Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

A Distributed Optimization Approach to Constrained OSNR Problem ⋆

Yan Pan * **Tansu Alpcan** ** **Lacra Pavel** *

* *Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: {yanpan,pavel}@control.utoronto.ca)*
** *Deutsche Telekom Laboratories, Technische Universitat Berlin Emst-Reuter-Platz 7, 10587, Germany (e-mail: tansu.alpcan@telekom.de)*

Abstract: This paper studies constrained optical signal-to-noise ratio (OSNR) problem via a distributed optimization approach. In multi-channel optical systems, the signal over an optical link can be regarded as an interfering noise for others, which leads to OSNR degradation. Regulating the input optical power at the *Source* (transmitter) aims to achieve satisfactory OSNR level at the *Destination* (receiver) for each channel. Moreover, because all wavelength-multiplexed channels in a link share the same optical fiber, the total input power in a link has to be below the nonlinearity threshold, which corresponds to a *link capacity constraint.* We formulate the OSNR optimization problem as one of utility maximization with the objectives of achieving an OSNR target level for each channel while minimizing the interference and also satisfying the link capacity constraint. We derive conditions for the existence of a unique optimal solution, leading to a basis for an admission control scheme. By using a Lagrangian relaxation approach we propose two distributed update algorithms: a primal algorithm and a dual algorithm, and study their convergence properties both theoretically and numerically.

1. INTRODUCTION

In optical networks, multi-channel optical systems are realized by wavelength division multiplexing (WDM), which consists of several sources multiplexed in wavelength domain and transmitted over the same optical fiber. Control of optical networks via an optimization-based approach arises in the context of evolution of optical communications from statically designed point-to-point links, to reconfigurable WDM networks. A reconfigurable optical network operates dynamically, with existing channels being continuously served while network reconfiguration (e.g., channel added/dropped) is being performed. Essential research topics in this area include optimization of channel performance with general topologies and online reconfiguration, Mukherjee [2000].

At the physical transmission level, channel performance (quality of service, QoS) is directly determined by the bit-error rate (BER), which in turn, depends on OSNR, dispersion and nonlinear effects, Agrawal [2002]. OSNR is considered as the dominant performance parameter in link optimization, with dispersion and nonlinearity being limited with proper link design, Forghieri et al. [1998].

In multi-channel optical systems, a signal over the same optical link can be regarded as an interfering noise for others, which leads to the quality degradation in service, i.e., OSNR degradation. Regulating the input optical power per channel at *Source* aims to achieve a satisfactory OSNR level at *Destination.* Particularly in optical networks, because all wavelength-multiplexed channels in a link share the optical fiber, the total input power in a link has to be below the nonlinearity threshold, which can be regarded as the *link capacity constraint.* The OSNR optimization problem with link capacity constraint has been studied in Pan and Pavel [2007b], Pavel [2007] via a Nash game theoretic approach, Basar and Olsder [1999]. However, the desired OSNR target for each channel can not be guaranteed in all of these cases.

We consider a central OSNR optimization problem in optical links, formulated as one of utility maximization. Each channel obtains at the minimum its individual fixed target OSNR level, and beyond that optimizes its input optical power, according to its objective performance function, while respecting OSNR constraints (i.e., target OSNR levels) of all other channels and link capacity constraint. The individual objective performance function, which we assume to be strictly convex and continuously differentiable, can be defined, for example, as the difference between a convex increasing cost on channel's input optical power and a strictly concave utility-like function representing its willingness to increase its input optical power with the aim of attaining better individual OSNR level.

This problem is similar to Pavel [2006], in which the network OSNR optimization problem was formulated such that all channels maintain a desired individual OSNR level, while input optical power is minimized. In this paper, the OSNR optimization problem is subject to more physical constraints. Moreover, the objective function of this optimization problem is more general than the one in Pavel [2006]. A similar team optimization problem has

⋆ This work was supported in part by the Natural Sciences and Engineering Research Council of Canada.

been considered in Alpcan et al. [2007] for power control in wireless networks, as well as in Kelly et al. [1998], Srikant [2004] for congestion control. However, there are several differences that make the OSNR optimization problem more challenging: a more complex mathematical OSNR model due to cascaded optical amplified spans with a typical automatic power control (APC) operation mode and self-generation and accumulation of optical amplified spontaneous emission (ASE) noise in optical amplifiers, as well as specific constraints in optical networks imposed by dispersion and nonlinear effects.

The paper is organized as follows. We review an OSNR model for links in Section 2. In Section 3, we formulate the central OSNR optimization problem. We present a relaxation of the original system problem and develop two algorithms in Section 4. In Sections 5 and 6, we provide simulation results and concluding remarks, respectively.

2. LINK OSNR MODEL

We present a link OSNR model similar to the one in Pavel [2006]. Consider a single point-to-point optical link (Fig.1), composed of N cascaded optical amplifiers (OAs). A total set of channels, $\mathcal{M} = \{1, \dots, m\}$, corresponding to a set of wavelengths, are transmitted over the link from *Source* (transmitter site, Tx) to *Destination* (receiver site, Rx).

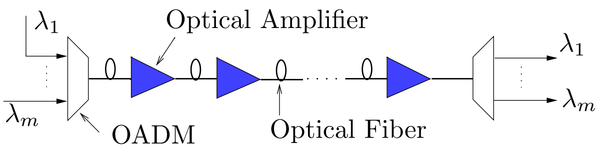

Fig. 1. A single point-to-point optical link

For the i^{th} channel, we denote by $u_i(n_i^0)$ and $p_i(n_i^{out})$ the optical power of the input signal (noise) (at Tx) and output signal (noise) (at Rx), respectively. We also denote by $\mathbf{u} = [u_1, \dots, u_m]^T$, the vector of input signal powers, $\mathbf{n}^0 = [n_1^0, \dots, n_m^0]^T$, the vector of input noise powers for all channels and let $\mathbf{u}_{-i}$ denote the vector obtained from $\mathbf{u}$ by deleting the i^{th} element, i.e., $\mathbf{u}_{-i} = [u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_m]^T$.

Cascaded OAs are used to amplify optical power of all channels simultaneously and introduce amplified spontaneous emission (ASE) noise for each channel. We denote for the i^{th} channel the ASE noise self-generated in the k^{th} OA as $ASE_{k,i}$, which is wavelength-dependent. Moreover, we assume that all cascaded optical spans have the same length and all OAs have the same spectral shape. Therefore each i^{th} channel experiences a different wavelength-dependent gain, G_i. Each OA is operated in the automatic power control (APC) mode, such that a constant total power target at each span is maintained, which we assume to be same for each span, denoted by P_0, which is selected below the threshold for nonlinear effects, Mecozzi [1998].

The OSNR for the i^{th} channel at the output of a link defined as $OSNR_i = p_i/n_i^{out}$ is given next, Pavel [2006].

Lemma 1. In a single optical link, the OSNR of the i^{th} channel at the output of the link is

$$OSNR_i = \frac{u_i}{n_i^0 + \sum_{j \in \mathcal{M}} \Gamma_{i,j} u_j} \quad (1)$$

where $\mathbf{\Gamma} = [\Gamma_{i,j}]$ is the system-related matrix with

$$\Gamma_{i,j} = \sum_{r=1}^{N} \frac{(G_j)^r}{(G_i)^r} \frac{ASE_{r,i}}{P_0}, \ \forall i, j \in \mathcal{M}$$

Remark 1. Mathematically, the OSNR model is similar to the wireless signal to interference ratio (SIR) model, Alpcan et al. [2007], except that the system-related matrix has non-zero diagonal elements, while the off-diagonal elements are dependent on specific optical network parameters, such as optical span gains, self-generated ASE noise, etc. Another specific feature is the fact that cross-coupling terms in OSNR not only due to crosstalk as in the wireless case. More precisely, even with no crosstalk (in the single point to point link case), ASE self-generation and accumulation in optical links lead to cross-coupling terms.

3. OSNR OPTIMIZATION PROBLEM

The OSNR optimization problem is subject to the following specific constraints:

(C.1) The OSNR set of all channels is above a predefined set of OSNR targets. We let $\widehat{\gamma}_i$ be the i^{th} channel's target OSNR and the corresponding vector form is $\widehat{\gamma} = [\widehat{\gamma}_1, \dots, \widehat{\gamma}_m]^T$. Thus the *target OSNR constraint* can be written as

$$OSNR_i \geq \widehat{\gamma}_i, \ \forall i \in \mathcal{M} \quad (2)$$

(C.2) A physical constraint is the non-negativity of input optical power (*non-negativity constraint*), i.e.,

$$u_i \geq 0, \ \forall i \in \mathcal{M} \quad (3)$$

(C.3) The total power constraint, or *link capacity constraint*, is imposed on input optical power as well,

$$\textstyle\sum_{i \in \mathcal{M}} u_i \leq P_0 \quad (4)$$

where P_0 is the total power target.

3.1 System problem

Let $C_i(u_i)$ be the i^{th} channel's individual cost function. Typically, we make the following assumptions:

(A.1) $C_i(u_i)$ is a strictly convex and continuously differentiable function with respect to u_i, $\forall i \in \mathcal{M}$.
(A.2) $C_i(u_i) \to +\infty$ as $u_i \to 0$

The constrained system problem is formulated as the minimization of the central cost function, i.e., $\min_{\mathbf{u}} \sum_{i \in \mathcal{M}} C_i(u_i)$, such that $\forall i \in \mathcal{M}$, $u_i \geq 0$, $OSNR_i \geq \widehat{\gamma}_i$ and $\sum_{i \in \mathcal{M}} u_i \leq P_0$. By using the OSNR model in (1), we can rewrite the set of conditions, $OSNR_i \geq \widehat{\gamma}_i$, $\forall i \in \mathcal{M}$, as

$$u_i + \textstyle\sum_{j \in \mathcal{M}} (-\widehat{\gamma}_i \Gamma_{i,j}) u_j \geq n_i^0 \widehat{\gamma}_i$$

in a vector form, $\mathbf{Tu} \geq \mathbf{b}$ with

$$\mathbf{T} = \begin{bmatrix} 1-\widehat{\gamma}_1\Gamma_{1,1} & -\widehat{\gamma}_1\Gamma_{1,2} & \cdots & -\widehat{\gamma}_1\Gamma_{1,m} \\ -\widehat{\gamma}_2\Gamma_{2,1} & 1-\widehat{\gamma}_2\Gamma_{2,2} & \cdots & -\widehat{\gamma}_2\Gamma_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ -\widehat{\gamma}_m\Gamma_{m,1} & -\widehat{\gamma}_m\Gamma_{m,2} & \cdots & 1-\widehat{\gamma}_m\Gamma_{m,m} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} n_1^0\widehat{\gamma}_1 \\ n_2^0\widehat{\gamma}_2 \\ \vdots \\ n_m^0\widehat{\gamma}_m \end{bmatrix} \quad (5)$$

The link capacity constraint, (4) can also be written in a vector form, $\mathbf{1}^T\mathbf{u} \leq P_0$, where $\mathbf{1}$ is a vector with all elements equal to 1. Thus the constrained system optimization problem is formulated by

$$\min_{\mathbf{u}} \textstyle\sum_{i \in \mathcal{M}} C_i(u_i) \quad (6)$$

such that

$$\mathbf{u} \in \Omega := \{\mathbf{u} \in \mathbb{R}^m : \mathbf{T}\mathbf{u} \geq \mathbf{b}, \mathbf{1}^T\mathbf{u} \leq P_0,\ u_i \geq 0, \forall i \in \mathcal{M}\}$$

Assumption (A.2) ensures that the solution to the system optimization problem, (6) does not hit $u_i = 0$, $\forall i \in \mathcal{M}$.

This system problem is more general than the one in Pavel [2006], where the individual cost function is selected exactly as the individual optical input power. Moreover, the total power constraint is not considered in Pavel [2006].

3.2 Solution

Theorem 1. Consider a constrained optimization problem defined in (6). The constraint set Ω in (6), is non-empty, and there exists a unique positive solution, $\mathbf{u}^*$ if $\forall i \in \mathcal{M}$,

$$\widehat{\gamma}_i < \frac{1}{\sum_{j \in \mathcal{M}} \Gamma_{i,j}} \tag{7}$$

where $\Gamma_{i,j}$ is defined in (1) and

$$\mathbf{1}^T \cdot \widetilde{T}(\widehat{\gamma}) \cdot \mathbf{b}(\widehat{\gamma}) \leq P_0 \tag{8}$$

with $\mathbf{b}(\widehat{\gamma}) = [n_1^0\widehat{\gamma}_1, \ldots, n_m^0\widehat{\gamma}_m]^T$ and $\widetilde{\mathbf{T}}(\widehat{\gamma}) = \mathbf{T}^{-1}(\widehat{\gamma})$.

Proof: We first show that the constraint set Ω in (6) is non-empty.

The system-related matrix, $\mathbf{\Gamma}$, in the OSNR model, (1), is a positive matrix, such that if (7) is satisfied, we have

$$1 - \widehat{\gamma}_i\Gamma_{i,i} > \widehat{\gamma}_i \textstyle\sum_{j \neq i, j \in \mathcal{M}} \Gamma_{i,j} > 0$$

i.e.,

$$1 - \widehat{\gamma}_i\Gamma_{i,i} > \textstyle\sum_{j \neq i, j \in \mathcal{M}} |-\widehat{\gamma}_i\Gamma_{i,j}|$$

such that the matrix $\mathbf{T}$ defined in (5) has positive main diagonal entries. Moreover, $\mathbf{T}$ is strictly diagonally dominant. According to Gershgorin's Theorem, Horn and Johnson [1999], all eigenvalues of $\mathbf{T}$ have positive real part. Moreover, all off-diagonal elements of $\mathbf{T}$ are non-positive such that $\mathbf{T}$ is an M-matrix, Fiedler [1986], which has the following properties: $\mathbf{Tu} \geq \mathbf{b} > 0$ implies $\mathbf{u} \geq 0$, $\mathbf{T}$ is non-singular and the inverse of $\mathbf{T}$ is non-negative. Thus

$$\mathbf{u} \geq \mathbf{T}^{-1}\mathbf{b} := \widetilde{\mathbf{T}}\mathbf{b} \tag{9}$$

such that $\mathbf{1}^T \cdot \mathbf{u} \geq \mathbf{1}^T \cdot \widetilde{T} \cdot \mathbf{b}$, i.e., $\mathbf{1}^T \cdot \mathbf{u} \geq \mathbf{1}^T \cdot \widetilde{T}(\widehat{\gamma}) \cdot \mathbf{b}(\widehat{\gamma})$.

Recall the total input power constraint in (6), $\mathbf{1}^T \cdot \mathbf{u} \leq P_0$. Then if the condition (8) is satisfied, i.e., $\mathbf{1}^T \cdot \widetilde{T}(\widehat{\gamma}) \cdot \mathbf{b}(\widehat{\gamma}) \leq P_0$, the set $\{\mathbf{u} \in \mathbb{R}^m : \mathbf{Tu} \geq \mathbf{b}, \mathbf{1}^T\mathbf{u} \leq P_0\}$ is non-empty. Recall that $\mathbf{Tu} \geq \mathbf{b} > 0$ implies $\mathbf{u} \geq 0$, therefore, we prove that if $\widehat{\gamma}$ is selected such that the conditions of (7) and (8) are satisfied, the constraint set, Ω, is non-empty.

Moreover, the constraint set, Ω, is convex, (Bertsekas [1999], pp.689). We note that based on the link capacity constraint and non-negativity constraint, we have $0 \leq u_i \leq P_0$, $\forall i \in \mathcal{M}$, such that Ω is bounded. In addition, it is also closed since it consists of intersection of half-spaces (the simplest case is shown in Fig.2). Thus this constrained system optimization problem is a strictly convex optimization problem on a convex, bounded and closed constraint set, which always admits a unique globe minimum, $\mathbf{u}^*$, Bertsekas [1999]. ■

We illustrate the constraint set, Ω, and the conditions, (7) and (8), with a simple example when $m = 2$ in Fig.2.

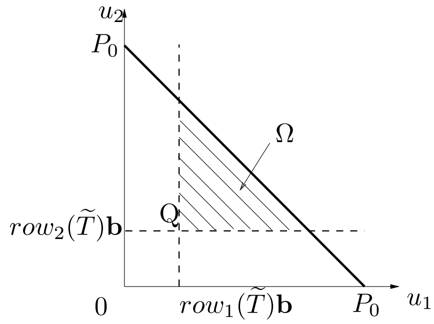

Fig. 2. Illustration of the constraint set, Ω.

From (9), we have $\left[\begin{smallmatrix} u_1 \\ u_2 \end{smallmatrix}\right] \geq \begin{bmatrix} row_1(\widetilde{T}) \cdot \mathbf{b} \\ row_2(\widetilde{T}) \cdot \mathbf{b} \end{bmatrix}$, where $row_i(A)$ is defined as the i^{th} row of the matrix A. From Fig.2, it's ready to see that the intersection point, Q, lies in the set of the total input power constraint, i.e.,

$$\left(row_1(\widetilde{T}) \cdot \mathbf{b}, row_2(\widetilde{T}) \cdot \mathbf{b}\right) \in \{\mathbf{u} | u_1 + u_2 \leq P_0\}$$

if $\sum_{i=1}^{2} row_i(\widetilde{T})\mathbf{b} \leq P_0$. Therefore Ω is non-empty.

Let us take a close look at the second condition, (8). In a real network system, it is always a question how to express the conditions under certain physical constraints.

(a). The input noise negligible: $n_i^0 \approx 0, \forall i \in \mathcal{M}$ Recall that $\mathbf{n}^0$ denotes the input noise power at Tx and can be considered to include some external noise, such as thermal noise. If the input noise is neglected, $\mathbf{n}^0$ includes only some other external noise, which is also negligible, Forghieri et al. [1998]. In this case, we get $\mathbf{b} = [\, n_1^0\widehat{\gamma}_1,\ \ldots,\ n_m^0\widehat{\gamma}_m\,]^T \approx \mathbf{0}$. Then $P_0 \geq \mathbf{1}^T \cdot \widetilde{T} \cdot \mathbf{b} \approx 0$, which means the constraint set is non-empty with only the first condition, (7). Thus the OSNR target, $\widehat{\gamma}_i$, can be selected in a distributed way based on the first condition, (7).

(b). Maximum OSNR target: $\widehat{\gamma}_{max}$ We know from the proof of Theorem 1 that the matrix $\mathbf{T}$ defined in (5) is strictly diagonally dominant, which means all eigenvalues of $\mathbf{T}$ are inside the unit disk, Horn and Johnson [1999], so that the spectral radius of $\mathbf{T}$, $\rho(\mathbf{T})$, satisfies $\rho(\mathbf{T}) < 1$. Moreover, $\mathbf{T}^{-1} = (\mathbf{I} - diag(\widehat{\gamma}_i)\mathbf{\Gamma})^{-1} = \sum_{k=0}^{\infty} diag((\widehat{\gamma}_i)^k)\mathbf{\Gamma}^k$ exists, where $diag(a_i)$ is a diagonal matrix whose diagonal entries are $a_1, \ldots, a_M$, and $\mathbf{T}^{-1}$ is positive component-wise, Pavel [2006]. We can rewrite (8) as

$$\mathbf{1}^T \cdot \sum_{k=0}^{\infty} diag((\widehat{\gamma}_i)^k)\mathbf{\Gamma}^k \cdot diag(\widehat{\gamma}_i)\mathbf{n}^0 \leq P_0 \tag{10}$$

such that if $\widehat{\gamma}_i$ increases (given $\widehat{\gamma}_j$, $j \neq i$), the left-hand side (LHS) of (10), i.e., LHS of (8), will increase. Thus when $\widehat{\gamma}_i \leq \widehat{\gamma}_{max}$, $\forall i \in \mathcal{M}$, we can find $\widehat{\gamma}_{max}$ by using the following equation:

$$\widehat{\gamma}_{max} \cdot \mathbf{1}^T \cdot (\mathbf{I} - \widehat{\gamma}_{max}\mathbf{\Gamma})^{-1} \cdot \mathbf{n}^0 = P_0 \tag{11}$$

It can be seen from the link OSNR model (1) that the channel performance is interference limited. In addition, (11) shows that the OSNR target levels significantly affect the capacity of an optical link: the link decides the OSNR threshold, $\widehat{\gamma}_{max}$, by using (11), and any new channel which has a required OSNR level no more than $\widehat{\gamma}_{max}$ will be admitted to transfer over the link. This idea can be used for a link to develop channel admission control schemes.

4. THE RELAXED SYSTEM PROBLEM AND DISTRIBUTED ALGORITHMS

In this section, we use a penalty or barrier function, Bertsekas [1999], to relax the original system optimization problem, (6), into a an equivalent unconstrained problem, such that decentralized power update schemes can be developed and used. We first build a relaxed system optimization problem and study its unique solution. We show that the solution of the relaxed system problem is arbitrarily close to the one of the original system optimization problem. After that, we study two distributed algorithms (primal and dual) and show that they converge to the unique solution of the relaxed system problem.

We rewrite the constraints in the system problem, (6), the link capacity constraint and the target OSNR constraint in a vector form, as

$$\widehat{\mathbf{T}}\mathbf{u} \geq \widehat{\mathbf{b}} \tag{12}$$

where the augmented matrix $\widehat{\mathbf{T}} = \left[\begin{smallmatrix}\mathbf{T}\\ -\mathbf{1}^T\end{smallmatrix}\right]$ and $\widehat{\mathbf{b}} = \left[\begin{smallmatrix}\mathbf{b}\\ -P_0\end{smallmatrix}\right]$. Thus we rewrite the system problem in the following way:

$$\min_{\mathbf{u}\geq 0} \textstyle\sum_{i\in\mathcal{M}} C_i(u_i) \tag{13}$$

subject to $\widehat{\mathbf{T}}\mathbf{u} \geq \widehat{\mathbf{b}}$.

We have established in Theorem 1 that there exists an optimal solution, $\mathbf{u}^*$, to the constrained system problem, (13). Moreover, the cost function is strictly convex and the constraints are linear such that there is no duality gap and dual optimal prices (Lagrange multipliers), exist, Bertsekas [1999]. Thus it readily follows that there exist a Lagrange multiplier vector, λ, such that $\mathbf{u}^*$ minimizes the following differentiable Lagrangian function,

$$L(\mathbf{u};\lambda) = \textstyle\sum_{i\in\mathcal{M}} C_i(u_i) - \lambda^T(\widehat{\mathbf{T}}\mathbf{u} - \widehat{\mathbf{b}}) \tag{14}$$

We let

$$q_i = row_i(\widehat{\mathbf{T}}^T)\lambda \tag{15}$$

Thus we rewrite (14) as

$$L(\mathbf{u};\lambda) = \textstyle\sum_{i\in\mathcal{M}} (C_i(u_i) - q_i u_i) + \lambda^T\widehat{\mathbf{b}} \tag{16}$$

The unique optimal solution, $\mathbf{u}^*$, is given by $\frac{\partial}{\partial u_i}L(\mathbf{u}^*;\lambda) = C_i'(u_i^*) - q_i = 0$, i.e.,

$$u_i^* = C_i^{'-1}(q_i) \tag{17}$$

where $C_i^{'-1}$ exists because C_i' is strictly monotonic, due to the strict convexity of C_i.

The dual problem of (13) is given by $\max_{\lambda\geq 0} D(\lambda)$, where $D(\lambda)$ is defined as $D(\lambda) = \min_{\mathbf{u}\geq 0} L(\mathbf{u};\lambda)$. Based on(15), (16) and (17), the associated dual problem is given by

$$\max_{\lambda\geq 0} \sum_{i\in\mathcal{M}} \left(C_i(C_i^{'-1}(q_i)) - q_i C^{'-1}(q_i)\right) + \lambda^T\widehat{\mathbf{b}} \tag{18}$$

where q_i is defined as in (15).

4.1 A barrier function

Basically, the penalty or barrier function method is used to add a term to the objective function to penalize any violation of the constraints. Here we select a continuous barrier function, $B_i(\cdot)$, with the following properties:

(P.1) $B_i(x)$ is non-increasing with respect to x_i.

(P.2) $B_i(x)$ attains the value 0 if $x_i > \widehat{b}_i$, where $\widehat{b}_i$ is defined as in (12).

(P.3) The following statement is generally true for $B_i(\cdot)$:

$$\lim_{u_i\to\infty} \int_{\widehat{b}_i}^{y_i(\mathbf{u})} B_i(x)dx \to -\infty,\ \forall i \in \mathcal{M} \tag{19}$$

where

$$y_i(\mathbf{u}) := row_i(\widehat{\mathbf{T}})\mathbf{u} \tag{20}$$

An ideal construction of $B_i(\cdot)$ is continuously non-negative over the feasible region and $B_i(\cdot)$ approaches ∞ as the boundary is approached from the interior of the feasible region.

4.2 A primal algorithm

Based a selected barrier function, $\lambda(\mathbf{u})$, which satisfies the properties (P.1)-(P.3), we define a primal algorithm as the set of differential equations, $\forall i \in \mathcal{M}$,

$$\dot{u}_i(t) = \frac{d}{dt}u_i(t) = g_i(u_i, q_i) = -k_i\left(C_i'(u_i(t)) - q_i(t)\right) \tag{21}$$

where $k_i > 0$ is the user-defined step-size constant and $q_i(t)$ is defined as in (15):

$$q_i(t) = row_i(\widehat{\mathbf{T}}^T)\lambda(\mathbf{y}(\mathbf{u})) \tag{22}$$

with a barrier function vector, $\lambda(\cdot) = [\lambda_1(\cdot), \ldots, \lambda_{m+1}(\cdot)]^T$ and variables, $y_i(\mathbf{u})$, defined in (20).

Remark 2. The primal update algorithm leads us to implement it in a distributed way. The link as the coordinator calculates the vector $\mathbf{q}(t)$ in (22) based on the received channel input power, channel performance preference and link constraint, then feeds this updated information back to each channel. The channel adjustment algorithm (21) is completely distributed and can be implemented by individual channels using only local information. Fig.3 depicts the information flow and the primal update algorithm is also represented in terms of blocks in Fig.4.

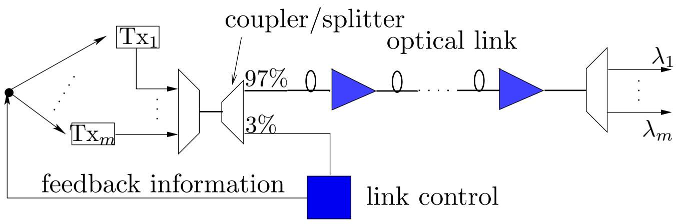

Fig. 3. The primal update algorithm: information flow

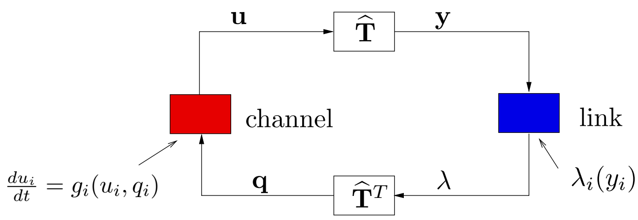

Fig. 4. The primal update algorithm: block representation

Now we establish that with the properties of the barrier function, λ_i, we can construct a Lyapunov function,

$$V_p(\mathbf{u}) = \sum_{i\in\mathcal{M}} C_i(u_i) - \sum_{i\in\mathcal{M}} \int_{\widehat{b}_i}^{y_i(\mathbf{u})} \lambda_i(x)dx \tag{23}$$

for the system of differential equations, (21) to analyze its convergence properties.

The following theorem states that the unique equilibrium, $\mathbf{u}^*$, is a stable point of the system.

Theorem 2. (**asymptotically global stability**) The unique solution, $\mathbf{u}^*$, to the primal algorithm, (21), minimizing $V_p(\mathbf{u})$, is a stable point, to which all trajectories starting from any initial condition converge.

Proof: Recall that the cost function, $C_i(u_i)$ is strictly convex such that $\sum_{i\in\mathcal{M}} C_i(u_i)$ is also strictly convex, thus together with the second assumption for $C_i(u_i)$, (A.2), and the non-increasing property of the barrier function, $V_p(\mathbf{u})$ defined in (23) is strictly convex, Srikant [2004], and

$$\frac{\partial V_p(\mathbf{u})}{\partial u_i} = C_i'(u_i) - row_i(\widehat{\mathbf{T}}^T)\lambda(\mathbf{u}) = 0$$

Solve the set of above equations we obtain

$$u_i^* = C_i^{'-1}(row_i(\widehat{\mathbf{T}}^T)\lambda(\mathbf{u}^*)), \forall i \in \mathcal{M}$$

which is the unique solution of the primal algorithm, (21).

We construct a modified Lyapunov function given as

$$\overline{V}_p(\mathbf{u}) = V_p(\mathbf{u}+\mathbf{u}^*) + C$$

where $\mathbf{u}^*$ is the unique solution and C is a constant such that $\overline{V}_p(\mathbf{u}) > 0$, $\forall \mathbf{u} \neq 0$ and $\overline{V}_p(\mathbf{0}) = 0$. Moreover, $\forall \mathbf{u} \neq \mathbf{u}^*$, we have

$$\dot{V}_p(\mathbf{u}) = \sum_{i\in\mathcal{M}} \left(\frac{\partial}{\partial u_i} V_p(\mathbf{u}) \cdot \dot{u}_i\right)$$
$$= -k_i \sum_{i\in\mathcal{M}} \left(C_i'(u_i) - row_i(\widehat{\mathbf{T}}^T)\lambda(\mathbf{u})\right)^2 < 0$$

Since $\dot{\overline{V}}_p(\mathbf{u}) = \dot{V}_p(\mathbf{u}+\mathbf{u}^*)$, we have $\dot{\overline{V}}_p(\mathbf{u}) < 0$, $\forall \mathbf{u} \neq 0$. Recall the properties of the barrier function, (19), we have

$$\overline{V}_p(\mathbf{u}) \to \infty, \text{as } \|\mathbf{u}\| \to \infty$$

Thus all of the conditions of Barbashin-Krasovskii theorem (Khalil [2002], pp124) are satisfied and the equilibrium point $\mathbf{u}^*$ is globally asymptotically stable. ■

Based on Lyapunov function $V_p(\mathbf{u})$, we establish a relaxed system problem for the original system problem, (13):

$$\min_{\mathbf{u}\geq 0} V_p(\mathbf{u}) \qquad (24)$$

The barrier function, $\lambda_i(\cdot)$, can be selected such that the unique solution of (24) may arbitrarily closely approximate the optimal solution of the original system problem, (13).

4.3 A dual algorithm

In the primal algorithm, the differential equations show that $u_i(t)$ varies gradually given the barrier function $\lambda_i(\mathbf{u})$. Next, we consider another algorithm, where $\lambda_i(t)$ varies gradually, with u_i given as the function of $\lambda_i(t)$, which we refer as a dual algorithm:

$$\dot{\lambda}_i(t) = \frac{d}{dt}\lambda_i(t) = k_i\left(\widehat{b}_i - row_i(\widehat{\mathbf{T}})\mathbf{u}(t)\right) \qquad (25)$$

where

$$u_i(t) = C_i^{'-1}(row_i(\widehat{\mathbf{T}}^T)\lambda(t)) \qquad (26)$$

Recall the definition for q_i in (15), $q_i(t) = row_i(\widehat{\mathbf{T}}^T)\lambda(t)$, We can rewrite the dual algorithm, (25) and (26), in a distributed way:

$$\text{Link: } \dot{\lambda}_i(t) = k_i\left(\widehat{b}_i - row_i(\widehat{\mathbf{T}})\mathbf{u}(t)\right)$$
$$q_i(t) = row_i(\widehat{\mathbf{T}}^T)\lambda(t),\ i \in \mathcal{M} \qquad (27)$$
$$\text{Channel: } u_i(t) = C_i^{'-1}(q_i(t)),\ i \in \mathcal{M} \qquad (28)$$

The dual algorithm can be implemented in the distributed way similar to the case of the primal update algorithm: the link varies the vector $\mathbf{q}$ in accordance with (27) and provides channel the value of q. Each channel calculates its individual input power based on the received q_i.

Remark 3. Here q_i is acting as a *price*, as in Kelly et al. [1998], charging channel i to pay for creating interference to other channels, as well as utilizing the link source. From this point of view, the dual problem, (18) is solved to obtain revenue-maximizing prices for the link.

The following theorem establishes the global convergence.

Theorem 3. The dual algorithm, (27)-(28), is globally asymptotically stable and converges to a unique solution that solves the dual problem, (18).

Proof: The proof follows the one in Srikant [2004]. Suppose $\widetilde{\mathbf{u}}$ is the unique solution of the original constrained system problem, (13). Then $\widetilde{\mathbf{q}}$ is also a unique solution since $\widetilde{q}_i = C_i'(\widetilde{u}_i)$. Moreover, recall that $\widehat{\mathbf{T}} = \begin{bmatrix} \mathbf{T} \\ -\mathbf{1}^T \end{bmatrix}$ with rank($\mathbf{T}$) $= m$ ($\mathbf{T}$ is invertible), such that $\widehat{\mathbf{T}}^T$ has full row rank. Thus $\widetilde{\lambda}$ is a unique optimal solution according to $\widetilde{\mathbf{q}} = \widehat{\mathbf{T}}^T\widetilde{\lambda}$. From Karush-Kuhn-Tucker (KKT) conditions, Bertsekas [1999], there exists a unique $\widetilde{\mathbf{u}}$ and $\widetilde{\lambda}$ satisfying the following conditions: $\widehat{b}_j = row_j(\widehat{\mathbf{T}})\widetilde{\mathbf{u}}$ or $row_j(\widehat{\mathbf{T}})\widetilde{\mathbf{u}} > \widehat{b}_j$ and $\widetilde{\lambda}_j = 0$. Next we consider the following Lyapunov function,

$$V_d(t) = \sum_{j=1}^{M+1}(row_j(\widehat{\mathbf{T}})\widetilde{\mathbf{u}} - \widehat{b}_j)\lambda_j - \sum_{i=1}^{M}\int_{\widetilde{q}_i}^{q_i}(\widetilde{u}_i - C_i^{'-1}(\xi))d\xi \qquad (29)$$

By using $\mathbf{q} = \widehat{\mathbf{T}}^T\lambda$, which leads to $\dot{\mathbf{q}} = \widehat{\mathbf{T}}^T\dot{\lambda}$, we get

$$\frac{dV_d(t)}{dt} = (\widehat{\mathbf{T}}\mathbf{u} - \widehat{\mathbf{b}})^T\dot{\lambda}$$

Using the vector form of the dual algorithm, (25), i.e., $\dot{\lambda} = diag(k_i)(\widehat{\mathbf{b}} - \widehat{\mathbf{T}}\mathbf{u})$, we obtain

$$\frac{dV_d(t)}{dt} = (\widehat{\mathbf{T}}\mathbf{u} - \widehat{\mathbf{b}})^T diag(k_i)(\widehat{\mathbf{b}} - \widehat{\mathbf{T}}\mathbf{u}) \leq 0$$

with $\frac{dV_d(t)}{dt} = 0$ only when the condition, $\widehat{b}_j = row_j(\widehat{\mathbf{T}})\widetilde{\mathbf{u}}$, is satisfied. It follows from the Lyapunov's stability theorem, Khalil [2002], that the unique solution is globally asymptotically stable. ■

5. SIMULATION RESULTS

The algorithms are simulated in MATLAB for a single point-to-point optical link in Fig.1. The number of channels will not affect the performance of the algorithms. Thus we consider that the link is with six channels and the link capacity threshold is $P_0 = 2.5$ mW. Within the set of channels, there are two levels of target OSNR, a 26 dB level desired on the first three channels and a 22 dB level on the next three channels. We note that the conditions on the target OSNR, (7) and (8), are satisfied such that the feasible constraint set is non-empty. The cost function for channel i is defined as

$$C_i(u_i) = \alpha_i u_i^2 - \beta_i \ln u_i,\ i = 1, \cdots, 6 \qquad (30)$$

where $\alpha_i > 0$ and $\beta_i > 0$. The selected cost function $C_i(u_i)$ is obviously strictly convex and continuously differentiable and $C_i(u_i) \to +\infty$ as $u_i \to 0$.

Due to space limitations, we present the simulation results of the primal algorithm (21) only. The coefficients in (30) is selected as $\alpha = 10^{-3} \times [3, 3, 3, 5, 5, 5]^T$ and $\beta = 10^{-3} \times [1.25, 1.5, 1.75, 0.9, 1, 1.1]^T$. The step-size is fixed for each channel with $k_i = 0.1$. The barrier function for channel i is selected as

$$\lambda_i(u_i) = 1000 \left(\max\{0, \hat{b}_i - row_i(\widehat{\mathbf{T}})\mathbf{u}\} \right)^6$$

such that $\lambda_i(u_i)$ is zero when the constraints are satisfied and there is a penalty with any violation of the constraints. The total power and channel OSNR vs time are shown in Fig.5 and Fig.6, respectively. By using the primal algorithm to adjust all channel powers, the channel OSNR levels converge to the corresponding desired values with the total power not exceeding the link capacity constraint.

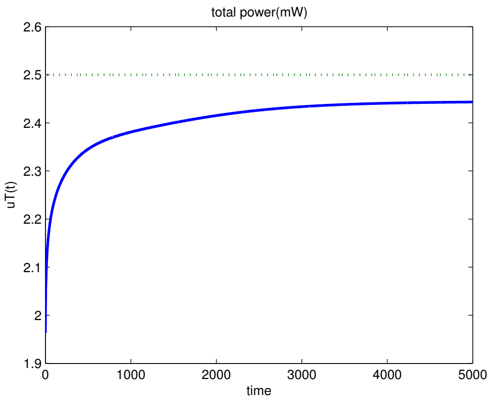

Fig. 5. The primal algorithm: Total Power vs Time

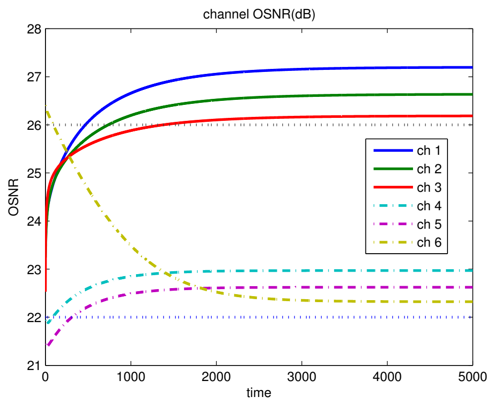

Fig. 6. The primal algorithm: Channel OSNR vs Time

6. CONCLUSION

We have studied a constrained OSNR optimization problem in optical links via a utility maximization approach. Each channel obtains at the minimum its individual fixed target OSNR level, and beyond that optimizes its input optical power level, according to its objective performance function, while respecting target OSNR levels of all other channels and link capacity constraint. The system optimization problem admits a unique solution given target OSNR levels for all channels which satisfy certain conditions. By using a barrier function, we have transformed the original constrained problem into an unconstrained optimization problem that can be solved using two algorithms, a primal algorithm and a dual algorithm, in a decentralized and iterative way. By using the Lyapunov function, we have established the global asymptotical stability of these algorithms. Extension of the results from single link case to network case is an interesting future research direction. One of the difficulties is that in the network case, the constraint set is no longer automatically convex, due to the power propagation, Pan and Pavel [2007a].

REFERENCES

G. P. Agrawal. *Fiber-optic communication systems*. John Wiley, 3rd edition, 2002.

T. Alpcan, X. Fan, T. Basar, M. Arcak, and J. T. Wen. Power control for multicell CDMA wireless networks: A team optimization approach. *Wireless Networks (Online first), http://dx.doi.org/10.1007/s11276-006-0006-5*, 2007.

T. Basar and G. J. Olsder. *Dynamic noncooperative game theory*. SIAM Series Classics in Applied Mathematics, 2nd edition, 1999.

D. P. Bertsekas. *Nonlinear programming*. Athena Scientific, 2nd edition, 1999.

M. Fiedler. *Special matrices and their applications in numerical mathematics*. Martinus Nijhoff Publishers, 1986.

F. Forghieri, R.W. Tkach, and D.L. Favin. Simple model of optical amplifier chains to evaluate penalties in WDM systems. *IEEE/OSA Journal of Lightwave Technology*, 16(9):1570–1576, 1998.

R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, 1999.

F. P. Kelly, A. K. Maulloo, and D. Tan. Rate control for communication networks: Shadow prices, proportional fairness and stability. *Journal Oper. Res. Soc.*, 49(3): 237–252, 1998.

H. K. Khalil. *Nonlinear systems*. Prentice Hall, 3rd edition, 2002.

A. Mecozzi. On the optimization of the gain distribution of transmission lines with unequal amplifier spacing. *IEEE Photonics Technology Letters*, 10(7):1033–1035, 1998.

B. Mukherjee. WDM optical communication networks: Progress and challenges. *IEEE Journal on Selected Areas in Communications*, 18(10):1810–1824, 2000.

Y. Pan and L. Pavel. Games with coupled propagated constraints in optical networks: the multi-link case. *Proc. 46th IEEE Conf. Decision and Control*, pages 3443–3449, Dec. 2007a.

Y. Pan and L. Pavel. Global convergence of an iterative gradient algorithm for the Nash equilibrium in an extended OSNR game. *Proc. IEEE INFOCOM*, pages 206–212, May 2007b.

L. Pavel. OSNR optimization in optical networks: modeling and distributed algorithms via a central cost approach. *IEEE Journal on Selected Areas in Communications*, 24(4):54–65, 2006.

L. Pavel. An extension of duality to a game-theoretic framework. *Automatica*, 43(2):226–237, 2007.

R. Srikant. *The mathematics of Internet congestion control*. Systems & Control: Foundations & Applications. Birkhauser, Boston, 2004.