Extremum seeking control of cascaded Raman optical amplifiers

Peter M. Dower, *Member, IEEE,* Peter M. Farrell, and Dragan Nešić, *Senior Member, IEEE*

***Abstract*— This paper considers the control of a particular type of optical amplifier that finds application in long-haul wavelength division multiplexed optical communications systems. The objective of this consideration is to demonstrate an application of extremum seeking to the regulation of amplifier output signal power across a range of signal wavelengths, where limited control authority is available. Although such amplifiers are nonlinear and distributed parameter devices, an extremum seeking design is demonstrated to be a promising approach for achieving the stated amplifier control objectives.**

***Index Terms*— Extremum seeking, optical amplifier control, signal power regulation, nonlinear PDEs.**

I. Introduction

A Raman amplifier is a distributed electro-optical device that finds application in submarine DWDM (dense wavelength division multiplexed) optical communications systems where improved signal to noise ratio is required and when amplifier spacings are greater than that suitable for erbium doped fiber amplifiers (EDFAs) [1], [2].

Such amplifiers impart broad spectrum optical gain over a range of modulated optical signal wavelengths through a photon coupling process known as stimulated Raman scattering (SRS) [3]. In that process, mixing of short and long wavelength molecular vibrations generated by the absorption of pairs of high and low energy photons results in the emission of pairs of long wavelength photons (and heat quanta, called optical phonons), thereby amplifying the long wavelength signal [3], [4], [5]. In Raman amplifiers, this coupling mechanism is exploited by the injection of short wavelength *pump* photons (via a set of pump lasers), which then interact with the longer wavelength signal photons in a spatially distributed fashion along the fiber to yield the optical (Raman) signal gain mentioned. By this same process, coupling across signal wavelengths also occurs, giving rise to a common source of crosstalk in optical communication systems.

In implementation, a Raman amplifier consists of a set of pump lasers (typically with 2 to 8 different wavelengths) coupled to a optical fiber span of hundreds of kilometres in length. Propagating through this fiber span are typically 40 to 160 DWDM signal channels (wavelengths), all subject to wavelength dependent loss, crosstalk, propagation delay and so on. With the objective of approximately inverting the response of the signal channels, output signal power regulation across the entire signal spectrum is highly desirable. This output power regulation is typically implemented in open-loop, or using ad-hoc control strategies, through the measurement and tuning of the Raman gain via adjustment of the pump laser powers. Naturally, this means that the output signal power obtained in practical implementations can suffer from significant degradation due to upstream signal power variations and other external influences. The potential and utility for more rigorous feedback control design of such amplifiers is thus clear: to facilitate robust signal power regulation in the presence of upstream signal power and other uncertainties.

A general approach to solving the problem of robust optimal control in optical amplifiers consists of two steps. First, we need to optimize the operation of the system at its equilibrium so that the optimal operation of the system in steady state is guaranteed. Second, a robust controller needs to be designed to regulate the transients towards the desired optimal equilibrium. The purpose of this paper is to investigate the use of extremum seeking control in addressing the first step of this approach. Transient control for Raman amplifiers via the loop shaping for the linearized finite dimensional model is addressed in [6].

Extremum seeking control (ESC) is a paradigm whose principal aim is finding an extremum of an unknown reference-to-output steady state map [7]. This problem arises in a range of different engineering applications, including control of electromechanical valve actuators and axial flow compressors [8], [9]. While ESC is quite old, its rigorous stability analysis was only recently addressed and established. Indeed, local stability properties of a class of extremum seeking algorithms were proved for the first time in [10] and, more recently, non-local stability was shown under stronger conditions in [11]. The schemes presented in [10], [11] are adaptive in nature and their analysis is done via the averaging and singular perturbation techniques. Stability of a very different type of ESC that is more related to the classical nonlinear optimization and nonlinear programming methods was presented in [12]. This ESC was shown to yield stability with extremum seeking under very weak conditions allowing, for instance, non-smooth reference-to-output maps, infinite dimensional attractors for the system dynamics and multiple control inputs. The results in [12] provide a theoretical framework for the ESC design presented in this paper.

The main contribution of this paper is the implementation of an ESC algorithm to optical Raman amplifiers that is designed using the framework proposed in [12] to provide on-line optimization of the steady state operation for these devices [13]. In particular, the objective considered in this paper is to minimize the variation of output amplifier optical signal powers from a desired power over the entire operating

The authors are with the Department of Electrical & Electronic Engineering, University of Melbourne (e-mail: {p.dower, p.farrell, d.nesic}@ee.unimelb.edu.au).

signal wavelength band by adjusting the optical powers and/or wavelengths of the available optical pump lasers (actuators). We emphasize that our method is applicable to other types of optical amplifiers, such as erbium doped fibre amplifiers (EDFAs) [14], as well as other objective functions that are not considered here for reasons of brevity. Moreover, this method applies to co-propagating and counter-propagating situations, as well as discrete or distributed Raman amplifiers (see [1]). We emphasize that a theoretical justification and stability analysis of this type of ESC is given in [12] and is not considered in this paper. Instead, we emphasize the problem formulation, the algorithm itself and the simulations that illustrate that the desired closed loop performance is indeed achieved.

In terms of organization, Section II describes the operation of Raman amplifiers and provides a PDE model (with boundary conditions) that captures that operation. Section III then describes in detail the general formulation of ESC and its application to Raman amplifier control. Finally, Section IV illustrates the operation of the resulting closed loop system by means of simulation.

II. Raman amplifier operation

A. Physics

Optical signals propagating in a single mode optical fibre are subject to attenuation, dispersion and nonlinear distortion [15]. Attenuation in optical fibres occurs as a result of scattering from unavoidable inhomogeneities in the glass and from electronic and vibrational resonances due to the glass itself and to impurities. The combination of these gives rise to a series of low loss windows in optical fibre including one in the region of 1550 nm of particular interest here. The absorption resonances also give rise to chromatic dispersion which combines with the waveguide dispersion of the fibre. The effect of this is that light waves at different wavelengths or frequency will travel at different velocities giving rise to pulse broadening, inter-symbol interference in digital systems, and phase effects in nonlinear energy transfer. A variety of nonlinear effects occur in optical fibres through coupling of light fields to the fibre materials. For the purposes of this work, the Raman effect allows the transfer of optical power from a light field at shorter wavelength (and higher photon energy) to a light field at longer wavelength [16] through an energy dissipative, photon conserving, scattering process called stimulated Raman scattering, or SRS. By manipulating optical power propagating at shorter wavelengths, it is possible to increase the optical power propagating at longer wavelengths. Locating the signal spectrum at wavelengths longer than that of a controllable "pump" laser spectrum, the signal spectrum can thus be influenced. This process yields the signal gain fundamental to Raman amplifier operation. This same process also gives rise to inter-signal and inter-pump cross-talk, generating a common source of noise in Raman amplifier systems (the other is ASE, or amplified stimulated emission).

B. Physical setup

The Raman amplifier considered here operates on an optical fibre link that consists of the following components, systems and signals:

1) A *distributed* optical fibre span $\mathcal{O}^-$ of length L^-.
2) A *discrete* optical fibre span $\mathcal{O}^+$ of length L^+.
3) A collection of $n_p = n_p^- + n_p^+$ pump lasers (actuators) with individual power control inputs given by the elements of the vector $u^\mp \in \mathbf{R}^{n_p^\mp}$, centered at a set of discrete, non-overlapping wavelengths
$$\Lambda^{p\mp} = \left\{ \lambda_i^{p\mp} \,|\, i = 1, 2, \ldots, n_p^\mp \right\}.$$
4) A collection of $n_p^\mp$ optical coupling devices, collectively denoted $B^\mp$, to launch the various pump wavelengths into the fibre spans at particular locations.
5) A collection of n_s exogenously generated and applied signals (data), each with an individual power given by the elements of the vector $d \in \mathbf{R}^{n_s}$, centered at a set of discrete, non-overlapping wavelengths
$$\Lambda^s = \{ \lambda_i^s \,|\, i = 1, 2, \ldots, n_s \}.$$
6) A collection of ${n_s}^\mp$ optical coupling devices and sensors, collectively denoted by $H^\mp$ to measure the propagating signal power vectors $y^\mp$, indexed by signal wavelength, at two locations along the fibre.

Remark 1: "Distributed" here means that the amplification process occurs in a fibre that is also used for transmission, and hence the fibre ends are located tens or hundreds of kilometres apart. "Discrete" refers to the case where the amplifying fibre remains on a spool such that both ends of the fibre are colocated. The distributed or discrete nature of a span is essential in determining whether signal measurements at one end of a fibre can be fed back sufficiently fast for compensation purposes.

The components and systems listed above can be assembled into the Raman amplifier setup $\mathcal{L}$ shown in Figure 1, which consists of a distributed *backward pumped* amplifier $\mathcal{L}_b$ and a discrete *forward pumped* amplifier $\mathcal{L}_f$ in cascade. This setup attempts to combine the low noise advantages of backward pumping with the speed of forward pumping [2]. $\mathcal{L}_b$ is chosen to be discrete so as to ensure colocation of all sensors and actuators, thereby avoiding any networked control issues.

In this setup, signal photons are injected via d by some external network or device, and propagate downstream from left to right. Along the way, these signal photons may be either absorbed by a photo-diode to generate an electrical signal $y^\mp$, lost due to Rayleigh scattering in the fibre span $\mathcal{O}_a$ or $\mathcal{O}_b$, or undergo a change in wavelength due to simulated Raman scattering [5]. Those remaining signal photons leave the amplifier via s.

Pump photons, on the other hand, can be injected by some external network or device $v^\mp$, or by a local pump laser $u^\mp$. These pump photons propagate in their respective launch directions (either upstream or downstream), and are subject to Rayleigh and Raman scattering in an analogous manner to signal photons. Those remaining pump photons leave the amplifier at either the upstream end via w^- or the downstream end via w^+.

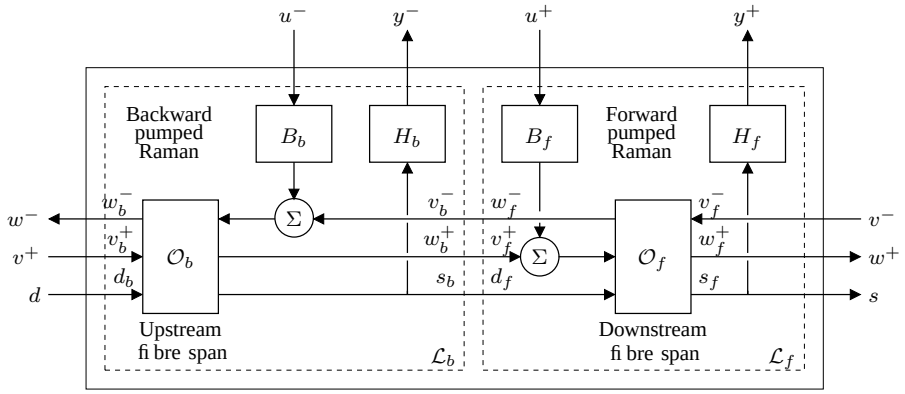

Fig. 1. A forward and backward pumped Raman amplifier cascade.

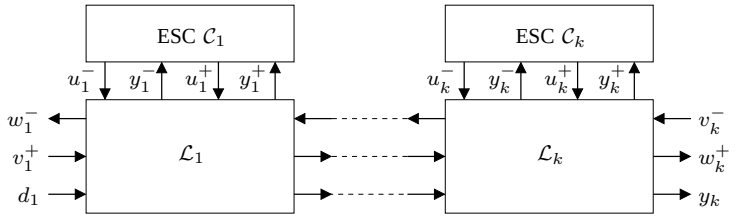

Fig. 2. An optical link consisting of a cascade of cascades of forward / backward pumped Raman amplifiers.

This interpretation of inputs and outputs can be summarized by an input / output representation of $\mathcal{L}$ of the form

$$\left[\begin{array}{c}\text{propagating}\\ \text{outputs}\\ \hline \text{measured}\\ \text{outputs}\end{array}\right] = \mathcal{L}\left(\left[\begin{array}{c}\text{propagating}\\ \text{inputs}\\ \hline \text{control}\\ \text{inputs}\end{array}\right]\right)$$

so that

$$\left[\begin{array}{c}w^-\\ w^+\\ s\\ \hline y^-\\ y^+\end{array}\right] = \mathcal{L}\left(\left[\begin{array}{c}v^-\\ v^+\\ d\\ \hline u^-\\ u^+\end{array}\right]\right) \tag{1}$$

The partitioning of propagating and control inputs employed in (1) emphasizes the distinction between photons propagating from elsewhere in the optical network (considered as disturbances) and photons injected locally by the controllable pump lasers. An analogous distinction is made with respect to outputs.

C. Power dynamics across a fibre span

In Raman amplifier fibre spans, SRS is the dominant physical process. Consequently, modelling of the fibre spans $\mathcal{O}^{\mp}$ utilized in the amplifier of Figure 1 must incorporate SRS dynamics. As the amplifier spans differ in specification only (not in the underlying physical process), it is sufficient to consider a single span $\mathcal{O}$, as shown in Figure 3.

Although the complete dynamics of SRS requires an application of the nonlinear Schrödinger equation, a simpler average field power model [2] suffices for Raman amplifier fibre spans. This model is described by the first order nonlinear

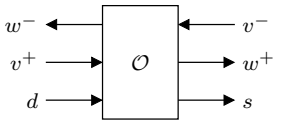

Fig. 3. A Raman amplifier fibre span $\mathcal{O}$.

vector PDE

$$\tau\frac{\partial \mathbf{P}}{\partial t} + \mu\frac{\partial \mathbf{P}}{\partial z} = -A\mathbf{P} + \mathrm{diag}(\mathbf{P})C\mathbf{P} \tag{2}$$

where $\mathbf{P} \equiv \mathbf{P}(t,z)$ is the augmented vector of backward propagating pump powers $\mathbf{P}^{p-}$, forward propagating pump powers $\mathbf{P}^{p+}$, and forward propagating signal power $\mathbf{P}^{s}$, all at time t and distance z from the upstream end of the span. Explicitly,

$$\mathbf{P}(t,z) = \left[\begin{array}{c}\mathbf{P}^{p-}(t,z)\\ \mathbf{P}^{p+}(t,z)\\ \mathbf{P}^{s}(t,z)\end{array}\right], \tag{3}$$

where the units are those of power. In (2), τ, μ, A, $C \in \mathbf{R}^{N\times N}$, $N := n_p + n_s$, represent respectively matrices of propagation delays per unit length, propagation directions, losses per unit length, and coupling coefficients per unit length. τ, μ and A are diagonal matrices, with the i^{th} diagonal entry corresponding to the wavelength of the i^{th} entry in $\mathbf{P}$. The units of τ and A are respectively $\mu s\,(km)^{-1}$, and $(km)^{-1}$. Matrix μ is dimensionless, with

$$\mu_{ii} := \begin{cases} -1 & \mathbf{P}_i \text{ counter-propagating,}\\ +1 & \mathbf{P}_i \text{ co-propagating.}\end{cases}$$

The ij^{th} entry of C denotes the coupling coefficient from the j^{th} to the i^{th} wavelength of $\mathbf{P}$. In the coordinates given, C

is not skew-symmetric (although it is when $\mathbf{P}$ is transformed to coordinates of photon flux). The operator $\mathrm{diag} : \mathbf{R}^N \rightarrow \mathbf{R}^{N\times N}$ denotes the mapping from vectors to diagonal matrices. The numerical computation of the matrices τ, A and C is covered briefly in Appendix I.

In order to finalize a model for span $\mathcal{O}$, it is important to note that PDEs of the form of (2) require the specification of one spatial and one temporal boundary condition per entry in $\mathbf{P}$. The spatial boundary conditions are specified by the vector of backward propagating pump powers v^-, the vector of forward propagating pump powers v^+, and the vector of forward propagating signal powers d. The temporal initial condition is specified by an initial spatial power distribution $\xi \in \mathcal{B}[0, L_\mathcal{O}]$, where this space is that of bounded functions defined over the span. Consequently, the span $\mathcal{O}$ can be modelled by a parameterized flow $\mathcal{S}$ and output map h, where

$$\begin{aligned}
\mathcal{S}_\mathcal{O}(\xi,\theta) \;&:=\; \left\{ \begin{array}{c} \text{Solutions of PDE (2) on span } \mathcal{O}, \\ \text{initialized with power distribution} \\ \mathbf{P}(0,\cdot) = \xi \in \mathcal{X}_\mathcal{O}, \text{ with boundary} \\ \text{conditions} \\ \left[\begin{array}{c} \mathbf{P}^{p-}(\cdot, L_\mathcal{O}) \\ \mathbf{P}^{p+}(\cdot, 0) \\ \mathbf{P}^{s}(\cdot, 0) \end{array} \right] \equiv \theta \\ v^{\mp} : \mathbf{R} \rightarrow \mathbf{R}^{n_p^{\mp}},\; d : \mathbf{R} \rightarrow \mathbf{R}^{n_s} \end{array} \right. \\
h_\mathcal{O}(\phi) \;&:=\; \left[\begin{array}{c} \phi^{p-}(\cdot, 0) \\ \phi^{p+}(\cdot, L_\mathcal{O}) \\ \phi^{s}(\cdot, L_\mathcal{O}) \end{array} \right],\; \phi = \left[\begin{array}{c} \phi^{p-} \\ \phi^{p+} \\ \phi^{s} \end{array} \right] \in \mathcal{S}_\mathcal{O}(\xi,\theta)
\end{aligned} \tag{4}$$

Here, $\xi \in \mathcal{X}_\mathcal{O} := (\mathcal{B}[0, L_\mathcal{O}], \|\cdot\|_1)$ is the initial spatial power distribution,

$$\|\xi\|_1 \;:=\; \int_0^{L_\mathcal{O}} |\xi(z)|_1 \, dz \,,$$

and $L_\mathcal{O}$ is the length of span $\mathcal{O}$. Assuming steady state conditions for $t < 0$, a candidate initial state $\xi \in \mathcal{X}_\mathcal{O}$ can be determined by solving the boundary constrained ODE

$$\mu \frac{\partial \xi}{\partial z} \;=\; -A\xi + \mathrm{diag}(\xi) C \xi \quad s.t. \quad \left\{ \begin{array}{c} \xi^{p-}(L_\mathcal{O}) = v_\circ^- \\ \xi^{p+}(0) = v_\circ^+ \\ \xi^{s}(0) = d_\circ \end{array} \right. \tag{5}$$

where $v_\circ^-$, $v_\circ^+$ and $d_\circ$ represent the steady state backward pump, forward pump and upstream signal powers prevailing for $t < 0$. There, the $p^{\mp}$ and s superscripts denote the components of ξ defined as per the components of (3).

Remark 2: [12] Where the state of all members of a parameterized family of systems can be represented as elements of a Banach space, the trajectories of these systems can be described collectively in terms of a flow. In general, a flow $\mathcal{S}$ on a Banach space $\mathcal{X}$ and a parameter space $\mathbf{R}^q$ ($q \geq 1$ fixed) defines a set of continuous functions indexed by an initial condition $\xi \in \mathcal{X}$ and a parameter $\theta \in \mathbf{R}^q$. In particular, a trajectory $\phi(\cdot, \xi, \theta) : \mathbf{R} \rightarrow \mathcal{X}$ of a system selected via a choice of parameter $\theta \in \mathbf{R}^q$ and initialized at $\xi \in \mathcal{X}$ must be an element of $\mathcal{S}(\xi, \theta)$. Furthermore, any such function $\phi(\cdot, \xi, \theta) \in \mathcal{S}(\xi, \theta)$ must satisfy the following two conditions: (i) $\phi(0, \xi, \theta) = \xi$ for all $\xi \in \mathcal{X}$ (the initial condition property); (ii) $\phi(t_1 + \cdot, \xi, \theta) = \phi(\cdot, \phi(t_1, \xi, \theta), \theta)$ for all $t_1 \geq 0$, $\xi \in \mathcal{X}$ and $\theta \in \mathbf{R}^q$ (the semi-group property). This is consistent with the interpretation of $\phi(t, \xi, \theta)$ as the solution or state of the particular system, as selected by parameter θ, at time t, starting from the initial condition ξ. Where the state of a family of systems with trajectories defined by flow $\mathcal{S}$ is measured via some output map $h : \mathcal{X} \rightarrow \mathbf{R}^p$, that family can be represented by the flow / output map pair $(\mathcal{S}, h)$.

Example: A flow and output map pair $(\mathcal{S}_\mathcal{O}, h_\mathcal{O})$ as per (4) is used to model the fibre span $\mathcal{O}$ of Figure 3. With a given initial spatial power distribution ξ, a known backward pump power v^-, zero forward pump power $v^+ \equiv 0$, and a known signal power d, a trajectory $\phi(\cdot, \xi, \theta)$ can be selected from the flow $\mathcal{S}_\mathcal{O}(\xi, \theta)$ by fixing

$$\theta \;:=\; \left[\begin{array}{c} v^- \\ 0 \\ d \end{array} \right].$$

Applying an output map h that selects one downstream signal channel as the output, this trajectory yields the spatial and temporal output transient illustrated in Figure 4. Here, the slow backward propagating transient observed on the left is due to a variation in v^-, whilst the fast forward propagating transient on the right is due to a variation in d.

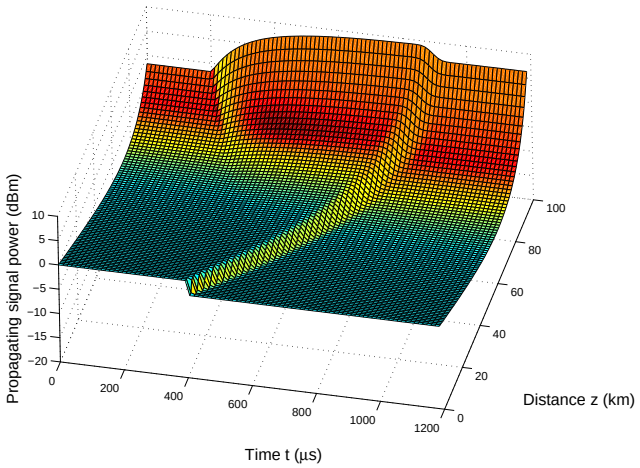

Fig. 4. Example of a spatially distributed signal power transient arising in a fibre span.

D. Raman amplifiers and cascades

Using the fibre span concept modelled by (4), it is possible to formally construct the component amplifiers and cascades shown in Figures 1 and 2.

Backward pumped amplifier $\mathcal{L}_b$: The amplifier $\mathcal{L}_b$ consists of the span $\mathcal{O}_b$ of length L_b, modelled by a flow and output map pair $(\mathcal{S}_b, h_b)$, and the input and output operators B_b and H_b that represent actuation and measurement respectively.

There,

$$\mathcal{S}_b\left(\xi_b,\left[\begin{array}{c} v_b^- \\ v_b^+ \\ \hline d_b \\ \hline u^- \end{array}\right]\right) := \mathcal{S}_{\mathcal{O}_b}\left(\xi_b,\left[\begin{array}{c} v_b^- + B_b(u^-) \\ \hline v_b^+ \\ \hline d_b \end{array}\right]\right) \ni \phi_b\,,$$
$$h_b(\phi_b) := \left[\begin{array}{c} h_{\mathcal{O}_b}(\phi_b) \\ \hline H_b(\phi_b^s(\cdot, L_b)) \end{array}\right], \tag{6}$$

$\xi_b \in \mathcal{X}_b$ is the initial state of $\mathcal{L}_b$, and $(\mathcal{S}, h)$ is as per (4). The partitions shown are as per (1). This captures the operation of $\mathcal{L}_b$ in Figure 1.

Forward pumped amplifier $\mathcal{L}_f$: The amplifier $\mathcal{L}_f$ utilizing the span $\mathcal{O}_f$ of length L_f is modelled in a similar way,

$$\mathcal{S}_f\left(\xi_f,\left[\begin{array}{c} v_f^- \\ v_f^+ \\ \hline d_f \\ \hline u^+ \end{array}\right]\right) := \mathcal{S}_{\mathcal{O}_f}\left(\xi_f,\left[\begin{array}{c} v_f^- \\ \hline v_f^+ + B_f(u^+) \\ \hline d_f \end{array}\right]\right) \ni \phi_f\,,$$
$$h_f(\phi_f) := \left[\begin{array}{c} h_{\mathcal{O}_f}(\phi_f) \\ \hline H_f(\phi_f^s(\cdot, L_f)) \end{array}\right] \tag{7}$$

where $\xi_f \in \mathcal{X}_f$ is the initial state of $\mathcal{L}_f$, and $(S_{\mathcal{O}_f}, h_{\mathcal{O}_f})$ is as per (4). As in the backward pumped case, this captures the operation of $\mathcal{L}_f$ in Figure 1.

Amplifier cascade $\mathcal{L}$: The Raman amplifier cascade $\mathcal{L}$ can then be modelled by combining (6) and (7), yielding the product flow and output map pair $(\mathcal{S}_\times, h_\times)$, where

$$\mathcal{S}_\times\left((\xi_b,\xi_f),\left[\begin{array}{c} v^- \\ v^+ \\ \hline d \\ \hline u^- \\ u^+ \end{array}\right]\right)$$
$$:= \left\{(\phi_b,\phi_f)\;\middle|\;\begin{array}{rcl} \phi_b & \in & \mathcal{S}_b\left(\xi_b,\left[\begin{array}{c} h_f(\phi_f)|_{w_f^-} \\ v^+ \\ \hline d \\ \hline u_- \end{array}\right]\right) \\ \phi_f & \in & \mathcal{S}_f\left(\xi_f,\left[\begin{array}{c} v^- \\ h_b(\phi_b)|_{w_b^+} \\ \hline h_b(\phi_b)|_{s_b} \\ \hline u^+ \end{array}\right]\right) \end{array}\right\}$$
$$h_\times\left((\phi_b,\phi_f)\right) := \left[\begin{array}{c} h_b(\phi_b)|_{w_b^-} \\ h_f(\phi_f)|_{w_f^+} \\ \hline h_f(\phi_f)|_{d_f} \\ \hline H_b(\phi_b^s(\cdot, L_b)) \\ H_f(\phi_f^s(\cdot, L_f)) \end{array}\right] \tag{8}$$

The fact that $\mathcal{S}_\times$ describes a flow follows immediately as the defining interconnections are the result of a spatial cascade of two PDEs. The mappings from parameter to output generated by (8) are then of the form (1). Modelling of other cascades, such as the optical link $\mathcal{L}_1, \cdots, \mathcal{L}_k$ of Figure 2, follows in a similar fashion.

E. Comparison with EDFAs

Erbium Doped Fibre Amplifiers are much more commonly used than Raman amplifiers in optical fibre communication systems and their control has aroused much more interest. From the perspective of the issues dealt with in this paper, the major differences between EDFAs and Raman amplifiers are as follows:

(a) EDFAs are almost universally discrete amplifiers, consisting of a few tens of metres of erbium doped glass fibre, while Raman amplifiers are usually distributed, utilizing up to one hundred kilometres or more of standard transmission fibre;
(b) The dynamics of Raman amplifiers are dominated by the transmission and time of flight, whereas EDFA dynamics are dominated by the dynamics of the erbium ions doped in the glass fibre. There, the Er^{3+} excited state which supplies the gain has a lifetime of around ten milliseconds, whilst the excited state is populated or depleted by absorption and stimulated emission processes that occur over timescales in the range of microseconds to tens of milliseconds [14].
(c) Signal spectral shape in EDFAs can be controlled via the simultaneous actuation of a number of pump lasers that deliver power to different sections of a multistage amplifier. An extreme case of this spectral shaping is presented in [17]. In Raman amplifiers, the same objective can be achieved via the simultaneous actuation of several co-located or distributed pump lasers that impart gain to particular bands of the signal spectrum, forming the basis of this paper.

Whilst both types of amplifiers are naturally modelled via partial differential transport equations, the short fibre lengths utilized in EDFAs allows a lumped approximation of that model, yielding a much simpler ODE structure. Thus, the application of ESC to EDFA control would in principle be simpler, as the conditions to be checked in [12] are significantly simpler.

Furthermore, the discrete nature of EDFAs allows access for essentially simultaneous measurement to both the input and output end of the amplifier, allowing simpler gain control, spectral gain flattening and transient suppression. Such access becomes problematic for distributed Raman amplifiers since the results of the measurement at the remote end cannot be received at the proximal end significantly before the arrival of the amplified signal itself. These distinguishing aspects of Raman amplifiers recommended themselves to the authors as additionally challenging compared to EDFAs.

III. Extremum seeking

A. Background material

Recent efforts [12] in extremum seeking have yielded a design framework and closed loop stability analysis results for ESC under quite general conditions. As this framework is applicable to the optical amplifier control problem considered here, a summary of the relevant details follows as a precursor to consideration of the application at hand.

Consider a system that is representable by a flow $\mathcal{S}$ and an output map $h : \mathcal{X} \to \mathbf{R}^p$, so that a trajectory of the system state and output are given respectively by $\phi \in \mathcal{S}(\xi, \theta)$ and $h(\phi)$, where $\xi \in \mathcal{X}$ is the initial system state in the Banach space $\mathcal{X}$, and $\theta \in \mathbf{R}^p$ is the parameter (or input value) that selects a particular trajectory (or subset of trajectories).

Let $J : \mathbf{R}^p \to \mathbf{R}$ denote a continuous "cost" function, selected by the user, that quantifies some trajectory-based measure of performance of interest. Suppose that for any fixed parameter value $\theta \in \mathbf{R}^q$, the output of the system converges to a limit that is uniquely determined by the particular input. Continuity of J then implies that the following input-to-output map, called a "readout map", is well defined:

$$g(\theta) := \lim_{t \to \infty} J \circ h(\phi(t, \xi, \theta)) \,. \tag{9}$$

Suppose that the map g defines a non-empty set Θ^* of minimizers (or maximizers), denoted by

$$\Theta^* := \{\theta \in \mathbf{R}^q \,|\, g(\theta) \leq g(\omega) \,\forall\, \omega \in \mathbf{R}^q\} \neq \emptyset \,.$$

Then, under some additional technical assumptions [12], the ESC design framework presented therein yields a controller that minimizes (or maximizes) the readout map g. The steps that generate this ESC design are as follows:

0) Problem formulation,
1) Optimization algorithm design,
2) Readout map approximation,
3) ESC algorithm construction.

These steps are elucidated in the remainder of this section.

0) Problem formulation: The "zeroth" step in the ESC design approach considered in [12] is to identify the parameterized flow $\mathcal{S}$, the output map h, and an appropriate cost function J consistent with whatever control objective is of interest. (Naturally, this is largely application specific.)

1) Optimization algorithm design: With the problem formulated as described above, the next step is to design or select an algorithm $\mathcal{A}$ that optimizes the static readout map $g : \mathbf{R}^q \to \mathbf{R}$, where knowledge of g is limited to a fixed number m of evaluations per iteration. (It is important to note that, by definition (9), algorithm $\mathcal{A}$ necessarily ignores transients in the flow ϕ.) In terms of generality, the results in [12] apply to algorithms in difference equation (or inclusion) form,

$$\theta_{k+1} = F(\theta_k, G(\theta_k)) \,. \tag{10}$$

Here, $F : \mathbf{R}^q \times \mathbf{R}^m \to \mathbf{R}^q$ defines a parameter update process, whilst $G : \mathbf{R}^q \to \mathbf{R}^m$ defines the vector of m allowable evaluations of g per iteration. More specifically, [12] requires that the measurement function G be defined by

$$G(\theta) := \begin{pmatrix} g(\theta + v_1(\theta)) \\ \cdots \\ g(\theta + v_m(\theta)) \end{pmatrix}$$

where v_i, $i = 1, 2, \ldots, m$, defines m "dither" functions. These functions can be used to capture the m readout map evaluations required for approximate gradient calculations, the conduct of line searches, and any other processes employed in the numerical optimization of the static map g. Indeed, with mild assumptions on g and F, G can be selected [12] such that the generated sequence of parameters $\{\theta_k\}$ converges to the set of minimizers Θ^*.

2) Readout map approximation: Although algorithm $\mathcal{A}$ necessarily optimizes the readout map g as $k \to \infty$ in (10), the limit definition (9) of g implies that $\mathcal{A}$ cannot be implemented as is in real time. Instead, the infinite horizon limit must be approximated over a finite horizon, defined by a "waiting time" denoted T. This yields the approximate readout map

$$g_T(t, \theta) := h(\phi(T + t, \xi_t, \theta)), \tag{11}$$

where $\xi_t = \phi(t, \xi, \theta_{[0,t]})$, and $\theta_{[0,t]}$ is the parameter selection applied over the preceding interval $[0, t]$. Subsequent evaluations of g required in algorithm $\mathcal{A}$ can then be approximated (see [12, Assumption 3]) by

$$G_T(t, \theta) := \begin{pmatrix} g_T(t, \theta + v_1(\theta)) \\ g_T(t + T, \theta + v_2(\theta)) \\ \cdots \\ g_T(t + (m-1)T, \theta + v_m(\theta)) \end{pmatrix} . \tag{12}$$

3) ESC algorithm construction: An ESC algorithm $\mathcal{C}$ can then be constructed by combining algorithm $\mathcal{A}$ with the approximation (12), yielding the iteration

$$\theta_{k+1} = F\left(\theta_k, G_T(kmT, \theta_k)\right) , \quad k = 0, 1, 2, \ldots \tag{13}$$

where $\theta_\circ$ is the initial parameter value used. It is worth emphasizing that the parameter update process F issues a sequence of m parameter step commands as dictated by G_T, yielding the sequence of parameter selections

$$\begin{aligned} &\theta_\circ + v_1(\theta_\circ), \ldots, \theta_\circ + v_m(\theta_\circ), \\ &\quad \theta_1 + v_1(\theta_1), \ldots, \theta_1 + v_m(\theta_1), \\ &\qquad \vdots \\ &\qquad \theta_k + v_k(\theta_k), \ldots, \theta_k + v_m(\theta_k), \\ &\qquad\quad \vdots \end{aligned}$$

each of which give rise to transients in ϕ that are allowed to decay over T seconds before the corresponding readout is measured. Under very general technical conditions, it was shown in [12] that this ESC algorithm yields semi-global practical convergence to the set of minimizers Θ^*, where the parameter that needs to be adjusted in the ESC algorithm is the waiting time T. We note that this approach provides a framework for ESC design since the numerical optimization algorithm (10), the dither functions v_i and the waiting time T are to be chosen by the designer depending on the plant dynamics and the cost function that needs to be minimized. We will discuss below the particular choices made in our ESC algorithm.

B. Application to Raman amplifier control

The control topology of interest is shown in Figure 2. The control objective is to shape the downstream signal power spectrum according to some pre-specified function. For simplicity, we consider the case where that function represents a flat signal spectrum. This is to be achieved via the feedback of the downstream signal power measurements y^- and y^+ through a controller $\mathcal{C}$ to the pump power inputs u^- and u^+,

where $\mathcal{C}$ utilizes the ESC algorithm described in the preceding section. The construction of $\mathcal{C}$ is documented in the remainder of this section. A brief discussion of the associated technical conditions [12] is included in Appendix II.

0) Problem formulation: The parameterized flow and output map pair $(\mathcal{S}, h)$ are given by

$$\begin{aligned} \mathcal{S}(\xi,\theta) &:= \mathcal{S}_\times\left(\xi, \left[\begin{array}{c} v^- \\ v^+ \\ \hline d \\ \theta \end{array}\right]\right) \ni \phi_\times \\ h(\phi_\times) &:= \left[\begin{array}{c|cc} 0_{n_s\times N} & 0_{n_s\times n_s} & I_{n_s\times n_s} \end{array}\right] h_\times(\phi_\times) \end{aligned}$$

where v^-, v^+, d are fixed apriori, $0_{n\times m}$ is the zero $n\times m$ matrix, $I_{n\times n}$ is the $n\times n$ identity matrix, and $\xi \in \mathcal{X}_b \times \mathcal{X}_f$ is the initial condition of the combined spans $\mathcal{O}_b$ and $\mathcal{O}_f$. The output map $h : \mathcal{X}_b \times \mathcal{X}_f \to \mathbf{R}^{n_s}$ selects the downstream signal power propagating through $\mathcal{O}_f$ as the only measurement.

The required control objective stated above can be implemented via minimization of the cost function

$$J(h) := \|h - h^*\|_Q^2 + \sum_{i,j=1, i\neq j}^{n_s} \left(\frac{h_i - h_j}{\lambda_i^s - \lambda_j^s}\right)^2$$

where $h^* \in \mathbf{R}^{n_s}$ is the reference downstream signal power, $\|h\|_Q := \sqrt{h'Qh}$, and $Q \in \mathbf{R}^{n_s\times n_s}$ is positive semi-definite. Here, the first term penalizes departures in output signal power from the reference h^*, whilst the second term penalizes ripple in the signal spectrum.

1) Optimization algorithm design: A standard steepest descent algorithm $\mathcal{A}$ is selected [18] for the minimization of the static readout map g defined by (9). In particular, $\mathcal{A}$ is comprised of the following steps:

Step (i): Gradient approximation. A finite difference approximation of the normalized gradient $\hat{n}$ of the readout map is computed via evaluation of the readout map on a simplex of radius δ centred on θ_k. This defines the dither functions v_i, $i = 1, \ldots, 2n_p + 1$.

Step (ii): Line search. A line search in the direction $\hat{n}$ of Step (i) is then conducted, requiring a further $n_{LS} \leq \log_2\left(\frac{|\hat{n}|}{TOL_{LS}}\right) + 1$ evaluations of the readout map, and defining the additional dither functions v_i, $i = 2n_p + 2, \ldots 2n_p + n_{LS} + 1$. The line search tolerance TOL_{LS} is fixed larger than the machine tolerance apriori. The line search minimizer θ_{LS} obtained from these n_{LS} evaluations is returned.

Step (iii): Interpolation. Using θ_{LS} of Step (ii), a quadratic function defined on $\mathbf{R}$ is fitted in the direction $\hat{n}$ of Step (i) through the three points $(\theta_k, g(\theta_k))$, $\left(\frac{\theta_k+\theta_{LS}}{2}, g\left(\frac{\theta_k+\theta_{LS}}{2}\right)\right)$, and $(\theta_{LS}, g(\theta_{LS}))$, defining an additional two dither functions, v_i, $i = 2n_p + n_{LS} + 2, 2n_p + n_{LS} + 3$. The minimizer, denoted θ_I, of this quadratic is returned.

Step (iv): Update. The parameter θ_k is updated to $\theta_{k+1} = \mathrm{argmin}_{\zeta\in\{\theta_k,\theta_{LS},\theta_I\}} g(\zeta)$. This requires no additional dither function definitions as the readout map has been evaluated at these parameter values previously.

Steps (i) to (iv) thus define the functions F and G of algorithm $\mathcal{A}$, utilizing $m := 2n_p + n_{LS} + 3$ dither functions.

2) Readout map approximation: This step requires the selection of a sufficiently long waiting time T such that g_T approximates sufficiently closely the static readout map g. A suitable under-bound for T can be obtained from examining the time constant T_* of an approximation of the flow $\mathcal{S}_\times$ used to model the amplifer (see [6]). In particular, it can be shown that

$$T_* := \max(T_b, T_f), \tag{14}$$

where $T_\bullet$, $\bullet \in \{b, f\}$, is defined by

$$T_\bullet := \max_{\substack{i=1\ldots n_s \\ j=1\ldots n_p}} \max\left(\tau_i^s L_\bullet, \tau_j^p L_\bullet, \frac{\tau_i^s \pm \tau_j^p}{\alpha_j^p}\right)$$

A guide to waiting time selection is thus to choose $T \gg T_*$.

3) ESC algorithm construction: The ESC algorithm $\mathcal{C}$ follows immediately from iteration (13).

Remark 3: It is important to note that due to the generality of the approach used here, ESC can readily be applied to other optical amplifier types with minimal additional work. In particular, the authors note that the inclusion of Erbium-doped fibre amplifiers (EDFAs) [14], [19] in the links described is immediately possible, as EDFA dynamics can also be modelled via a flow and an output map.

Remark 4: It has been assumed that measuring devices are available from which we can obtain optical power for each signal channel on a time scale commensurate with the waiting time T. This may not be practical due to cost, and the modification of the algorithm under conditions of incomplete information will be the subject of further work. An alternative regulation strategy is to equalize the optical signal to noise ratio across the signal channels instead of the signal power with a view to minimizing the bit error rate over all channels. This approach will also be considered in future work.

IV. Simulations

A. A backward pumped Raman amplifier

An extremum seeker that implements output signal power regulation for a backward pumped Raman amplifier $\mathcal{L}_b$ can be constructed as documented above. This is a special case of the control setup shown in Figures 1 and 2, where the absence of the forward pumped amplifier $\mathcal{L}_f$ is codified by selecting $B_f \equiv 0$, $L_f \equiv 0$, and $H_f \equiv 0$. This corresponds to removing the forward pump laser and the discrete downstream fibre span $\mathcal{O}_f$, thereby combining the backward and forward pump measurements $y^\mp$. This measurement is assumed (for simplicity) to be exact, so that $H_b \equiv 1$ (the identity operator). The precise details of the upstream fibre span $\mathcal{O}_b$ specifications are detailed below. The values used do not describe a particular amplifier implementation, but rather reflect typical values that might be used [2], [1], [20], [5], [15].

$$\begin{aligned} n_p &= 2 \quad (n_p^- = 2,\, n_p^+ = 0) \\ n_s &= 4 \\ \Lambda^{p-} &= \{1442,\, 1490\}\ nm \\ \Lambda^s &= \{1530,\, 1550,\, 1570,\, 1590\}\ nm \end{aligned}$$

$$\tau = \quad \text{diag}\left(\begin{bmatrix} 4.8767 \\ 4.8775 \\ 4.8783 \\ 4.8787 \\ 4.8791 \\ 4.8796 \end{bmatrix}\right) \mu s(km)^{-1}$$

$$A = \quad \text{diag}\left(\begin{bmatrix} 0.0642 \\ 0.0485 \\ 0.0415 \\ 0.0411 \\ 0.0431 \\ 0.0479 \end{bmatrix}\right) (km)^{-1}$$

$$C = \begin{bmatrix} 0 & -0.15 & -0.40 & -0.36 & -0.19 & -0.09 \\ 0.15 & 0 & -0.15 & -0.17 & -0.31 & -0.42 \\ 0.38 & 0.15 & 0 & -0.12 & -0.15 & -0.16 \\ 0.34 & 0.17 & 0.12 & 0 & -0.12 & -0.15 \\ 0.18 & 0.29 & 0.15 & 0.12 & 0 & -0.12 \\ 0.09 & 0.39 & 0.16 & 0.15 & 0.12 & 0 \end{bmatrix} (W\,km)^{-1}$$

$$L_b = \quad 100\,km$$
$$L_f = \quad 0\,km$$

This amplifier specification implies a time constant (14) of $T_* \approx 200\,\mu s$. To ensure that all amplifier transients have decayed sufficiently between successive measurements, a waiting time of $T = 1200\,\mu s$ is utilized here.

Combining the backward pumped Raman amplifier $\mathcal{L}_b$ and the ESC algorithm $\mathcal{C}$ obtained from the design steps in Section III-B yields a closed loop system as per Figure 2, where only $k = 1$ amplifier cascade is present. Recall that whilst the amplifier maps pump inputs to downstream signal powers, the extremum seeker and readout map together map downstream signal power measurements to amplifier pump inputs. Simulating the span $\mathcal{O}_b$ using a finite difference solver, Figure 5 illustrates the dither generated by the extremum seeker in the execution of the gradient descent algorithm, the resulting transient in the downstream signal power, and the convergent trend observed in the downstream signal power. Note that the elapsed time between transients observed is due to the waiting time implemented. Given that the pump and signal powers are observed to converge to constant values during that interval, it is clear that the waiting time selected is sufficiently long.

In demonstrating that the ESC implemented does minimize the cost function J, Figure 6 illustrates that the pump power trajectory achieved does converge to an element of Θ^*. Whilst illustrative from a simulation point of view, it is important to note that the cost surface shown may not be known (or readily evaluated) in practice. Figures 7 and 8 illustrate that downstream signal power spread and deviation prior to and after application of ESC, showing a clear reduction in those quantities.

B. Two pump, four signal backward and forward pumped Raman amplifier

A cascaded amplifier consisting of both backward and forward pumped components $\mathcal{L}_b$ and $\mathcal{L}_f$ is now considered. Here, the forward pump lasers and discrete downstream fibre span are reinstated, with $B_f \equiv 1$ (again, the identity operator),

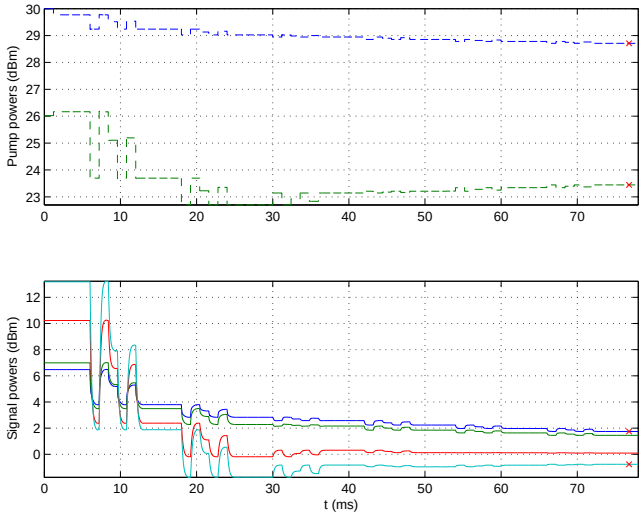

Fig. 5. Input dither and amplifi er output transients generated by the extremum seeker.

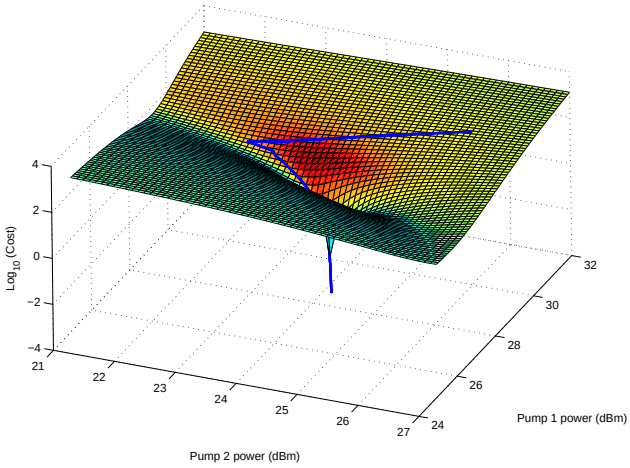

Fig. 6. Readout map is minimized by the generated amplifi er input power trajectory.

whilst the measurement utilized is shifted from y^- to y^+ (that is, $H_b \equiv 0$, $H_f \equiv 1$). Otherwise, the specification of the cascaded amplifiers are identical to that of Section IV-A with the exception of the following:

$$n_p = \quad 2 \quad (n_p^- = 1,\, n_p^+ = 1)$$
$$L_b = \quad 100\,km$$
$$L_f = \quad 100\,km$$

Using the same simulation technique as in Section IV-A, the corresponding transient response of the closed loop (amplifiers and ESC) can be obtained. The resulting steady state pump and signal power distribution before and after application of ESC is shown in Figures 9 and 10. Note that $z = 100\,km$ corresponds to the location of the pump lasers (backward and forward), sensors and the controller $\mathcal{C}$ implementing the ESC algorithm. It is evident that the signal power spread and deviation from the set-point at the downstream end of the discrete downstream fibre span has been reduced.

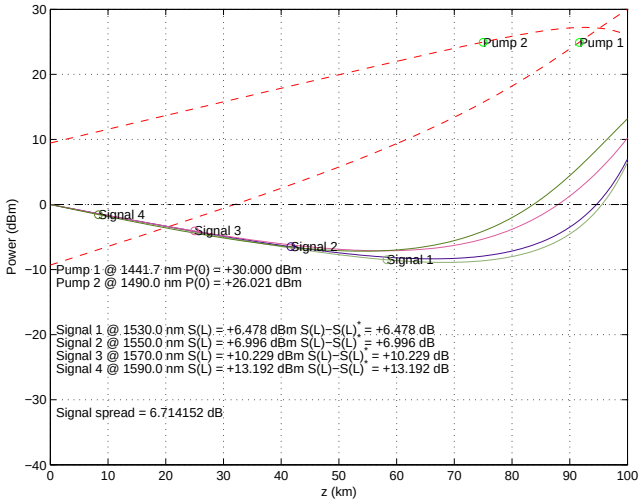

Fig. 7. Spatial power distribution prior to ESC.

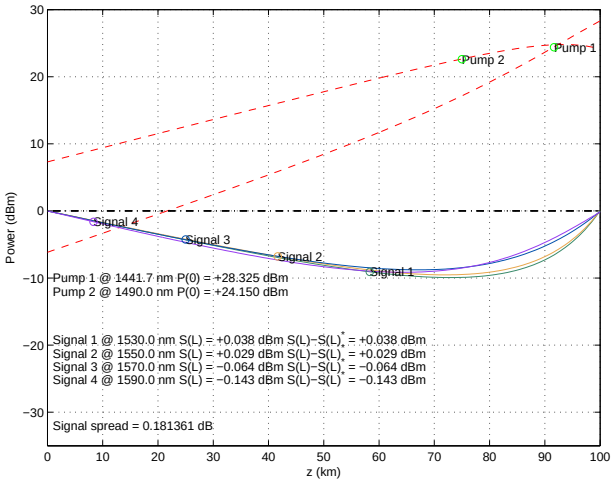

Fig. 8. Spatial power distribution after ESC.

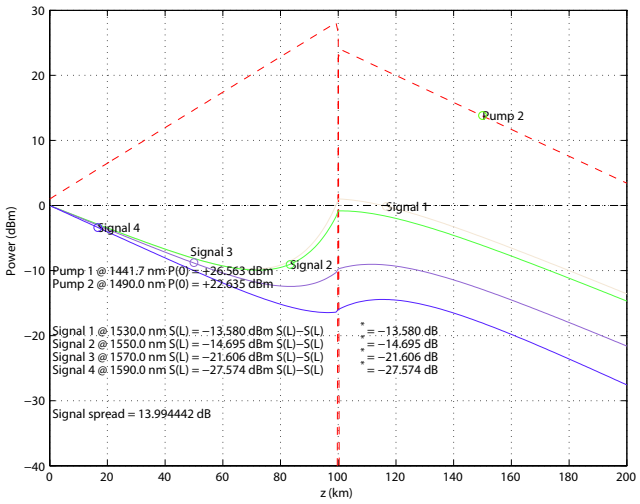

Fig. 9. Open-loop pump and signal power steady state.

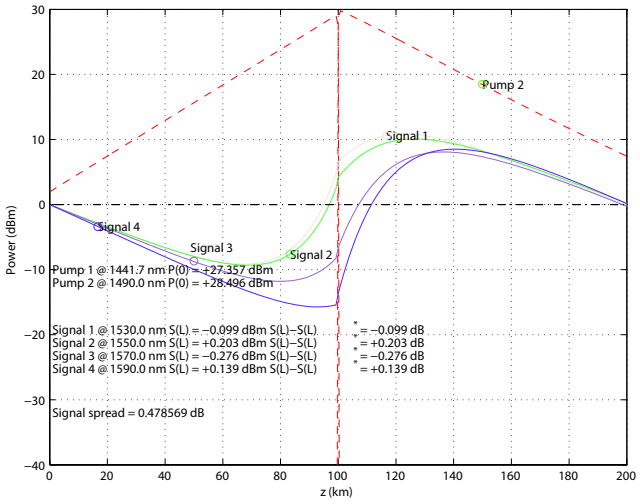

Fig. 10. Closed-loop pump and signal power steady state.

C. Four pump, forty signal backward and forward pumped Raman amplifier

A more realistic amplifier specification with 4 pumps controlling 40 signal channels can also be considered. The required specification changes are as follows:

$$\begin{aligned} n_p &= \quad 4 \quad (n_p^- = 2,\, n_p^+ = 2) \\ n_s &= \quad 40 \\ L_b &= \quad 50\,km \\ L_f &= \quad 50\,km \end{aligned}$$

Additional signal wavelengths utilize a $100\,GHz$ spacing, starting at $1531\,nm$. The 44×44 matrices τ, A and C follow from application of the interpolants of Appendix I. Prior to application of ESC, the signal spectrum is as shown in Figure 11.

Due to the increased number of pumps, more combinations of pump dithers are required in order to evaluate the gradient of the cost function during ESC. Figure 12 illustrates this process, and shows the downstream signal channel powers converging to the required level of $-10\,dBm$. Here, the MATLAB implementation of the Nelder-Mead simplex method is used as the optimization algorithm. The resulting signal spectrum and spatial power distribution is shown in Figures 13 and 14 respectively.

V. Conclusion

In this paper, the general extremum seeking framework of [12] was applied to the problem of signal power regulation in cascaded (optical) Raman amplifiers. Using a finite difference based PDE solver to simulate such cascades and optical links, the ESC strategy was found to significantly reduce both signal spread and signal power deviation from a desired set-point, across all propagating signal wavelengths. Due to the generality of the approach used, these results can readily be extended to other types of optical amplifiers (such as EDFAs).

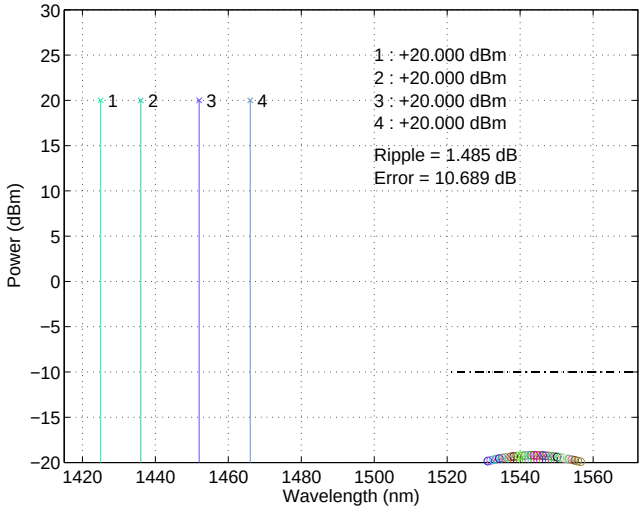

Fig. 11. Signal spectrum prior to ESC.

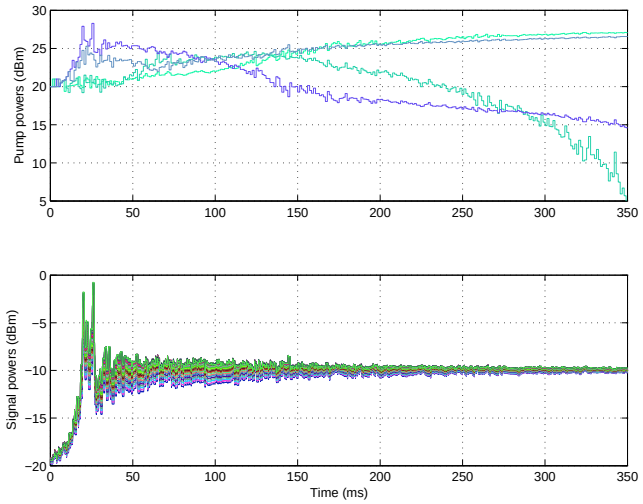

Fig. 12. Transients during ESC.

APPENDIX I
NUMERICAL COEFFICIENTS OF (2)

Where required, the coefficient matrices are approximated via interpolation of experimental data (see for example [1], [20], [21]). In particular, the entries of the matrices τ, A and C were determined numerically as follows:

$$\begin{aligned} \tau_{ii} &= \tau(\lambda_i) \\ A_{ii} &= a(\lambda_i) \\ C_{ij} &= c(\lambda_i, \lambda_j) \end{aligned} \tag{15}$$

where

$$\begin{aligned} \tau(\lambda) &= \textstyle\sum_{i=1}^{4} \hat{\eta}_i \left(\lambda - \hat{\lambda}_i\right)^i , \\ a(\lambda) &= \textstyle\sum_{i=1}^{4} \hat{\alpha}_i \lambda^{i-1} , \\ c(\lambda,\eta) &= \gamma(\lambda,\eta)\, \Gamma\left(\operatorname{sgn}(\lambda-\eta)\, \rho(\lambda,\eta)\right) \end{aligned} \tag{16}$$

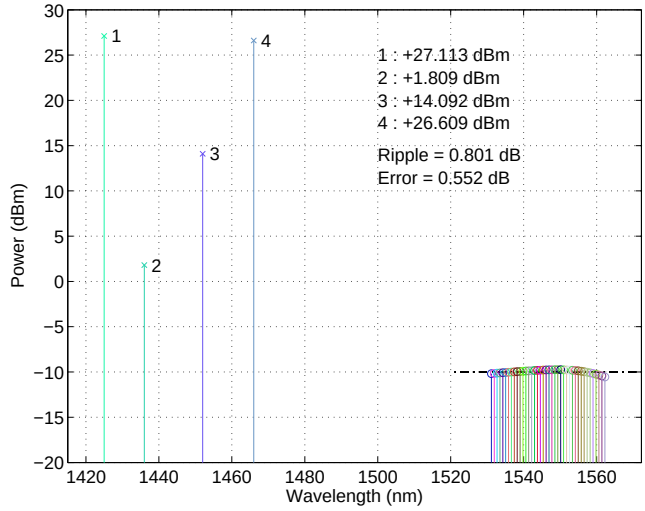

Fig. 13. Signal spectrum after ESC.

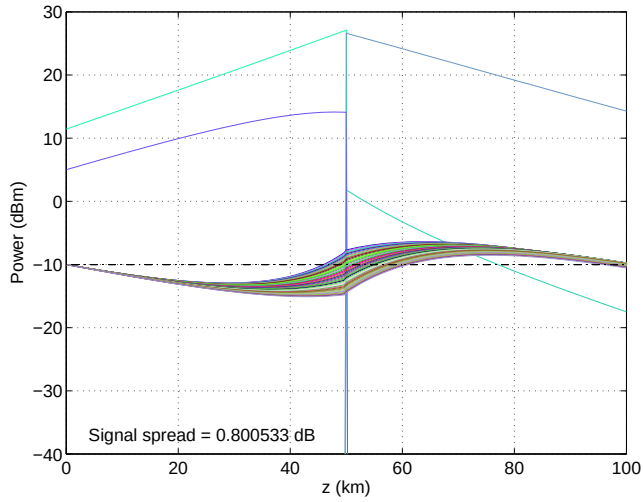

Fig. 14. Spatial power distribution after ESC.

$$\begin{aligned} \gamma(\lambda,\eta) &= \begin{cases} \frac{1}{4} & \text{if } \lambda > \eta \\ 0 & \text{if } \lambda = \eta \\ -\frac{\eta}{4\lambda} & \text{otherwise} \end{cases} \\ \Gamma(\zeta) &= \textstyle\sum_{i=1}^{3} \hat{\kappa}_i \exp\left(-\hat{\beta}\left(\zeta - \hat{\zeta}_i\right)^2\right) \\ \rho(\lambda,\eta) &= \hat{c}\left(\frac{1}{\eta} - \frac{1}{\lambda}\right) \end{aligned}$$

$$\hat{\lambda} = \begin{bmatrix} -26.097 \\ 0.1500 \\ 0.1500 \\ 0.1500 \end{bmatrix} \times 10^4$$

$$\hat{\eta} = \begin{bmatrix} 1.8583 \times 10^{-5} \\ 3.4249 \times 10^{-8} \\ 1.3836 \times 10^{-11} \\ 1.4457 \times 10^{-14} \end{bmatrix}$$

$$\hat{\alpha} = \begin{bmatrix} -19.966 \\ 4.3466 \times 10^{-2} \\ -3.1125 \times 10^{-5} \\ 7.3616 \times 10^{-9} \end{bmatrix}$$

$$\hat{\kappa} = \begin{bmatrix} 1.5513 \\ 5.4599 \times 10^{-1} \\ 2.3692 \times 10^{-1} \\ 1.2848 \times 10^{-1} \end{bmatrix} \quad \hat{\beta} = 5.3173 \times 10^{-2}$$
$$\hat{\zeta} = \begin{bmatrix} 4.1843 \\ 21.574 \end{bmatrix} \quad \hat{c} = 2.9979 \times 10^{5}$$

The units for $\hat{\lambda}_i$, $\hat{\eta}_i$, $\hat{\alpha}_i$, $\hat{\kappa}_i$, $\hat{\zeta}_i$, $\hat{\beta}$ and $\hat{c}$ are (respectively) nm, $\mu s\,(km)^{-1}\,(nm)^{-i}$, $(km)^{-1}\,(nm)^{-(i-1)}$, $(W\,km)^{-1}$, $km(s\,nm)^{-1}$, $(s\,nm)^{2}\,(km)^{-2}$ and $km\,s^{-1}$. The choice of the interpolants (16) was made on a largely qualitative basis in view of widely available [21] graphical and tabular representations of the wavelength dependencies (15) of commonly used optical fibres.

APPENDIX II
ASSUMPTIONS IN [12]

The main result of [12] (Theorem 2) relies on a number of technical assumptions. Our ability to apply this main result relies upon the validation of these assumptions for the Raman amplifier ESC problem at hand. This is a difficult task as a number of these assumptions involve trajectory based estimates and Lyapunov functions for the PDE (2). Rather than attempting an exhaustive test, we present a physical rationalization for these assumptions. Where possible, notation is taken from [12].

Assumption 1:

1. With the cascade amplifier pump input vector θ fixed, the spatial power distribution supported by such amplifiers readily converges to a steady state provided $v^{\mp}$ and d are fixed. This steady state is an attractor for the amplifier dynamics. We note that the steady state ODE (5) has a unique solution, so that the attractor is a singleton.
2. The output map $J \circ h$ is continuous in this case, so that the readout map g given by (9) is well defined.
3. For forward pumped amplifiers, it is possible to show explicitly that the readout map g is locally Lipschitz and has a unique global minimum. We expect this to also be true in the backward pumped case.
4. The non-depleted pump approximation [2] facilitates the approximate evaluation of transients in Raman amplifiers. That evaluation demonstrates the existence of an appropriate waiting time $T > T_*$, where T_* is given by (14).

Assumption 2:

$F(\theta, G(\theta))$ is defined by the optimization algorithm and the amplifier model, which in this case is not set-valued. Local convergence of the numerical algorithm follows from standard numerical analysis [18]. Since F is continuous in θ, convergence of the discrete time system defined by F and G is guaranteed by the converse Lyapunov results of [22].

Assumption 3:

The dither commands are implemented by the optimization algorithm. The waiting time is selected as per Assumption 1.4.

Assumption 4:

1. Closeness of the cost function and the readout map does appear to be bounded by deviations in the initial spatial power distribution ξ.
2. Closeness of steady states in amplifier operation also appears to be bounded by deviations in the pump power vector θ.
3. The pump dither is totally bounded by the optimization algorithm.
4. The optimization algorithm naturally generates parameter updates θ^+ that deviate little with small deviations in the measured cost.

Assumption 5:

In the forward pumped case, the set of minimizers Θ^* is a singleton. We expect the same to hold for the cascaded case.

REFERENCES

[1] M. Islam, "Raman amplifiers for telecommunications," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 8(3), pp. 548–559, 2002.

[2] J. Bromage, "Raman amplification for fibre communications systems." *Journal of Lightwave Technology*, vol. 22(1), 2004.

[3] C. Raman and K. Krishnan, "A new type of secondary radiation," *Nature*, vol. 121, p. 501, 1928.

[4] R. Stolen and E. Ippen, "Raman gain in glass optical waveguides," *Appl. Phys. Lett.*, vol. 22, pp. 276–278, 1973.

[5] G. Agrawal, *Nonlinear fiber optics*. San Diego, CA, USA: Academic Press, 2001.

[6] P. Dower and P. Farrell, "On linear control of backward pumped raman amplifiers," *Proc. 14^{th} IFAC Symposium on System Identification*, pp. 547–552, 2006.

[7] K. Ariyur and M. Krstic, *Real-time optimization by extremum-seeking control*. Wiley, 2003.

[8] K. S. Peterson and A. G. Stefanopoulou, "Extremum seeking control for soft landing of an electromechanical valve actuator," *Automatica*, vol. 40, pp. 1063–1069, 2004.

[9] H. Wang, S. Yeung, and M. Krstić, "Experimental application of extremum seeking on an axial-flow compressor," *IEEE Transactions on Control Systems Technology*, vol. 8, pp. 300–309, 2000.

[10] M. Krstić and H. H. Wang, "Stability of extremum seeking feedback for general nonlinear dynamic systems," *Automatica*, vol. 36, pp. 595–601, 2000.

[11] Y. Tan, D. Nešić, and I. Mareels, "On non-local stability properties of extremum seeking control," *Automatica*, vol. 42, pp. 889–903, 2006.

[12] A. Teel and D. Popović, "Solving smooth and nonsmooth multivariable extremum seeking problems by the methods of nonlinear programming," *Proc. American Control Conference*, pp. 2394–2399, 2001.

[13] P. Dower, P. Farrell, and D. Nesic, "An application of extremum seeking in cascaded optical amplifier control," *Proc. IEEE Conference on Decision and Control (San Diego)*, pp. 4472–4477, 2006.

[14] E. Desurvire, *Erbium-doped fiber amplifiers: principles and applications*. Wiley, 1994.

[15] J. Palais, *Fiber optic communications*, 4th ed. Prentice-Hall, 1998.

[16] H. Kidorf, K. Rottwitt, M. Nissov, M. Ma, and E. Rabarijaona, "Pump interactions in a 100-nm bandwidth raman amplifier," *IEEE Photonics Technology Letters*, vol. 11, no. 5, pp. 530–532, 1999.

[17] H. Ono, M. Yamada, and M. Shimizu, "S-band erbium-doped fiber amplifier with a multistage configuration configuration - design, configuration and gain tilt compensation," *IEEE J. Lightwave Technology*, vol. 21, no. 10, pp. 2240–2246, 2003.

[18] R. Burden and J. Faires, *Numerical analysis*. PWS Kent, 1989.

[19] N. Stefanovic and L. Pavel, "ℓ_2 nonlinear control of edfa system with amplified spontaneous emission," *Proc. IEEE Conference on Control Applications*, pp. 289–296, 2005.

[20] S. Nakamura, Y. Koyamada, N. Yoshida, N. Karasawa, H. Sone, M. Ohtani, Y. Mizuta, R. Morita, H. Shigekawa, and M. Yamashita, "Finite-difference time-domain calculation with all parameters of sellmeier's fitting equation for 12-fs laser pulse propagation in a silica fiber," *IEEE Photonics Technology Letters*, vol. 14, no. 4, 2002.

[21] R. Ramaswami and K. Sivarajan, *Optical networks: A practical perspective*, 2nd ed. Morgan Kaufmann, 2002.

[22] C. Kellett, "Advances in converse and control lyapunov functions," Ph.D. dissertation, University of California, Santa Barbara, 2002.