Time-Varying Spectrum Allocation Policies and Blocking Analysis in Flexible Optical Networks

Konstantinos Christodoulopoulos, Ioannis Tomkos, and Emmanouel Varvarigos

Abstract—**We consider the problem of serving traffic in a spectrum-flexible optical network, where the spectrum allocated to an end-to-end connection can change so as to adapt to the time-varying required transmission rate. In the proposed framework, each connection is assigned a route and is allocated a reference frequency over that route, using an appropriate Routing and Spectrum Allocation (RSA) algorithm, but the spectrum it utilizes around the reference frequency is allowed to expand and contract to match source rate fluctuations. We propose and analyze three spectrum expansion/contraction (SEC) policies for modifying the spectrum allocated to each connection. The first policy, named the Constant Spectrum Allocation (CSA) policy, allocates a number of spectrum slots for exclusive use by each connection. We also present two policies that enable the dynamic sharing of spectrum slots among connections, named the Dynamic High Expansion-Low Contraction (DHL) and the Dynamic Alternate Direction (DAD) policy. We give exact formulas for calculating the blocking probability for a connection and for the whole network under the CSA policy and provide corresponding approximate analyses under the DHL and DAD policies. We also present a simple iterative RSA algorithm that uses the developed blocking models so as to minimize the average blocking of the network.**

Index Terms—**Spectrum-flexible networks, time-varying traffic, spectrum sharing, spectrum expansion/contraction policies, routing and spectrum allocation, blocking probability.**

I. Introduction

INCREASING the capacity and improving the efficiency of optical transport networks has been an important research challenge for many years. To cope with traffic increases of almost 34% per year [1], extensive research efforts have been devoted on advanced modulation formats and digital equalization in the electronic domain to enable per-channel rates of 40 and 100 Gbps with improved transmission distance in traditional Wavelength Division Multiplexing (WDM) systems [2]. However, although wavelength routed WDM networks offer well-known advantages, their rigid and coarse granularity leads to inefficient capacity utilization, a problem expected to become more severe with the deployment of higher channel rate systems.

In order to address the issues of stranded and underutilized bandwidth, low agility, inefficient resource utilization due to over-provisioning, and wasted capital/operational expenses, a new networking approach is required. Optical Burst Switching (OBS) and Optical Packet Switching (OPS) utilize the time domain to enable statistical multiplexing gains through the sharing of the network resources. OBS prototypes and commercial products (e.g. by Matisse, Huawei, Intune) have reached the market for metro-ring networks, but still have not met commercial success, while there are a number of studies that question the benefits of OBS when used for transport networks [4]. OPS is considered a long term solution as its enabling technologies are still immature [5]. In addition to the time domain, exploited in the OBS and OPS paradigms, the frequency is another domain that can be harvested to provide improved flexibility, granularity and efficiency for the optical networks. Typically, wavelength routed WDM networks, as well as OBS and OPS, operate over the ITU-T grid, that is, a constant 50-GHz frequency spaced grid. Taking a different approach, recent research efforts have focused on networks that support variable spectrum connections to obtain flexibility and statistical multiplexing gains in the spectrum domain [3]. Spectrum-flexible, elastic, flexgrid are few examples of the terms used to describe these architectures [6]-[15].

A spectrum-flexible network has spectrum granularity finer than that of standard 50-GHz ITU grid WDM systems and can also combine the spectrum units (slots) to create wider channels on an as needed basis. To establish a connection in a spectrum-flexible network, every spectrum-flexible optical cross connect (OXC), or flexible channel bandwidth OXC, as referred to in [12], on the connection's route allocates sufficient spectrum to create an appropriately sized end-to-end optical path. Standardization of a grid with granularity 6.25 and 12.5 GHz is under discussion. It stands to reason that, from a networking perspective, the optimal slot size depends on the traffic characteristics, but, for realistic networks and traffic loads, the related studies show that significant benefits can be achieved with the standardized granularities.

A variety of spectrum-flexible systems have been described in the literature. The Spectrum-sLICed Elastic optical path network (SLICE) [6],[8] utilizes optical Orthogonal Frequency Division Multiplexing (OFDM) to enable spectrum-flexible transmissions. Optical OFDM distributes the data on several low data rate subcarriers (multi-carrier system). The spectrum of adjacent subcarriers can overlap, since they are orthogonally modulated, increasing spectral efficiency. A bandwidth-variable OFDM transponder generates an optical signal using just enough spectrum with appropriately modulated subcarriers to serve the client demand. Single-carrier systems are also

Manuscript received 15 December 2012; revised 1 September 2012. Parts of this paper were presented at the 16th Inter. Conf. on Optical Networking Design and Modeling (ONDM) University of Essex, UK, 2012.

K. Christodoulopoulos is with the School of Computer Science and Statistics, Trinity College Dublin, Ireland (e-mail: christok@tcd.ie).

I. Tomkos is with the Athens Information Technology Center, Athens, Greece (e-mail:itom@ait.edu.gr).

E. Varvarigos is with the Department of Computer Engineering and Informatics, University of Patras, Greece (e-mail: manos@ceid.upatras.gr).

Digital Object Identifier 10.1109/JSAC.2013.130103.

envisioned to operate in a spectrum-flexible manner, as e.g. the Flexible-WDM (FWDM) architecture considered in [7].

The introduction of spectrum-flexible networks and non-constant spectrum connections pose new challenges on the networking level, as traditional algorithms designed for fixed-grid WDM systems are no longer applicable. The resource allocation problem in spectrum-flexible networks is usually referred to as the Routing and Spectrum Allocation (RSA) problem. The planning of spectrum-flexible networks has been considered in [6],[7],[14], where static or offline RSA algorithms were proposed. In the planning problem, a traffic matrix with the requested transmission rates of all connections is given (corresponding to their peak rate typically multiplied by a certain oversubscription ratio). The solution consists of the paths and spectrum allocation that should be used by the connection demands so as to minimize a performance metric, such as the utilized spectrum. Connections are served for their requested capacity and no spectrum overlapping is allowed among them, resulting in a waste of spectrum when connections do not fully utilize their allocated spectrum.

After planning and installing the network we enter the operational phase, in which the rates of the connections fluctuate with time and, depending on the granularity of the system, at certain instants changes have to be performed to accommodate the traffic variations. The relatively few works that have considered the online problem in spectrum-flexible networks [8]-[11], consider traffic changes as new connection requests and terminations. For example, in [10], scenarios of establishing and releasing new connections of 10 up to 400 Gbps are considered. Assuming that the network employs high capacity flexible optical transponders, relatively large periods of time will pass until an additional connection needs to be established or released. We adopt a different approach by focusing on the short- to medium-term traffic fluctuations that occur in reality.

We envision an elastic network where nodes communicate over adjustable-rate end-to-end connections, without establishing new or releasing old connections unless specifically required. Changes (usually smooth) in the requested rate happen dynamically, and can be absorbed by the flexible transponders by changing their utilized spectrum in real time. This is done by expanding (if feasible) or contracting the continuous spectrum allocated to the existing connections. No spectrum overlapping among connections is allowed at any given time, but the spectrum can be shared among connections at different time instants. Spectrum sharing among connections enables the statistical multiplexing of traffic over the same network, yielding gains similar to those obtained by the time-sharing of resources in OBS/OPS networks. A similar approach was followed in [13], where connections with negatively correlated rates were placed adjacent so as to obtain statistical multiplexing gains. The dynamic spectrum sharing policies we propose are more general than those in [13], enabling the sharing of spectrum among more than just two adjacent connections. Moreover, our methods work for general and uncorrelated traffic.

Specifically, the goal of this paper is to develop a framework for serving time varying traffic in a spectrum-flexible network that enables the dynamic spectrum sharing among connections, propose basic sharing policies, and examine the blocking performance of the resulting network. The techniques and analyses to be given apply to an OFDM-based or any other type of spectrum-flexible optical network, as long as dynamic spectrum-adjustable transponders are employed in the system. In our previous work [15], we used simulation experiments to get a first estimation on the performance improvements that can be obtained when connections can adapt their utilized spectrum to follow their traffic variations. In this paper we present a more complete and detailed solution.

In the proposed framework, each connection is assigned a route and a reference frequency. The connection is allowed to expand and contract the spectrum used around this reference frequency, as long as it does not overlap with that used by adjacent connections at any particular time. We propose and compare three spectrum expansion/contraction (SEC) policies. The first is a simple Constant Spectrum Allocation (CSA) policy that defines the exclusive use of a set of spectrum slots to a connection. This policy, which is adaptive but offers no sharing and no statistical gains among connections, is compared to two dynamic policies, called the Dynamic High Expansion-Low Contraction (DHL) policy and the Dynamic Alternate Direction (DAD) policy, which enable the dynamic sharing of spectrum slots among connections. We give an exact analysis for calculating the blocking probability of each connection and of the whole network for the CSA, the DHL and the DAD policies. However, since the calculations for the DHL and DAD cases are computationally demanding we also present approximation models. We also present a simple iterative Routing and Spectrum Allocation (RSA) algorithm that takes into account these blocking models to allocate routes and reference frequencies to connections so as to minimize the network blocking.

We perform simulation experiments to evaluate the proposed framework, the blocking performance of the SEC policies and the accuracy of the developed blocking models. The analytical results obtained are in very close agreement with corresponding simulations. Our results show that the DHL and DAD policies significantly outperform the CSA policy, by enabling the sharing of spectrum between adjacent connections and the corresponding statistical multiplexing gains.

II. Spectrum-Flexible Network and Time-Varying Traffic

We consider a spectrum-flexible (flexgrid) optical network, where the spectrum is divided into constant spectrum slots with granularity C GHz finer than the typical 50-GHz grid used in WDM systems (e.g. C=12.5 or 6.25 GHz). The switching granularity of the nodes and the transponders is one spectrum slot. The network supports a restricted number of spectrum slots T, determined by the switching window of the OXCs. A spectrum slot is identified by its starting frequency $F_0 + F \cdot C$, F=0,1,...,T-1, where F_0 is the lowest frequency supported in the system. To simplify notation, we will use a quantized frequency axis and characterize each frequency slot with an integer F, F=0,1,...,T-1.

A connection is served by a specific path and a set of continuous spectrum slots on all the links of that path (satisfying the related spectrum continuity constraint). Guardband G (in spectrum slots) is used between spectrum-adjacent connections to ensure that the interference between them is acceptable. The traffic of the connection varies as a function of time (a birth-death traffic model will be used in Section III to analyze the blocking performance), and the connection can dynamically increase/decrease the spectrum slots it utilizes around its reference frequency, so as to follow the variations in the requested traffic rate. The way connections adapt their utilized spectrum to their instantaneous traffic rate is called the *spectrum expansion/contraction* (abbreviated SEC) policy. To support dynamic operation we assume that the network employs spectrum-flexible transponders capable of dynamically adapting the spectrum they utilize [2],[3]. The transponders in the envisioned network may either (a) expand/contract the spectrum used in both directions non-symmetrically, or (b) they may do so symmetrically and be able to dynamically adjust their central frequency. (We note that the central frequency referred here differs from the reference frequency used in our analysis). In the latter case, techniques, such as the push-pull technique in [11], can be used to avoid traffic interruption. Signaling extensions should be provided for the basic expansion/contraction of the spectrum used by a connection, and, if needed, for the change in its central frequency. These can be done with relatively simple extensions to the RSVP-TE protocol, which is used for Label Switched Path (LSP) signaling in the GMPLS control plane. Finally, the proposed SEC policies need also to be extended with contention resolution rules in order to resolve the resource allocation between conflicting connections requesting to adapt their allocated resources simultaneously. Addressing these issues falls outside the scope of the current paper.

Network performance does not only depend on the SEC policy used, but also on the arrangement of the connections in the space (routing) and frequency (spectrum allocation) domains, which depends itself on the RSA solution. Fig. 1 presents the process of serving time-varying connections in the envisioned spectrum-flexible network. The RSA under time-varying traffic algorithm serves the connection request, by assigning a path and a reference frequency to it so as to minimize the average network blocking, taking into account the dynamic nature of the connections and the specific SEC policy used. After the RSA decision has been made, the SEC policy is responsible for accommodating the dynamic traffic variations of the connection by expanding/contracting the spectrum that it uses at spectrum-slot level granularity. The RSA algorithm is used again when a connection cannot, on a regular basis, obtain additional spectrum slots (high blocking), or when the requested rate exceeds the transponder capabilities. Then, RSA is called to route the excess traffic over a different spectrum-path, or reroute the entire connection (to save in guardbands).

Our emphasis in this paper is more on the general framework for serving time-varying traffic and on the SEC policies that can be used, and less on the RSA under time-vayring traffic algorithm. Static RSA algorithms [6],[7],[14] assign paths-spectrum to static (constant-rate) connections, but they can also be used to serve time-varying traffic. This can be done by using as input to the static RSA algorithms static demands

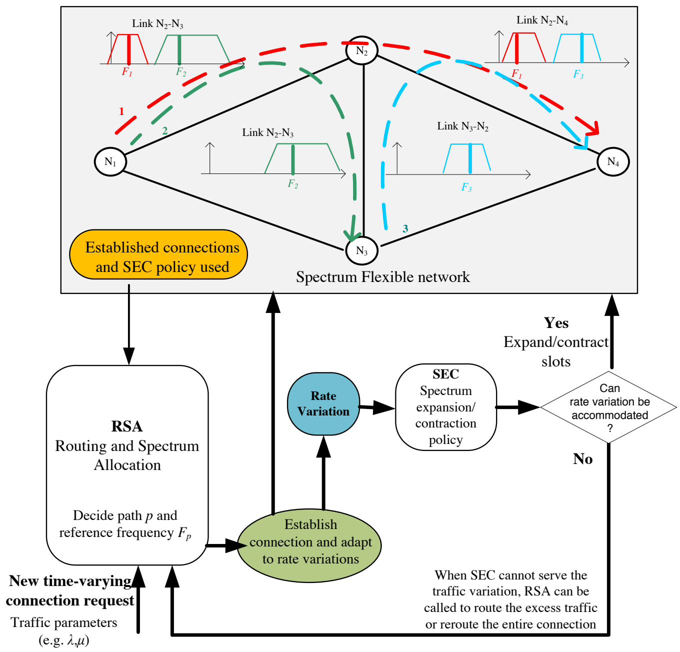

Fig. 1. Flow diagram for serving a connection request. RSA is used to determine the path and reference frequency and SEC policy takes care of traffic variations. When the SEC policy cannot accommodate a traffic variation, or if the rate exceeds the capability of the transponder, RSA can be triggered to set up an additional connection or reroute the existing connection.

that are however estimated using probabilistic models that account for the dynamic characteristics of the connections and the SEC policy used. For the sake of being specific, we outline in Section IV an RSA under time-varying traffic algorithm that uses blocking models obtained by our analysis and assigns paths and reference frequencies to the connections so as to minimize the overall blocking in the network.

A. Spectrum-flexible framework for serving dynamic traffic

The optical network is represented as a graph ($\boldsymbol{V}$,$\boldsymbol{E}$), where $\boldsymbol{V}$ is the set of nodes and $\boldsymbol{E}$ the set of fiber links. We consider a connection that has been processed by the RSA algorithm and has been assigned a specific path p and a reference spectrum slot F_p, and utilizes n_p^H and n_p^L spectrum slots higher and lower than F_p, respectively. Therefore, a total of $n_p = n_p^H + n_p^L$ spectrum slots, $[F_p + n_p^L, F_p - n_p^H - 1]$, have been allocated to the connection on all the links of path p, so as to satisfy the spectrum continuity constraint. We will call F_p the reference frequency and not the starting frequency, since the connection can utilize spectrum slots lower than that. The total spectrum n_p (in slots) used by the connection is adjusted as a function of time to follow the traffic fluctuations, but no two connections can utilize the same spectrum slots over any link at any given time. This non-overlapping spectrum assignment constraint in spectrum-flexible optical networks corresponds to the single wavelength assignment constraint of traditional WDM networks.

Fig. 2 presents the spectrum slot utilization of two links, l and l', on path p. We denote by $U(p,l)$ and by $B(p,l)$ the upper and bottom spectrum-adjacent connections, respectively, of connection p on link l. We also denote by $F_{U(p,l)}$ and by

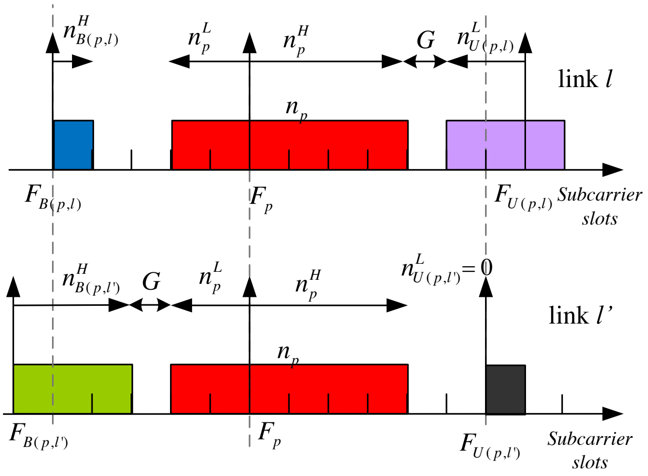

Fig. 2. Example of spectrum slot utilization on two links l and l'. The figure illustrates the spectrum allocated to connection p and its upper and bottom spectrum-adjacent connections $U(p,l), U(p,l')$ and $B(p,l), B(p,l')$ on links l and l', respectively.

$n^L_{U(p,l)}$ the reference frequency and the number of lower slots utilized by the upper spectrum-adjacent connection, and by $F_{B(p,l)}$ and $n^H_{B(p,l)}$ the reference frequency and the number of higher slots utilized by the bottom spectrum-adjacent connection on link l. We also denote by G the guardband in spectrum slots used to enable the switching of the connections with low and acceptable interference (G=1 slot in Fig. 2).

According to the expansion policy, if the traffic of connection p increases, requiring the allocation of an additional spectrum slot to it, we increase either its higher or its lower spectrum slots. Similarly, the used policy decreases the higher or lower allocated slots when the transmission rate of the connection decreases. The non-overlapping spectrum assignment requirement constrains the number of higher and lower spectrum slots that can be utilized by the connection:

$$\begin{aligned} 0 \le n^H_p \le \min_{l\in p}(F_{U(p,l)} - n^L_{U(p,l)}) - F_p - G, \text{ and} \\ 0 \le n^L_p \le F_p - \max_{l\in p}(F_{B(p,l)} + n^H_{B(p,l)}) - G. \end{aligned} \tag{1}$$

Blocking will occur when a connection requires an additional slot (due to an increase in its transmission rate) and there are no available slots to accommodate it. In particular, for blocking to occur for p, an additional slot should be requested when the right hand sides of Eq. (1) are met with strict equalities,

$$n^H_p = \min_{l\in p}(F_{U(p,l)} - n^L_{U(p,l)}) - F_p - G, \text{ and} \tag{2}$$

$$n^L_p = F_p - \max_{l\in p}(F_{B(p,l)} + n^H_{B(p,l)}) - G. \tag{3}$$

According to the flow diagram in Fig. 1, if blocking happens too frequently, meaning that the connection needs more spectrum resources than those it shares with its spectrum-adjacent connections, the RSA algorithm could be triggered to route the excess traffic or reroute the entire connection.

Under the proposed spectrum sharing framework, a connection shares the spectrum slots with its upper and bottom spectrum-adjacent connections. In general, spectrum sharing can be performed in a number of ways and the proposed framework can be extended to include additional rules. For example, we can allocate a number of slots exclusively to each connection and then use a policy for sharing the remaining slots among the connections. These options are not ruled out and are left for future studies. The SEC policy that defines how the spectrum expands/contracts affects significantly the network's operation and blocking performance. In the next subsection we present three such SEC policies.

B. Spectrum Expansion/Contraction (SEC) policies

1) Constant spectrum allocation (CSA) policy: In the first policy, a connection is assigned a path p and reference frequency F_p and has exclusive use of all spectrum slots higher than F_p, up to the reference frequency of its closest upper spectrum-adjacent connection. That is, connection p can expand and use $n^H_p, 0 \le n^H_p \le N^H_p$, higher spectrum slots, where

$$N^H_p = \min_{l\in p}(F_{U(p,l)}) - F_p - G. \tag{4}$$

Under the CSA policy, connection p uses slots $[F_p, F_p + N^H_p - 1]$ independently of what other connections do, while it cannot use lower spectrum slots at all.

This policy does not enable the sharing of spectrum slots among the connections, it resembles the static spectrum allocation for connections of given rate, and is proposed mainly for comparison purposes. In such a system, blocking will occur for connection p when it requires an additional slot, and it has already used up all its higher slots, that is, when $n^H_p = N^H_p$. This is a special case of the general bounding conditions described by Eq. (2) and (3), after setting n^L=0 for all connections.

2) Dynamic high expansion-low contraction (DHL) policy: We now present a dynamic SEC policy, called *Dynamic High Expansion-Low Contraction* (DHL), that enables the sharing of the spectrum among the connections. With DHL, a connection p wishing to increase its transmission rate, first uses its higher spectrum slots, increasing n^H_p until it reaches a slot already occupied by an upper spectrum-adjacent connection on some link of p, that is, until the bounding condition of Eq. (2) is met. Then, if additional bandwidth is needed, it expands its lower spectrum slots, increasing n^L_p until it reaches a slot that is occupied by some bottom adjacent connection on some link, that is, until the bounding condition of Eq. (3) is met. If the connection needs to increase further its rate and there is no higher or lower free slot space, blocking occurs (for the excess rate). Note that there may be cases where even though at previous instants there were no free higher spectrum slots and the connection had to utilize some lower spectrum slots, the constraining upper spectrum-adjacent connection has freed some of its used slots, creating free higher slot space. Such cases can create slot voids, a problem that can be addressed, if desired, by a defragmentation process that might run in between arrivals. The DHL policy performs indirectly slot defragmentation, since it always searches first for free higher slots, even if lower spectrum slots have already been used by that connection, filling the free higher spectrum slots in every chance it gets. Note that spectrum defragmentation has been raised as an issue in other works [8]-[10]. When a connection decreases its spectrum slots due to a reduction in its rate, we first release lower spectrum slots and, if these have been reduced to zero, we release higher spectrum slots.

3) Dynamic Alternate Direction (DAD) policy: The third examined SEC policy, called *Dynamic Alternate Direction* (DAD), also enables the sharing of the spectrum among the connections and aims at the symmetrical use of spectrum around the reference frequencies. With DAD, a connection *p* wishing to increase its transmission rate, alternates between using its higher and lower spectrum slots starting from its higher slots, increasing n_p^H and n_p^L alternately until it reaches a slot already occupied by an upper or bottom, respectively, spectrum-adjacent connection, that is, until the bounding condition of Eq. (2) or (3) is met. Then, if additional slots are needed, it expands towards the other direction, in which case the symmetry is lost. When expanding the spectrum of a connection, the DAD policy always examines if it can expand towards the direction that uses fewer slots. Thus, slots that were previously bounding at one direction and are freed at a later time have priority and are firstly used if needed. In this way voids are filled and spectrum defragmentation, as in the DHL policy, is indirectly achieved. Blocking occurs if the connection needs more slots and there is no higher or lower free slot space, that is, if the bounding conditions of Eq. (3) and (2) are both met. When a connection decreases its slots due to a reduction in its rate, we first release spectrum slots from the direction that has used more slots, and once we have an equal number of higher and lower slots, we decrease the lower spectrum slots. Thus, both expansion and contraction processes are designed to yield symmetrical spectrum utilization so that the central frequency of the connections does not frequently change and remains close to the related reference frequency when congestion is low.

III. Analyzing the Performance of the Spectrum Expansion/Contraction Policies

In this section we present analytical methods for calculating the average blocking probability of the spectrum expansion/contraction (SEC) policies described in Section II.B.

As already mentioned, each connection is assigned by the RSA algorithm a specific path *p* and a reference frequency F_p. The traffic rate of the connection, however, may fluctuate dynamically with time, and the same applies to the number of spectrum slots allocated to it. In the analysis to be proposed, the requested slots of a connection are assumed to follow a birth-death Markovian model. In particular, we assume that requests for additional spectrum slots over path *p* are generated according to a Poisson process of rate λ_p and their holding times are exponentially distributed with mean $1/\mu_p$, corresponding to a traffic load $\rho_p = \lambda_p/\mu_p$ for that connection. These traffic parameters are taken into account in the RSA algorithm when assigning routes and reference frequencies to the connections. We also assume that the additional slot requests are independent for different connections. If the slot requests of different connections were correlated, we could design an RSA algorithm that would exploit this correlation information to obtain gains that are more significant than those described in our performance results.

In practice, in core networks, a connection over path *p* will carry traffic from many users aggregated at the corresponding ingress/source node and destined for users at the specific egress/destination node. Since the traffic is aggregated over many users and the spectrum is allocated in the form of slots of a given granularity, rate fluctuations will be relatively smooth, and a birth-death model for the number of slots looks appropriate. This is a well-studied traffic model that has been widely used for circuit switched traffic [16]-[18] and for fixed-grid WDM networks [19]-[20] (in which cases, circuit or wavelength requests are created instead of slot requests). Our model can also be related to other models, e.g., models assuming that user sessions arrive and depart from the system, since by knowing the capacity supported by a slot we can transform the user session arrival model to a slot birth-death model.

A. Computing the blocking of an expansion/contraction policy

The state of the network can be formulated as a *d*-dimensional continuous time Markov chain, where d=$2\cdot|\boldsymbol{V}|\cdot(|\boldsymbol{V}|-1)$ and $|\boldsymbol{V}|$ is the number of nodes. This can be seen by noting that there are $|\boldsymbol{V}|\cdot(|\boldsymbol{V}|-1)$ possible connections in the network, and for each connection we need to record both the number of higher and lower utilized slots to characterize it. If some pairs of nodes never communicate with each other or some pairs utilize more than one connection, the dimensions of the system can be reduced/increased accordingly. We denote by $\boldsymbol{n} = (n_1^H, n_1^L, n_2^H, n_2^L, ..., n_{d/2}^H, n_{d/2}^L)$ the state of the network at any given time, and by $P(\boldsymbol{n})$ the steady-state probabilities. The set of feasible states is described by Eq. (1). The transitions between states depend on the SEC policy used. Since the blocking states are given by Eq. (2) and (3), the blocking probability of connection *p*, denoted by $P_{bl}(p)$, and the network blocking probability averaged over all connections, denoted by P_{av}, are given by:

$$P_{bl}(p) = \sum_{\boldsymbol{n}:\text{Eq.(2) and (3) are met}} P(\boldsymbol{n}), \tag{5}$$

$$P_{av} = \frac{\sum_p \lambda_p \cdot P_{bl}(p)}{\sum_p \lambda_p}. \tag{6}$$

To calculate the average network blocking probability P_{av}, in addition to the traffic parameters, we need to know the utilized paths and the reference frequencies selected by the RSA algorithm for all connections.

To calculate the probabilities $P(\boldsymbol{n})$ for the blocking states of $\boldsymbol{n}$ defined by Eq. (2) and (3), we need to solve the global balance equations for the *d*-dimensional Markov chain that describes the system for the particular SEC policy used, something that is not trivial in most cases. Since in the proposed policies spectrum sharing is possible only between spectrum-adjacent connections, the blocking states for a connection depend on the utilization of its bottom and upper spectrum-adjacent connections over all the links of the path it follows. In turn, the utilization of these connections depends on the utilization of their bottom and upper spectrum-adjacent connections, and so on. Thus, the interdependence among the connections is quite complicated and cannot be simplified in the general case. In the following subsections we analyze the CSA, DHL and DAD SEC policies presented in Sections II.B.1, 2 and 3, and provide analytical expressions to calculate, exactly for the CSA policy and approximately for the DHL and DAD policies, their corresponding blocking probabilities.

1) Exact model for blocking computation for the CSA policy: In the CSA policy, no spectrum is shared among the connections, and the Markov chain describing the network state is greatly simplified. This is because the independence among the connections under the CSA policy makes possible the decomposition of the *d*-dimensional Markov chain into *d* separate 1-dimensional chains, each corresponding to an M/M/m/m queuing model and describing one connection. Thus, the blocking probability for new slot requests over path *p* is given by the Erlang-B formula:

$$P_{bl}^{CSA}(p) = \frac{(\rho_p)^{N_p^H}/N_p^H!}{\sum_i (\rho_p)^i/i!}, \tag{7}$$

where $\rho_p = \lambda_p/\mu_p$, and N_p^H is the number of slots allocated exclusively to connection *p* [calculated by Eq. (4)]. The network blocking probability is then obtained from Eq. (6), using Eq. (7) for the blocking probability of each connection. Note that Eq. (7) is a special case of the general Eq. (5), after considering the particularities of the CSA policy, which removes the interdependence among the connections.

2) Model for blocking computation for the DHL policy: We now analyze the performance of the DHL SEC policy described in Section II.B.2. The state transitions for connection *p* are given in Table 1. In this table we denote by $\boldsymbol{e}_p^L$ (or $\boldsymbol{e}_p^H$) the vector with all zero elements apart from element n_p^L (or n_p^H, respectively) which equals to one. Using the transition rules of DHL as reported in Table 1, we can define the corresponding *d*-dimensional Markov chain, compute the steady state probabilities from the global balance equations, and the blocking probabilities for the connections by Eq. (5). Note that dependencies exist between the state of connection *p* and that of all its bottom and upper spectrum-adjacent connections, on all the links traversed by *p*, which are in turn dependent on their bottom and upper spectrum-adjacent connections, and so on. So the interdependence of the connections goes deep and the computation of the exact state probabilities is feasible only for small networks, and is intractable for more realistic situations. In particular, since the way the variables n_p^H and n_p^L characterizing a connection increase/decrease depends on each other and on the utilization of the upper adjacent connections, the system cannot be decomposed and we cannot express the stationary distributions in product forms and follow an analysis like [16]. Since the calculation of the exact blocking of the DHL is very complicated and can be performed only for small problem instances, we present in the next subsection an approximate model for estimating it.

a) Approximate model for computing the blocking of the DHL policy for large networks

We focus on a particular connection *p* and we want to find the probability that a new spectrum slot request over *p* will be blocked under the DHL policy. For the given connection *p*, we define a hypothetical (approximate) system $\widetilde{\text{DHL}}$ in which the connection *p* follows the DHL policy while all other connections expand their spectrum only towards their higher slots. We denote the *d*/2-dimensional state of the approximate system $\widetilde{\text{DHL}}$, by $\tilde{\boldsymbol{n}} = (\tilde{n}_1, ..., \tilde{n}_{B(p,l)}, \tilde{n}_p, \tilde{n}_{U(p,l)}, ..., \tilde{n}_{d/2})$, where $B(p,l)$ and $U(p,l)$ are the bottom and upper spectrum-adjacent connections of *p* on link *l* of *p*. We make the following approximating assumptions for calculating the slot request blocking probability of connection *p*:

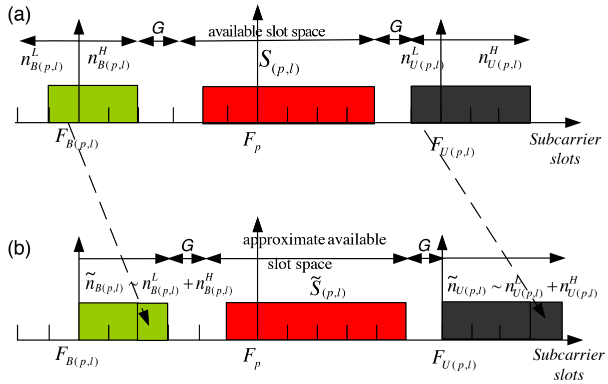

Fig. 3. (a) Upper and lower spectrum-adjacent connections following the DHL policy, (b) approximate system $\widetilde{\text{DHL}}$ where the bottom and upper spectrum-adjacent connections use only their higher spectrum slots.

Approximating Assumption A1: the occupancy distribution for the total number of slots of the connections in the $\widetilde{\text{DHL}}$ policy is approximately equal to the corresponding occupancy distributions for the total number of slots (sum of lower and upper spectrum slots) of the connections under the DHL policy. That is,

$$\tilde{P}(\tilde{\boldsymbol{n}}) \approx \sum_{i=1,2,\ldots,d/2} \sum_{\tilde{n}_i = n_i^H + n_i^L} P(\boldsymbol{n}), \tag{8}$$

which means that $\tilde{n}_i \sim n_i^H + n_i^L$, where $\tilde{n}_i$ represents the total number of slots utilized by connection *i* in the $\widetilde{\text{DHL}}$ policy and "$\sim$" signifies that they follow approximately the same distribution.

Approximating Assumption A2: the variables $n_{B(p,l)}^L$ and $n_{U(p,l)}^L$, for all $l \in p$, follow approximately the same distribution ($n_{B(p,l)}^L \sim n_{U(p,l)}^L$) so that:

$$\begin{aligned} P(..., n_{B(p,l)}^H, n_{B(p,l)}^L, n_p^H, n_p^L, n_{U(p,l)}^H, n_{U(p,l)}^L, ...) \approx \\ P(..., n_{B(p,l)}^H, n_{U(p,l)}^L, n_p^H, n_p^L, n_{U(p,l)}^H, n_{U(p,l)}^L, ...). \end{aligned} \tag{9}$$

To visualize the above approximations, Fig. 3 displays an example of the spectrum utilization of a link *l*. Fig. 3a shows the actual case where the upper and bottom spectrum-adjacent connections of *p* follow the DHL policy, while Fig. 3b shows the approximate $\widetilde{\text{DHL}}$ system.

The available slot space $S(p,l)$ for link $l \in p$ in the DHL system is

$$S_{p,l} = F_{U(p,l)} - F_{B(p,l)} - 2 \cdot G - n_{B(p,l)}^H - n_{U(p,l)}^L, \tag{10}$$

while the corresponding approximate available slot space in

TABLE I
OUTGOING STATE TRANSITIONS FROM STATE $\boldsymbol{n}$ UNDER THE DHL POLICY

state $\boldsymbol{n}$:	$n_p^H = n_p^L = 0$, and Eq. (2) not met	$n_p^H > 0$, Eq. (2) not met, and $n_p^L > 0$	$n_p^H > 0$, Eq. (2) not met, and $n_p^L = 0$	$n_p^H = n_p^L = 0$, Eq. (2) met, and Eq. (3) not met	Eq. (2) met, $n_p^L > 0$, and Eq. (3) not met	$n_p^H > 0$, Eq. (2) met, $n_p^L = 0$, and Eq. (3) not met	$n_p^H = 0$, Eq. (2) met, $n_p^L = 0$, and Eq. (3) met	Eq. (2) met, $n_p^L > 0$, and Eq. (3) met	$n_p^H > 0$, Eq. (2) met, $n_p^L = 0$, and Eq. (3) met
Outgoing transition from state $\boldsymbol{n}$ with rate λ to state	$\boldsymbol{n} + \boldsymbol{e}_p^H$	$\boldsymbol{n} + \boldsymbol{e}_p^H$	$\boldsymbol{n} + \boldsymbol{e}_p^H$	$\boldsymbol{n} + \boldsymbol{e}_p^L$	$\boldsymbol{n} + \boldsymbol{e}_p^L$	$\boldsymbol{n} + \boldsymbol{e}_p^L$	Blocked	Blocked	Blocked
Outgoing transition from state $\boldsymbol{n}$ with rate $(n_p^H + n_p^L) \cdot \mu_p$ to state		$\boldsymbol{n} - \boldsymbol{e}_p^L$	$\boldsymbol{n} - \boldsymbol{e}_p^H$		$\boldsymbol{n} - \boldsymbol{e}_p^L$	$\boldsymbol{n} - \boldsymbol{e}_p^H$		$\boldsymbol{n} - \boldsymbol{e}_p^L$	$\boldsymbol{n} - \boldsymbol{e}_p^H$

the $\widetilde{\text{DHL}}$ system is

$$\begin{aligned}\tilde{S}_{p,l} = F_{U(p,l)} - F_{B(p,l)} - 2 \cdot G - \tilde{n}_{B(p,l)} \sim \\ F_{U(p,l)} - F_{B(p,l)} - 2 \cdot G - n^H_{B(p,l)} - n^L_{B(p,l)}\end{aligned} \tag{11}$$

Under assumptions A1 and A2, the available slot space of the approximate system follows the same probabilistic distribution with the available slot space in the actual system (that is $S_{p,l} \sim \tilde{S}_{p,l}$), since A2 states that the lower spectrum slots utilized by the bottom spectrum-adjacent connection follows the same distribution with the lower spectrum slots utilized by the upper adjacent connection ($n^L_{B(p,l)} \sim n^L_{U(p,l)}$), while A1 states that the total number of slots used by each connection in both systems is probabilistically the same ($\tilde{n}_p \sim n_p^H + n_p^L, \tilde{n}_{B(p,l)} \sim n^H_{B(p,l)} + n^L_{B(p,l)}, \tilde{n}_{U(p,l)} \sim n^H_{U(p,l)} + n^L_{U(p,l)}$).

The justification for these approximating assumptions is as follows. Under light load, very few slot requests are blocked in both the DHL and the $\widetilde{\text{DHL}}$ system, and the total number of slots (sum of lower and upper spectrum slots) used by each connection in both systems will be similar. Also, under light load the connections would mainly utilize their higher spectrum slots and very seldom utilize their lower spectrum slots. Connection p, on which we focus, follows the same (DHL) policy in both systems, exploiting both the higher and lower slots and will have approximately the same blocking probability under both systems. Connections other than p follow different sharing policies in the DHL and in the $\widetilde{\text{DHL}}$ system and the corresponding blocking probabilities in the two systems will be different (but low for light load), but this is not important since our focus is on connection p. Under heavy load, the majority of the connections will use almost completely their higher spectrum slots and will not find free lower spectrum slots. So in both cases the upper and bottom adjacent connections will utilize in a similar way their lower spectrum slots, which are the parameters of interest in our approximation. So, unless the spectrum allocation for the connections in the network is unbalanced (with some of them utilizing heavily their lower spectrum slots and some others lightly), the proposed approximation will give good results.

Under the proposed approximation (system $\widetilde{\text{DHL}}$), the connection p under study can utilize exclusively all the higher spectrum slots N_p^H up to the closest upper spectrum-adjacent connection [see Eq. (4)], and shares its lower spectrum slots with its bottom spectrum-adjacent connections. In the approximate $\widetilde{\text{DHL}}$ system, we can describe connection p with only a single parameter $\tilde{n}_p$, without distinguishing between its higher and lower spectrum slots. Thus, the transition rules of Table 1 are simplified. The total number of slots $\tilde{n}_p$ of p can be expressed by a birth-death process with birth rate λ_p and death rate μ_p, something that does not hold separately for the n_p^H and n_p^L variables in the actual DHL system.

We assume that connection p has b different bottom spectrum-adjacent connections, and, to simplify the notation, we denote by B_i, *i*=1,2,...,*b*, the *i*-th bottom spectrum-adjacent connection, dropping the notation of $B(p,l)$ that indexes the links of path p. Thus, we will denote by F_{B_i} the reference frequency slot, by $N^H_{B_i}$ the maximum number of slots that can be utilized, and by $\tilde{n}_{B_i}$ the number of slots utilized by the *i*-th bottom spectrum-adjacent connection. Since we assume that the bottom adjacent connections utilize only their higher spectrum slots in the approximate $\widetilde{\text{DHL}}$ system, $\tilde{n}_{B_i}$ follows a birth-death model, with corresponding parameters λ_{B_i} and μ_{B_i}, $i = 1, 2, ..., b$. The proposed approximation simplifies the interdependence among the connections. In particular, since connection p is affected only by its b bottom spectrum-adjacent connections, only b+1 dimensions of $\tilde{\boldsymbol{n}}$ play role in calculating the blocking of p in the approximate $\widetilde{\text{DHL}}$ system. The remaining dimensions (other connections) do not affect connection p or its bottom spectrum-adjacent connections and can thus be excluded from the blocking probability calculations. Fig. 4a presents the notation used for the b bottom adjacent connection of p and Fig. 4b shows an example of the slot utilization for the $\widetilde{\text{DHL}}$ system.

The set **A** of feasible states in the approximate $\widetilde{\text{DHL}}$ system is described by

$$\begin{aligned}\mathbf{A} : 0 \le \tilde{n}_{B_i} \le \min(N^H_{B_i}, F_p - F_{B_i} + N_p^H - \tilde{n}_p - G), i = \\ 1, 2, .., b,\ 0 \le \tilde{n}_p \le F_p + N_p^H - \max_i(F_{B_i} + \tilde{n}_{B_i}) - G.\end{aligned} \tag{12}$$

A new spectrum slot request for the connection p under study is blocked when the following holds:

$$\tilde{n}_p = F_p + N_p^H - \max_i(F_{B_i} + \tilde{n}_{B_i}) - G. \tag{13}$$

The approximate system $\widetilde{\text{DHL}}$ can be analyzed as a system that consists of b+1 independent queues, each described by a birth-death Markov chain [18]. For this system:

- Balance equations $\lambda_i \cdot \tilde{P}(\tilde{\boldsymbol{n}}) = (\tilde{n}_i + 1) \cdot \mu_i \cdot \tilde{P}(\tilde{\boldsymbol{n}} + \boldsymbol{e}_i)$ hold for all variables i in $\tilde{\boldsymbol{n}}$ and all pairs of adjacent states (reversible Markov chain).

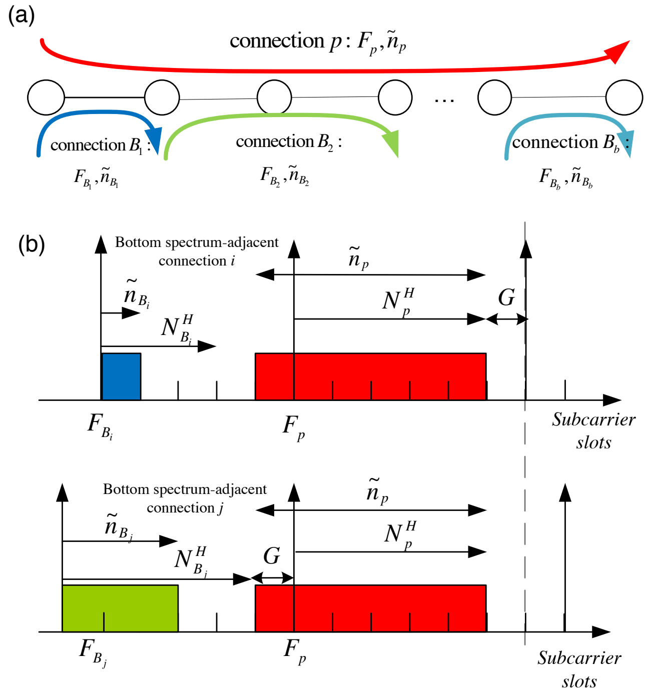

Fig. 4. (a) The approximate system $\widetilde{\text{DHL}}$ simplifies the interdependence among the connections. The connection p under study is affected only by its b bottom spectrum-adjacent connections, which are noted here by $B_i, i = 1, \ldots, b$, (b) An example of slot utilization in the $\widetilde{\text{DHL}}$ system. Connection p under study has exclusive use of its higher spectrum slots N_p^H and competes only with its bottom spectrum-adjacent connections for spectrum slots.

- The stationary distribution can be expressed in product form, $\tilde{P}_p(\tilde{\boldsymbol{n}}) = \tilde{P}_p(\tilde{n}_p) \cdot \tilde{P}_{B_1}(\tilde{n}_{B_1}) \cdot \ldots \cdot \tilde{P}_{B_b}(\tilde{n}_{B_b})$, where $\tilde{P}(\tilde{n}_p)$, and all $\tilde{P}_{B_i}(\tilde{n}_{B_i})$ depend only on the number $\tilde{n}_p$ and $\tilde{n}_{B_i}$ of the utilized slots.

Thus, according to [18], the stationary distribution of this truncated system has the product form

$$\tilde{P}(\tilde{\boldsymbol{n}}) = \frac{\frac{(\rho_p)^{\tilde{n}_p}}{\tilde{n}_p!} \cdot \prod_i \frac{(\rho_{B_i})^{\tilde{n}_{B_i}}}{\tilde{n}_{B_i}!}}{Q}, \tag{14}$$

$$\text{where } Q = \sum_{\tilde{n} \in \mathbf{A}} \frac{(\rho_p)^{\tilde{n}_p}}{\tilde{n}_p!} \cdot \prod_i \frac{(\rho_{B_i})^{\tilde{n}_{B_i}}}{\tilde{n}_{B_i}!}, \tag{15}$$

is a normalizing constant and $\mathbf{A}$ is the set of feasible states [see Eq. (12)]. The approximate blocking probability for p is

$$P_{bl}^{\widetilde{\text{DHL}}} = \sum_{\tilde{\boldsymbol{n}}:\text{ Eq. (13) is met}} \tilde{P}(\tilde{\boldsymbol{n}}). \tag{16}$$

The corresponding state probabilities $\tilde{P}(\tilde{\boldsymbol{n}})$ can be calculated easily using Eq. (14), even for large networks, since the state probabilities are written in product form. The approximate network blocking probability is then calculated by Eq. (6), using Eq. (16) for approximating the blocking probabilities of the connections.

3) Model for blocking computation for the DAD policy: We now analyze the performance of the DAD SEC policy described in Section II.B.3. The state transitions for connection p are given in Table 2, and can be used to define the corresponding d-dimensional Markov chain, compute the steady state probabilities from the global balance equations, and the blocking probabilities for the connections by Eq. (5). As with the DHL policy, dependencies between the state of connection p and that of all the other connections make the computation of the exact state probabilities intractable for realistic problem sizes. For this reason in the next subsection we present an approximate model for estimating the blocking probability of the DAD policy.

a) Approximate model for computing the blocking of the DAD policy for large networks

To calculate the approximate blocking of the DAD policy we follow an approach similar to that used in the approximation model $\widetilde{\text{DHL}}$ for the DHL policy. We focus on a particular connection p and we find the probability that a new spectrum slot request over p will be blocked under the DAD policy. For the given connection p, we define a hypothetical (approximate) system $\widehat{\text{DAD}}$ in which the connection p follows the DAD policy while its upper and bottom spectrum adjacent connections achieve a perfect symmetrical use of their spectrum. We denote the (d-1)-dimensional state of the approximate system $\widehat{\text{DAD}}$ by $\hat{\boldsymbol{n}} = (\hat{n}_1^H, \hat{n}_1^L, \ldots, \hat{n}_{B(p,l)}^H, \hat{n}_{B(p,l)}^L, \hat{n}_p, \hat{n}_{U(p,l)}^H, \hat{n}_{U(p,l)}^L, \ldots, \hat{n}_{d/2}^H, \hat{n}_{d/2}^L)$. We make the following approximating assumptions for calculating the slot blocking probability of p:

Approximating Assumption B1: the same as the approximating assumption A1 for connection p, that is, $\hat{n_p} \sim n_p^H + n_p^L$ so that slot utilization of p can be expressed with a single variable.

Approximating Assumption B2: the variables $n_{B(p,l)}^L$ and $n_{B(p,l)}^H$ are independent and each follows a birth-death model with arrival rate $\lambda_{B(p,l)}/2$ and death rate $\mu_{B(p,l)}$, and so do $n_{U(p,l)}^L$ and $n_{U(p,l)}^H$, for all $l \in p$.

The justification for approximating assumption B2 is as follows. Under light load, very few connections are blocked in both the DAD and the $\widehat{\text{DAD}}$ system, and the connections expand their spectrum symmetrically towards their higher and lower slots. Splitting the traffic of each connection and serving alternately its slot requests with its higher and lower slots, is well approximated at light load by having two birth-death processes with arrival rates equal to half the total arrival rate of the original process. Under heavy load, the majority of the connections will be limited by both their upper and lower spectrum-adjacent connections. If the spectrum allocation for the connections in the network is balanced and connections have symmetrical distance from their adjacent connections they will utilize their spectrum symmetrically, which depends on the RSA solution. So, for light and heavy load the upper and bottom adjacent connections of p will utilize in a symmetrical way their higher and lower spectrum slots, which are the parameters of interest in our approximation, and the proposed approximation will give good results.

We let b and u be the number of different bottom and upper spectrum-adjacent connections of p, respectively. Under our approximation assumptions, the bottom adjacent connections of p utilize their higher and lower spectrum slots independently and these are expressed by a birth-death process with corresponding parameters $\lambda/2$ and μ. The same holds for the upper spectrum-adjacent connections. The proposed

TABLE II
OUTGOING STATE TRANSITIONS FROM STATE $\boldsymbol{n}$ UNDER THE DAD POLICY

state $\boldsymbol{n}$:	$n_p^H = n_p^L = 0$, and Eq. (2) not met	$n_p^H = n_p^L > 0$, and Eq. (2) not met	$n_p^H > n_p^L > 0$, Eq. (2) not met, and Eq. (3) met	$n_p^H = n_p^L = 0$, Eq. (2) met, and Eq. (3) not met	$n_p^H = n_p^L > 0$, Eq. (2) met, and Eq. (3) not met	$n_p^H > n_p^L > 0$, Eq. (2) met, and Eq. (3) not met	$n_p^H = n_p^L = 0$, Eq. (2) met, and Eq. (3) met	$n_p^H = n_p^L > 0$, Eq. (2) met, and Eq. (3) met	$n_p^H > n_p^L$, Eq. (2) met, and Eq. (3) met
Outgoing transition from state $\boldsymbol{n}$ with rate λ to state	$\boldsymbol{n} + \boldsymbol{e}_p^H$	$\boldsymbol{n} + \boldsymbol{e}_p^H$	$\boldsymbol{n} + \boldsymbol{e}_p^H$	$\boldsymbol{n} + \boldsymbol{e}_p^L$	$\boldsymbol{n} + \boldsymbol{e}_p^L$	$\boldsymbol{n} + \boldsymbol{e}_p^L$	Blocked	Blocked	Blocked
Outgoing transition from state $\boldsymbol{n}$ with rate $(n_p^H + n_p^L) \cdot \mu_p$ to state		$\boldsymbol{n} - \boldsymbol{e}_p^L$	$\boldsymbol{n} - \boldsymbol{e}_p^H$		$\boldsymbol{n} - \boldsymbol{e}_p^L$	$\boldsymbol{n} - \boldsymbol{e}_p^H$		$\boldsymbol{n} - \boldsymbol{e}_p^L$	$\boldsymbol{n} - \boldsymbol{e}_p^H$

approximation removes the interdependence among the connections, since connection p is affected only by the higher slot utilization of its b bottom and the lower slot utilization of its u upper spectrum-adjacent connections, but these are independent from all the other connections. Thus, only $b+u+1$ dimensions of $\hat{\boldsymbol{n}}$ play role in calculating the blocking of p in the approximate $\widehat{\text{DAD}}$ system. Following a similar analysis with the one used previously for the $\widehat{\text{DHL}}$ approximate system (not presented here due to length limitations), the $\widehat{\text{DAD}}$ system consists of $b+u+1$ independent queues, each described by a birth-death Markov chain. So we can express its stationary distribution in product form and obtain steady state probability expressions for system $\widehat{\text{DAD}}$ similar to Eq. (14)-(15). Then we can calculate the corresponding state probabilities $\hat{P}(\hat{\boldsymbol{n}})$, similar to Eq. (16), and use it to calculate the approximate network blocking probability by Eq. (6).

4) Analytical blocking models complexity: The model of the CSA policy involves the application of Erlang-B formula for each connection, which takes little computational effort. The hardest calculation in the approximation model of the DHL policy for a connection is in the calculation of Q in Eq. (15), the denominator of Eq. (14). If b is the number of different bottom spectrum-adjacent connections of p, the number of required calculations may increase as fast as exponentially with b [16],[17]. Since b can be as high as the longest path used in the network, for realistic transport networks that are designed with mesh topologies and typically have bounded hop paths this calculations can be done very fast, as the running times reported in our experiments indicate. For example, for random graphs of average connectivity degree k the diameter of the network (longest distance between two nodes) increases as $log_k|\boldsymbol{V}|$ [21], and assuming that connections are routed over shortest paths, the calculation of the blocking performance is polynomial to the number of network nodes $|\boldsymbol{V}|$. Even if we were able to obtain the exact model for the DHL policy in product form, its calculation would be exponential to the number of nodes $|\boldsymbol{V}|$ (since there is interdependence among all connections), which would make it infeasible even for medium-sized networks [16]. Similarly, for the DAD policy the calculations involved in the approximated model are exponential to *b+u*. The proposed approximations result in product form expressions for the blocking probability and also reduce the size of the system; this approach is similar to the conditional decomposition technique used in [17].

IV. ROUTING AND SPECTRUM ALLOCATION

In the dynamic scenario with time-varying traffic considered in this study, the connections expand/contract their utilized spectrum so as to follow the traffic variations, in the way determined by the SEC policy used. As explained earlier, network performance does not only depend on the SEC policy, but also on the Routing and Spectrum Allocation (RSA) algorithm used. The role of the RSA under time-varying traffic is to assign the routes and reference frequencies to the connections within the available *T* slots so as to minimize the average blocking of the network.

To solve the dynamic (time-varying) RSA problem we transform it into a *static* RSA problem, solve the static problem, apply the blocking models presented above to calculate the network blocking under dynamic traffic of this RSA solution, and iteratively search for solutions with better blocking performance. The term static is used here to refer to the problem that takes as input a traffic matrix with specific number of required slots for the connections, as if the traffic did not change dynamically with time.

Consider the case where we want to establish a set of new time-varying connections (the case of a single connection is a subcase). Previously established connections can be left with the same starting frequencies or be reconfigured, according to our policy. To formulate the related static problem, we initially assume that the CSA policy is used. The blocking performance of an RSA solution for the used SEC policy will always be better or equal to that of the CSA policy, since the latter is the simplest policy and does not permit the sharing of spectrum slots among connections. We assume we are given a blocking threshold *B* that is considered acceptable (e.g., $B = 10^{-6}$). We use Eq. (7) to calculate for each connection p in the network the number N_p of spectrum slots for which the CSA blocking is acceptable, that is, $P_{bl}^{CSA}(p)(\rho_p, N_p) < B$. We use the computed set of N_p values for all connections, as the traffic matrix in a static Routing and Spectrum Allocation (RSA) algorithm to find paths and reference frequencies slot for the connections. We denote by $T* = \max_p(F_p + N_p)$ the highest slot allocated to a connection by the static algorithm. If the system can support *T** subcarriers slots $(T* < T)$, the algorithm finishes and we have found an acceptable solution with the CSA policy that does not require spectrum sharing at all (any policy that enables sharing will exhibit better performance). Note that finding a static RSA solution within

T slots is NP-hard, since it is the decision problem of the corresponding minimization problem proven to be NP-hard in [14]. In this study we use the heuristic algorithm based on Simulated Annealing of [14] to obtain the static RSA solutions. The values N_p used as input to this static RSA algorithm correspond to the minimum distance of the reference frequency of connection p from its upper spectrum-adjacent connections [N_p^H in Eq. (4)]. If the RSA algorithm does not find a solution within T slots, we iteratively increase the acceptable blocking threshold B and use Eq. (7) to obtain new values for the number N_p of slots required by each p. A solution is acceptable if it utilizes fewer than the T slots supported by the system. After obtaining an acceptable static RSA solution we take into account the specific SEC policy used and apply the corresponding blocking model to calculate the average blocking of that RSA solution. In particular, for the CSA policy we use Eq. (7), for the DHL policy we use the approximation of Eq. (16) and for the DAD policy the corresponding equations. Then we apply Eq. (6) to obtain the network blocking probability. However, we do not stop the first time we find an acceptable solution within T slots, but we search for different static RSA solutions with the same numbers of required slots (using e.g. Simulated Annealing) or keep decreasing the number of required slots until we find, say, K solutions that are acceptable. We select from the K solutions, the one with the lowest network blocking probability. Fig. 5 presents the flow diagram of the proposed RSA algorithm under time-varying traffic.

It is clear that the problem of finding the RSA solution that minimizes network blocking is very complicated. In the general case, the network blocking depends on the paths, the reference frequencies and ordering of the connections in the spectrum domain, on the SEC policy, and on the traffic parameters. The proposed algorithm solves the RSA problem under time-varying traffic *indirectly*. It solves a related static problem considering constant rate connections, or equivalently, the use of the basic CSA policy that allocates a constant number of slots to each connection, which is not accurate for any dynamic sharing policy but gives a rough estimation (lower bound) of the blocking. Then it applies the blocking model developed for the particular SEC policy used to estimate the performance under dynamic traffic and searches for lower blocking solutions by iterating the above process.

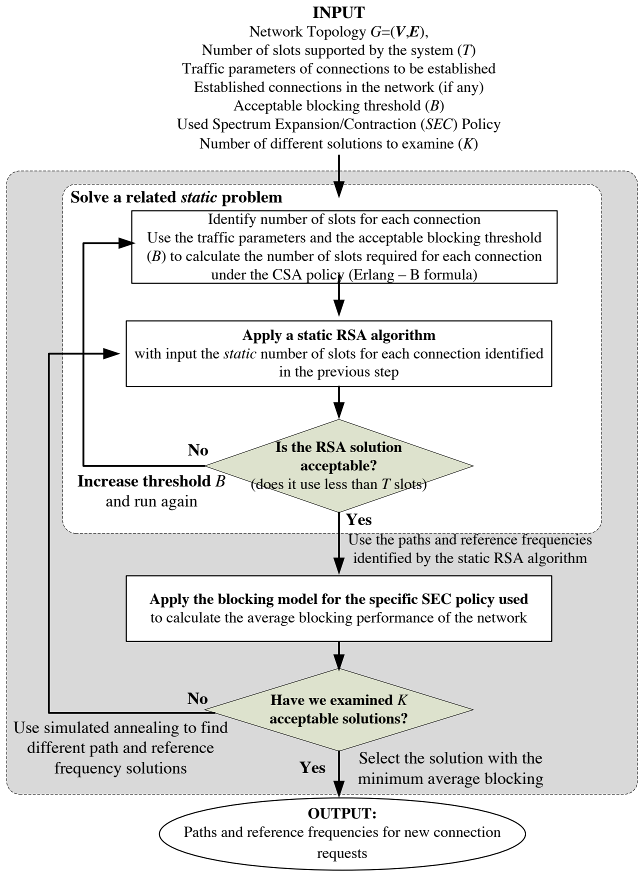

Fig. 5. Block diagram of the proposed RSA algorithm to assign paths and reference frequencies to a set of time-varying connection requests.

V. Performance results

In this section we present performance evaluation results for serving traffic with time-varying rates in a spectrum-flexible network under the proposed framework, for the SEC policies presented in Section II.B and the RSA algorithm described in Section IV. We also study the accuracy of the developed analytical blocking models (Section III) by comparing them against corresponding simulation results.

We performed experiments using a realistic topology based on the 14-node generic Deutsche Telekom (DT) network (Fig. 6). We assumed that communication is performed among all source-destination pairs in the network. Spectrum slot requests for each source-destination pair p are generated according to a Poisson process of rate λ_p and their duration is exponentially distributed with mean $1/\mu_p = 1$. The arrival rate λ_p for the slot requests of each connection p is drawn from an exponential distribution with mean λ. In these experiments we assumed that we are given the traffic parameters for all source-destination pairs and we establish the related connections at start time. Thus, $\lambda \cdot |\boldsymbol{V}| \cdot (|\boldsymbol{V}| - 1)$ is the total average network load in Erlangs, where $|\boldsymbol{V}|$ is the number of network nodes. We are also given the number T of spectrum slots supported by the system, which remains fixed during the execution of each experiment instance, and we measure the blocking performance of the connections slot requests. The results do not depend on the slot size C chosen, as the traffic requests are also expressed in slots.

We graph the blocking performance of the CSA, the DHL, and the DAD policies as calculated using the developed analytical models (exact for the CSA policy, and approximate for the DHL and DAD policies). For the same traffic scenarios, we conducted full network simulation experiments and we also graph the corresponding blocking probability returned by the simulations for 10^7 slot requests. For comparison purposes, we also present the blocking performance of a network that does not follow the framework and SEC policies presented here, but supports the full sharing of all spectrum slots among the connections. This type of network can be viewed as a typical WDM network with spectrum slots corresponding to wavelengths, with the additional constraint of having to use spectrum guardbands between spectrum-adjacent connections. This reduces to a WDM network with $T/(1+G)$ wavelengths.

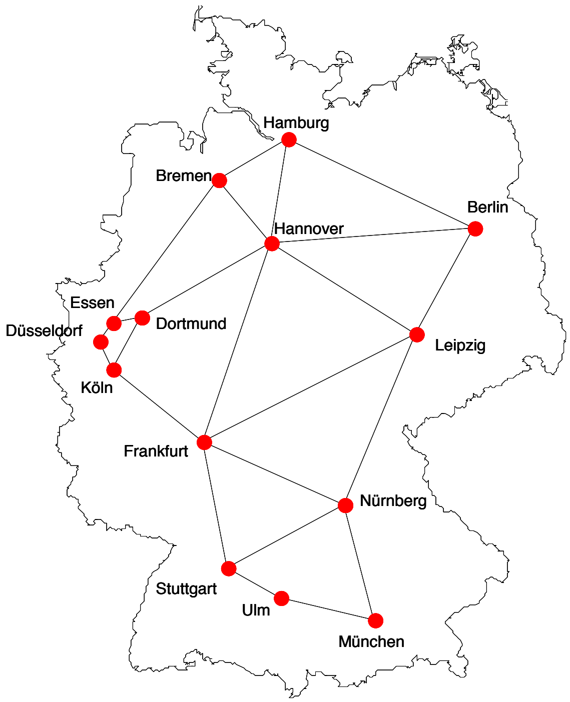

Fig. 6. The generic DT network topology, with 14 nodes and 23 undirected links used in the performance experiments.

We used the analytical models developed in [20] to compute the blocking of such a WDM network under the fixed alternate routing and random wavelength allocation policy, and we also performed the related simulation experiments. We have set G=1 slot, which corresponds to the minimum guardband requirement; the performance of the WDM network for higher values of G is expected to be worse.

In Fig. 7a we graph the network blocking performance as a function of the total load (in Erlangs) assuming T=250 spectrum slots are available in the network. We observe that the proposed analytical models for calculating the blocking probabilities of the CSA, DHL and DAD policies are in very close agreement with the corresponding simulation results. The blocking model for the CSA policy, which reduces to the Erlang-B formula, does not involve any approximating assumptions, and thus we expected to have such a good accuracy. The simulation experiments show that the analytical model developed for the DHL policy, using the approximating assumptions A1 and A2 is very accurate. In the case of the DAD policy, the approximate model slightly underestimates the average blocking probability for heavy load, and slightly overestimates it for light load. Under the DAD policy, connections tend to utilize their spectrum in a symmetric way, but connections that are instantaneously loaded use more spectrum and push their adjacent connections to utilize spectrum in the opposite directions. At light load this does not increase the blocking of the remaining connections. So, at light load, a connection can access more spectrum with smaller effects than if its adjacent connections utilized their spectrum symmetrically, and thus the approximate model calculates slightly higher average blocking than the actual one. At heavy load, connections frequently utilize more spectrum than they were assigned by the RSA algorithm, and spectrum usage around the central frequency becomes non-symmetric relatively often, leading to less efficient utilization than the ideal case assumed by the approximation model. However, under both the heavy and the light load cases, the calculated performance is very close to that obtained by the simulations, which indicates that the approximation model is quite accurate.

The calculations of the average network blocking performance (for all connections) for a given RSA solution for the DHL policy took around 10 sec, and 30 sec for the DAD policy, while the related simulations required around 15 minutes. Our simulation experiments verify that our approximating analysis for the DHL and DAD policies are able to calculate the connection and network blocking probabilities in an accurate and quick manner.

The network blocking probability of the DAD and DHL policies is lower than that obtained for the CSA policy by more than one order of magnitude, in most cases. This is the gain that we obtain by enabling spectrum slot sharing among spectrum-adjacent connections, as done with these dynamic policies. The DAD policy achieves better performance than the DHL policy. In the DHL policy each connection utilizes more heavily its higher spectrum slots, and the statistical multiplexing gains are achieved by sharing the spectrum mainly with its bottom adjacent connections. The DAD policy is more symmetric and connections share spectrum more efficiently with both their upper and bottom adjacent connections achieving higher multiplexing gains. For example, using the approximation blocking models we can find that for a connection it is more efficient to share the same amount of spectrum under the DAD policy with an upper and a bottom spectrum-adjacent connection that have half the arrival rates than under the DHL policy with a single bottom adjacent connection at full rate. The path length has both positive and negative effects in the blocking performance. The longer paths deteriorate the spectrum allocation process and result in higher fragmentation of spectrum (similar to the effect of the wavelength continuity constraint in WDM systems). The dynamic spectrum sharing performed by the DHL and DAD policies, resolves partly the defragmentation problem and improves the performance of the system. However, the longer the path becomes, the more connections compete for the same (spectrum) resources, and after a point efficient spectrum sharing becomes difficult. So the placement of adjacent connections in the network (role of RSA algorithm) becomes very important on networks with longer paths. The performance of the WDM network and the corresponding RWA algorithms utilizing $T/2$ wavelengths (remember that G=1) is worse than the solutions that follow the proposed framework. The used analytical model for calculating the RWA blocking [20] is very accurate for high traffic loads, but its accuracy deteriorates slightly for lower load values.

Fig. 7b presents the performance of the network as a function of the number of spectrum slots T supported, assuming a total network load of 1000 Erlangs. We observe that the proposed analytical models for calculating the blocking performance of the CSA, DHL and DAD policies are quite accurate. Again we observe that the performance of the DHL and DAD policies is superior to that of CSA, which shows that

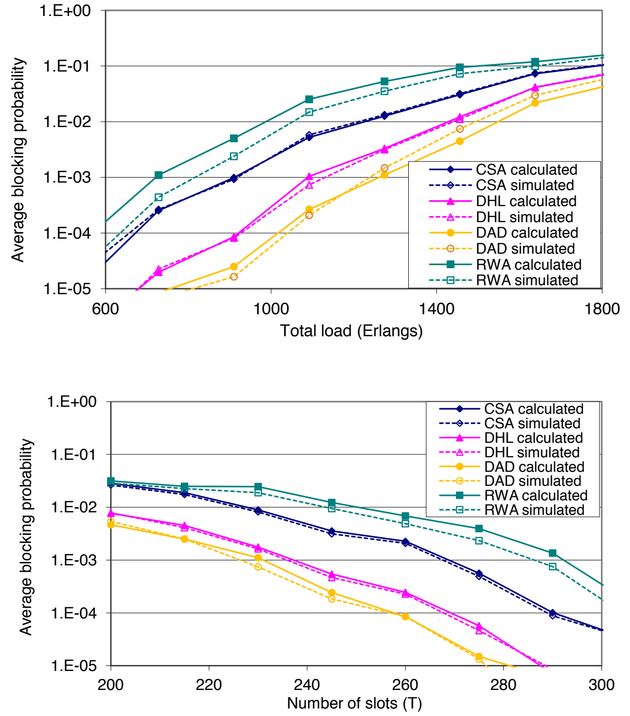

Fig. 7. Average blocking probability for the DT network as a function of (a) the total load, assuming T=250 slots, and (b) the number of spectrum slots T, assuming 1000 Erlangs total load.

important performance gains can be obtained though spectrum sharing among connections.

VI. Conclusion and Future Directions

We considered the problem of serving dynamic traffic in a spectrum-flexible optical network, where the spectrum allocated to a connection varies so as to follow the time-varying required transmission rate. We presented a general framework for serving dynamic traffic in such a network that assigns to each connection a route and a reference frequency. The connection is allowed to expand and contract the spectrum that it utilizes around this reference frequency. We proposed three spectrum expansion/contraction policies, the CSA policy, which exclusively allocates a number of slots to a connection, and the DHL and DAD policies, which enable the dynamic sharing of spectrum slots among spectrum-adjacent connections. We analyzed the network blocking performance under the proposed framework, obtaining exact formulas for the CSA policy and accurate analytical approximations for the DHL and DAD policies. We presented a simple iterative RSA algorithm and applied the proposed analytical models to calculate the dynamic network blocking performance. Our performance results showed that the proposed analytical models are in close agreement with corresponding simulation results. Our proposed framework for serving time-varying traffic enables the dynamic spectrum sharing and is shown to achieve significant statistical multiplexing gains, reducing the blocking probability and increasing the efficiency of the system. The DAD policy achieved the best blocking performance since it enables better sharing of spectrum and higher efficiency.

Future directions include the development of more sophisticated SEC policies that take into account the correlations among the connections' traffic variations (provided they are known), RSA algorithms that will directly incorporate the policy blocking models in their formulation, and also methods to reroute connections and defragment the spectrum so as to further improve the performance of the spectrum-flexible network under dynamic traffic.

References

[1] Cisco white paper, "Cisco Visual Networking Index: Forecast and Methodology, 2010-2015"
[2] C. Glingener, "Optical Networking Trends and Evolution", invited, OFC 2011
[3] O. Gerstel, M. Jinno, A. Lord, S. J. Ben Yoo, "Elastic optical networking: a new dawn for the optical layer?", IEEE Com. Mag., 50(2), 2012
[4] P. P. Marino, F. Neri, "On the Myths of Optical Burst Switching", IEEE Trans. on Com., 2011
[5] J. Gripp, J. E. Simsarian, J. D. LeGrange, P. G. Bernasconi, and D. T. Neilson, "Architectures, components, and subsystems for future optical packet switches," IEEE J. Sel. Topics Quantum Electron., 16(5), 2010
[6] M. Jinno, H. Takara, B. Kozicki, Y. Tsukishima, Y. Sone, Matsuoka, "Spectrum-efficient and scalable elastic optical path network: architecture, benefits, and enabling technologies" IEEE Com. Mag., 47(11), 2009
[7] A. Patel, P. Ji, J. P. Jue, T. Wang, "Routing, Wavelength Assignment, and Spectrum Allocation in Transparent Flexible Optical WDM (FWDM) Networks", Photonics in Switching, PDPWGI, 2010
[8] T. Takagi, H. Hasegawa, K. Sato, Y. Sone, B. Kozicki, Hirano, M. Jinno, "Dynamic Routing and Frequency Slot Assignment for Elastic Path Networks that Adopt Distance Adaptive Modulation", OFC 2011
[9] A. Castro, L. Velasco, M. Ruiz, M. Klinkowski, J. P. Fernndez-Palacios, D. Careglio, "Dynamic Routing and Spectrum (Re)Allocation in Future Flexgrid Optical Networks", Computers Networks, 56, 2012.
[10] A. Patel, P. Ji, J. P. Jue, T. Wang, "Defragmentation of Transparent Flexible Optical WDM (FWDM) Networks", OFC 2011
[11] F. Cugini, M. Secondini, N. Sambo, G. Bottari, G. Bruno, P. Iovanna, P. Castoldi, "Push-Pull Technique for Defragmentation in Flexible Optical Neworks", OFC 2012
[12] S. Gringeri, B. Basch, V. Shukla, R. Egorov, T. Xia, "Flexible Architectures for Optical Transport Nodes and Networks", IEEE Com. Mag., 48(7), 2010
[13] G. Shen, Q. Yang, S. You, W. Shao, "Maximizing time-dependent spectrum sharing between neighboring channels in CO-OFDM optical networks", ICTON 2011
[14] K. Christodoulopoulos, I. Tomkos, E. Varvarigos, "Elastic Bandwidth Allocation in Flexible OFDM-based Optical Networks", IEEE/OSA J. Ligh. Tech. , 29(9), 2011
[15] K. Christodoulopoulos, I. Tomkos, E. Varvarigos, "Dynamic Bandwidth Allocation in Flexible OFDM-based Networks", OFC, 2011
[16] F.P. Kelly, "Blocking Probabilities in Large Circuit Switched Networks," Advances in Applied Probability, 18, pp. 473-505, 1986
[17] G. L. Choudhurym, K. K. Leung, W. Whitt, "An Algorithm to Calculate Blocking Probabilities in Multi-Rate Multi-Class Multi-Resource Loss Models", Advances in Applied Probability, vol. 27, 1995
[18] D. Bertsekas, R. Gallager, "Data Networks", 2^{nd} ed., Prentice Hall, 1992
[19] A. Birman, "Computing approximate blocking probabilities for a class of all-optical networks", IEEE J. Selected Areas in Communication, 1996
[20] A. Sridharan, K. Sivarajan, "Blocking in All-Optical Networks", IEEE/ACM Trans. on Netw., 12 (2), 2004
[21] R. Albert, A. L. Barabasi. "Statistical mechanics of complex networks", Reviews of Modern Physics, 2002

Konstantinos (Kostas) Christodoulopoulos is a post-doc researcher at the School of Computer Science and Statistics of Trinity College Dublin, working on a project co-funded by IBM Ireland. He received a Diploma in Electrical and Computer Engineering from the National Technical University of Athens, Greece, in 2002, the M.Sc in Advance Computing from Imperial College of London, U.K., in 2004, and the Ph.D degree from the Department of Computer Engineering and Informatics, University of Patras, Greece, in 2009. He then worked for one year in the same department as an adjunct professor. His research interests are in the areas of algorithms and protocols for optical networks, interconnection networks, distributed and parallel computing.

Ioannis Tomkos is with AIT since Sep 2002. In the past, he was a senior scientist (1999 - 2002) at Corning Inc. USA and research fellow (1995 - 1999) at University of Athens, Greece. At AIT he founded the "High Speed Networks and Optical Communication (NOC)" Research Group that was/is involved with leading role in many EU funded research projects (including 8 currently active). Dr. Tomkos has received in 2007 the prestigious title of "Distinguished Lecturer" of IEEE Communications Society. Together with his colleagues and students he has co-authored about 450 peer-reviewed journal/conferences articles. He has served as the Chair of the International Optical Networking Technical Committee of IEEE Communications Society (2007-2008) and the Chairman of the IFIP working group on "Photonic Networking" (2008-2009). He is currently the Chairman of the OSA Technical Group on Optical Communications (2009-2012) and the Chairman of the IEEE Photonics Society Greek Chapter (2010-2012). He has been Chair or/and member of various committees for the major conferences (e.g. OFC, ECOC, GlobeCom, ICC, ICTON, ONDM, BroadNets, etc.) in the area of telecommunications/networking (more than 100 conferences/workshops). In addition, he is a member of the Editorial Boards of the IEEE/OSA Journal of Lightwave Technology, the IEEE/OSA Journal of Optical Communications and Networking and the IET Journal on Optoelectronics. Among many guest editorials for special issues in journals, he was the Chief Editor for a 2012 special issue on "The evolution of Optical Networking" for the prestigious "Proceedings of the IEEE".

Emmanouel (Manos) Varvarigos received a Diploma in ECE from the National Technical University of Athens in 1988, and the M.S. and Ph.D. degrees in EECS from the Massachusetts Institute of Technology in 1990 and 1992, respectively. He has held faculty positions at the University of California, Santa Barbara and Delft University of Technology. In 2000 he became a Professor of Computer Engineering and Informatics (CEID) at the University of Patras, Greece, where he founded the Communication Networks Lab. He is currently the Director of the Hardware, Communications and Networking Division of CEID and the Scientific Director of the Greek School Network Sector at the Computer Technology Institute (CTI). He has participated in more than 30 US- and EU-funded research projects in optical networking and grid computing and in many national projects. He has over 200 publications in international journals and conferences. His research activities are in the areas of optical networking, grid and cloud computing, and wireless ad hoc networks.