forall k \in K, \, \ell \in \mathcal{L}_{s_k d_k} \tag{29}$$

$$y_p \in \mathbb{Z}_+ \quad \forall p \in \mathcal{P}, \tag{30}$$

where $\phi(y) = \infty$ if y, the optical hop establishment decisions, cannot be associated with a wavelength assignment that satisfies clash constraints and zero otherwise. If we omit the term $\phi(y)$, formulation [BGPR] amounts to a relaxation of the [OC] formulation given in (6-11) where the wavelength clash constraints

(8) are ignored. Then, the wavelength index on the y variables is no longer required. Note that, having dropped the wavelength indexing, the master does not suffer the symmetry in wavelength assignment that was present in the [OC] formulation given in (6-11). The LP relaxation of [BGPR] without the term $\phi(y)$ is not very strong as it is a relaxation of the linear program associated with the [OC] formulation. To improve the model, one can add further necessary (but not sufficient) conditions for the existence of a feasible wavelength assignment:

$$\sum_{p \in \mathcal{P}} z_a^p \, y_p \leq W \quad \forall a \in A \tag{31}$$

that state that, on each link, there are at most W optical hops, where W is the number of available wavelengths. Given a feasible master integer solution $(\bar{x}, \bar{y})$ to (26-30), the feasibility check entails solving a sub-problem to find a **Feasible Wavelength Assignment** (FWA):

$$\begin{aligned} \text{[FWA]} \quad \sum_{\lambda \in \Lambda} y_{p\lambda} &= \bar{y}_p \quad \forall p \in \mathcal{P} && (32) \\ \sum_{p \in \mathcal{P}} z_a^p \, y_{p\lambda} &\leq 1 \quad \forall \lambda \in \Lambda, \, a \in A && (33) \\ y_{p\lambda} &\in \{0,1\} \quad \forall p \in \mathcal{P}, \, \lambda \in \Lambda. && (34) \end{aligned}$$

This second stage problem captures the wavelength assignment decisions (FWA) of the GRWA problem. Observe that the [FWA] 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 (33).

Benders' approach entails a polyhedral characterization of the set $\{y : \phi(y) = 0\}$. The master should be iteratively augmented with feasibility cuts when its solution $\bar{y}$ does not lead to a feasible second stage problem. 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). 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 [34]. However, Farkas Lemma does not apply when the second stage problem is an integer program as in our case (for examples of Benders' approaches with integer subproblems, see [14, 15, 23, 31]). In practice, we ignore the term $\phi(y)$. Solving this relaxed problem provides a valid dual bound. Its primal solution must be checked for feasibility by solving the associated FWA subproblem. One can develop an exact approach by integrating the feasibility check, [FWA], in an implicit enumeration approach to Benders' master integer program (a branch-and-bound for instance): each time a primal master solution is encountered and fails the feasibility check, it is discarded and the procedure proceeds to enumerating other primal solutions to Benders' master. We do not use this exact approach but instead we apply LP-based primal heuristics to Benders' master. When a primal solution to the first stage problem is found, we check whether the associated FWA subproblem is feasible. Otherwise, we discard it.

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

If we aggregate optical hops with the same end-nodes in formulation (26-30), we obtain a formulation in terms of variables y_{ij} representing the number of optical hops established between nodes i and j of the *connectivity network*. Lightpaths with the same end-nodes and O/E/O nodes can also be aggregated. The set of aggregate lightpaths is denoted by $\tilde{\mathcal{L}}$. The resulting formulation is referred to as **Bender's Virtual Routing** formulation:

$$\min \sum_{(i,j)\in V^2} y_{ij} + \phi(y) \tag{35}$$

$$[\text{BVR}] \quad \sum_{\ell \in \tilde{\mathcal{L}}_{s_k d_k}} x_{\ell k} = D_k \quad \forall k \in K \tag{36}$$

$$\sum_{k\in K} \sum_{\ell \in \tilde{\mathcal{L}}_{s_k d_k}} t_k\, y^{\ell}_{ij}\, x_{\ell k} \leq U\, y_{ij} \quad \forall (i,j) \in V^2 \tag{37}$$

$$x_{\ell k} \in \mathbb{Z}_+ \quad \forall k \in K,\, l \in \tilde{\mathcal{L}}_{s_k d_k} \tag{38}$$

$$y_{ij} \in \mathbb{Z}_+ \quad \forall (i,j) \in V^2. \tag{39}$$

Now we show that both the LP relaxation of formulation [BGPR] and [BVR] without the term $\phi(y)$ provides the same dual bound as the LP relaxation of the [OC] formulation given in (6-11). Moreover, these bounds are not better than the trivial combinatorial bound that can be computed a priori on the number of required optical hops: indeed, given wavelength capacity, the number of optical hops is at least $\lceil \sum_k t_k D_k / U \rceil$.

Proposition 2. *The dual bounds obtained from the linear relaxations of the [OC] formulation given in (6-11), and from both (26-30) and (35-39) where the term $\phi(y)$ is omitted, are all equal to the trivial dual bound $\sum_{k\in K} t_k D_k / U$.*

Proof: An optimal solution to the LP relaxation of (35-39) 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 (35-39) is a relaxation of (26-30), itself a relaxation the [OC] formulation given in (6-11), its LP solution $\sum_{k\in K} t_k D_k / U$ is also a lower bound for the [OC] formulation. Thus, to obtain the bound result for the [OC] formulation, 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 the [OC] formulation given in (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 $\bar{y}_p = \frac{\sum_{\ell,k} y^{\ell}_p\, t_k\, \bar{x}^k_{\ell}}{U}$. 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 $\bar{x}^k_p$ 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. ∎

The second stage problem differs from [FWA] given in (32-34). 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 (35-39) 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. In the first step, we solve a subproblem to identify a min cost **Feasible Grooming and Path Assignment** problem:

$$\min \sum_{p \in \mathcal{P}} y_p \tag{40}$$

$$\text{[FGPA]} \quad \sum_{p \in \mathcal{P}_{ij}} y_p \geq \bar{y}_{ij} \quad \forall (i,j) \in V^2 \tag{41}$$

$$\sum_{\ell \in \mathcal{L}_{s_k d_k}} x_{\ell k} = D_k \quad \forall k \in K \tag{42}$$

$$\sum_{k \in K} \sum_{\ell \in \mathcal{L}_{s_k d_k}} y_p^\ell \, t_k \, x_{\ell k} \leq U \, y_p \quad \forall p \in \mathcal{P} \tag{43}$$

$$x_{\ell k} \in \mathbb{Z}_+ \quad \forall k \in K, \, \ell \in \mathcal{L}_{s_k d_k} \tag{44}$$

$$y_p \in \mathbb{Z}_+ \quad \forall p \in \mathcal{P}. \tag{45}$$

Constraints (41) 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 the [BGPR] formulation (26-30). Note that we allow a modification of the routing of the flows to get more flexibility. Then, if problem [FGPA] as defined in (40-45) is feasible, we solve the feasibility check [FWA] problem (32-34) in a second step. Stage 2 and 3 together define a global feasibility check. Given that our procedure consists in two hierarchical stages, it is not exact: if we fail to find a feasible solution, it does not mean that no feasible solution exists. In our numerical experiment, the procedure never failed to obtain a feasible solution.

4. Hybridization of Dantzig-Wolfe reformulation and Benders' Decomposition

Benders' decomposition can also be applied to the [DGP] formulation given in (13-18): in the first stage, we fix grooming pattern selection; in the second stage, we assign wavelengths. This leads to the master problem that we call the **Hybrid Grooming Pattern** formulation:

$$\min \sum_{g\in\mathcal{G}} \sum_{p\in\mathcal{P}} y_p^g \mu_g + \phi(\mu) \tag{46}$$

$$\begin{aligned} \text{[HGP]} \quad \sum_{g\in\mathcal{G}} x_k^g \, \mu_g \ &= \ D_k \quad \forall k \in K & (47)\\ \mu_g \ &\in \ \mathbb{Z}_+ \quad \forall g \in \mathcal{G}, & (48)\end{aligned}$$

Let $\phi(\mu) = \infty$ if the optical hops implied by solution μ 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 [FWA] given in (32-34), where $\bar{y}_p = \sum\limits_{g\in\mathcal{G}} y_p^g \bar{\mu}_g$.

Alternatively, in formulation [HGP] given in (46-48), one can also aggregate the optical hop configurations that are logically equivalent (same end nodes and O/E/O nodes) leading to the set $\tilde{\mathcal{O}}$ of aggregated optical hop configurations and associated set $\tilde{\mathcal{G}}$ of aggregated grooming patterns. This relaxation leads to **Hybrid Aggregated Grooming Pattern** formulation:

$$\min \sum_{g\in\tilde{\mathcal{G}}} \sum_{p\in\mathcal{P}} y_p^g \mu_g + \phi(\mu) \tag{49}$$

$$\begin{aligned} \text{[HAGP]} \quad \sum_{g\in\tilde{\mathcal{G}}} x_k^g \, \mu_g \ &= \ D_k \quad \forall k \in K & (50)\\ \mu_g \ &\in \ \mathbb{Z}_+ \quad \forall g \in \tilde{\mathcal{G}}. & (51)\end{aligned}$$

Given a fixed master integer solution $\bar{\mu}$ to (49-51), $\phi(\bar{\mu}) = 0$ if there exists associated physically routed grooming configurations and a feasible wavelength assignment. Let $\bar{\kappa}_o = \sum_{g\in\tilde{\mathcal{G}}(o)} \bar{\mu}_g$ 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 the search for a **Feasible Grooming Pattern and Wavelength Assignment** (FGPWA):

$$\begin{aligned} \text{[FGPWA]} \quad \sum_{\lambda\in\Lambda} y_{p\lambda} \ &= \ \sum_{o\in\mathcal{O}} y_p^o \kappa_o \quad \forall p \in \mathcal{P} & (52)\\ \sum_{o'\in\mathcal{O}: o'\equiv o} \kappa_{o'} \ &= \ \bar{\kappa}_o \quad \forall o \in \tilde{\mathcal{O}} & (53)\\ \sum_{p\in\mathcal{P}} z_a^p \, y_{p\lambda} \ &\leq \ 1 \quad \forall \lambda \in \Lambda, \ a \in A & (54)\\ \kappa_o \ &\in \ \mathbb{Z}_+ \quad \forall o \in \tilde{\mathcal{O}} & (55)\\ y_{p\lambda} \ &\in \ \{0,1\} \quad \forall p \in \mathcal{P}, \ \lambda \in \Lambda. & (56)\end{aligned}$$

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 (52) state that the number of wavelengths reserved for physical path p depends on the selected optical hop configurations. Constraints (53) enforce a selection of optical hop configurations that reproduce the connectivity selection.

5. Implementation Features

The above models share some similarities. In this section, we present some formulation strengthening through partial reformulation, variable domain reduction, a 0-1 discretization, and valid inequalities that can be useful for several models. These techniques are essential features for our solution approaches. They contributed to make our models tractable. 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 improving our primal heuristics. We also detail the pricing procedures for the column generation approaches.

5.1 Enhanced Demand Covering Constraints

Due to the divisibility of the granularities, there 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 an 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 $\frac{t}{t'} \in \mathbb{Z}_+$. 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 bandwidth 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:

$$\underbrace{\left\lfloor \sum_{k' \in K_{s_k d_k}} \frac{t_{k'}}{t_k} D_{k'} \right\rfloor}_{\text{LHS}_k} \geq \sum_{\ell \in \mathcal{L}_{s_k d_k}} \sum_{\substack{k' \in K_{s_k d_k}: \\ t_{k'} \geq t_k}} \frac{t_{k'}}{t_k} x_{\ell k'} \geq \underbrace{\sum_{\substack{k' \in K_{s_k d_k}: \\ t_{k'} \geq t_k}} \frac{t_{k'}}{t_k} D_{k'}}_{\text{RHS}_k} \quad \forall k \in K. \tag{57}$$

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

Observation 2. *Under Assumption 2, any solution to (6, 57, 8-11) can be transformed into a solution of the [OC] formulation given in (6-11) by disaggregating granularity reservations that exceed their associated demands in order to cover lower granularity demands.*

5.2 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 by $\underline{b}_{\ell k}$ and $\overline{b}_{\ell k}$, $\underline{b}_k$ and $\overline{b}_k$, respectively, with

$$\underline{b}_{\ell k} \leq \underline{b}_k \text{ and } \overline{b}_{\ell k} \leq \overline{b}_k \quad \forall k \in K,\ \ell \in \mathcal{L}_{s_k d_k}. \tag{58}$$

Moreover,

$$\underline{b}_k \geq \sum_{\ell \in \mathcal{L}_{s_k d_k}} \underline{b}_{\ell k} \text{ and } \overline{b}_k \leq \sum_{\ell \in \mathcal{L}_{s_k d_k}} \overline{b}_{\ell k} \quad \forall k \in K. \tag{59}$$

At the outset, the lower bounds are set to 0 and the upper bounds to ∞. Strengthening one of these bounds may imply strengthening the others through relations (58-59). We also define a domain of discrete values with the same index convention. The default domain for $x_{\ell k}$ is $Q_{\ell k} = \{\underline{b}_{\ell k}, \ldots, \overline{b}_{\ell k}\}$ 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, one can assume, w.l.o.g., that at least $y_p - 1$ wavelengths are used at their full capacity U. Moreover, one can 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, one can restrict the problem 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 2-hop lightpaths:

Observation 5. (Non-degenerated 2-hops) *The maximum (s, d) reservation routed over a 2-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 2-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:

$$\overline{b}_{\ell k} \le \frac{t_k^+}{t_k} - 1 \text{ , if } t_k < t_{s_k d_k}^{\max} \tag{60}$$

$$\overline{b}_{\ell k} \le \frac{U}{t_k} - 1 \text{ , if } \ell \text{ is a 2-hop lightpath} \tag{61}$$

$$\overline{b}_k \le \left\lfloor \sum_{\substack{k' \in K_{s_k d_k}: \\ t_{k'} \le t_k}} \frac{t_{k'}}{t_k} D_{k'} \right\rfloor . \tag{62}$$

Bound (60) 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 (61) considers the case of a 2-hop lightpath where the capacity reservation cannot cover the wavelength capacity U as stated in Observation 5. Bound (62) specifies that a capacity reservation cannot exceed the cumulative demand for that granularity and lower granularities. Regarding lower bounds, we can state:

$$\underline{b}_k = D_k, \text{ if } t_k = t_{s_k d_k}^{\max}; \tag{63}$$

$$\underline{b}_k = \max\left\{0, D_k - (\text{LHS}_{k^+} - \text{RHS}_{k^+})\frac{t_k^+}{t_k}\right\}, \text{ if } t_k < t_{s_k d_k}^{\max},\ k^+ = (s_k, d_k, t_k^+). \tag{64}$$

Aggregate bound (63) 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 (64) 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 (58-59).

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 $D_k > \overline{b}_k$, then some fraction of the demand k must be covered using larger granularities. Hence, $\left\lceil \frac{t_k(D_k - \overline{b}_k)}{t_k^+} \right\rceil$ can be added to demand (s_k, d_k, t_k^+) and $\frac{t_k^+}{t_k} \left\lceil \frac{t_k(D_k - \overline{b}_k)}{t_k^+} \right\rceil$ 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 2 gives an example of demand and the associated bounds, as well as left and right hand-side values of the aggregate demand constraints (57). Note that all these bounds are also valid when dealing with aggregate lightpath set $\tilde{L}$ used in virtual routing formulation ([BVR]) of (35-39). Upper bounds (61) remain valid when using classic demand constraints (7).

Table 2: An example of demand vector for different granularities and associated capacity reservation bounds.

				single-hop lightpath	2-hop lightpath		
t_k	D_k	$\overline{b}_k$	$\underline{b}_k$	$\overline{b}_{\ell k}$	$\overline{b}_{\ell k}$	RHS$_k$	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

5.3 Zero-One Discretization

Zero-One extended formulations resulting from unary decomposition of traffic have been a source of progress in handling network design applications [9, 10, 11, 12]. For the [OC] formulation given in (6-11) such discretization enables us to tighten the model with valid inequalities that have been proposed for capacitated network design problems in the binary case [2]. We shall use these tightening when dealing with model [BVR] given in (35-39) and model [BGPR] given in (26-30). The 0-1 reformulation arises from the unary decomposition of capacity reservation variables $x_{\ell k}$. We implement the change of variables:

$$x_{\ell k} = \sum_{q \in Q_{\ell k}} q\, x_{\ell k q} \text{ with } \sum_{q \in Q_{\ell k}} x_{\ell k q} = 1,\ x_{\ell k q} \in \{0,1\}\ \forall q \in Q_{\ell k}, \tag{65}$$

where $x_{\ell kq} = 1$ 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 $q \in Q_{\ell k} \setminus \{0\}$, as $x_{\ell k} = 0$ can be achieved by setting all $x_{\ell kq} = 0$. In the sequel, when dealing with 0-1 transformations, we shall consider that 0 has therefore been omitted from the value domain, i.e., $\tilde{Q}_{\ell k} = Q_{\ell k} \setminus \{0\}$ and $\sum_{q \in \tilde{Q}_{\ell k}} x_{\ell kq} \leq 1$.

Note that for a single-hop aggregate lightpath $\ell \in \tilde{\mathcal{L}}$ and $t_k = t^{\max}_{s_k d_k}$, $\overline{b}_{\ell k}$ can be quite large as it is not constrained by bounds (60) or (61). To avoid dealing with too many $x_{\ell kq}$ 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 $x'_{\ell k}$ as the residual reservation. Then, $x_{\ell k} = x'_{\ell k} + x''_{\ell k}$ with

$$x'_{\ell k} \in Q'_{\ell k} = \{0, \ldots, \frac{U}{t_k} - 1\} \text{ and } x''_{\ell k} \in Q''_{\ell k} = \{0, \frac{U}{t_k}, 2\frac{U}{t_k}, \ldots, \left\lfloor \frac{\overline{b}_{\ell k}\, t_k}{U} \right\rfloor \frac{U}{t_k}\}. \tag{66}$$

We then apply the classical unary decomposition (65) to $x'_{\ell k}$ and $x''_{\ell k}$ separately. Note that we can omit the distinction x' and x'' variables and use notation $x_{\ell kq}$ for $q \in Q'_{\ell k} \cup Q''_{\ell k}$ without any confusion as $Q'_{\ell k} \cap Q''_{\ell k} = \{0\}$, i.e., the 0-1 transformation is $x_{\ell k} = \sum_{q \in Q'_{\ell k} \cup Q''_{\ell k}} q\, x_{\ell kq}$, $\sum_{q \in Q'_{\ell k} \cup Q''_{\ell k}} x_{\ell kq} = 1$. 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, $\tilde{Q}'_{\ell k}$ is either empty or a singleton.

5.4 Valid Inequalities

We make use of valid inequalities that are traditionally used to improve the formulation of network design problems. Here we simply show how they can be adapted to our model. The **cut set** inequalities [1, 6, 21] are well-known valid inequalities for the network loading problem. 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:

$$\sum_{d \in V \setminus \{s\}} \sum_{p \in \mathcal{P}_{sd}} y_p \geq \left\lceil \frac{\sum_{d \in V \setminus \{s\}} D_{sd}}{U} \right\rceil \quad \forall s \in V\ , \qquad \sum_{s \in V \setminus \{d\}} \sum_{p \in \mathcal{P}_{sd}} y_p \geq \left\lceil \frac{\sum_{s \in V \setminus \{d\}} D_{sd}}{U} \right\rceil \quad \forall d \in V. \tag{67}$$

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:

$$\sum_{\ell \in \mathcal{L}_{sd}} \left\lceil \frac{\sum_{k \in K_{sd}} t_k\, x_{\ell k}}{U} \right\rceil \geq \left\lceil \frac{D_{sd}}{U} \right\rceil \quad \forall (s,d) \in V^2. \tag{68}$$

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:

$$\sum_{\ell \in \mathcal{L}_{sd}} \left\lceil \frac{t_k\, x_{\ell k}}{U} \right\rceil \geq \left\lceil \frac{t_k\, \underline{b}_k}{U} \right\rceil \quad \forall k \in K. \tag{69}$$

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

$$\overline{\alpha}_{sd} = \min\{u \in [1,U] : u\,(1 + \max\{\left\lfloor \frac{D_{sd}}{U} \right\rfloor, 1\}) > D_{sd}\} = \left\lfloor \frac{D_{sd}}{\max\{\left\lfloor \frac{D_{sd}}{U} \right\rfloor, 1\} + 1} \right\rfloor, \tag{70}$$

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:

$$\sum_{\ell \in \mathcal{L}_{sd}} \left(\left\lfloor \frac{\sum_{k \in K_{sd}} t_k \, x_{\ell k}}{U} \right\rfloor + \delta\big((\sum_{k \in K_{sd}} t_k \, x_{\ell k}) \bmod U > \overline{\alpha}_{sd} \big) \right) \leq \max\{ \left\lfloor \frac{D_{sd}}{U} \right\rfloor, 1 \} \quad \forall (s,d) \in V^2, \tag{71}$$

where $\sum\limits_{\ell \in \mathcal{L}_{sd}} \left\lfloor \frac{\sum_{k \in K_{sd}} t_k \, x_{\ell k}}{U} \right\rfloor$ 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 ℓ to carry larger than average bandwidth.

Using a 0-1 discretization of the traffic reservation variables x allows us to enforce tighter relations with design variables. The so-called **strong linking** inequalities [10] 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, $y_p = \sum_{\lambda \in \Lambda} y_{p\lambda}$, are not binary but general integer, hence strong linking inequalities take the form of GUB constraints:

$$\sum_{q \in \tilde{Q}_{\ell k}} \left\lceil \frac{q \, t_k}{U} \right\rceil x_{\ell k q} \leq y_p \quad \forall k \in K, \; p \in \mathcal{P}, \; \ell \in \mathcal{L}_{s_k d_k} : \; y_p^\ell = 1. \tag{72}$$

One can also derive Gomory cuts and cover cuts from knapsack type constraints, such as the flow bound constraints $\underline{b}_k \leq \sum\limits_{\ell \in \mathcal{L}_{s_k d_k}} \sum_{q \in \tilde{Q}_{\ell k}} q \, x_{\ell k q} \leq \overline{b}_k$. The so-called k**-demand cuts** [1], take the form

$$\sum_{\ell \in \mathcal{L}_{s_k d_k}} \sum_{q \in \tilde{Q}_{\ell k}} \left\lfloor \frac{q}{\epsilon} \right\rfloor x^k_{\ell q} \leq \left\lfloor \frac{\overline{b}_k}{\epsilon} \right\rfloor \quad \text{and} \quad \sum_{\ell \in \mathcal{L}_{s_k d_k}} \sum_{q \in \tilde{Q}_{\ell k}} \left\lceil \frac{q}{\epsilon} \right\rceil x^k_{\ell q} \geq \left\lceil \frac{\underline{b}_k}{\epsilon} \right\rceil, \tag{73}$$

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 (57) (see [35]). We also use **Lifted Knapsack Cover** and ϵ**-split and** c**-strong** inequalities proposed by [2]. Separation is done through enumeration since all the cut classes that we consider are of polynomial size. Details and further inequalities are presented in [35], such as generalized upper bound constraints. Note that the domain reduction techniques of Section 5.3, also results in tighter domain for integer design variables such as y_{ij}, and hence make the valid inequalities of this section even be more efficient (in particular strong linking inequalities).

5.5 Optimizing Bandwidth Reservation on a Grooming Pattern

The pricing problem (19) or (25) relies on a core sub-problem consisting in making optimal bandwidth reservation for a fixed single-hop lightpath. Given the symmetry breaking restrictions as defined 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 does not carry more traffic than the aggregate demand bound defined in (57).

Generating proper columns is known to strengthen the column generation formulation as shown in [33]. Thus, given a lightpath ℓ on a (s,d) pair, a bandwidth reservation $b \leq U$ is feasible if the following integer polyhedron is not empty:

$$X_{sd}(b) = \{x_k \in \mathbb{Z}_+{}^{|K_{sd}|} : \sum_{k \in K_{sd}} t_k\, x_k \;=\; b \tag{74}$$

$$\sum_{k' \in K_{sd}: t_{k'} \geq t_k} \frac{t_{k'}}{t_k} x_{k'} \;\leq\; \text{LHS}_k \quad \forall k \in K_{sd} \tag{75}$$

$$x_k \;\leq\; \frac{t_k^+}{t_k} - 1 \quad \forall k \in K_{sd} :\; t_k \neq t_{sd}^{\max}\}. \tag{76}$$

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 π, computed from the dual variables associated with demand constraints (57), the function

$$\psi^{sd} : u \in \{0,\ldots,U\} \to \psi^{sd}(u) = \max\{\sum_{k \in K_{sd}} \pi_k\, x_k :\; 0 \leq b \leq u,\; X_{sd}(b) \neq \emptyset,\; x \in X_{sd}(b)\}, \tag{77}$$

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 ψ^{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.

5.6 Generating Wavelength Configurations

The pricing problem of wavelength routing configurations that is given in (24) 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, the wavelength routing configuration pricing problem (24) can be written as:

$$\min\{\sum_{o \in \mathcal{O}} \bar{c}_{g_o^\star} \, \mu_{g_o^\star} : \sum_{o \in \mathcal{O}} \sum_{p \in \mathcal{P}} z_a^p \, y_p^o \, \mu_{g_y^\star} \leq 1 \quad \forall a \in A\}, \tag{78}$$

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 LHS_k defined in (57). If the solution of (78) does not satisfy the proper column bound constraints:

$$\sum_{k' \in K_{s_k d_k} : t_{k'} \geq t_k} \frac{t_{k'}}{t_k} \, x_{k'}^c \leq \text{LHS}_k \quad \forall k \in K, \tag{79}$$

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 (79): a greedy heuristic and an exact method making use of an in-situ column generation technique [24].

The greedy algorithm is a standard procedure which is presented in [35]. The in-situ column generation [24] 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). 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 (79). Let $\bar{K}$ be the set of demands for which constraints (79) are violated. Let $\tilde{\mathcal{O}} \subset \mathcal{O}$ be the optical configurations for which $g_o^\star$ is involved in some violated constraints (79), and $\bar{\mathcal{O}} = \mathcal{O} \setminus \tilde{\mathcal{O}}$. For each $o \in \tilde{\mathcal{O}}$, we define an indicator variable $\kappa_o = 1$ if we chose to use a grooming pattern (defined by x_k^o) for this optical hop configuration. Then, the problem of **Pricing of Wavelength Configuration** can be written as:

$$\min \sum_{o\in\bar{\mathcal{O}}} \bar{c}_{g^\star_o}\, \mu_{g^\star_o} + \sum_{o\in\tilde{\mathcal{O}}} (\sum_{p} y^o_p \kappa_o - \sum_{k\in K} \pi_k\, x^o_k) \tag{80}$$

$$[\text{PWC}] \sum_{o\in\bar{\mathcal{O}}} \sum_{p\in\mathcal{P}} z^p_a\, y^o_p\, \mu_{g^\star_o} + \sum_{o\in\tilde{\mathcal{O}}} \sum_{p\in\mathcal{P}} z^p_a\, y^o_p \kappa_o \;\leq\; 1 \quad \forall a \in A \tag{81}$$

$$\sum_{o\in\tilde{\mathcal{O}}} \sum_{k'\in K: t_{k'}\geq t_k} \frac{t_{k'}}{t_k} x^o_k \;\leq\; \text{LHS}_k \quad \forall k \in \bar{K} \tag{82}$$

$$\kappa_o \;\in\; \{0,1\} \quad \forall o \in \tilde{\mathcal{O}} \tag{83}$$

$$(x^o_k, \kappa_o) \;\in\; \mathcal{G}(o) \quad \forall k \in K,\; o \in \tilde{\mathcal{O}} \tag{84}$$

$$\mu_{g^\star_o} \;\in\; \{0,1\} \quad \forall o \in \bar{\mathcal{O}}. \tag{85}$$

The mathematical programming formulation involves replacing $(x^o_k, \kappa_o) \in \mathcal{G}(o)$ by the polyhedral description of $\mathcal{G}(o)$ as given in a specific example in (12) where bound constraints are transformed into variable bounds implying setup variable κ_o.

6. Selected Models and Algorithms

Four alternative solution approaches have been developed. They correspond to solution to the four models that we have found the most tractable for computational purposes. Table 3 summarises the formulations that we introduced. Their abbreviation starts with 'O' for the original formulation, 'D' for a Dantzig-Wolfe reformulation, 'B' for Benders, 'P' for a pricing subproblem, and 'F' for a feasibility subproblem. The columns of the Table provide the model's name, its reference, the formulation from which it is derived – or is a subproblem of –, the assumptions underlying the model, and the theoretical comparative value of the LP relaxation (omitting the term $\phi(y)$ in Benders' models): 0 refers to the trivial bound of Proposition 2, while number 1, 2, 3 and 4 refer to the ordering of the LP bounds in a sorting by non-decreasing value. The four models that we solved are marked with *** signs: they are [DWRC], [BGPR], [BVR], and [HGP]. We shall also consider the 0-1 discretization of [BGPR], [BVR]. The remaining formulations have served to derive the selected models or they appear as a subproblem in the proposed approaches. We now describe briefly the overall algorithm used for each of the four selected models. We discuss how the elements of the previous section are used in combination and the selected implementation strategies.

6.1 A Column Generation Approach for the Wavelength Routing Configuration Formulation, [DWRC]

The Wavelength Routing Configuration formulation (20-23) is denoted by [DWRC]. Its LP relaxation is solved using a nested column generation approach which requires a large computing time. The enhanced

Table 3: Formulations for the GRWA

Logo	**Name**	**Ref**	**From**	**Assump.**	**LP**
Compact Formulation					
[OC]	Compact Original formulation	(6-11)		1-4	0
Formulations derived from a Dantzig-Wolfe decomposition principle					
[DGP]	D-W Grooming Pattern	(13-18)	[OC]	1-5	3
[DWRC]	*** D-W Wavelength Routing Configuration ***	(20-23)	[OC]	1-6	4
Formulations derived from a Benders decomposition principle					
[BGPR]	*** Benders' Grooming and Physical Routing ***	(26-30)	[OC]	1-4	0
[BVR]	*** Benders' Virtual Routing ***	(35-39)	[BGPR]	1-4	0
Formulations derived from a Benders' decomposition applied to a D-W reformulation					
[HGP]	Hybrid Grooming Pattern	(46-48)	[DGP]	1-5	1
[HAGP]	*** Hybrid Aggregated Grooming Pattern ***	(49-51)	[HGP]	1-5	2
Formulations of Dantzig-Wolfe pricing subproblems					
[PGC]	Pricing Grooming Configuration	(19)	[DGP]	1-5	
[PWC]	Pricing Wavelength Configuration	(24) and (80-85)	[DWRC]	1-6	
Formulations of second or third-stage Benders Feasibility subproblems					
[FWA]	Wavelength Assignment	(32-34)	[BGPR]	1-4	
[FGPA]	Feasible Grooming and Path Assignment	(40-45)	[BVR]	1-4	
[FGPWA]	Grooming Pattern and Wavelength Assignment	(52-56)	[HAGP]	1-5	

demand covering constraints take the form:

$$\sum_{c \in \mathcal{C}} \sum_{\substack{k' \in K_{s_k d_k}: \\ t_{k'} \geq t_k}} \frac{t_{k'}}{t_k} \, x^c_{k'} \, \nu_c \geq \mathrm{RHS}_k \quad \forall k \in K. \tag{86}$$

Aggregate bounds $\underline{b}_k$ and $\overline{b}_k$ and LHS_k are not explicitly formulated in [DWRC] because they induce difficulties for finding a feasible solution of the LP relaxation, although enforcing these bounds would have result in tighter 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 (67) that takes the form

$$\sum_{d \in V \setminus \{s\}} \sum_{c \in \mathcal{C}} \sum_{p \in \mathcal{P}_{sd}} y_p^c \, \nu_c \geq \left\lceil \frac{\sum_{d \in V \setminus \{s\}} D_{sd}}{U} \right\rceil \quad \forall s \in V \tag{87}$$

and

$$\sum_{s \in V \setminus \{d\}} \sum_{c \in \mathcal{C}} \sum_{p \in \mathcal{P}_{sd}} y_p^c \, \nu_c \geq \left\lceil \frac{\sum_{s \in V \setminus \{d\}} D_{sd}}{U} \right\rceil \quad \forall d \in V, \tag{88}$$

where indicator y_p^c is one if configuration c uses path p. We also use the (s,d)-lightpath inequalities (68): $\sum_{c \in \mathcal{C}} y_{sd}^c \, \nu_c \geq \left\lceil \frac{D_{sd}}{U} \right\rceil \quad \forall (s,d) \in V^2$, where indicator y_{sd}^c is equal to the number of grooming patterns on c that carry (s,d) traffic, and k-lightpath cuts (69): $\sum_{c \in \mathcal{C}} y_k^c \, \nu_c \geq \left\lceil \frac{t_k \, \underline{b}_k}{U} \right\rceil \quad \forall k \in K : \underline{b}_k > 0$, 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 (71) 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 α_{sd} unit of (s,d) flow. The dual variables associated with these cuts either result in modified dual values π_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 for a given set of cuts, we enumerate the valid inequalities and add the violated ones.

The primal bounds 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.

6.2 A 2-stage Hierarchical Optimization for the Grooming and Physical Routing Formulation, [BGPR]

We apply a hierarchical optimization approach based on Benders' Grooming and Physical Routing formulation (26-31) denoted by [BGPR]. We solve [BGPR] ignoring the term $\phi(y)$ that models the feasibility of the second stage problem [FWA] given in (32-34). The resulting dual bound is that of Proposition 2. We considered this relaxation of Benders' first stage problem because generating Benders' cuts cannot be done in a standard way given that feasibility subproblem [FWA] is not an LP. Note that solving the LP relaxation of this sub-problem would be 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 it would never return feasibility cuts. In [17], a sufficient property on the set of selected optical hops for the feasibility of the [FWA] problem has been presented. They might be used to generate Benders' feasibility cuts, but we did not investigate this issue.

In fact, we have obtained better results using the 0-1 discretization of Section 5.3, a model that we denote by [BGPR 0-1]. This extended formulation is solved using CPLEX and we apply cutting plane methods to improve the dual bound. We restrict solutions by fixing bandwidth reservation on a dedicated single-hop lightpath as presented in Observation 4, and it slightly improves the dual bound. Primal solutions are obtained using CPLEX default primal heuristics. We verify a posteriori the true feasibility of the proposed primal solution by solving the [FWA] problem (32-34) using CPLEX.

6.3 A 3-stage Hierarchical Optimization for the Virtual Routing Formulation, [BVR]

The Virtual Routing formulation (35-39) is denoted by [BVR]. Again we consider a relaxation of this first stage of the Benders' approach, where we ignore the term $\phi(y)$. Its LP optimum provides a valid dual bound, although it is the trivial bound of Proposition 2. We tackle (35-39) directly by branch-and-bound using CPLEX with default cut generation and primal heuristics. We also test the impact of the cut set constraints (67). However, a cutting plane approach is more effective on the 0-1 discretization variant denoted by [BVR 0-1] that makes use of the unary decomposition presented in Section 5.3. Then, we use the cut set constraints (67), strong linking cuts (72), Gomory cuts (73), lifted cover cuts on constraints and ϵ-split c-strong cuts on constraints with $1 \leq \epsilon \leq 4$ [2, 13]. Cut set, strong linking and Gomory cuts can be enumerated at each iteration of the cutting plane method, whereas for lifted knapsack cover and ϵ-split c-strong inequalities, we use the separation method presented in [2, 13]. 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 $x_{\ell kq}$. For instance, the strong linking cuts takes the form: $\sum_{q \in \tilde{Q}_k^{a\ell}} x_{\ell kq} \leq y_{ij} \quad \forall (i,j) \in V^2,\ k \in K,\ \ell \in \mathcal{L}_{s_k d_k} :\ y_{ij}^{\ell} = 1$. Primal solutions to [BVR] or [BVR 0-1] that are obtained with CPLEX, need to be checked for feasibility using the two-step procedure described in Section 3.2. We first solve (40-45) and then (32-34). We observed that adding constraints to enforce request-commodity upper bounds (62) helps CPLEX primal heuristic to find better primal solutions.

6.4 A Column Generation Approach to the Hierarchical Optimization of the Aggregated Grooming Pattern Formulation, [HAGP]

Benders' reformulation (49-51) of the Aggregated Grooming Pattern formulation is denoted by [HAGP]. 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: $\sum_{g \in \tilde{\mathcal{G}}} \sum_{\substack{k' \in K_{s_k d_k}: \\ t_{k'} \geq t_k}} \frac{t_{k'}}{t_k} \, x^g_{k'} \, \mu_g \leq \mathrm{LHS}_k \; \forall k \in K$. The solution of the grooming pricing problem was discussed in Section 5.5. At each iteration of the column generation procedure, we add multiple columns to the master (all the grooming patterns with a negative reduced cost). To improve the dual bound, we use the same valid inequalities as for [DWRC] formulation given in (87-88). 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 [DWRC]. 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 [FGPWA] problem defined in (52-56).

7. Numerical Results and Comparison

The four above approaches have been tested and compared on realistic size data sets. For model [BVR] and [BGPR], one can either consider the integer formulation or its 0-1 discretization, denoted with a 0-1 postfix. We report results on both [BVR] and [BVR 0-1] as each model has its pros and cons, while we only report results with [BGPR 0-1] that are better than with [BGPR]. Thus, in total, we deal with 5 models (or their 0-1 variant) from Table 3: (i) [DWRC], (ii) [BGPR 0-1], (iii) [BVR], (iv) [BVR 0-1], and (v) [HAGP]. For the tests, we generated data instances that represent "realistic" data for four different instances for the NSF network (14 nodes and 21 edges) [22] and the EON network (20 nodes and 39 edges) [27]. Details on the generation of these instances can be found in [35]. Table 4 summarizes the characteristics of each instance and gives the total number of requests, the overall required bandwidth, and 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, $csDB$:

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

Table 4 also provides our best dual bound under Assumptions 1-4, DB, our best primal bounds over all four approaches, PB, and the single-hop routing solution primal bound, $shPB$. The numbers between parentheses next to the best bounds refer to the model under Assumptions 1-4 that releases this best bound (i.e. (i) [DWRC] and (v) [HAGP] are not considered as they rely on further Assumptions 5-6).

In Table 5, 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 with our separation

Table 4: 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 (iii)	209 (iii) (5.5)	269 (35.8)
NSF 2	1,332	26,918	13	141	149	159 (ii), (iii), (iv)	170 (iv) (6.9)	182 (14.4)
NSF 3	1,949	40,064	25	209	216	227 (iii), (iv)	246 (iii) (8.3)	364 (60.3)
NSF 4	60,303	114,328	47	596	603	614 (iii)	624 (iii) (1.6)	685 (11.5)
EON 1	6,639	78,837	33	411	420	461 (ii)	484 (iii) (4.9)	605 (31.2)
EON 2	2,292	36,547	18	191	201	237 (ii), (iv)	258 (iii) (8.8)	380 (60.3)
EON 3	3,667	79,592	36	415	425	468 (ii)	493 (iv) (5.3)	741 (58.3)
EON 4	120,676	217,337	67	1,132	1,143	1,188 (ii)	1211 (iii) (1.9)	1,329 (11.8)
average value	24,816	78,454	32.2	409.1	417.7	444	461.8 (5.4)	569.3 (35.5)

Table 5: Comparing dual bound at the root node.

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

Table 6: Comparing dual bounds obtained after running CPLEX default branch-and-cut for 1 hour.

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

Table 7: Comparing primal bounds PB obtained with the primal heuristics of this paper without adding cutting planes before hand.

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

Table 8: Comparing primal bound PB obtained with the primal heuristics of this paper applied after running the cutting plane procedure of this paper.

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

procedure. There is no time limit set for our own separation routines, which is a brute force enumeration of valid inequalities. The last row provides the average computation time in seconds. We re-emphasize that our models might not be strictly comparable because they may rely on different assumptions sets as it was clearly indicated in Table 3, even if most instances may not be affected by such restrictive assumptions. Here, a "*" indicates that the bound is only valid under restrictive Assumptions 5 or 6. In Table 6, 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, $babDB$, and those obtained when running CPLEX default branch-and-cut with a time limit of 1 hour after adding all the cuts of Section 5.4 at the root node, $bacDB$ (again there is no time limit for our own separation routines). This can only be done for the direct formulations that do not require dynamic column generation. In Tables 7 and 8, 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 7, we do not use the cuts of Section 5.4, while the results of Table 8 are obtained making use of our cutting plane procedure. We give the average computational time on the 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. For the Wavelength Routing Configuration formulation (20-23), [DWRC], the LP relaxation is very hard to solve to optimality. The computational times are very large. The resulting dual bound, although it is the strongest of all our models in theory, it is not as good as that of formulation [HAGP], given in (46-48), because we do not use all the possible formulation strengthening such as the aggregate flow upper bound LHS_k. The time reported in Table 5 includes the time consuming cutting plane method. The primal bounds have a poor quality because fixing wavelength configurations 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.)

2. For the Grooming and Physical Routing formulation (26-30) in its 0-1 form, [BGPR 0-1], the LP relaxation bound is better than the trivial bound. This is due to the domain restrictions on 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 amongst the models that do not make restrictive Assumptions 5-6. However, with this formulation, very few primal bounds were obtained and, moreover, their costs are not as good as for the [BVR] formulation. This can be explained by the fact that [BGPR 0-1] has many more variables than [BVR], which makes CPLEX heuristics less efficient. The computing times of the root dual bound with our cutting plane method are quite large: [BGPR 0-1] involves more constraints and more induced cuts than [BVR]. 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 used all the families of valid inequalities of Section 5.4 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).

3. With Benders Virtual Routing formulation (35-39), [BVR], 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. 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 6,000 (resp. 7,000) seconds on

average when solving the three stage procedure without (resp. with cut set) inequalities. The valid solution obtained in the second stage has often the same cost as the first stage incumbent.

For [BVR 0-1], we observe that the LP relaxation value is better than the trivial dual bound thanks to our preprocessing of variable domains. The dual bounds obtained applying CPLEX default branch-and-cut to the original formulation (without the cuts of Section 5.4) are not better than for the integer version, denoted [BVR]. However, adding the cuts of Section 5.4 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 our cutting plane procedure and we consider all the feasible inequalities of Section 5.4 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. Although the comparison is unfair given the different time limit, it indicates the interest of the cuts of Section 5.4. Combining these cuts with those generated with the built-in cutting plane method of CPLEX yields extra improvements. The dual bounds obtained from the [BVR 0-1] formulation are very close to those obtained from [BVR]. Because of the large number of variables in [BVR 0-1], CPLEX primal heuristics tend to be less effective (we are not always able to find integer solutions to the first stage problem [BVR 0-1]). However, it gives two of the best primal bounds. For many instances, no primal solutions were found at the first stage; then the second stage problem was not called. Hence, in Table 7, the average computational time is smaller for [BVR 0-1] than for [BVR]. However, when using <all the families of valid inequalities of Section 5.4, 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.

4. For the Benders reformulation (49-51) of the Grooming Pattern formulation, [HAGP], 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 [BVR], 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 5,500 seconds using all the cuts of Section 5.4; 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.

In summary, our numerical experiments show that, under a computing time limit, the best bounds are not derived from the theoretically strongest formulation as they are compared in Table 3, but by applying branch-and-cut approaches to weak formulations exploiting problem specific cuts and CPLEX built-in cut generators as well as CPLEX primal heuristics: on average the best dual bounds are obtained with [BGPR 0-1], while the best primal bounds are derived from the [BVR]. However, if we do not use CPLEX branch-and-bound, its built-in cut generators, and primal heuristics (all of which are only suitable for compact formulations), the best dual bounds are obtained with the hybrid approach relying on both Dantzig-Wolfe and Benders' decomposition: formulation [HAGP] provides the best dual bounds at root node on average; although [HAGP] relies on restrictive Assumptions 5. The hybrid Dantzig-Wolfe / Benders' approach also provides good primal solutions and does so for all instances unlike [BGPR 0-1] and [BVR 0-1]. Our pure column generation approach, using [DWRC], is comparatively much more time consuming. Zero-One discretizations provide better dual bounds at the root node and more opportunity to derive cutting planes, but they demand (much) larger computing time (although our computing times could probably be much improved by developing more efficient separation procedure for the cuts of Section 5.4). Moreover, primal heuristics are not as efficient in these larger variable spaces.

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 primal and dual bounds we have developed and compared (a) a column generation approach for a Dantzig-Wolfe reformulation, (b) hierarchical optimization approaches based on Benders' decompositions, and (c) an original hybridization of these two techniques. Important features in our approaches are symmetry breaking reformulations that avoid wavelength indexing or allow for request exchanges. We also break symmetries by restricting the solution set to solutions involving large granularity reservations and favoring single-hop routing. Further enhancements are obtained by adapting 0-1 discretization and cutting plane procedures that can be found in the network design literature. The 0-1 discretization also lead to an original variable domain reduction: our refinement (66) for this problem was quite helpful computationally. The primal bounds were obtained using either our own rounding procedure (equivalent to a depth first dive into a branch-and-price tree) for column generation formulations, or CPLEX built-in primal heuristics for the compact formulations. For both dual and primal bounds, we manage to exploit CPLEX capabilities rather than compete with it.

We obtained provably good solutions for the GRWA problem (an average gap of around 5%), while the gap reported in the literature for this problem are generally relatively large. Beyond this, our main contributions consisted in analyzing the comparative advantages of different formulations and associated solutions methods. The hybridization of Dantzig-Wolfe and Benders' decomposition is an originality of this study along specific domain reduction techniques that can be useful in similar applications. Some of our implementations involve advanced techniques (such as nested decomposition, multi-stage hierarchical optimization, in-situ column generation, or symmetry breaking features) that might inspire other work. Further developments would 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). We see as the most promising research direction the further development of the hybrid Dantzig-Wolfe / Benders' approach on formulation [HAGP]: one could project its solution in the compact space of the x and y variables and derive cutting planes or apply primal heuristics on this projection. The current picture given by our numerical results is biased in favour of the pure hierarchical approaches because they can take advantage of the professional implementation of CPLEX for both primal and dual bounds, while the numerical comparisons without CPLEX branch-and-cut and primal heuristics are an indication of the potential of the hybrid Dantzig-Wolfe / Benders' approach.

Acknowledgments

We express deep thanks to the anonymous referees for their in-depth report and constructive remarks that have contributed to improve our submission.

References

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

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

[3] 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.

[4] 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.

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

[6] B. Brockmüller, 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.

[7] 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.

[8] B. Chen, G.N. Rouskas, and R. Dutta. Clustering methods for hierarchical traffic grooming in large-scale mesh WDM networks. *Journal of Optical Communications and Networking*, 2(8):502–514, August 2010.

[9] I. Correia, L. Gouveia, and F. Saldanha d Gama. Discretized formulations for capacitated location problems with modular distribution costs. *European Journal of Operational Research*, 204:237–244, 2010.

[10] A. Frangioni and B. Gendron. 0-1 reformulations of the multicommodity capacitated network design problem. *Discrete Applied Mathematics*, 157(6):1229–1241, 2009.

[11] Antonio Frangioni and Bernard Gendron. A stabilized structured Dantzig-Wolfe decomposition method. *Mathematical Programming*, 140(1):45–76, 2013.

[12] Bernard Gendron and Gouveia Luis. Reformulations by discretization for piecewise linear integer multicommodity network flow problems. *forthcoming in Transportation Science*, 2015.

[13] 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.

[14] M. Hamdouni, 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.

[15] J.N. Hooker and G. Ottosson. Logic-based benders decomposition. *Mathematical Programming*, 96(1):33–60, 2003.

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

[17] 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.

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

[19] 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.

[20] 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.

[21] 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.

[22] R.M. Krishnaswamy and K.N. Sivarajan. Design of logical topologies: A linear formulation for wavelength-routed optical networks with no wavelength changers. *IEEE/ACM Transactions on Networking*, 9(2):186–198, april 2001.

[23] Gilbert Laporte and François V. Louveaux. The integer l-shaped method for stochastic integer programs with complete recourse. *Operations Research Letters*, 13(3):133 – 142, 1993.

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

[25] J. F. Lutz. Delay limitations when extending metro ethernet into the wide area. In *Optical Fiber Communication Conference and Exposition and The National Fiber Optic Engineers Conference, Technical Digest (CD)*. Optical Society of America, 2005.

[26] B. Mukherjee. *Optical WDM Networks*. Springer, 2006.

[27] M.J. O'Mahoney, D. Simeonidou, A. Yu, and J. Zhou. The design of a european optical network. *Journal of lightwave technology*, 13(5):817–828, May 1993.

[28] 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.

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

[30] 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.

[31] Suvrajeet Sen and Hanif D. Sherali. Decomposition with branch-and-cut approaches for two-stage stochastic mixed-integer programming. *Mathematical Programming*, 106(2):203–223, 2006.

[32] F. Sumika, H. Souza, D.L. Guidoni, and G.R. Mateus. A column generation-based heuristic for the GRWA with protection and QoS in WDM optical networks. In *IEEE Symposium on Computers and Communications (ISCC)*, pages 922–927, 2013.

[33] 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.

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

[35] B. Vignac. *Reformulation and Decomposition Based Methods for a Multi-Layer Capacitated Network Design Problem*. PhD thesis, Université de Montréal / Université Bordeaux 1, 2010.

[36] H. Wang and G.N. Rouskas. Traffic grooming in optical networks: Decomposition and partial linear programming (LP) relaxation. *Journal of Optical Communications and Networking*, 5(8):825–835, August 2013.

[37] H. Wang and G.N. Rouskas. Hierarchical traffic grooming: A tutorial. *Computer Networks*, 69:147–156, 2014.

[38] 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.

[39] 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.

[40] 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.

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

[42] A. Zymolka. *Design of Survivable Optical Networks by Mathematical Optimization*. Phd thesis, TU Berlin, 2007.