Optical Packet Network with Limited Range Wavelength Conversion: a novel formalization of the optimal scheduling problem

Andrea Detti, Giorgia Parca, Valeria Carrozzo, Silvello Betti

***Abstract*— We consider synchronous optical packet networks formed by switches equipped with a complete-set of limited-range wavelength converters. On these networks, we dealt with scheduling algorithm that maximizes the switch throughput.**

So far, previous literature works have formalized this scheduling problem as *the finding of a maximum bipartite matching* (MBM) in a convex graph. The MBM formalization has collected several follow-ups, mainly focused on measuring switch-level performance. We revise the MBM formalization by measuring network-level performance. Surprisingly, we find out that when optical switches are cascaded, MBM formalization has two not negligible lacks: i) a useless degradation of optical signal quality; ii) a tendency in shifting optical packets toward lower wavelengths, thus increasing the occurrence of wavelength contention.

To solve these issues, we propose a novel formalization of the scheduling problem as *the finding of a maximum bipartite matching with minimum edges weights* (MW-MBM). We show that MW-MBM outperforms MBM in terms of both network throughput and optical signal to noise ratio. Performance evaluation is carried out by means of NS2 simulator that we extend to toughly model optical components (e.g., SOA FWM wavelength converter). The simulator is provided as open-source.

***Index Terms*—Limited range wavelength conversion, optical packet switching, wavelength division multiplexing network, optical switch, optimum scheduling, Hungarian method.**

INTRODUCTION

All optical packet switching is an attractive integration of IP and WDM technologies for devising terabits routers with limited dimension, weight and power consumption [1].

An optical packet is an IP packet or an aggregate of them (so called, *burst*). Optical packets are transferred on different wavelengths through optical fibers that interconnect optical switches. The packet payload ever remains in the optical domain; while, switch-by-switch, the header undergoes an O/E/O conversion for electronic processing.

Packet routing is the first processing task that a switch performs by elaborating the header of an entering packet. Packet routing finds the output fiber over which the packet should be forwarded in order to reach the destination. After the routing, there could be many entering packets contending the use of the same output fiber. Since those packets can demand more optical resources than available ones, a scheduling algorithm carries out a selection. The scheduler defines which packets can be actually forwarded on the output fiber and which are their output wavelengths. Finally, the forwarding is accomplished by opportunely controlling the switch optical devices, e.g. space matrix, optical gates, wavelength converters, etc.

The rate of packets forwarded by a switch (i.e. the *throughput*) is strongly enhanced by the use of wavelength converters [2]. Without wavelength converter, when more packets with the same input wavelength are routed on the same output fiber, a "wavelength contention" occurs and the scheduler can forward only one packet[1]. With wavelength converters, the scheduler is able to suitable change the input wavelengths, making possible the forwarding of more packets.

The networking benefits of wavelength conversion could be paid at physical layer with a degradation of the signal quality, especially harsh when the wavelength detuning is wide [4][5]. Moreover, a wide detuning may slow down the switching speed due to the longer time needed to tune a wavelength over a wider range [7]. For these reasons, in practical systems, the conversion range is limited to a feasible value and literature refers to this capability as "limited range wavelength conversion".

The set of wavelength converters available in a switch may be "complete" or "shared" [3]. In case of complete-set there is a wavelength converter for each output wavelength; thus, any packet can be relayed on an output wavelength, independently of its input wavelength. On the contrary, the shared-set architecture provides a restricted number of wavelength converters that are commonly used by incoming packets according to a scheduling decision. The complete-set architecture is the most expensive and provides maximum throughput performance. The shared-set architecture is cheaper and, suitably dimensioned, may furnish performance near to the maximum one.

All optical packet network can be slotted (i.e., synchronous) or un-slotted (i.e., asynchronous) [1]. In the slotted case the time axis is organized as a succession of fixed time-slots, the packet duration is equal to the time slot and

Andrea Detti and the other Authors are with the Electronic Engineering Department, University of Rome, Tor Vergata, Italy (e-mail: {name.surname}@uniroma2.it).

[1] We are assuming the absence of optical buffering. Indeed, current buffer technology is based on fiber delay line which makes too much cumbersome the optical switch.

packets are temporary aligned at the switch enter. Consequently, slot-by-slot a set of packets coming from different wavelengths/fibers have to be mapped on output wavelengths and there are "several" mapping possibilities; therefore, it is worth investigating optimal scheduling algorithms [8]. In an un-slotted network, packet duration is variable and packets are not temporary aligned at the switch enter. Therefore, the scheduler handles only "one" packet at a time and scheduling algorithm simply consists in forwarding the packet on an output wavelength currently free.

In this paper, we deal with scheduling algorithms that maximize switch throughput. The network is slotted, the switches are equipped with a complete-set of limited range wavelength converters and there is only best-effort traffic. With this regard, the two papers [8] [9] have formalized the scheduling problem as *the finding of a Maximum Bipartite Matching (MBM) in a convex graph* with un-weighted edges[2]. The MBM formalization has received several literature follows-ups (e.g., [10][11][12][13][14]), mostly focused on switch-level performance. Moreover, the first-available-algorithm (FAA) [9] is a low complexity algorithm to find an MBM.

We have built in NS2 [22] an optical switch module that uses the FAA scheduler. The simulation code models both network and physical layer aspects. We have cascaded a set of such switches and have surprisingly observed that

i) *the scheduler excessively deteriorates the optical signal quality by performing useless wavelength conversions ;*
ii) *switch-after-switch the useless wavelength conversions tends to aggregate optical packets on lower WDM channels, with a consequent increase in the occurrence of wavelength contention.*

We repeat the simulations substituting the FAA algorithm with the seminal tool for finding MBM: the Glover algorithm (1967). We observe exactly the same problems. We argue that the observed issues are not related to the specific algorithm, but they derive from the fact that MBM formalization has not been meant to limit wavelength detuning[6].

Motivated by this problem, we propose to improve the formalization of the scheduling problem presented in [8] [9] by weighting the graph edges. Upon the weighted graph, the optimal scheduler choice is *a Maximum Bipartite Matching with Minimum Weight* (MW-MBM). In doing so, we avoid wavelength conversion when it is not useful to solve wavelength contention. When wavelength contention occurs, we limit the wavelength detuning.

We provide the effectiveness of our novel MW-MBM formulation by devising a scheduling algorithm based on the seminal *Hungarian method* (1955). We are aware that Hungarian method has a high complexity; nevertheless, we do not care of that issue in the paper.

The paper is organized as follows; in section I we revise the MBM formalization. In section II we discuss why such a formalization has the two lacks previously stated. In section III we describe our novel MW-MBM formalization and in section III.A we devise a possible scheduling algorithm. Finally in section IV we carry out the performance evaluation by means of NS2 (v2.31) simulator. For the physical layer aspects, we have considered wavelength converters based on Four Wave Mixing (FWM) in Semiconductor Optical Amplifiers (SOAs) [5]. All the simulation code and running scripts are open-source available at Reference [23], thus our results can be reproduced by the other researchers and our coding efforts can be exploited by the research community.

I. THE MBM FORMALIZATION

In this section we revise the MBM formalization of the scheduling problem developed in [8] [9]. All parameters that we defined are summarized in Tab. 1.

TAB. 1 - DEFINITIONS

Parameter	***Definition***
N	Symbolic variable indicating the number of packets routed on an output fiber in a generic time-slot
M	Number of wavelengths of the WDM system
d	Available conversion range; i.e., maximum up or down detuning available through wavelength conversion, measured in WDM channels
p_k	k-th packets routed on an output fiber in a time-slot, $0 \le k < N-1$
λ_h	Wavelength of the h-th channel, $0 \le h < M-1$
i_k	The index of the wavelength of the k-th packet when it enters the switch
$begin(p_k)$	Lower index of the output wavelength where the packet p_k can be relayed
$end(p_k)$	Greater index of the output wavelength where the packet p_k can be relayed
$w(k,h)$	Weight of forwarding p_k on λ_h
Δ	Weight of an unfeasible forwarding

We consider a WDM network with M channels, allocated on M wavelengths $\{\lambda_0, \lambda_1, ..., \lambda_{M-1}\}$. Let us focus our attention on an output fiber of a switch. At the start of a time-slot there is a set of N packets $\{p_0, p_1, ..., p_{N-1}\}$ directed to the output fiber. The generic k-th packet arrives from a wavelength with index i_k and it can be detuned of at most d adjoining (lower or upper) WDM channels; i.e., its output wavelength index belongs to the range $begin(p_k) \div end(p_k)$, defined in Eqs. (1). We refer the parameter d as *available conversion range*.

$$begin(p_k) = \max(0, i_k - d) \qquad end(p_k) = \min(M-1, i_k + d) \qquad (1)$$

Let us sort the entering packets in ascending order versus their input wavelength index. The sorted sequence of packets forms the left vertices of a bipartite graph, considering one vertex for packet. The right vertices are the sequence of output wavelengths $\{\lambda_0, \lambda_1, ..., \lambda_{M-1}\}$, considering one vertex for wavelength. Graph edges are feasible forwardings of input packets on output wavelengths. If the k-th packet can be forwarded on the wavelength λ_h (i.e. $begin(p_k) \le h \le end(p_k)$)

[2] We observe that in [8] the Authors consider a weighted graphs for QoS purpose, but weights are placed on the vertices and not on the edges.

then there is an edge between the k-th packet and λ_h. The resulting graph is named *conversion-graph* and has the property to be *convex* [3].

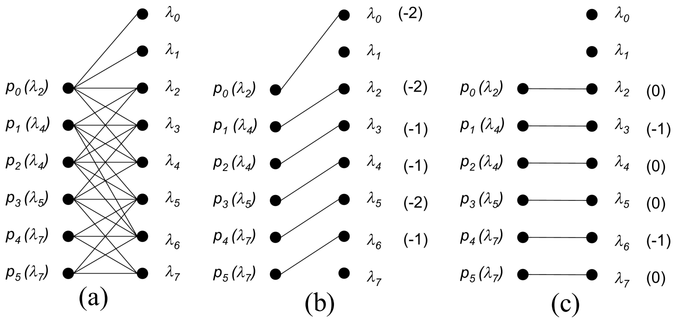

Fig. 1 – Conversion-graph for an 8 wavelengths WDM system with available conversion range equals to 2 channels (a), the Glover/FAA maximum bipartite matching (b), empirical (HSA) maximum bipartite matching (c)

Fig. 1a is the conversion graph that we would have in case of a WDM system with 8 wavelengths, available conversion range equals to 2 channels and six entering packets. The label of a left vertex is $p_k(\lambda i_k)$.

The conversion-graph contains all the possibilities that the scheduler has for mapping input packets on output wavelengths. A feasible mapping is a sub-graph for which there is at most one edge for vertex; indeed, an output wavelength can transfer only a packet, and a packet can only be transferred on an output wavelength. Such a sub-graph is called "matching" and it is a possible scheduler choice. The number of edges of a matching is equal to the number of packets that are forwarded by the switch in a time-slot; thus, the throughput maximization consists in finding a matching that has the greater number of edges; i.e. a Maximum Bipartite Matching (MBM). We highlight that, in general, there exists more than one maximum matching. As instance, Fig. 1b and Fig. 1c reports two possible maximum matchings of the graph in Fig. 1a [4].

In the 1967, Glover [15] found out that the research of a maximum bipartite matching in a convex graph can be solved with a greedy algorithm; i.e. an algorithm that locally performs optimal choice at each stage. The Glover algorithm processes iteratively the right vertices $\{\lambda_0, \lambda_1, ..., \lambda_{M-1}\}$. At the j-th iteration, among the remaining left-vertices (i.e., packets) that have an edge with λ_j, the algorithm selects the k-th left vertex (i.e., p_k) with the minimum value of $end(p_k)$. We observe that among the possible candidate left-vertices, the Glover algorithm performs the most "precautionary" choice, since it selects the left-vertex that has the smallest flexibility in being matched in the next steps[5].

The Glover algorithm has a computational complexity in the order of the number of edges, however literature provides lower complex algorithms [16][17] that return the same matching. With respect to the specific optical scheduling issue, a relevant improvement in terms of complexity has been carried out by the Authors of reference [8]. They have argued that the convex bipartite graph faced in the WDM scheduling matter has the additional property to be *ordered* [6]. This property strongly simplifies the complexity of finding the maximum bipartite matching; consequently, they [8] have proposed an efficient algorithm named First Available Algorithm (FAA), whose output matching is however equal to the Glover one.

II. Understanding the issues of the MBM formalization

In this section, we explain why the MBM formalization of the scheduling problem leads to the couple of issues mentioned in the Introduction. For ease of illustration, we assume to find a maximum bipartite matching through the classical Glover algorithm. Anyway, at the end of section we extend the conceptual findings to any solver algorithm; i.e. to the whole MBM formalization.

Let us consider a simple situation for which there is only one input packets p_1 to be forwarded (e.g., Fig. 2a). The packet enters the switch on wavelength λ_4. The available conversion range d is 2 WDM channels. It is straightforward to understand that the best scheduling to do is to relay p_1 without any wavelength conversion, thus avoiding the signal degradation introduced by wavelength converter. This said, it is interesting to observe what would have performed a scheduler based on the Glover algorithm: at the first iteration step the algorithm selects the first output wavelength, λ_0, and tries to fill that channel with a packet that can be relayed on λ_0 (and that has the minimum $end(p_k)$ value). This packet does not exist, the iteration loops, and the same occurs until output wavelength λ_2 is reached. Then, the algorithm tries to fill the channel with a packet that can be relayed on λ_2. This packet is just p_1. Thus, the scheduler based on the Glover algorithm has useless detuned p_1 from λ_4 to λ_2; i.e., the packet experiences a input-output detuning of 2 channels. Moreover, it is instructive to observe that if the available conversion range d had been equal to 4 channels, then packet p_1 would be relayed on λ_0; i.e., the packet would experience a input-output detuning of 4 channels. Finally, we point out that same scheduling outputs would be obtained by using FAA [8] instead of Glover.

Generalizing the previous example, we argue that *the usage of Glover algorithm, or of any own optimization (e.g. FAA), tends to shift packets toward lower wavelengths, even when such a shift is not necessary for solving wavelength contention. Moreover, the amount of input-output detuning increases with the increase in the available conversion range.*

This behavior leads to a useless degradation of the optical signal because of unnecessary wavelength conversions. In addition, even if the wavelength conversion was noiseless, we

[3] That is, for a given left vertex the related subset of right vertices is continuous.

[4] On the right vertices of Fig. 1b and Fig. 1c we also report the amount of detuning that packets underwent, measured in terms of WDM channels. E.g., the value -2 means that the packet has been converted from λ_k to λ_{k-2}.

[5] In case of more left vertices with the same *end* value, the choice among them can be arbitrary.

[6] Since the available detuning d is equal for each packet, the resulting graph has the property of being *ordered*; i.e., given two left edges p_i, p_k with $i<k$, then begin(p_i)≤ begin(p_k) and end(p_i)≤end(p_k).

point out a critical "networking" problem: switch-after-switch network packets tend to be aggregated on lower wavelengths and this aggregation increases the occurrence of wavelength contention.

Fig. 1b reports the maximum bipartite matching singled out by the Glover (or FAA) algorithm on the graph of Fig. 1a. We observe the tendency of Glover algorithm in shifting packets toward lower wavelengths; indeed, all packets are downshifted. On the contrary, the Fig. 1c represents another possible maximum bipartite matching that we have empirically derived attempting to minimize the wavelength detuning. We achieve a mapping with only 2 wavelength conversion, just performed to resolve wavelength contention among (p_1, p_2) and (p_4, p_5). The empirical matching outperforms the Glover one both in terms of signal quality and in terms of spreading of wavelengths usage. This empirical approach has taught us that *an optimal scheduling algorithm should be able to differentiate the available choices also in terms of involved detuning.* Unfortunately, MBM formalization prevents this modus-operandi, since the "cost" of detuning a packet is not taken into account.

III. A NOVEL FORMALIZATION OF THE SCHEDULING PROBLEM: MINIMUM WEIGHT - MAXIMUM BIPARTITE MATCHING

To overcome the issues of the MBM formalization we propose that *the primary goal of throughput maximization should be combined with a secondary goal that is the minimization of the wavelength detuning*. The secondary goal limits signal quality degradation and avoids unnecessary wavelength conversion.

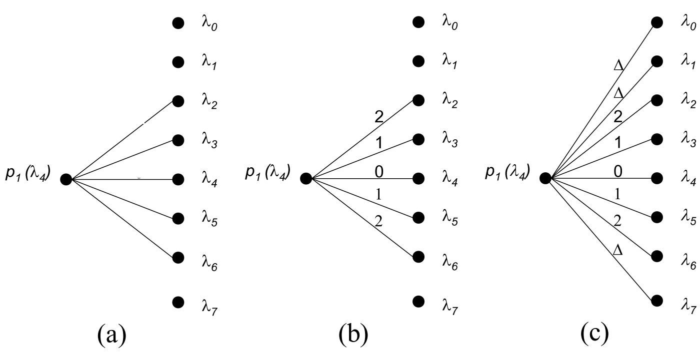

Fig. 2 – Conversion-graph for an 8 wavelengths WDM system with available conversion range equals to 2 WDM channels in case of a single incoming packet p_1 on wavelength λ_4 (a), associated conversion-weighted-graph (b) and complete-conversion-weighted-graph.

To cope with this combined optimization, we revise the conversion-graph described in section I by adding the weight $w(k,h)$ to the edge (p_k, λ_h). The value of $w(k, h)$ is defined by Eq.(2) and it is linearly proportional to the amount of input-output detuning measured in terms of WDM channels [7]. We named this new graph as *conversion-weighted-graph*.

[7] We observe that this simple weighting should be further tuned, in agreement with the physical behavior of the optical devices. As instance, if down-conversions yielded a lower signal impairment than up-conversion, then the down-conversion should be weighted lower than up-conversions.

$$w(k,h) = abs(i_k - h) \tag{2}$$

As instance, Fig. 2b reports the conversion-weighted-graph that extend the conversion-graph of Fig. 2a.

Upon the conversion-weighted-graph, we formalize the scheduling problem as *the finding of a maximum bipartite matching, whose sum of edges weights is minimum* (*MW-MBM*). This means that among all the possible maximum matchings (i.e., scheduler choices that maximize the switch throughput), we single out one of them providing the minimum "overall" detuning.

A. An MW-MBM scheduler based on Hungarian Algorithm

In this sub-section, we derive a scheduling algorithm that finds a MW-MBM of the conversion-bipartite-graph. The scheduler design is meant at providing the performance improvement that the MW-MBM formalization would give versus the MBM one. Thus, we do not tackle computational complexity issue, and defer the design of a low complex scheduler to future works.

We face the research of a MW-MBM of the conversion-weighted-graph by exploiting the Hungarian method (1955) [19][20]. Given a "complete" bipartite weighted graph where each edge has a weight $w(k,h)$, the Hungarian algorithm returns a maximum matching with the minimum cost, i.e. with the minimum sum of edge weights.

A graph is complete when there is an edge between any left and right vertex. The conversion-weighted-graph that we handle is not "complete"; it does not contains edges related to not feasible wavelength conversions (e.g., in Fig. 2b the edge (p_1,λ_7) does not exist). Therefore, we need to complete carefully the conversion-weighted-graph to exploit Hungarian method.

We define *complete-conversion-weighted-graph* the graph formed by the same vertices and edges of the conversion-weighted-graph, plus an additional set of "unfeasible" edges, weighted by Δ, that make complete the graph. Summarizing, the weight $w(k,h)$ of the edge (p_k , λ_h) of the complete-conversion-weighted-graph is defined in Eq. (3).

$$w(k,h) = \begin{cases} abs(i_k - h) & \text{if } begin(p_k) \le h \le end(p_k) \\ \Delta & \text{otherwise} \end{cases} \tag{3}$$

Fig. 2c is the complete-conversion-weighted-graph related to the conversion-weighted-graph of Fig. 2b. We point out that the additional edges (weighted by Δ) are related to unfeasible wavelength conversions, since the associated detuning is greater than d.

Let we use the complete-conversion-weighted-graph as the input of the Hungarian algorithm, thus obtaining an MW-MBM of such a complete graph. We name this MW-MBM as A. The matching A is a formed by E=min(N,M) edges, where F are feasible edges and the remaining U are unfeasible edges (i.e. F+U=E).

We point out that the matching A is not exactly the scheduler goal. Indeed, the scheduler has to find a MW-MBM of the (incomplete) conversion-weighted-graph. Nevertheless, in the next we show that:

1. a MW-MBM of the conversion-weighted-graph can be derived by a MW-MBM (A) of the complete-conversion-weighted-graph by simply removing from A the U unfeasible edges, providing that A contains the maximum number of feasible edges F;
2. by means of the Hungarian algorithm, a proper dimension of Δ makes possible to derive a MW-MBM (A) with the maximum number of feasible edge F.

Lemma 1. *If a MW-MBM (A) of the complete-conversion-weighted-graph contains the maximum number F of feasible edges, then the matching B derived by A removing the U unfeasible edges is a "maximum" matching of the conversion-weighted-graph.*

Proof. We face the proof by contradiction. Let us assume that A contains the maximum number F of feasible edges and B is *not* a maximum matching of the conversion-weighted-graph. Let be B' a maximum matching of the conversion-weighted-graph that contains F' edges, with $F' > F$. Let us build the maximum matching A' of the complete-conversion-weighted-graph formed by the edges of B' plus an additional set of E-F' unfeasible edges. In doing so, A' contains more feasible edges than A, but this contradicts the assumption of the proof. Therefore, it is not possible that B was not a maximum matching of the conversion-weighted-graph

Theorem 1. *If a MW-MBM (A) of the complete-conversion-weighted-graph contains the maximum number F of feasible edges, then the matching B derived by A removing the U unfeasible edges is also a MW-MBM of the conversion-weighted-graph.*

Proof. Given Lemma 1, it remains to demonstrate that the maximum matching B it is also the one with the minimum sum of edge weights. We proceed by contradiction. Let us assume that A be a MW-MBM of the complete-conversion-weighted-graph that contains the maximum number F of feasible edges, B is a maximum (lemma 1) matching of the conversion-weighted-graph, but B is *not* a minimum weight matching. Therefore, there is a maximum matching B' of the conversion-weighted-graph, which sum of edge weights is lower than the one of B. We point out that the number of edge of B and B' is equal to F, since they are both maximum matchings. Let us build the maximum matching A' of the complete-conversion-weighted-graph through the edges of B' plus an additional set of E-F unfeasible edges. In doing so, A' is ever a maximum matching with the maximum number of F of feasible edges but contains cheaper edges than A. This contradicts the assumption of the proof, thus proofing by contradiction that B is also a minimum weight matching.

Following, Theorem 1 we strived to dimension Δ so that the Hungarian algorithm applied on the complete-conversion-weighted-graph returns a MW-MBM (A) with the maximum number of feasible edges.

To achieve this goal, it is enough that any maximum matching of the complete-conversion-weighted-graph formed by x+1 feasible edges is ever less costly than a maximum matching formed by only x feasible edges.

The greatest cost (GC) of a maximum matching with $x+1$ feasible edges and the cheapest cost (CC) of a maximum matching with x feasible edges can be written as

$$GC(x+1) = (x+1)d + (E-(x+1))\Delta \ , \quad CC(x) = (E-x)\Delta \tag{4}$$

In the GC computation, we considered that the $x+1$ feasible edges had the maximum weight d, while in the CC computation the x feasible edges had the minimum weight, i.e. 0. In both cases, the remaining unfeasible edges have a weight Δ.. Therefore, the parameter Δ should guaranteed the following inequality,

$$GC(x+1) < CC(x) \quad \text{for } 0 \le x \le E-1 \tag{5}$$

A sufficient condition guaranteeing previous inequality is $\Delta > d$ E. Considering that d and E are at most equals to M, a value of Δ independent of d that satisfies the inequality in (5) is $\Delta = M^2 + 1$.

TAB. 2 - HUNGARIAN-BASED SCHEDULING ALGORITHM (HSA)

```
# Build of the complete-conversion-weighted-graph matrix (CCWG)

for k=0 to N-1 do
  for h=0 to M-1 do
    CCWG[k][h]=w(k,h);
  end for
end for

# Computation of MW-MBM of the complete-conversion-weighted-graph through Hungarian method

WMATCH[] = Hungarian.solver(CCWG)

# Removal of unfeasible edges to obtain MW-MBM of the conversion-weighted-graph

for h=0 to M-1 do
  k= WMATCH[h];
  if ((h<begin(p_k) or h> end(p_k)) and k!=void)
    WMATCH[h]=void;
  end if
end for
```

Tab. 2 summarizes the scheduling algorithm that we devised. The algorithm is based on the Hungarian method, thus we named it *Hungarian-based Scheduling Algorithm* (HSA). The first operation accomplished by HSA is the building of the $CCWG$ matrix, which represents the complete-conversion-weighted-graph aforementioned. The matrix has a row for each left vertex (i.e. packet) and a column for each right vertex (i.e. wavelength). The value of $CCWG$ [k][h] is the weight of the edge (p_k , λ_h) as defined in Eq. (3). The next step of the HSA algorithm is the derivation of the minimum weight - maximum bipartite matching by means of a Hungarian.solver, i.e. a routine that implements the Hungarian algorithm (e.g., [21]). The output of the Hungarian.solver is the vector $WMATCH$[] of size M that contains for each output wavelength the index of the packet selected to be relayed on;

i.e. $WMATCH[h]=k$ means the packet p_k has been selected for forwarding on the wavelength λ_h. If none packet has been selected for λ_h , then $WMATCH[h]=void$.

The last step the HSA algorithm removes from the *WMATCH*[] vector the unfeasible edges, thus obtaining a MW-MBM of the conversion-weighted-graph.

We conclude the section by observing that the empirical maximum matching reported in Fig. 1c is the same that would be obtained by using HSA.

IV. PERFORMANCE EVALUATION

This section reports the performance comparison between a scheduler based on the MBM formalization and a scheduler based on the MW-MBM formalization. We will discuss the simulation tool, the simulation settings and the numerical results.

A. *The simulation tool*

We have carried out performance evaluation by means of NS2 [22]. We have extended the Tcl and C++ code of NS2 in order to simulate network and physical layer aspects of an optical synchronous WDM packet network (simulator code available at [23]).

Regarding networking functionalities, we have implemented FAA [8], Glover and HSA algorithms. The FAA algorithm has been selected to realize the MBM based scheduler[8]. The HSA algorithm has been selected to realize the MW-MBM based scheduler.

Regarding the physical layer modeling, we have extended the NS2 *simplex-link* class in order to simulate the parallel convey of packets on different WDM channels. Moreover, we have extended the NS2 packet header (i.e., the common-header) in order to include the wavelength number, the signal and noise power of the packet. These parameters are updated switch-after-switch according to the scheduler algorithm and to the physical models hereafter detailed.

1) *Physical model of the optical cross connect*

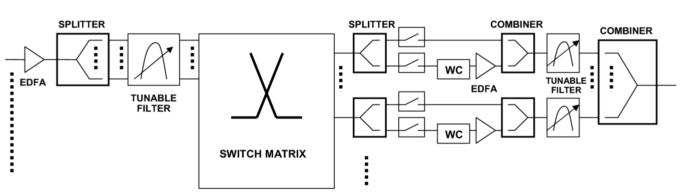

Fig. 3 – Optical switch architecture

Fig. 3 reports the data-plane of the optical cross connect architecture that we have modeled in the simulator [4]. For ease of illustration, we have not reported control-plane modules (i.e., packet header O/E/O conversions, routing agent, scheduler module, control lines, etc.) and we have drawn only one input and output fiber.

At the switch ingress, a set of Erbium Doped Fiber Amplifiers (EDFA) compensates for fiber attenuation and OXC power losses. We name such EDFAs as "line-EDFAs". After the line-EDFA, splitters and filters separate the M WDM channels. A switch matrix accomplishes the packet forwarding by temporary switching an input channel to an output channel, based on the scheduling algorithm decision. At the exit of the switching matrix, we have two branches. A branch is used when wavelength conversion is needed; otherwise, a no-conversion branch is available. The scheduler controls the used branch by triggering optical gates. Moreover, within the wavelength converter branch we use an additional EDFA (named "WC-EDFA") to averagely compensate conversion power loss. Subsequently, tunable filters select output signal and suppress optical pump and original signal if wavelength conversion has been occurred. Finally, combining function is performed to collect all the channels on a single output fiber.

Hereafter we describe the relevant formulas that we have used to model signal losses and noise contributions. We assume that the major sources of noise are EDFAs and wavelength converters.

The EDFA noise has been evaluated as (see Tab. 3 for parameters description):

$$N_{EDFA} = 2n_{sp}hc(G-1)Bw/\lambda_s \tag{6}$$

We have considered a wavelength conversion process based on Four Wave Mixing (FWM) effect in Semiconductor Optical Amplifier (SOA). This nonlinear phenomenon takes place through the injection of two optical signal at different wavelengths into the SOA and gives rise to a third signal on a certain wavelength, such as the three wavelengths satisfy a phase matching condition.

We assume that the conversion frequency response is almost constant in the signal bandwidth; hence, the conversion process is transparent to modulation formats and bit-rate.

Wavelength conversion introduces a loss of power on the signal. This degradation could be measured by the conversion-efficiency $\eta(\Omega)$; i.e. the power ratio between converted and input signal. The parameter Ω is the frequency-detuning, measured as the difference between the frequency of the converted signal (ω_c) and the one of the input signal (ω_s). We have considered that two adjoining WDM channels differ of 0.8 nm.

In addition to the power loss, wavelength conversion introduces noise due to Amplified Spontaneous Emission (ASE). We have evaluated the conversion-efficiency and the ASE spectrum S_n through the modelling results of D'Ottavi et al. [6]. For the sake of completeness, we report hereafter the relevant formulas (see Tab. 3 for parameters description). Moreover, the SOA linear gain G is evaluated as analytical solution of Eq. (9).

Fig. 4 reports the efficiency and the noise power introduced by a wavelength conversion. The detuning increasing leads to a suddenly degradation of the conversion efficiency, while noise power only decreases of some decimals points. The combination of these two effects involves that the greater the detuning, the more the degradation between the input and the output value of OSNR.

[8] We also test Glover algorithm, nevertheless obtaining the same results of FAA. Thus, the choice between of Glover and FAA was arbitrary; we have chosen FAA since it is more cited in the optical scheduling literature.

$$\eta(\Omega)=\frac{1}{4}\left[\frac{P_p}{P_s+P_p}\right]^2 P_\sigma^2 G\left[\ln\left(\frac{G_0}{G}\right)\right]^2 \cdot \left|\frac{(1-i\alpha)}{P_\sigma}H_c+\varepsilon'_{hb}H_{hb}+\varepsilon'_{ch}(1-i\beta)H_{ch}+ + \frac{\Gamma g_0}{\gamma}\left[\frac{P_s+P_p}{P_\sigma}\frac{(1-G)}{\ln(G_0/G)}+1\right]\varepsilon^0_{ch}(1-i\beta)H_{ch}\right|^2 \quad (7)$$

$$H_c=1/(1+j\Omega\tau_s);\ H_{hb}=1/(1+j\Omega\tau_e);\ H_{ch}=1/(1+j\Omega\tau_e)(1+j\Omega\tau_p)$$

$$S_n(\Omega)=\left\{\hbar\omega_c\frac{\Gamma g(0)}{\Gamma g(0)-\gamma}\left[F(G-1)\ +\frac{P_s+P_p}{P_\sigma}(F-1)G\ln(G)\right]\right\}\cdot \left[1+\frac{\omega_s}{\omega_c}\frac{\eta(\Omega)}{G}+4\frac{\omega_p}{\omega_c}\frac{\eta(\Omega)}{G}\frac{P_s}{P_p}\right] \quad (8)$$

$$\frac{P_s+P_p}{P_\sigma}=\frac{\Gamma g_0-\gamma}{\gamma}\left[\frac{1-(G/G_0)^{\gamma/\Gamma g_0}}{G-(G/G_0)^{\gamma/\Gamma g_0}}\right] \quad (9)$$

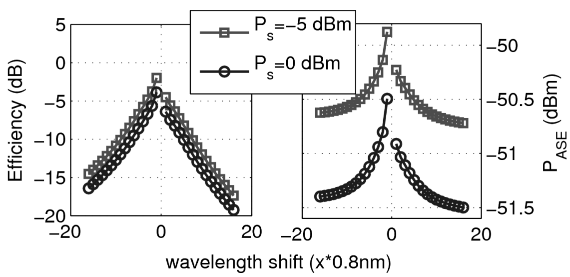

Fig. 4 – Conversion efficiency and noise power versus detuning measured in WDM channels of 0.8 nm in case of two values of input signal power (P_s)

TAB. 3 - PHYSICAL PARAMETERS OF OXC

SOA parameters		
Parameter	**Definition**	**value**
λ_s	Signal wavelength	1550 nm
P_σ	Saturation Power	10 dBm
P_p	Pump Power	5 dBm
g_0	Unsaturated local gain	140 cm^{-1}
G_0	Unsaturated device gain	24 dB
γ	Scattering loss	10 cm^{-1}
Γ	Confinement factor	0.2
τ_s	Carrier lifetime	200 ps
τ_p	Carrier-phonon scattering time	650 fs
τ_e	Carrier-carrier scattering time	100 fs
α	Linear linewidth enhancement factor	6
β	Non-linear linewidth enhancement factor	2
ε^0_{ch} ε'_{ch}	CH non-linear coefficients	0.5 W^{-1} 5 W^{-1}
ε'_{hb}	SHB non-linear coefficient	1 W^{-1}
F	Inversion factor	2

OXC, Fiber, EDFA parameters		
L_{OXC}	OXC overall loss	15 dB
L_f	Fiber loss (10 km)	10 dB
G	WC-EDFA gain	5 dB
G	Line-EDFA gain	25 dB
n_{sp}	Spontaneous emission rate	1
Bw	Filters bandwidth	5 GHz

Finally, Tab. 4 shows the algorithm that we have implemented in the simulator to update the signal power P_s and the noise power P_n of a packet that traverses a switch; Tab. 3 reports all physical parameters and respective values that we have used.

TAB. 4 - SIGNAL AND NOISE POWER UPDATING ALGORITHM

P_n=P_n+ $N_{Line\text{-}EDFA}$;
if (wavelength conversion)
  $P_n = P_n * \eta(\Omega) + S_n(\Omega) * Bw$;
  $P_s = P_s * \eta(\Omega)$;
  $P_n = P_n * G_{WC\text{-}EDFA} + N_{WC\text{-}EDFA}$;
  $P_s = P_s * G_{WC\text{-}EDFA}$;
end

B. *Simulation settings*

The simulation time is suitable regulated to transfer about 10^7 packets.

We have considered two network topologies. A first network topology is the single-switch topology, that we have used to analyze MBM and MW-MBM switch-level performances. The topology is formed by 8 traffic sources S_i attached to an 8x1 switch. The 8x1 switch is connected to a common data receiver R through an optical fiber supporting 32 WDM channels.

The second network topology is the "aggregating-chain" topology (Fig. 5) that we have used to analyze MBM and MW-MBM network-level performances. We have suitably designed this topology to highlight the MBM network issues. The topology is formed by a sequence of "switching-stages". Each switching-stage is formed by a set of 2x1 switches that aggregate traffic coming from the previous stage. A number of K sources are attached to the first stage. The number Z of stages is Z=$ceil(log_2K)$. All optical fiber support 32 WDM channels.

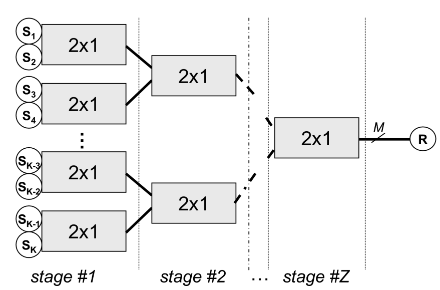

Fig. 5 – Aggregating-chain topology in case of K sources

Any source generates packets of fixed length, equal to the duration of a time slot (e.g., 20 ms). The inter-arrival time between consecutive packets approximately follows a negative exponential distribution (i.e. Poisson traffic); time samples are computed by a negative exponential random number generator and are suitably increased to align packets with time-slot boundary. When generated, each packet has an initial signal power of 1mW and an initial OSNR of 30 dB. The wavelength index of a generated packet is a uniform random variable in the set 0,1,…M-1. To concisely represent the amount of traffic fed in the network by all the sources, we define the *traffic load* as the ratio between the i) average number of packets offered by all the sources in a time-slot and ii) the number of wavelength *M*.

C. Numerical results

1) Single-switch analysis

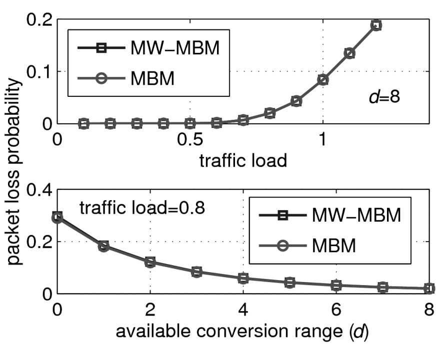

Fig. 6 – Packet loss probability in case of single-switch topology and WDM system with 32 wavelengths (*M*=32) versus traffic load and available conversion range (*d*) for MBM and MW-MBM based schedulers

This section reports the result obtained in case of single-switch topology. The upper plot of Fig. 6 shows the packet loss probability versus traffic load with an available conversion range (*d*) equal to 8 channels. The lower plot of Fig. 6 shows the packet loss probability versus *d* with a traffic load equal to 0.8. We observe that MBM and MW-MBM schedulers provide the same packet loss probability; i.e., the same switch-level throughput. In both cases the traffic offered to the switch is the same and both formalizations are meant to maximize switch throughput, therefore the schedulers achieve the same loss performance.

The numerical behaviors have straightforward explanations: i) by increasing the traffic load, the number of wavelength contentions that can not be resolved increases as well, with a consequence increase in packet loss probability; ii) by increasing the available conversion range, the scheduler flexibility in resolving wavelength contention increases as well, with a consequent decrease in packet loss probability.

Up to now, we have shown that *MBM and MW-MBM offer the same switch-level throughput performance, providing that offered traffic is the same*. In the next, we show that they profoundly differ in terms of detuning performance. To highlight better those differences we set the traffic load to 0.1 and vary the available conversion range.

Fig. 7 reports the average detuning and the conversion probability suffered by forwarded packets. The average detuning is measured in terms of WDM channels; if a packet has input wavelength λ_i and it is forwarded on λ_h, the measured detuning is then *abs(i-h)*. Although throughput performances are the same, we observe that in case of MBM about the 95% of forwarded packets suffer for a wavelength conversion; whereas in case of MW-MBM this percentage drastically falls down to 5%. This is an evidence that *MBM scheduler performs wavelength conversions independently of their effectiveness in resolving wavelength contentions*.

We also observe that in case of MBM the average detuning increases at the increase in available conversion range *d*, whereas this does not occur for MW-MBM. Therefore, *even the "amount" of detuning carried out by an MBM scheduler is uncorrelated with wavelength contention phenomena*.

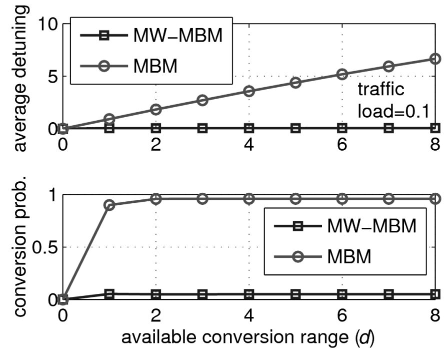

Fig. 7 – Average detuning measured in WDM channels and conversion probability in case of single-switch topology, WDM system with 32 wavelengths (*M*=32), traffic load = 0.1 versus available conversion range (*d*) for MBM and MW-MBM based schedulers.

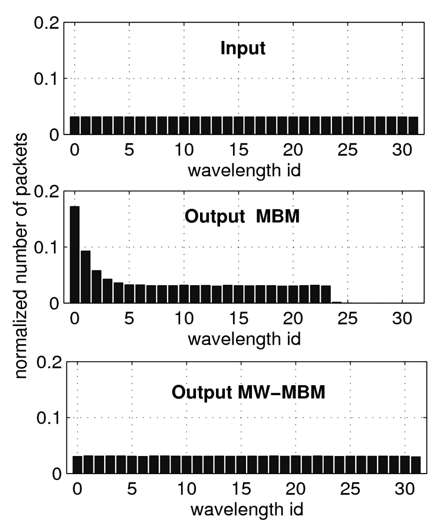

Fig. 8 – Input/output wavelengths usage and wavelength conversions occurrence in case of single-switch topology, WDM system with 32 wavelengths (*M*=32), available conversion range equals to 8 channels (*d=8*), traffic load = 0.1, for MBM and MW-MBM based schedulers.

Fig. 8 reports a measure of the occupancy of the input and output wavelengths of the switch in case of available conversion range equals to 8 channels. The upper subplot is the histogram of the occupancy of input wavelengths. The height of the *x*-th bar is the ratio between i) the number of packets that have entered the switch on the wavelength λ_x and ii) the overall number of packets that have entered the switch. The source model is devised to spread uniformly generated packets on WDM channels, thus we have a uniform usage of input wavelengths. The middle subplot is the histogram of the occupancy of the output wavelengths in case of MBM scheduler. The height of the *x*-th bar is the ratio between i) the number of packets that are forwarded on the wavelength λ_x

and ii) the overall number of forwarded packets. We observe that wavelengths with lower indexes are more used than others; i.e., the scheduler tends to shift packets on lower wavelengths. The lower plot is the histogram of the occupancy of the output wavelengths in case of MW-MBM scheduler. As expected, since only 5% of packets are converted and the related detuning is very limited (see Fig. 7), output packets then remain spread over the full set of WDM channels.

2) Network analysis

In this section we discuss the results obtained in case of aggregating-chain topology (see Fig. 5). We point that the single-switch analysis has highlighted that the distribution of the outgoing traffic upon the available wavelengths differs between MW-MBM and MBM (Fig. 8). This involves that in an aggregating-chain topology, apart from first stage switches that are directly connected with the sources, the distribution of the traffic offered to the remaining switches of the path is different between MW-MBM and MBM cases, for this reason we will expect different loss performance as well.

As "default" configuration we consider a available conversion range d equals to 6 channels, a topology with 5 switching-stages (i.e., 32 sources) and a traffic load equals to 0.8. In the next subsections, we analyze performances varying the aforementioned configuration parameters one at a time, while the not-varying ones are configured at the default value.

a) Analysis versus available conversion range

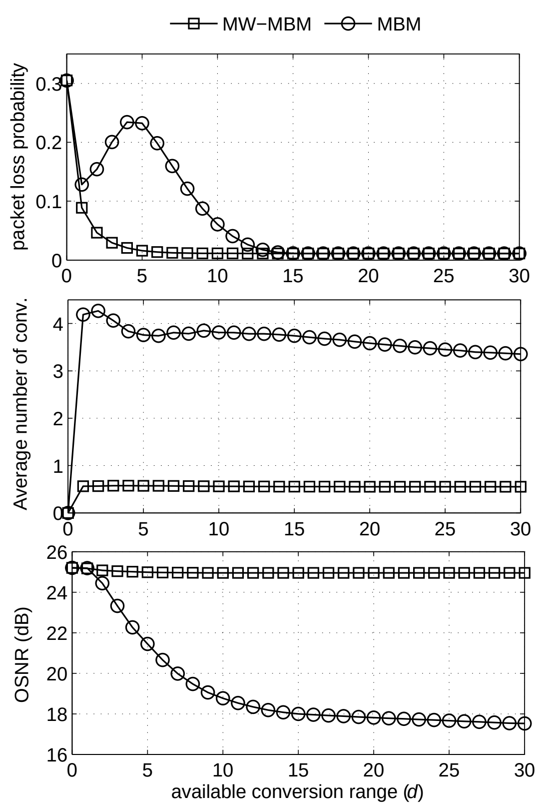

Fig. 9 – End-to-end packet loss probability (a), average number of conversion per packet (b) and end-to-end average OSNR per packet (c) versus available conversion range d, in case of aggregating-chain topology, WDM system with M= 32 wavelengths, 5 switching-stages, traffic load equal to 0.8

In Fig. 9a we plot the end-to-end packet loss probability. In the extreme cases both approaches offer the same performances, but the MW-MBM scheduler outperforms the MBM one up to d=15.

Looking at the performance provided by MBM, we observe a surprising result: the increase in d between 2 and 4 leads to an "increase" in the packet loss probability. As it occurs in the single-switch case, we would expect that the wider is the conversion range d, the lower is the packet loss probability, since the more are the scheduler possibilities in solving wavelength contention. Nevertheless, this actual *pro* is counterbalanced by the following *con*: the greater is the value of d, the higher is degree of packets aggregation on lower wavelengths[9]. Hence, a greater number of packets enter (on different fibers) the "next" switch on the same wavelength and more wavelength contentions and packet losses occur.

As instance, Fig. 10 reports the usage of the output wavelengths of a switch of the 4-th switching-stage in case of d=2 and d=4. In case of d=4 there is a stronger aggregation of packets on lower wavelengths, more wavelength contentions therefore occur in the switch of the 5-th switching-stage.

MW-MBM scheduler does not suffer the aforementioned *con*, since it uniformly spreads packets on WDM channels. Therefore, by increasing the available conversion range we have a monotone decreasing behavior of packet loss probability.

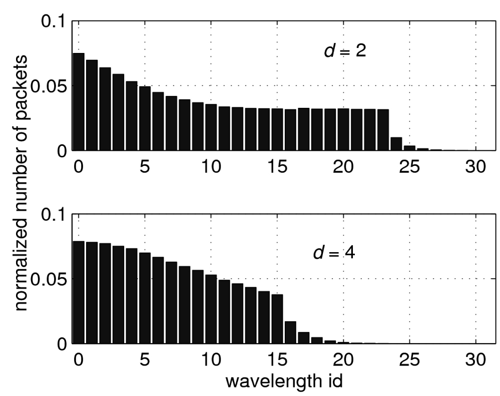

Fig. 10 – Output wavelengths usage of a switch of the 4-th switching-stage for two values of available conversion range d in case of aggregating-chain topology, WDM system with M=32 wavelengths, traffic load = 0.8 and MBM based scheduler

Fig. 9b reports the average number of wavelength conversions suffered by a packet from source to destination. We observe that the MW-MBM approach strongly reduces the occurrence of wavelength conversions with respect to MBM. In both cases, the average number of wavelength conversions is quite independent of d, nevertheless the motivations of this equivalent behavior are different.

In case of MBM, each switch tends to convert packet

[9] This con can be explained as follow: the increase of d leads to an increase of the average detuning (see Fig. 7) suffered by packets. Since most of wavelength conversions are down-conversions, an higher average detuning involves an higher degree of packets aggregation on lower wavelengths. Moreover, we point out that such a degree of aggregation further increase switch-after-switch.

toward lower wavelengths. Consequently, the number of wavelength conversions mainly depends on the number of switches traversed by a packet from source to destination (see next section IV.C.2.c).

In case of MW-MBM, the amount of wavelength conversions is mainly related to the amount of wavelength contentions; i.e. to the traffic load (see next section IV.C.2)b).

Fig. 9c reports the average OSNR of a packet measured at the destination. We remind that the OSNR inversely depends both on the number of wavelength conversions suffered by a packet (Fig. 9b) and on the amount of detuning associated to each conversion [6].

In case of MBM, the OSNR decreases at the increase in *d*. In fact, although the number of wavelength conversion (Fig. 9b) is quite constant, the wavelength detuning increases at the increase in *d* (Fig. 7).

In the MW-MBM case, the OSNR is practically constant; indeed, both the number of wavelength conversions and the detuning are practically constant.

b) Analysis versus traffic load

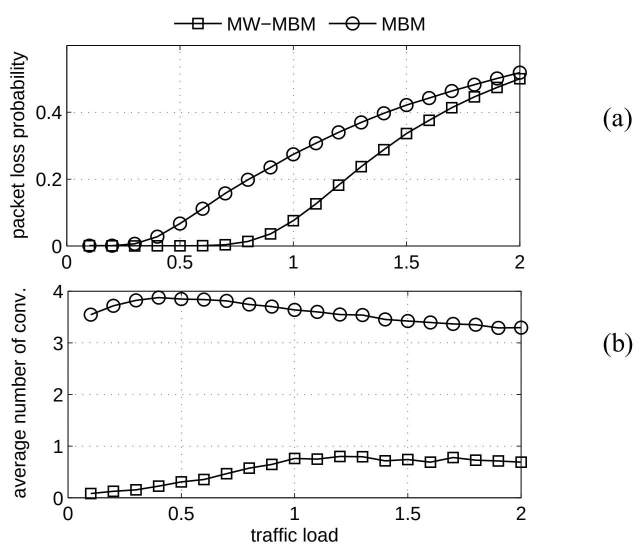

Fig. 11 – End-to-end packet loss probability (a) and average number of conversions per packet (b) versus traffic load, in case of aggregating-chain topology, WDM system with *M*=32 wavelengths, 5 switching-stages, available conversion range *d*=6

Fig. 11a reports the packet loss probability versus traffic load. As expected, in the extreme cases the two approaches offer the same loss performances. In the middle cases, MW-MBM outperforms MBM thanks to the wider spreading of packets on WDM channels, which reduces the occurrence of wavelength contentions in the network.

Fig. 11b reports the average number of wavelength conversions suffered by a packet. As previously stated commenting Fig. 9b, the number of conversions in case of MBM scheduler mainly depends on the number of switches traversed by a packet; thus the number of conversions is quite independent of traffic load.

For the MW-MBM case, we observe a peculiar behavior. By increasing the traffic load we have an increase in number of wavelength conversions, since the more are the wavelength contentions to resolve. Nevertheless, in case of overload (i.e., traffic load > 1) this growing behavior is counterbalanced by the tendency of the scheduler in forwarding more and more packets without wavelength conversions. Indeed, the increase in traffic load increases the set of input packets on which the scheduler can operate to maximize throughput and minimize detuning. Thus, the scheduled will probably find a maximum-throughput mapping formed by a greater number of packets that must not be converted.

Concerning the OSNR performances (not reported in any figure), they are practically independent on traffic load: the OSNR of MBM scheduler is about 21 dB and the one of MW-MBM is about 25 dB.

c) Analysis versus the number of switching-stages

In this section we analyze the network performance versus the number of switching-stages (*Z*), i.e. the number of switch traversed by a packet from source to destination. We control such a parameter by suitably increasing the number of sources (*K*), since $Z=ceil(log_2K)$. Although we vary the number of sources, the amount of traffic that feeds the network does not vary.

Fig. 12a reports the packet loss probability performance. We observe that MW-MBM considerably outperforms MBM at the increase in the number of traversed switches. Switch-after-switch MBM aggregates more and more packets on lower wavelengths. This tendency, combined with the increase in the number of switches to traverse, drastically increases the end-to-end occurrence of wavelength contention, thus the packet loss probability. The MW-MBM packet loss probability is quite independent of the number of traversed switches; indeed, it depends on available conversion range (see Fig. 9a) and traffic load (see Fig. 11a).

Fig. 12b reports the average number of wavelength conversions suffered by a packet. The plot shows that in case of MBM the number of wavelength conversions linearly depend to the number of switches encountered by a packet from source to destination, since each switch tends to shift the packets toward lower wavelengths.

In case of MW-MBM the number of conversions mainly depend on traffic load and does not depend on number of traversed switch.

Fig. 12c reports the OSNR performance. Obviously, in both cases we have a decreasing behavior. Indeed, the more the traversed switches, the more the in-line EDFA encountered, thus the ASE noise.

We observe that in case of MBM the OSNR performance more rapidly falls down. The motivation of the different slope is the following: in addition to EDFA ASE noise, the MBM scheduler inserts more and more noise due to the increasing number of useless wavelength conversions. On the contrary, the MW-MBM scheduler does not perform useless wavelength conversions.

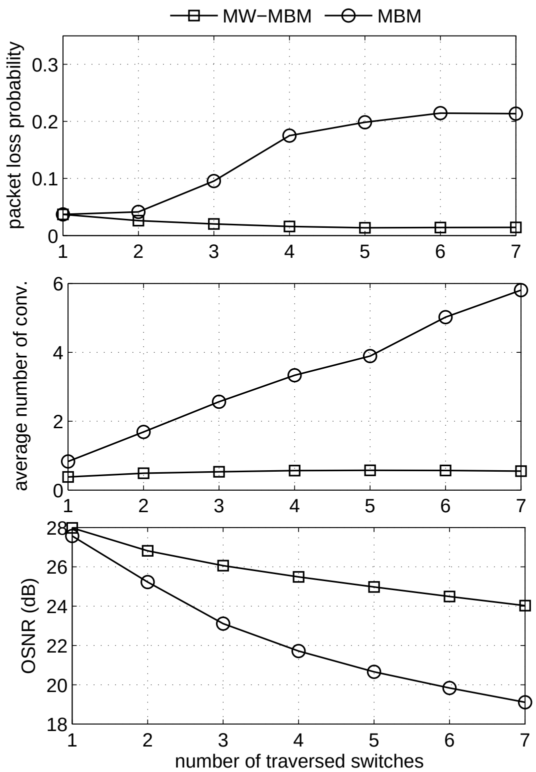

Fig. 12 – End-to-end packet loss probability (a), average number of conversion per packet (b) and end-to-end average OSNR per packet (c) versus the number of traversed switches, in case of aggregating-chain topology, WDM system with *M*=32 wavelengths, traffic load = 0.8, available conversion range *d*=6

V. Conclusion

The research carried out in this paper has highlighted that a WDM packet scheduler properly designed to maximize the throughput and, as secondary goal, to minimize the detuning strongly improves both network-level and physical-level performances, in comparison with a scheduler that does not account for the detuning minimization.

Such an optimal scheduler problem can be formalized as the finding of a maximum bipartite matching, whose sum of edges weights is minimum (MW-MBM).

We used the Hungarian algorithm to find the matching, but we envisage that lower complex or approximating algorithms need to be devised by further studies, since a high packet rate is expected for a WDM packet switch. We are investigating a low complex greedy algorithm that iteratively singles out the packet that has the highest probability to be discarded in the next iteration steps (for absence of a free feasible wavelength). Then the greedy algorithm schedules the forwarding of such a packet on the free output wavelength closest with the input one.

References

[1] S. J. Ben Yoo, "Optical Packet and Burst Switching Technologies for the Future Photonic Internet," J. Lightwave Technol. 24, 4468-4492 (2006)

[2] S.L. Danielsen, C. Joergensen, B. Mikkelsen, K.E. Stubkjaer, "Optical packet switched network layer without optical buffers," IEEE Photonics Technology Letters, vol.10, no.6, pp.896-898, Jun 1998

[3] V. Eramo, M. Listanti, M. Spaziani, "Resources sharing in optical packet switches with limited-range wavelength converters," Lightwave Technology, Journal of , vol.23, no.2, pp. 671-687, Feb. 2005

[4] E. Iannone, R. Sabella " Performance evaluation of an optical multi-carrier network using wavelength converters based on FWM in semiconductor optical amplifiers," IEEE Journal of Lightwave Technology, vol.13, no.2, pp.312-324, Feb 1995

[5] Geraghty, David F. and Lee, Robert B. and Verdiell, Marc and Ziari, Mehrdad and Mathur, Atul and Vahala, Kerry J. "Wavelength conversion for WDM communication systems using four-wavemixing in semiconductor optical amplifiers" IEEE Journal of Selected Topics in Quantum Electronics, Vol. 3, No. 5, 1997 pp. 1146-1155

[6] D'Ottavi, A.; Iannone, A.; Mecozzi, A.; Scotti, S.; Spano, P.; Dall'Ara, R.; Eckner, J.; Guekos, G., "Efficiency and noise performance of wavelength converters based on FWM in semiconductor optical amplifiers," Photonics Technology Letters, IEEE , vol.7, no.4, pp.357-359, Apr 1995

[7] Gangxiang Shen; Bose, S.K.; Tee Hiang Cheng; Chao Lu; Teck Yoong Chai, "Performance study on a WDM packet switch with limited-range wavelength converters," Communications Letters, IEEE , vol.5, no.10, pp.432-434, Oct 2001

[8] Zhenghao Zhang; Yuanyuan Yang, "Optimal scheduling algorithms in WDM optical interconnects with limited range wavelength conversion capability," Parallel and Distributed Systems, IEEE Transactions on , vol.15, no.11, pp. 1012-1026, Nov. 2004

[9] V. Eramo; M. Listanti; M. Di Donato, "Performance evaluation of a bufferless optical packet switch with limited-range wavelength converters," Photonics Technology Letters, IEEE , vol.16, no.2, pp.644-646, Feb. 2004

[10] Zhang, Z.; Liu, L.; Yang, Y., "Slotted Optical Burst Switching (SOBS) Networks," Network Computing and Applications, 2006. NCA 2006. Fifth IEEE International Symposium on , vol., no., pp.111-117, 24-26 July 2006

[11] Eramo, V.; Listanti, M.; Venuti, L.; Tarola, M., "Performance of scheduling algorithms in optical packet switches with limited-range wavelength converters," Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE , vol.4, no., pp.6 pp.-2051, 2-2 Dec. 2005

[12] Mohanty, S.R.; Bhuyan, L.N., "Lexicographic Fairness in WDM Optical Cross-Connects," INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE , vol., no., pp.197-205, 6-12 May 2007

[13] Liu, L. and Yang, Y. , "Optimal packet scheduling in output-buffered optical switches with limited-range wavelength conversion", in Proceedings of the 3rd ACM/IEEE Symposium on Architecture For Networking and Communications Systems (Orlando, Florida, USA, December 03 - 04, 2007). ANCS '07. ACM, New York, NY, 77-86.

[14] Eramo, V.; Listanti, M.; Germoni, A., "Cost Evaluation of Optical Packet Switches Equipped With Limited-Range and Full-Range Converters for Contention Resolution," Lightwave Technology, Journal of , vol.26, no.4, pp.390-407, Feb.15, 2008

[15] F. Glover, "Maximum matching in a convex bipartite graph", Naval Research Logistics Quarterly, vol. 14, 313-316, (1967)

[16] Lipski Jr., W., and Preparata, F. P., "Efficient algorithms for finding maximum matchings in convex bipartite graphs and related problems. Acta Informatica 15 (1981), 329-346"

[17] G. Steiner, J. S. Yeamans, "A Linear Time Algorithm for Maximum Matching in Convex, Bipartite Graphs", Elsevier Computerss Math. Applic. Vol. 31, No. 12, pp. 91-96, 1996

[18] West, Douglas Brent (1999), Introduction to Graph Theory (2nd ed.), Prentice Hall, ISBN 0-13-014400-2

[19] Harold W. Kuhn, "The Hungarian Method for the assignment problem", Naval Research Logistics Quarterly, 2:83–97, 1955

[20] F. Bourgeois, J. C. Lassalle, "An extension of the Munkres algorithm for the assignment problem to rectangular matrices", Comm of the ACM, vol. 14, pp. 802, 1971

[21] John Weaver, "Munkres algorithm code v2" available at http://johnweaver.zxdevelopment.com/wp-content/uploads/2007/05/munkres-v2.zip

[22] "Network Simulator 2" available at http://nsnam.isi.edu/nsnam/index.php/

[23] Andrea Detti, Ns 2.31 WDM extension, http://netgroup.uniroma2.it/Andrea_Detti/WDM/Simulator.tar.gz

APPENDIX I
LEAST FLEXIBLE FIRST

In this appendix we discuss an original greedy scheduling algorithm, named Least Flexible First (LFF), that provides a suboptimal solutions to the MW-MBM problem. The LFF algorithm can be summarized as follow:

1- step-by-step select the packet with smallest *flexibility* in being accomodated in the next step. The flexibility is the number of current free output wavelengths over which packet could be forwarded;
2- accomodate that packet on the free wavelength closer with its input wavelength and make busy the chosen output wavelength;
3- repeat previous steps until no more packet can be accomodated on output wavelengths.

Step 1 is meant at limiting the number of dropped packets; thus it copes with the primary goal of troughput maximization. In the next algorithm steps there will be lower free output wavelengths, thus step 1 pracautionary singles out the packet with lower chances to succeed the future. Step 2 is aimed at minimizing the wavelength detuning, thus it copes with the secondary goal of detuning minimization.

The computational burden of LFF algorithm is mainly due two step 1 and step 3. Step 1 is an iteration looping on the remaining packets; therefore the time complexity is in the order of $S*M$, where S is the number of input fiber of the optical switch and M is the number of output wavelengths. Indeed, $S*M$ is the maximum number of packets entering the switch. Step 3 provides an iteration looping at most M times. Overall the complexity of LFF is $O(S*M^2)$. We observe that the complexity of the Hungarian Method is well-known to be $O((S*M)^3)$, hence LFF provides a considerable complexity reduction.

We have integrated LFF in NS2 simulator and have evaluated its performance in the aggregating-chain topology (Fig. 5). The obtained results are reported in Fig. 13, Fig. 14 and Fig. 15. We observe that LFF performance is quite close with the optimum one, i.e. that achieved by MW-MBM scheduler based on the Hungarian Method (MW-MBM).

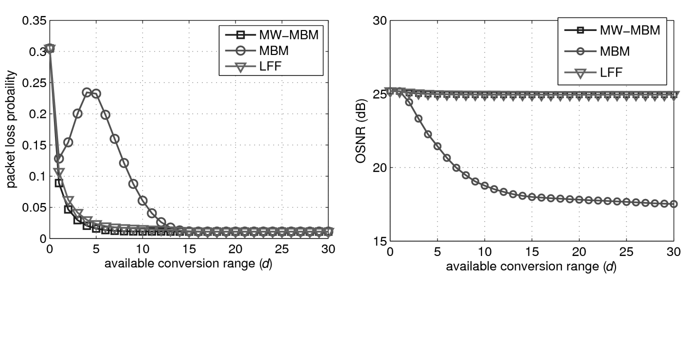

Fig. 13 – End-to-end packet loss probability and average OSNR per packet versus available conversion range (d), in case of aggregating-chain topology, WDM system with 32 wavelengths (M=32), 5 switching-stages, traffic load equal to 0.8

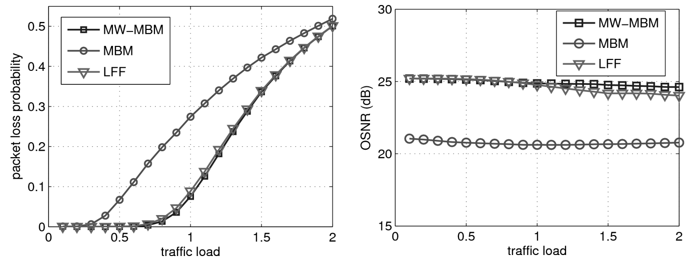

Fig. 14 – End-to-end packet loss probability and average OSNR per packet versus traffic load, in case of aggregating-chain topology, WDM system with 32 wavelengths (M=32), 5 switching-stages, available conversion range (d) equals to 6

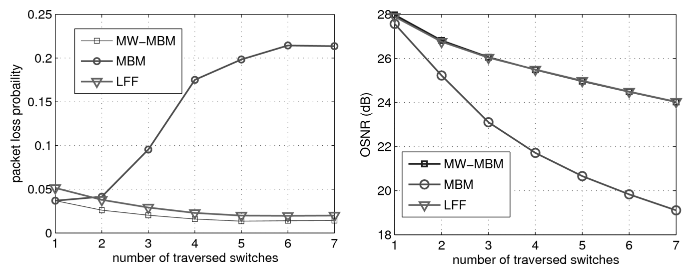

Fig. 15 – End-to-end average OSNR per packet versus the number of traversed switches, in case of aggregating-chain topology, WDM system with 32 wavelengths (M=32), traffic load = 0.8, available conversion range (d) equals to 6