Power Control for OSNR Optimization in Optical Networks: A Distributed Algorithm via A Central Cost Approach

Lacra Pavel
Department of Electrical and Computer Engineering
University of Toronto, Toronto, ON, Canada, M5S 3G4
Email: pavel@control.toronto.edu

Abstract— **This paper addresses the problem of optical signal-to-noise ratio (OSNR) optimization in optical networks. An analytical OSNR network model is developed for a general multi-link configuration, that includes the contribution of amplified spontaneous emission and crosstalk accumulation. An network OSNR optimization problem is formulated such that all channels maintain a desired individual OSNR level, while input optical power is minimized. An iterative, distributed algorithm for channel power control is proposed, which is shown to converge geometrically to the optimal solution. Convergence is proved for both synchronous and asynchronous operation.**

I. Introduction

Optical wavelength-division multiplexed (WDM) communication networks are evolving beyond the statically designed, point-to-point links. The goal is to realize reconfigurable networks, with arbitrary topologies, while at the same time maintaining network stability, optimal channel transmission performance and quality of service (QoS), [1]-[4].

At the physical transmission level, one parameter that directly determines channel performance and QoS is the bit-error rate (BER). BER in turn, depends on optical signal-to-noise ratio (OSNR), dispersion and nonlinear effects, [5]. Typically in link optimization, OSNR is considered as the dominant performance parameter, with dispersion and nonlinearity effects being kept low by proper link design. Therefore, OSNR optimization can be directly translated to QoS optimization. The dominant impairment in OSNR optimization is given by noise accumulation in chains of optical amplifiers and its effect on OSNR, [6]. A typical approach for OSNR optimization uses a static budget of device and fiber impairments along a link, with significant tolerance margins added such that at least a desired OSNR value is achieved. In [7], a heuristic algorithm was proposed for on-line OSNR equalization in single point-to-point links, but its convergence was not considered. In [6], OSNR equalization in a point-to-point link was formulated as a static optimization problem. The optimal channel optical power vector at the transmitter is found by solving an eigenvalue problem for the system transmission matrix, composed of link gains of all channels. The algorithm which is developed only for a single point-to-point link, is static and requires centralized, global link information.

The current OSNR optimization approaches, that are developed for single links, are not appropriate for reconfigurable optical networks, where the length of the links and the number of devices in a link are changing. Moreover, in these networks arbitrary topologies can be formed, different channels can travel via different optical paths, and also different channels can have different levels of QoS (OSNR) requirements. Therefore, it is desirable to adjust network parameters (optical power, gains) in an optimal way, based on on-line network feedback (from receiver to transmitter and various nodes). This dynamic optimization will result in increased network flexibility and capacity, so that the needed quality of service (QoS) is assured whenever a new channel is added/dropped.

These observations justify the need for on-line OSNR optimization algorithms, that have provable convergence properties for general network configurations. Moreover, particularly useful are decentralized algorithms, such that channel power at the transmitter (Tx), can be adjusted based on feedback from the corresponding receiver (Rx) only, plus other channel specific measurements. This is the problem we address in this paper: OSNR optimization in optical networks and development of iterative power control algorithms.

This problem is similar to power control in wireless communication systems, a topic which has been explored extensively, via either centralized approaches, [8], [9], or decentralized, noncooperative game approaches, [19],

[21]. There are several differences that make this dynamic network optimization problem more challenging in optical networks. In wireless networks, channels are characterized by static loss/gain only, with single Tx to Rx links, or multiple links with no gains (ATM networks). Optical networks bring new challenges: amplified spans, multiple links, accumulation and self-generation of optical (ASE) noise, as well as crosstalk generated at the routing elements. In [22] we used a game theoretic approach to address this problem, based on an OSNR model that considers with only ASE accumulation.

We consider this problem of developing algorithms for optimally re-adjusting network parameters such that all channel/routes maintain a desired QoS (OSNR), while at the same time taking into account specific constraints as imposed by dispersion and nonlinearity.

The paper is organized as follows. In Section 2 we develop the analytical model for OSNR in a generic multi-link configuration, both as recursive relations and end-to-end relations. In Section 3 we formulate the static OSNR optimization problem for a large-scale network, and characterize the optimal solution. In Section 4 we develop an iterative algorithm for adjusting power levels and we give conditions for its convergence to the optimal solution. The proposed algorithm is valid for general network configurations, and uses only local measurements or decentralized feedback. In Section 5 we prove that the algorithm is convergent also in the asynchronous case, which is particularly important for adaptation in a dynamic environment. A numerical example is given in Section 5 and conclusions are given in Section 6.

II. Network OSNR Model

Consider a generic optical network configuration (Figure 1), with a set of optical links $\mathcal{L} = \{1, ..., L\}$ connecting the optical nodes (for channel add/drop (OADM) or cross-connect (OXC)). A link l is composed of N_l equal length, cascaded optically amplified spans. A total set of channels, $\mathcal{M} = \{1, ..., m\}$, corresponding to a set of wavelengths, are transmitted across the network. This is realized by intensity modulation of the wavelengths, and then by multiplexing the wavelengths to be transmitted across the same link. We denote by $\mathcal{M}_l$ the set of channels transmitted over link l, $l \in \mathcal{L}$.

For a channel $i \in \mathcal{M}$, corresponding to wavelength λ_i, we denote by $\mathcal{R}_i$ its optical path, or collection of links, from source (Tx) to destination (Rx). For the i^{th} channel, we denote by u_i, s_i, and n_i the optical power of the input signal (Tx), output signal (Rx) and output noise (Rx), in the i^{th} channel bandwidth. We also denote by $\mathbf{u} = [u_1, \ldots, u_m]^T$ the vector of input powers for all channels.

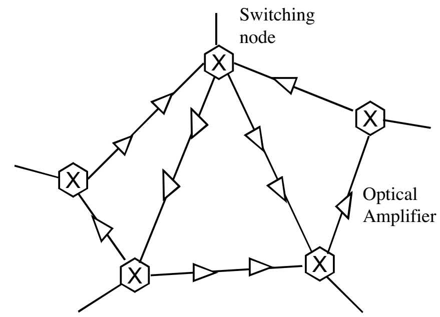

Fig. 1. Optical network configuration

In the following we develop a network OSNR model for a generic multi-link configuration, following the approach in [6]. As in [6], [15], dispersion and nonlinearity effects are considered to be limited, and the dominant impairment is amplified spontaneous emission (ASE) noise in amplifiers. Optical amplifiers are used to amplify the power of all channels in a link simultaneously. Because of their wavelength-dependent gain, ASE noise is also wavelength-dependent, and different channels will have different OSNR levels.

An optically amplified span is composed of optical fiber and an optical amplifier. For the k^{th} span, of the l^{th} link, the optical span transmission, $h_{l,k,i}$, for channel i is

$$h_{l,k,i} = G_{l,k,i} \, L_{l,k} \qquad \forall k = 1, \ldots, N_l \tag{1}$$

where $G_{l,k,i}$ is the amplifier gain (typically wavelength/ channel dependent), and $L_{l,k}$ is optical fiber loss (typically channel independent). For channel i, we denote by $p_{l,k,i}$ and $v_{l,k,i}$, the signal and the noise power, respectively, at the output of the k^{th} span on the l^{th} link.

We use the following assumptions.

(A.i.1): ASE noise power does not participate in amplifier gain saturation.

(A.i.2) All amplifiers in a link have the same spectral shape and are operated in automatic power control (APC), with the same total power target.

ASE noise power generated in the k^{th} amplifier on the l^{th} link, for the i^{th} channel, is

$$ASE_{l,k,i} = 2 n_{sp} \, [G_{l,k,i} - 1] \, h \nu_i \, B_o \tag{2}$$

where $n_{sp} > 1$ is the amplifier excess noise factor, h is the Planck constant, B_o is the optical bandwidth, and ν_i is the optical frequency corresponding to wavelength λ_i.

APC in (A.i.2) is the typical operation mode for amplifiers when the network is in steady-state. This ensures equal launching power into each span and compensates variations in fiber-span loss across a link, [6]. The total power target $P_{0,l}$, is selected to be bellow the threshold for nonlinear effects, [15]. Then at the output of each amplifier and at the input of the following span we have

$$\sum_{j \in \mathcal{M}_l} p_{l,k,j} = P_{0,l} \qquad \forall\, l \in \mathcal{R}_i \quad \forall k = 1, \ldots, N_l \tag{3}$$

The optical signal-to-noise ratio (OSNR) for the i^{th} channel, $i \in \mathcal{M}$, is defined as

$$OSNR_i = \frac{s_i}{n_i} \tag{4}$$

where s_i and n_i are the channel's signal, and noise power at the output (at Rx), respectively. We consider only forward propagation of signal and noise in steady-state, [6].

The following lemma generalizes the single link result in [6] to an arbitrary multi-link network.

Lemma 1: Let u_i, and $n_{0,i}$ be the i^{th} channel's signal, and noise power at Tx, respectively. Under (A.i.1)-(A.i.2) the i^{th} channel OSNR, along a path $\mathcal{R}_i$, is given as

$$OSNR_i = \frac{u_i}{n_{0,i} + \sum_{j \in \mathcal{M}} \Gamma_{i,j}\, u_j} \qquad \forall\, i \in \mathcal{M} \tag{5}$$

where the full (m x m) system matrix $\mathbf{\Gamma}$ is defined as $\mathbf{\Gamma} = [\Gamma_{i,j}]$ with

$$\Gamma_{i,j} = \sum_{l \in \mathcal{R}_i} \sum_{k=1}^{N_l} \frac{G_{l,j}^k}{G_{l,i}^k} \left(\prod_{q=1}^{l-1} \frac{\mathbf{T}_{q,j}}{\mathbf{T}_{q,i}} \right) \frac{ASE_{l,k,i}}{P_{0,l}}$$

if $j \in \mathcal{M}_l$, and $\Gamma_{i,j} = 0$, if $j \notin \mathcal{M}_l$, where $\mathbf{H}_{l,k,i} = \prod_{q=1}^{k} h_{l,q,i}$, and $\mathbf{T}_{l,i} = \mathbf{H}_{l,N_l,i}$, are the intermediary and, respectively, the full transmission for l^{th} link.

Proof: Appendix I.

In the following the OSNR model in Lemma 1 is extended to include crosstalk terms due to WDM components at the optical nodes (OADM or OXC), such as optical filters, demultiplexers, add/drop modules, routers or switches, [5]. We consider only the heterowavelength (out-of-band) crosstalk due to incomplete filtering. A similar approach could be used to include the homowavelength crosstalk (in-band) in the space switches following a model as in [2].

Let the l^{th} optical node, associated with the l^{th} link, be characterized by an insertion loss $L_{sw,l}$ (typically channel independent), and a crosstalk ratio, X_{he}. The crosstalk ratio X_{he} is a measure of the out-of-band crosstalk and represents the fraction of total power leaked into a specific channel from all the other channels, [5]. We assume also that an optical amplifier is present at the l^{th} optical node, characterized by a gain $G_{X,l,i}$ and an ASE noise power denoted by $ASE_{X,l,i}$. The following assumption will be used.

(A.i.3) For the accumulation of ASE noise, the contribution of in-line cascaded amplifiers is significantly higher than the ASE contribution of OXCs. This assumption is justified by the fact that each OXC stage follows after N_l (typically up to 10) in-line amplifiers.

Lemma 2: Under (A.i.1)-(A.i.3), the OSNR for the i^{th} channel, including both ASE and crosstalk accumulation effects, is given as

$$OSNR_i = \frac{u_i}{n_{0,i} + \sum_{j \in \mathcal{M}} \tilde{\Gamma}_{i,j}\, u_j} \qquad \forall\, i \in \mathcal{M} \tag{6}$$

where the system matrix $\tilde{\mathbf{\Gamma}}$ is given as $\tilde{\mathbf{\Gamma}} = [\tilde{\Gamma}_{i,j}]$

$$\tilde{\Gamma}_{i,j} = \sum_{l \in \mathcal{R}_i} \sum_{k=1}^{N_l} \frac{G_{l,j}^k}{G_{l,i}^k} \left(\prod_{q=1}^{l-1} \frac{\tilde{\mathbf{T}}_{q,j}}{\tilde{\mathbf{T}}_{q,i}} \right) \frac{ASE_{l,k,i}}{P_{0,l}} + \sum_{l \in \mathcal{R}_i} \left(\prod_{q=1}^{l} \frac{\tilde{\mathbf{T}}_{q,j}}{\tilde{\mathbf{T}}_{q,i}} \right) \mathbf{X}_{he}$$

if $j \neq i$, $j \in \mathcal{M}_l$, $\tilde{\Gamma}_{i,j} = 0$, if $j \neq i$, $j \notin \mathcal{M}_l$, and $\tilde{\Gamma}_{i,j} = \Gamma_{i,i}$, if $j = i$, and where $\tilde{\mathbf{T}}_{l,i} = \mathbf{T}_{l,i}\, G_{X,l,i} L_{sw,l}$.

Proof: Appendix I.

III. Network OSNR Optimization: Central Cost Formulation

In this section we use the general network OSNR model in Lemma 1 (or Lemma 2) to formulate the static OSNR optimization problem for a large-scale network.

A special OSNR optimization problem, of maximizing the minimum OSNR, was considered in [6] for a single link case. This is similar to the problem of maximizing the minimum signal-to-interference ratio (SIR) in wireless networks, [8], [12], [13]. The problem can be translated into an OSNR (SIR) equalization problem and formulated as an eigenvalue problem for the system matrix. However, thermal or external noise is neglected, and more importantly, the method is not appropriate for developing on-line algorithms.

In the following, we consider a different OSNR optimization problem, similar to the SIR optimization problem considered in [11], [10] for wireless networks. Specifically we formulate the performance objective to be that of ensuring that the OSNR set of all channels is above a predefined set of targets, as determined by QoS constraints.

This approach allows full flexibility in setting individual per channel targets and also recognizes the presence of external (input) noise. Moreover, if a feasible solution exists, then there exists a unique solution which minimizes transmitter power, obtained by solving a system of linear equations. A specific feature of this approach is the fact that it leads to iterative (on-line) algorithms with geometric convergence as we will show in the following sections.

We consider the problem of finding an input (transmitter) power vector so that the OSNR of all channels satisfy

$$OSNR_i \geq \hat{\gamma}_i \qquad \forall i \in \mathcal{M} \tag{7}$$

where the set of targets, $\hat{\gamma}_i$, can be predefined as needed by QoS requirements for each channel. Moreover, the minimum power satisfying this constraint is sought. Therefore, using (5) or (6), the optimization problem can be formulated as the minimization of the central cost function

$$\min \sum_{i \in \mathcal{M}} u_i \tag{8}$$

such that

$$\frac{u_i}{\sum_{j \in \mathcal{M}} \mathbf{\Gamma}_{i,j}\, u_j + n_{0,i}} \geq \hat{\gamma}_i \qquad \forall i \in \mathcal{M}$$

This requirement is translated into

$$u_i \geq \sum_{j \in \mathcal{M}} \hat{\gamma}_i \mathbf{\Gamma}_{i,j}\, u_j + \hat{\gamma}_i\, n_{0,i}, \qquad \forall i \in \mathcal{M}$$

Recall that $\mathbf{n}_0$ stands as the input noise power at the Tx, and also can be considered to include also some other external noise such as thermal noise. Equivalently, in matrix-vector form, from the above we need to have

$$\mathbf{u} \geq \widehat{\mathbf{\Gamma}}\, \mathbf{u} + \widehat{\mathbf{n}}_\mathbf{0} \qquad \text{and} \qquad \mathbf{u} \geq 0 \tag{9}$$

where

$$\widehat{\mathbf{\Gamma}} = \begin{bmatrix} \hat{\gamma}_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & \hat{\gamma}_m \end{bmatrix} \cdot \mathbf{\Gamma}$$

We call the set of OSNR targets $\hat{\gamma}_i$, $i = 1, \ldots, m$ a feasible set, if there is a nonnegative finite vector $\mathbf{u}$ that satisfies (9). In the foregoing $\widehat{\mathbf{\Gamma}}$ is the network transmission matrix, normalized (weighted) by $\hat{\gamma}_i$, while $\mathbf{u}$, and $\widehat{\mathbf{n}}_\mathbf{0}$ are the vectors of input signal and weighted noise power, respectively.

The existence of a nonnegative solution for (9) is specified by the following standard result, based on Perron's Theorem, [14], [25], which can be proved by using the Jordan canonical form of $\widehat{\mathbf{\Gamma}}$.

Theorem 1: [14] Let $\rho(\widehat{\mathbf{\Gamma}})$ be the spectral radius of the nonnegative matrix $\widehat{\mathbf{\Gamma}}$. Then, the following statements are equivalent:

(1) There exists a vector $\mathbf{u} > 0$ such that $\left(I - \widehat{\mathbf{\Gamma}} \right) \mathbf{u} \geq \widehat{\mathbf{n}}_\mathbf{0}$

(2) $\rho(\widehat{\mathbf{\Gamma}}) < 1$

(3) $\left(I - \widehat{\mathbf{\Gamma}} \right)^{-1} = \sum_{k=0}^{\infty} \widehat{\mathbf{\Gamma}}^k$ exists and is positive component-wise.

■

Moreover, if the targets are feasible, i.e., if $\rho(\widehat{\mathbf{\Gamma}}) < 1$ (by Theorem 1), it can be shown that the power vector that satisfies (9) with equality minimizes the sum of the power transmitted, [9]. Therefore, if $\rho(\widehat{\mathbf{\Gamma}}) < 1$, the optimal power vector $\mathbf{u}^*$ is the solution of

$$\mathbf{u} = \widehat{\mathbf{\Gamma}}\, \mathbf{u} + \widehat{\mathbf{n}}_\mathbf{0} \tag{10}$$

Note that setting $\mathbf{u}^*$ as in (10) is a static and centralized approach. In the next section we show how this formulation can be used to develop iterative and distributed algorithms.

IV. Network OSNR Optimization: Iterative Algorithm - Synchronous Case

In the following we use the static optimization problem formulated in the previous section, to develop an iterative network algorithm for adjusting power levels. The proposed algorithm is distributed and autonomous, i.e., uses decentralized feedback and only local measurements. We show that, for the synchronous updating of all channels' power levels, this distributed algorithm converges with a geometric rate to the optimal solution $\mathbf{u}^*$.

Recall (10) and consider the update equation

$$\mathbf{u}(n+1) = \widehat{\mathbf{\Gamma}}\, \mathbf{u}(n) + \widehat{\mathbf{n}}_\mathbf{0}(n) \tag{11}$$

where

$$u(n) = \begin{bmatrix} u_1(n) \\ \vdots \\ u_m(n) \end{bmatrix}$$

with $u_i(n)$ being the input optical power for the i^{th} channel at time index n.

This is a centralized algorithm that requires, at each iteration n, knowledge of the full matrix of the optical network $\widehat{\mathbf{\Gamma}}$, together with current power allocation $\mathbf{u}(n)$ for *all* channels.

In the following, a distributed algorithm is defined, that uses only individual, local, measurements for each channel

i^{th}. To show this, note that from (11), the updates for the ith channel can be written component-wise as

$$u_i(n+1) = \hat{\gamma}_i \left(\sum_{j \in \mathcal{M}} \mathbf{\Gamma}_{i,j}\, u_j(n) + n_{0,i} \right)$$

However from (5) we have that

$$\sum_{j \in \mathcal{M}} \mathbf{\Gamma}_{i,j}\, u_j(n) + n_{0,i} = \frac{u_i(n)}{OSNR_i(n)} \qquad (12)$$

Based on the foregoing consider then the following distributed algorithm.

Algorithm:

$$u_i(n+1) = (1-\mu)\, u_i(n) + \mu\, \hat{\gamma}_i\, \frac{u_i(n)}{OSNR_i(n)} \qquad (13)$$

where $\mu > 0$ is an update (adjustment) parameter.

Remark 1: Note that this algorithm is distributed and autonomous as it relies only on locally available information. By this updating rule, the current transmitter power should be adjusted to be proportional to the previous power, plus by a factor equal to the ratio between the target OSNR, $\hat{\gamma}_i$, and the measured OSNR at the receiver.

The following theorem establishes convergence conditions for algorithm (13).

Theorem 2: Assume that an optimal solution exists, i.e., $\rho(\widehat{\mathbf{\Gamma}}) < 1$. Then algorithm (13) is globally stable and converges to the optimal solution, $\mathbf{u}^*$, if

$$0 < \mu < \frac{2}{1 + \rho(\widehat{\mathbf{\Gamma}})} \qquad (14)$$

Moreover, the rate of convergence of the algorithm is geometrical.

Proof: The proof uses some matrix results reviewed in Appendix II, and is presented in Appendix III.

Note that a standard assumption in adaptive control is that system parameters are stationary between any updates, [24]. Similarly here, system parameters change at network reconfiguration, and updates have to be performed after all network topology changes have been propagated. Therefore, we assume that

(A.ii.1) Channel gains and the "interference" terms due to other channel's power variation are stationary between power updates.

Then the following result holds.

Lemma 3: Assume that (A.ii.1) holds. Then algorithm (13) minimizes the following quadratic, individual, cost function

$$V_i = \| OSNR_i - \hat{\gamma}_i \|^2 \qquad i = 1, \ldots, m \qquad (15)$$

Proof: Appendix III.

Remark 4: Note that for $\mu = 1$, (13) reduces to

$$u_i(n+1) = \hat{\gamma}_i\, \frac{u_i(n)}{OSNR_i(n)} \qquad (16)$$

which is similar to the wireless case, [11], [13], and, for the case of equal OSNR targets, to the algorithm proposed heuristically in [7].

V. Network OSNR Optimization: Iterative Algorithm - Asynchronous Case

In this section we prove that the algorithm (13) converges to the optimal solution in the asynchronous case also. This is particularly important for adaptation in dynamic, large-scale networks, whereby channels update their powers asynchronously.

We follow an approach as in [10], based on the worst case delay. We assume that asynchronous operation can occur in two cases. Firstly, all channels can have the same update rate, but not all channels are synchronously updated at the same instance. Secondly, different channels can have different update rates. Both cases fit within the formulation described bellow.

Formally, assume for example, that channel j^{th} makes an update followed by an update on channel i^{th} at a later time. Suppose that this later time is earlier than the time at which the effect of j^{th} channel update is propagated to the i^{th} channel. Therefore, effectively channel i^{th} is performing an update that depends on an outdated value of the j^{th} channel power. Thus, for the asynchronous case we have that the OSNR at instant (n), $OSNR_i(n)$, is a function of the delayed powers of the other channels, i.e.,

$$\frac{u_i(n)}{OSNR_i(n)} = \sum_{j \in \mathcal{M}} \mathbf{\Gamma}_{i,j}\, u_j(n - \tau_{i,j}(n)) + n_{0,i}$$

We will use this to study convergence in the asynchronous case. Let $\mathcal{A}(n)$ denote the collection of indices of channels making concurrent updates, i.e., a subset of $1, 2, \ldots, m$. Thus, using the foregoing relation, for the asynchronous operation, instead of (13), we have, if $i \in \mathcal{A}(n)$

$$\begin{aligned} u_i(n+1) = {} & (1-\mu) u_i(n) + \\ & \mu \hat{\gamma}_i \left(\sum_{j \in \mathcal{M}} \mathbf{\Gamma}_{i,j} u_j(n - \tau_{i,j}(n)) + n_{0,i} \right) \end{aligned} \qquad (17)$$

and $u_i(n+1) = u_i(n)$, if $i \notin \mathcal{A}(n)$.

The delay terms $\tau_{i,j}(n)$ are nonnegative, bounded integers, which can be appropriately defined to cover all the asynchronously cases discussed above. The following assumptions will be used.

(A.iii.1) The delay terms are uniformly bounded, i.e., there exists τ_0 such that

$$\tau_{i,j}(n) \leq \tau_0 < \infty \qquad \forall \quad n \geq 0, \quad i, j \in \mathcal{M}$$

(A.iii.2) There exists finite π such that every channel has an update at least once every π.

The following result gives convergence conditions for the asynchronous algorithm, (17).

Theorem 3: Assume that $\rho(\widehat{\mathbf{\Gamma}}) < 1$ holds in addition to (Aiii.1) and (Aiii.2). Then if

$$0 < \mu < \frac{2}{1 + \rho(\widehat{\mathbf{\Gamma}})} \tag{18}$$

then the asynchronous algorithm (17) converges to $\mathbf{u}^*$, with a geometrical rate, i.e.,

$$\| \, e_\tau(n) \, \|_{\tau,v,\infty} \leq \, \alpha^q \, \| \, e_\tau(0) \, \|_{\tau,v,\infty}$$

where

$$\alpha = |1 - \mu| + \mu \, \rho(\widehat{\mathbf{\Gamma}}), \quad \text{and} \quad q = \left\lfloor \frac{n}{\tau_0 + \pi} \right\rfloor$$

Proof: Appendix IV.

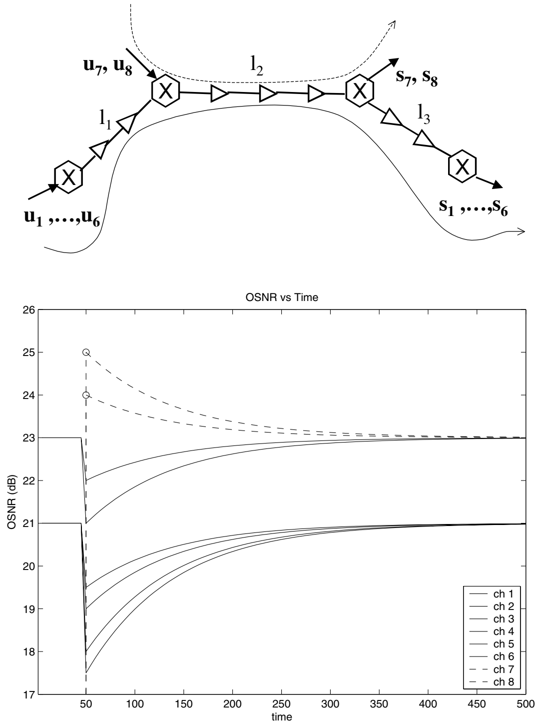

Fig. 2. (a) Example network configuration; (b) OSNR evolution in time

VI. Numerical Example

In this section we describe a MATLAB simulation used to exemplify the iterative algorithm. We considered a basic network configuration (Fig. 2 (a)), with three links, ten amplified spans per link, each optical amplifier having a parabolic spectral gain profile as in [6], and for simplicity we ignore the crosstalk. We assume that $m = 8$ channels can be transmitted, and initially only the first 6 channels were present. Within the set of 8 channels there are two levels of desired OSNR, a 21 dB level desired on the first 4 channels, and a 23 dB OSNR level on the other 4 channels. Initially, since only the first 6 channels are present, we set the optimal power vector (static) so that the desired levels are satisfied. At step $t = 100$ the network is reconfigured such that two new channels, 7 and 8, are added that pass only through link l_2. With transmitter powers maintained at the same level as before the add event, the OSNR for the existing channels has a sudden drop at $t = 100$ (see Fig. 2 (b)), due to the extra channels that share the link. Using the iterative algorithm and synchronous updating to adjust all channel powers, the channel OSNR levels converge back to the desired values (Fig. 2 (b)), of 21dB and 23dB, respectively.

VII. Conclusions

In this paper we considered OSNR optimization in optical networks and development of iterative algorithms for optimally re-adjusting network parameters, such that all channel/routes maintain a desired QoS (OSNR). We developed an analytical model for OSNR in a generic multi-link configuration, both as recursive relations and end-to-end relations. We formulated a static OSNR optimization problem for a generic large-scale network, and characterized the optimal solution. We developed an iterative distributed algorithm for adjusting channel power levels, that was shown to converge geometrically to the optimal solution. The algorithm is valid for general network configurations, and uses only local measurements or decentralized feedback. We proved that the algorithm is convergent also in the asynchronous case, which is particularly important for adaptation in a dynamic environment. An interesting future direction is extension of this approach to include the total power limit as a parameter to be optimized, as well as investigation of non-cooperative game theory approaches.

APPENDIX I
PROOF OF LEMMA 1 AND LEMMA 2

Proof of Lemma 1:

The proof follows by developing end-to-end (Tx to Rx) propagation relations for signal and noise, i.e., s_i and n_i. Recall that for the i^{th} channel, $p_{l,k,i}$ and $v_{l,k,i}$, denote the signal and the noise power, respectively, at the output of the k^{th} span on the l^{th} link. Using (1) it follows

$$\begin{aligned} p_{l,k,i} &= h_{l,k,i} \cdot p_{l,k-1,i} \\ v_{l,k,i} &= h_{l,k,i} \cdot v_{l,k-1,i} + ASE_{l,k,i} \end{aligned}$$

for $k = 1, \ldots, N_l$,with $ASE_{l,k,i}$, as in (2).

The total optical power for signal and noise in the bandwidth of channel i^{th} is given as

$$y_{l,k,i} = p_{l,k,i} + v_{l,k,i}$$

In (19), as in (**??**), the noise power at stage k is composed of: amplification of input noise and noise self-generated at each amplifier stage, $ASE_{l,k,i}$, as in (2).

Using these relations recursively after k we can write

$$p_{l,k,i} = p_{l,0,i} \prod_{q=1}^{k} h_{l,q,i} \quad (19)$$

$$v_{l,k,i} = v_{l,0,i} \prod_{q=1}^{k} h_{l,q,i} + \sum_{r=1}^{N_l} ASE_{l,r,i} \prod_{q=r+1}^{k} h_{l,q,i}$$

We will use the following notations and conventions. The signal power at the output of the l^{th} link, denoted by $s_{l,i}$, is taken by convention to be the same as the signal power at the output of the N_l^{th} span, i.e., $p_{l,N_l,i}$. A similar convention holds for noise, so that

$$s_{l,i} = p_{l,N_l,i} \quad \text{and} \quad n_{l,i} = v_{l,N_l,i} \quad (20)$$

Also, the signal power at the output of the $(l-1)^{th}$ link , $s_{l-1,i}$, is identical to the signal power at the input of the l^{th} link, which, by convention, is taken as the output of the 0^{th} span on the l^{th} link, i.e., , $p_{l,0,i}$. Therefore,

$$s_{l-1,i} = p_{l,0,i} \quad \text{and} \quad n_{l-1,i} = v_{l,0,i} \quad (21)$$

where the $(l-1)^{th}$ link is the preceding link on the $\mathcal{R}_i$ route of the i^{th} channel.

Therefore, for the signal and noise power at the output of the l^{th} link output, $s_{l,i}$ and $n_{l,i}$, we can write from (20, 19) and (21),

$$s_{l,i} = \mathbf{T}_{l,i}\, s_{l-1,i} \quad (22)$$

$$n_{l,i} = \mathbf{T}_{l,i}\, n_{l-1,i} + \sum_{r=1}^{N_l} \frac{\mathbf{T}_{l,i}}{\mathbf{H}_{l,r,i}} ASE_{l,r,i} \quad (23)$$

where the notations in the lemma were used. Recall that $u_i = s_{0,i}$ is the input signal power and $n_{0,i}$ is the input noise power. Then, by using (22) recursively after l, to obtain s_i and n_i at the end of the path $\mathcal{R}_i$, and using (4), yields

$$OSNR_i = \frac{u_i}{n_{0,i} + \sum_{l\in\mathcal{R}_i} \sum_{k=1}^{N_l} \frac{1}{\prod_{q=1}^{l-1} \mathbf{T}_{q,i}} \frac{ASE_{l,k,i}}{\mathbf{H}_{l,k,i}}} \quad (24)$$

From (19, 21) and (22) used recursively, we can write for $p_{l,k,i}$,

$$p_{l,k,i} = \left(\prod_{l-1\in\mathcal{R}_i} \mathbf{T}_{l-1,i} \right) \prod_{r=1}^{k} h_{l,r,i}\, u_i$$

Then, using the foregoing and the notations in the lemma, we see that (3) becomes

$$\sum_{j\in\mathcal{M}_l} \left(\prod_{q=1}^{l-1} \mathbf{T}_{q,j} \right) \mathbf{H}_{l,k,j}\, u_j = P_{0,l} \quad (25)$$

Since all amplifiers in a link have the same shape by (A.i.2), $G_{l,k,i}$ can be decomposed as

$$G_{l,k,i} = G_{l,i} \cdot \alpha_{l,k} \quad (26)$$

where $\alpha_{l,k}$ is the loss of a variable optical filter, adjusted to achieve constant total output power $P_{0,l}$, [6]. Then, using (1, 26) and the notations in the lemma we can write

$$\mathbf{H}_{l,k,i} = G_{l,i}^{k} \prod_{q=1}^{k} \tilde{\alpha}_{l,q}, \qquad \tilde{\alpha}_{l,q} = \alpha_{l,q}\, L_{l,q} \quad (27)$$

Substituting (27) for $\mathbf{H}_{l,k,i}$ into (25) we can find the wavelength independent part as

$$\prod_{q=1}^{k} \tilde{\alpha}_{l,q} = \frac{P_{0,l}}{\sum_{j\in\mathcal{M}_l} G_{l,j}^{k} \left(\prod_{q=1}^{l-1} \mathbf{T}_{q,j} \right) u_j} \quad (28)$$

for all $l \in \mathcal{R}_i, k = 1, \ldots, N_l$. Then, using (27, 28) into (24) yields after some manipulation,

$$OSNR_i = \frac{u_i}{n_{0,i} + \sum_{j\in\mathcal{M}} \mathbf{\Gamma}_{i,j}\, u_j} \qquad \forall i \in \mathcal{M}$$

with $\mathbf{\Gamma}$ defined as in (5).

■

Proof of Lemma 2: Let $\mathbf{x}_i$ be the crosstalk power, that falls within the bandwidth of the i^{th} channel, at the output (at Rx) . With the crosstalk included the OSNR, (4), is written as

$$OSNR_i = \frac{\mathbf{s}_i}{\mathbf{n}_i + \mathbf{x}_i} \quad (29)$$

Recall that in Lemma 1 any link-to-link connection was taken to be ideal. In order to include the crosstalk generated at OXC nodes, this has to be changed so that the recursive link relations for signal and noise power, (22), are modified as

$$\begin{aligned} s_{l,i} &= \tilde{\mathbf{T}}_{l,i} s_{l-1,i} \\ n_{l,i} &= \tilde{\mathbf{T}}_{l,i} n_{l-1,i} + \sum_{r=1}^{N_l} \tilde{\mathbf{T}}_{l,i} \frac{ASE_{l,r,i}}{\mathbf{H}_{l,r,i}} + ASE_{X,l,i} \end{aligned}$$

where $\tilde{\mathbf{T}}_{l,i} = \mathbf{T}_{l,i}\, G_{X,l,i} L_{sw,l}$ and where $ASE_{X,l,i}$ is the ASE generated at the l^{th} OXC.

Similalry, the crosstalk power term $\mathbf{x}_{l,i}$ added at the l^{th} OADM/OXC on the i^{th} channel is given as

$$\mathbf{x}_{l,i} = \tilde{\mathbf{T}}_{l,i}\, \mathbf{x}_{l-1,i} + \sum_{j \neq i, j \in \mathcal{M}_l} \mathbf{X}_{he} \tilde{\mathbf{T}}_{l,j}\, s_{l-1,j}$$

In the foregoing relation the first term is due to the crosstalk component propagated from the previous OADM/OXC node, while the second one represents the contribution due to the out-of-band crosstalk (incomplete filtering). With the foregoing relations, notice that the total optical power in the i^{th} channel's bandwidth, at the output of the l^{th} link is given as

$$\mathbf{y}_{l,i} = s_{l,i} + n_{l,i} + \mathbf{x}_{l,i}$$

where all contributions are included, i.e., signal, ASE noise and crosstalk.

Recall that $s_{0,i} = u_i$ and without loss of generality take $\mathbf{x}_{0,i} = 0$. Using the foregoing recursively after l, we can write for signal, noise and crosstalk terms, i.e., $\mathbf{s}_i$, $\mathbf{n}_i$ and $\mathbf{x}_i$, at the end of the path $\mathcal{R}_i$,

$$\begin{aligned} \mathbf{s}_i &= \prod_{l \in \mathcal{R}_i} \tilde{\mathbf{T}}_{l,i}\, u_i \\ \mathbf{n}_i &= \prod_{l \in \mathcal{R}_i} \tilde{\mathbf{T}}_{l,i}\, n_{0,i} + \\ &\quad + \sum_{l \in \mathcal{R}_i} \left(\prod_{q=l: q \in \mathcal{R}_i} \tilde{\mathbf{T}}_{q,i} \right) \sum_{k=1}^{N_l} \frac{1}{\mathbf{H}_{l,k,i}} ASE_{l,k,i} \\ &\quad + \sum_{l \in \mathcal{R}_i} \prod_{q=l+1: q \in \mathcal{R}_i} \tilde{\mathbf{T}}_{q,i}\, ASE_{X,l,i} \\ \mathbf{x}_i &= \sum_{j \neq i, j \in \mathcal{M}_l} \sum_{l \in \mathcal{R}_i} \left(\prod_{q=l+1: q \in \mathcal{R}_i} \tilde{\mathbf{T}}_{q,i} \right) \mathbf{X}_{he} \\ &\quad \left(\prod_{q=1:l} \tilde{\mathbf{T}}_{q,j} \right) u_j \end{aligned}$$

By (A.i.3), we can neglect the third term in $\mathbf{n}_i$, i.e., the ASE component due to the OXC.

Now using the foregoing relations into (29) and following the same approach as in the proof of Lemma 1, to replace for $\mathbf{H}_{l,k,i}$, we can rewrite $OSNR_i$ as

$$OSNR_i = \frac{u_i}{n_{0,i} + \sum_{j \in \mathcal{M}} \tilde{\mathbf{\Gamma}}_{i,j}\, u_j} \qquad \forall\, i \in \mathcal{M}$$

where the new matrix $\tilde{\mathbf{\Gamma}}$ is defined as in (6).

■

APPENDIX II
SOME USEFUL MATRIX RESULTS

In this appendix we review and prove some general matrix results. Let $\mathbf{A}$ be an (m x m) matrix that is component-wise non-negative, and let $\rho(\mathbf{A})$ be its spectral radius, [23], so that

$$\rho(\mathbf{A}) = \max_i |\lambda_i(\mathbf{A})| \tag{30}$$

Since $\mathbf{A}$ is nonnegative, by Perron-Frobenius Theorem, [14], it has a strictly positive eigenvalue λ_+, with the corresponding eigenvector $\mathbf{v}$, $\mathbf{v}^T = [v_1 \ldots v_m]$, having positive entries, $v_i > 0$, for all i. Moreover λ_+ is a *simple* eigenvalue and it is the largest eigenvalue in absolute value, i.e., $\rho(\mathbf{A}) = \lambda_+$. Therefore, we can write that

$$\mathbf{A}\,\mathbf{v} = \rho(\mathbf{A})\,\mathbf{v}$$

We define $\mathbf{D}_v$ as the following diagonal scaling matrix associated with $\mathbf{v}$

$$\mathbf{D}_v = \begin{bmatrix} \frac{1}{v_1} & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & \frac{1}{v_m} \end{bmatrix}$$

Then, for any m-dimensional vector $\mathbf{x}$, we denote by $\| \mathbf{x} \|_{v,\infty}$ the weighted l_∞ norm, related to the standard l_∞ vector norm, as

$$\| \mathbf{x} \|_{v,\infty} = \| \mathbf{D}_v\, \mathbf{x} \|_\infty \tag{31}$$

or, equivalently,

$$\| \mathbf{x} \|_{v,\infty} = \max_{1 \leq i \leq m} \frac{|x_i|}{v_i}$$

For matrix $\mathbf{A}$, we denote by $\| \mathbf{A} \|_{v,\infty}$ the corresponding induced matrix norm, [23],

$$\| \mathbf{A} \|_{v,\infty} = \max_{\mathbf{x} \neq 0} \frac{\| \mathbf{A}\,\mathbf{x} \|_{v,\infty}}{\| \mathbf{x} \|_{v,\infty}} \tag{32}$$

The following result can be proved.

Lemma 4: Let $\mathbf{A}$ be a component-wise non-negative matrix and $\mathbf{D}_v$ its corresponding scaling matrix . Then the following relations hold

$$\begin{array}{lll} (i) & \| \mathbf{A} \|_{v,\infty} & = \| \mathbf{D}_v \mathbf{A} \mathbf{D}_v^{-1} \|_\infty \\ (ii) & \rho(\mathbf{A}) & = \| \mathbf{D}_v \mathbf{A} \mathbf{D}_v^{-1} \|_\infty \\ (iii) & \| \mathbf{A} \mathbf{x} \|_{v,\infty} & \leq \rho(\mathbf{A}) \| \mathbf{x} \|_{v,\infty} \qquad \forall \mathbf{x} \end{array}$$

In the following we extend the weighted norm, (31), to a delayed vector. Let $\mathbf{x}(n)$ and $\mathbf{x}_\tau(n)$ denote an $m-$dimensional vector and its delayed version, respectively,

$$\mathbf{x}(n) = \begin{bmatrix} x_1(n) \\ \vdots \\ x_m(n) \end{bmatrix}, \quad \mathbf{x}_\tau(n) = \begin{bmatrix} x_1(n-\tau_1) \\ \vdots \\ x_m(n-\tau_m) \end{bmatrix}$$

where (n) is the time index and τ_i, are channel delays, assumed bounded, i.e., $\tau_i \leq \tau_0$, for all i. We denote by $\| \cdot \|_{\tau,v,\infty}$, the following norm,

$$\| \mathbf{x}_\tau(n) \|_{\tau,v,\infty} = \max_\tau \| \mathbf{x}_\tau(n) \|_{v,\infty} \qquad (33)$$

or, equivalently,

$$\| \mathbf{x}_\tau(n) \|_{\tau,v,\infty} = \max_{0 \leq \tau \leq \tau_0} \max_{1 \leq i \leq m} \frac{|x_i(n-\tau_i)|}{v_i}$$

Note that component-wise from the foregoing we have that

$$|x_i(n-\tau_i)| \leq v_i \| \mathbf{x}_\tau(n) \|_{\tau,v,\infty} \qquad (34)$$

for any $\tau_i \leq \tau_0$ and all i.

The following result extends Lemma 4, (iii).

Lemma 5: The following relation holds

$$\| \mathbf{A} \mathbf{x}_\tau \|_{\tau,v,\infty} \leq \rho(\mathbf{A}) \| \mathbf{x}_\tau \|_{\tau,v,\infty}$$

Appendix III
PROOF OF THEOREM 2 AND LEMMA 3

Proof of Theorem 2: Using (12) we see that (13) is equivalent to

$$u_i(n+1) = (1-\mu)\, u_i(n) + \mu \hat{\gamma}_i \left(\textstyle\sum_{j \in \mathcal{M}} \mathbf{\Gamma}_{i,j}\, u_j(n) + n_{0,i} \right)$$

or, in matrix-vector form,

$$\mathbf{u}(n+1) = (1-\mu)\mathbf{u}(n) + \mu \left(\widehat{\mathbf{\Gamma}} \mathbf{u}(n) + \widehat{\mathbf{n}}_\mathbf{0} \right) \qquad (35)$$

with $\widehat{\mathbf{\Gamma}}$ defined as in (9). Recall that $\mathbf{u}^*$ satisfies (10), and define

$$\mathbf{e}(n) = \mathbf{u}(n) - \mathbf{u}^*$$

We will prove that the vector $\mathbf{e}(n)$ converges to 0, for all initial conditions.

Using the foregoing relation and (10) into (35) yields, after some manipulation,

$$\mathbf{e}(n+1) = \mathbf{B}\, \mathbf{e}(n), \quad \text{with } \mathbf{B} = I - \mu \left(I - \widehat{\mathbf{\Gamma}} \right) \qquad (36)$$

Now the solution of (36) is

$$\mathbf{e}(n) = \mathbf{B}^n\, \mathbf{e}(0)$$

A necessary and sufficient condition such that $\mathbf{e}(n) \to 0$, for all initial conditions $\mathbf{e}(0)$, is that all eigenvalues of $\mathbf{B}$ are inside the unit circle, [23]. To show this, we will relate the eigenvalues of $\mathbf{B}$, (36), to those of $\widehat{\mathbf{\Gamma}}$, (9). Let $\lambda_i(\mathbf{B})$ be an eigenvalue of $\mathbf{B}$ and let ($\lambda_i(\widehat{\mathbf{\Gamma}})$, $\mathbf{w}$) be an eigenvalue, eigenvector pair of $\widehat{\mathbf{\Gamma}}$. Then using the definition of $\mathbf{B}$ in (36) we see that

$$\mathbf{B}\, \mathbf{w} = \mathbf{w} - \mu \left(\mathbf{w} - \lambda_i(\widehat{\mathbf{\Gamma}})\mathbf{w} \right)$$

where $\widehat{\mathbf{\Gamma}}\, \mathbf{w} = \lambda_i(\widehat{\mathbf{\Gamma}})\mathbf{w}$ was used. Equivalently,

$$\mathbf{B}\, \mathbf{w} = \lambda_i(\mathbf{B})\, \mathbf{w} \qquad (37)$$

$$\lambda_i(\mathbf{B}) = 1 - \mu\, (1 - \lambda_i(\widehat{\mathbf{\Gamma}}))$$

The foregoing shows that $\lambda_i(\mathbf{B})$, (37), is an eigenvalue of $\mathbf{B}$. Now, from (37), we can write

$$|\lambda_i(\mathbf{B})| \leq |1-\mu| + \mu\rho(\widehat{\mathbf{\Gamma}}) \qquad (38)$$

where the triangle inequality and (30) were used. Therefore $|\lambda_i(\mathbf{B})| < 1$, if the right hand side of (38) is less than 1, or equivalently, if

$$\mu\rho(\widehat{\mathbf{\Gamma}}) < \mu \qquad \text{and} \qquad \mu + \mu\rho(\widehat{\mathbf{\Gamma}}) < 2$$

Since $\rho(\widehat{\mathbf{\Gamma}}) < 1$ and $\mu + \mu\rho(\widehat{\mathbf{\Gamma}}) < 2$, by the conditions in the theorem, the foregoing are satisfied. Therefore $|\lambda_i(\mathbf{B})| < 1$, and hence $\mathbf{e}(n) \to 0$ as $n \to \infty$ for all initial conditions, proving convergence of the algorithm.

In the following we show that the rate of convergence is geometrical. Note that using Lemma 4, (iii), (Appendix II), applied to $\widehat{\mathbf{\Gamma}}$ we can write

$$\| \widehat{\mathbf{\Gamma}}\, \mathbf{e}(n) \|_{v,\infty} \leq \rho(\widehat{\mathbf{\Gamma}}) \| \mathbf{e}(n) \|_{v,\infty} \qquad (39)$$

Recall now (36) and take $\| \cdot \|_{v,\infty}$, as in (31), on both sides. Then using the triangle inequality it follows that

$$\| \mathbf{e}(n+1) \|_{v,\infty} \leq |1-\mu| \| \mathbf{e}(n) \|_{v,\infty} + \\ + \mu \| \widehat{\mathbf{\Gamma}} \mathbf{e}(n) \|_{v,\infty}$$

Using (39) in the foregoing inequality yields

$$\| e(n+1) \|_{v,\infty} \leq \alpha \| e(n) \|_{v,\infty}$$

for $\alpha = |1-\mu| + \mu\, \rho(\widehat{\mathbf{\Gamma}})$. Hence

$$\| e(n) \|_{v,\infty} \leq \alpha^n \| e(0) \|_{v,\infty}, \qquad \forall n \geq 0$$

Since from (38) we have $\alpha < 1$, this shows that the rate of convergence is geometrical.

■

Proof of Lemma 3: Since the cost function V_i, (15), is quadratic in $OSNR_i$, the following gradient-descent iterative algorithm, [24], minimizes V_i

$$OSNR_i(n+1) = OSNR_i(n) - \frac{1}{2}\,\mu\,\frac{\partial V_i}{\partial OSNR_i}$$

with $\mu > 0$ being an adjustment factor. Using the definition of V_i, (15), it follows that the foregoing is equivalent to

$$OSNR_i(n+1) = OSNR_i(n) - \mu(OSNR_i - \hat{\gamma}_i) \quad (40)$$

By (A.ii.1), between any power update as in (13), the denominator of $OSNR_i(n)$, (5),

$$n_{0,i} + \sum_{j \in \mathcal{M}} \mathbf{\Gamma}_{i,j}\, u_j(n)$$

does not change, so that

$$\frac{u_i(n+1)}{OSNR_i(n+1)} = \frac{u_i(n)}{OSNR_i(n)} \quad (41)$$

From (40, 41) it follows that

$$u_i(n+1) = u_i(n) - \mu(OSNR_i - \hat{\gamma}_i)\frac{u_i(n)}{OSNR_i(n)}$$

which is equivalent to (13) and the proof is complete.

■

APPENDIX IV
PROOF OF THEOREM 3

In this appendix we give a proof for Theorem 3, based on the worst case delay τ_0. Firstly, notice that if the update is made synchronously for all channels but with *old* or *outdated* information (delayed), than algorithm (17) is given similarly with $\mathcal{A}(n) = 1, \ldots, m$.

Therefore, using the definition of $\widehat{\mathbf{\Gamma}}$, from (17) we can write in matrix form

$$\mathbf{u}(n+1) = (1-\mu)\,\mathbf{u}(n) + \mu \left(\widehat{\mathbf{\Gamma}}\,\mathbf{u}_\tau(n) + \widehat{\mathbf{n}}_\mathbf{0} \right)$$

where $\mathbf{u}_\tau(n)$ is the delayed power vector,

$$\mathbf{u}_\tau(n) = \begin{bmatrix} u_1(n-\tau_1(n)) \\ \vdots \\ u_m(n-\tau_m(n)) \end{bmatrix}$$

with $\tau_j = \max_i \tau_{i,j}$, and $\tau_j \leq \tau_0, \forall j$ (by (A.iii.1)).

Define $\mathbf{e}(n) = \mathbf{u}(n) - \mathbf{u}^*$, with $\mathbf{u}^*$ as in (10). From the foregoing relations it can be shown that

$$\mathbf{e}(n+1) = (1-\mu)\,\mathbf{e}(n) + \mu\,\widehat{\mathbf{\Gamma}}\,\mathbf{e}_\tau(n)$$

Using the triangle inequality and Lemma 5 yields

$$\| \mathbf{e}(n+1) \|_{\tau,v,\infty} \leq |1-\mu| \; \| \mathbf{e}(n) \|_{\tau,v,\infty} + $$
$$+\mu\rho(\widehat{\mathbf{\Gamma}}) \| \mathbf{e}_\tau(n) \|_{v,\infty}$$

and the proof can follow along the same lines as for the synchronous case (see proof of Theorem 2).

In the following we give the proof for the totally asynchronous case, i.e., component-wise. Let

$$e_i(n) = u_i(n) - u_i^*$$

so that using (17, 10) we can write, if $i \in \mathcal{A}(n)$

$$e_i(n+1) = (1-\mu)e_i(n) + \mu \sum_{j \in \mathcal{M}} \widehat{\mathbf{\Gamma}}_{i,j} e_j(n - \tau_{i,j}(n))$$

and $e_i(n+1) = e_i(n)$, if $i \notin \mathcal{A}(n)$.

Then if $i \in \mathcal{A}(n)$, using the triangle inequality, we have

$$\begin{aligned} |\, e_i(n+1)\,| \;\; &\leq \;\; |1-\mu|\,|e_i(n)| \\ &\quad +\mu \sum_{j \in \mathcal{M}} \widehat{\mathbf{\Gamma}}_{i,j}\,|\, e_j(n-\tau_{i,j}(n))\,| \\ &\leq \;\; |1-\mu| \; \| e_\tau(n) \|_{\tau,v,\infty} \; v_i \\ &\quad +\mu \; \| e_\tau(n) \|_{\tau,v,\infty} \sum_{j \in \mathcal{M}} \widehat{\mathbf{\Gamma}}_{i,j}\, v_j \end{aligned}$$

where (34) (by (A.iii.1)) was used for both terms on the right-hand side.

Using the fact that $(\rho(\widehat{\mathbf{\Gamma}}), \mathbf{v})$ is an eigenvalue-eigenvector pair, from the foregoing it follows that

$$|\, e_i(n+1)\,|/v_i \leq \; \alpha \, \| e_\tau(n) \|_{\tau,v,\infty} \quad (42)$$

for $\alpha = |1-\mu| + \mu\,\rho(\widehat{\mathbf{\Gamma}})$, and $\alpha < 1$, by the conditions in the theorem.

Note also that if $i \notin \mathcal{A}(n)$

$$|\, e_i(n+1)\,|/v_i = |\, e_i(n)\,|/v_i \leq \| e_\tau(n) \|_{\tau,v,\infty} \quad (43)$$

or, using (34),

$$\| e_\tau(n+1) \|_{\tau,v,\infty} \leq \| e_\tau(n) \|_{\tau,v,\infty}$$

From (42,43) we see that $\| e_\tau(n) \|_{\tau,v,\infty}$ is a non-increasing function of n, that is also bounded from bellow by 0. Therefore $\| e_\tau(n) \|_{\tau,v,\infty}$ converges to 0.

In the following we show that the rate of convergence is geometrical. Let k be the index for the last update of channel i between n and $(n+\pi)$, (at least one exists by (Aiii.2)).

Then, for any $n' \geq n + (\tau_0 + \pi)$, we have to use the last updated value which is at the $(k+1)$ index (with τ_0 the maximum delay). Therefore we have

$$| e_i(n') | / v_i = |e_i(k+1)|/v_i \leq \alpha \| e_\tau(k) \|_{\tau,v,\infty}$$

where the last inequality followed from (42), taken at the two consecutive updates, k and $(k+1)$. Since k is the first update after index n, as in (43)

$$\| e_\tau(k) \|_{\tau,v,\infty} \leq \| e_\tau(n) \|_{\tau,v,\infty}$$

From the foregoing two relations we see that

$$| e_i(n') | / v_i \leq \alpha \| e_\tau(n) \|_{\tau,v,\infty}$$

for $\forall\, n' \geq n + (\tau_0 + \pi)$, so that

$$\| e_\tau(n') \|_{\tau,v,\infty} \leq \alpha \| e_\tau(n) \|_{\tau,v,\infty}$$

Using recursively the above relation it follows

$$\| e(n') \|_{\tau,v,\infty} \leq \alpha^q \| e(0) \|_{\tau,v,\infty}$$

with $q = \left\lfloor \frac{n'}{\tau_0 + \pi} \right\rfloor$, which, since $\alpha < 1$, shows that the algorithm converges at a geometrical rate.

■

REFERENCES

[1] Mukherjee, B.,'WDM optical communication networks: progress and challenges', *IEEE J. Selected Areas Commun.*, vol. **18**, 1810-1824, 2000.

[2] Ramamurthy, B., D. Datta, H. Feng, J. P. Heritage and B. Mukherjee, 'Impact of transmission impairments on the teletraffic performance of wavelength-routed optical networks', *IEEE J. Lightwave Tech.*, vol. **17**, 1713-1723, 1999.

[3] Pavel, L., 'Dynamics and stability in optical communication networks: A system theoretic framework', *Automatica*, *40(8)*, 1361-1370, Aug 2004.

[4] Pavel, L., 'A μ-Analysis Application to Stability of Optical Networks', in *Proc. American Control Conf.*, 3956-3961, July 2004.

[5] Agrawal, G.P., (2002). *Fiber-optic communication systems*, John Wiley & Sons.

[6] Forghieri, F., R.W. Tkach, D.L. Favin, 'Simple model of optical amplifier chains to evaluate penalties in WDM systems', *IEEE J. Lightwave Tech.*, vol. **16**, 1570-1576, 1998.

[7] Chraplyvy, A. R., J.A. Nagel, R.W. Tkach, 'Equalization in amplified WDM lightwave trasnmission systems', *IEEE Phot. T. Lett*, **4**, 920-922, 1992.

[8] Zander, J. 'Performance of optimum transmitter power control in cellular radio systems', *IEEE Trans. Vehicular Techn.*, vol. **41**, 305-311, 1992.

[9] Yates, R.D.,'A framework for uplink power control in cellular radio systems', *IEEE J. Selected Areas Commun.*, vol. **13**, 1341-1347, 1995.

[10] Mitra, D., 'An asynchronous distributed algorithm for power control in cellular radio systems', in *4th WINLAB Workshop on 3rd generation Wireless Info. Networks*, 177-186, Kluwer Academic Publishers, 1993.

[11] Foschini, G.J. and Z. Miljanic,'A simple distributed autonomous power control algorithm and its convergence', *IEEE Trans. on Vehicular Technology*, vol. **42**, 641-646, 1993.

[12] Grandhi, S.A, and R. Vijayan, D.J. Goodman, J. Zander, 'Centralized power control in cellular radio systems', *IEEE Trans. on Vehicular Technology*, vol. **42**, 466-468, 1993.

[13] Bambos, N., and S.C. Chen, G.J. Pottie, 'Radio link admission algorithm for wireless networks with power control and active link quality protection', in *Proc. IEEE INFOCOM 1995*, 97-104, 1995.

[14] Mync, H., *Nonnegative Matrices*, New York: Wiley, 1988.

[15] Mecozzi, A., 'On the optimization of the gain distribution of transmission lines with unequal amplifier spacing', *IEEE Phot. T. Lett.*, **10**, 1033-1035, 1998.

[16] Sun, Y., J. L. Zyskind, A. K. Srivastava, 'Average inversion level, modeling, and physics of Erbium-doped fiber amplifiers', *IEEE J. Sel. Top. Quantum Electr.*, vol. **3**, 991-1007, 1997.

[17] Pavel, L., 'Control design for transient power and spectral control in optical communication networks', in *Proc. IEEE Conf. Control Appl. CCA 2003*, Turkey, 415-422, June 2003.

[18] Falomari D., N. Mandayam and D. Goodman, 'A new framework for power control in wireless data networks: Games, utility and pricing', *Proc.Allerton Conf. on Communication, Control and Computing*, Illinois, 546-555, 1998.

[19] Alpcan, T., T. Basar, R. Srikant, E. Altman, 'CDMA uplink power control as a noncooperative game', in *Proc. 40th IEEE Conf. Decision and Control 2001*, 197-202, 2001.

[20] Chakrabarti and A. Mishra, 'QoS issues in ad hoc wireless networks', *IEEE Commun. Magazine*, 142-148, Feb 2001.

[21] Ji H., C.-Y. Huang, 'Non-cooperative uplink power control in cellular radio systems', *Wireless Networks*,**4** 233-240, 1998.

[22] Pavel, L., 'Power Control for OSNR Optimization in Optical Networks: A Noncooperative Game Approach', in *Proc. 43th IEEE Conf. Decision and Control 2004*, Dec. 2004.

[23] Horn, R. G. and C. R. Johnson, *Matrix Analysis*, Cambridge, University Press, 1999.

[24] Haykin, S., *Adaptive Filter Theory*, 3rd Ed., Prentice-Hall, 1996.

[25] MacCluer, C. R., 'The many proofs and applications of Perron's theorem', *SIAM Review*, vol. **42** (**3**), 487-498, 2000.