Efficient Discrete Rate Assignment and Power Optimization in Optical Communication Systems Following the Gaussian Noise Model

Ian Roberts, *Student Member, OSA*, and Joseph M. Kahn, *Fellow, IEEE*

***Abstract*—Computationally efficient heuristics for solving the discrete-rate capacity optimization problem for optical fiber communication systems are investigated. In the Gaussian noise nonlinearity model regime, this class of problem is an NP-hard mixed integer convex problem. The proposed heuristic minimizes the number of calls required to solve the computationally intensive problem of determining the feasibility of proposed discrete rate allocations. In a live system, optimization using this algorithm minimizes the number of potential discrete rate allocations tested for feasibility while additional discrete system capacity is extracted. In exemplary point-to-point links at 50 Gbaud with 50 Gb/s rate steps, the mean lost capacity per modem is reduced from 24.5 Gb/s with truncation to 7.95 Gb/s with discrete-rate optimization. With 25 Gb/s rate steps, the mean lost capacity is reduced from 12.3 Gb/s to 2.07 Gb/s. An unbiased metric is proposed to extend the capacity optimization objective from point-to-point links to mesh networks. A gain of 13% in distance-times-capacity metric is obtained from discrete-rate optimization with 50 Gb/s rate steps, and a 7.5% gain is obtained with 25 Gb/s rate steps.**

***Index Terms*—Discrete capacity optimization, gaussian noise model, network optimization, optical communications.**

I. Introduction

As TRAFFIC demands grow, increasing the utilization of existing fiber resources is often more economical than installing new fibers. Improvements in spectral efficiency allow system capacity gains to be obtained independently from increases in spectrum usage, and improve the utilization of existing terminal equipment. Coded modulation, shaping, and variable error-control coding can allow data rate adaptation to optimize channel usage [1]–[3]. Quantization of the bit rate carried by a channel, such as $n \times 100$ Gb/s, can leave significant capacity unrealized.

Communication modems implementing appropriate coding can deliver an arbitrary number of discrete communication rates. Sourcing multiplexed data streams of the precisely delivered modem rate leads to increased complexity upstream and downstream from the modem. Flex Ethernet [4] provides technologies to deliver varying data rates. In this paper, we examine the problem of optimizing with 10, 25, 50, and 100 Gb/s rate steps, assuming a symbol rate of 50 Gbaud.

In wireline or wireless multi-carrier systems, the imposition of a transmitter output power cap on a frequency-dependent channel leads to a bit rate allocation problem. For linear channels, this problem is solved efficiently via the Levin-Campello algorithm [5]. An important difference between the wireline or wireless problem and the WDM optical communication problem considered in this paper is that instead of the transmitter output power cap, there is an input power limit due to the nonlinear noise experienced in optical communication systems. In dispersion-uncompensated systems, this intra- and inter-channel nonlinear noise is well represented by the Gaussian noise model [6]–[8].

One can readily optimize the capacity using a continuous set of communication rates [9]. The capacity achieved sets an upper bound on the rate achievable using a discrete set of rates. When optimizing the total discrete rate, one attempts to approach this upper bound under the constraint of a discrete set of communication rates. One way to approach the continuous-rate bound is to use a finely spaced set of discrete communication rates and simply truncate the optimum continuous rate for each channel down to the nearest discrete rate. However, attempting to approach the continuous-rate optimum capacity with truncation requires a very large set of discrete rates.

An alternate approach to truncation is discrete-rate optimization, which involves selecting a set of channel rates, some above and others below the continuous-rate optima, such that the sum of rates is closer to the continuous-rate upper bound than the sum of truncated rates. Fig. 1 compares continuous-rate optimization, rate truncation, and discrete-rate optimization. As illustrated in Fig. 1, in discrete-rate optimization, only a fraction of the channels are able to have their rates upgraded above the continuous-rate optimum while maintaining feasibility, and the rest must receive truncated rates. Channels with upgraded rates have greater SNR requirements and require increased power assignments. Channels with higher powers cause greater nonlinear interference to their neighbors, reducing the likelihood of feasible rate upgrades to those neighbors. The problem of optimally allocating power to a set of channels with fixed SNR

Manuscript received May 30, 2017; revised July 25, 2017 and August 18, 2017; accepted August 22, 2017. Date of publication August 24, 2017; date of current version September 8, 2017. This work was supported in part by Ciena Corporation, and in part by Stanford Rambus Corporation Graduate Fellowship. *(Corresponding author: Ian Roberts.)*

The authors are with the E. L. Ginzton Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: iroberts@stanford.edu; jmk@ee.stanford.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2017.2744624

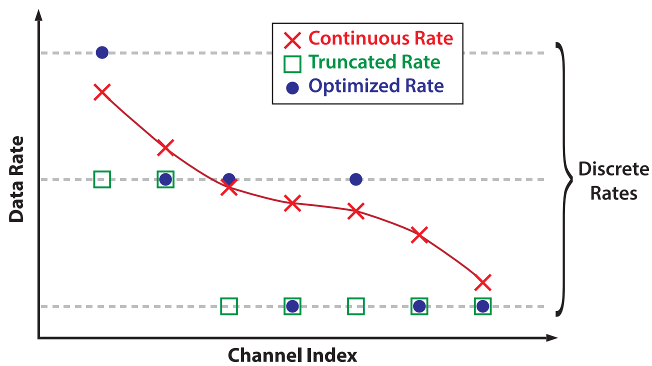

Fig. 1. An illustration of discrete-rate optimization in comparison to truncation of the continuous-rate solution.

requirements is defined as the minimum-margin problem [9], where the margin is the additional SNR provided to each channel in excess of the fixed requirements. In order to determine the feasibility of a proposed set of discrete rates, where some are greater than the continuous-rate capacity, solving the minimum-margin problem is required.

The point-to-point continuous-rate capacity problem and the minimum-margin problem have been demonstrated to have convex formulations using the Gaussian noise model [9]. Convex optimization is the study of the optimization of convex functions of the form $f(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y}) \leq \theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$ and $0 \leq \theta \leq 1$ [10]. Locally optimal solutions to convex optimization problems are also globally optimal, which is a very useful property for problems with high-dimensional search domains. For differentiable objectives, operations such as gradient descent offer guaranteed convergence to the global optimum [10]. Concave functions are those functions f where $-f$ is convex. Concave problems also fall under the classification of convex optimization, as they are solved with the same tools used for convex problems. The convexity of minimum margin and capacity problems is very important as it means that globally optimal solutions are readily obtainable even in problems with many dimensions. Computationally efficient optimal solutions to these and similar convex problems form the basis for approaching the problem of optimizing the total discrete rate.

Prior work has addressed various aspects of the discrete-rate optimization problem. Channel power optimization when combining multiple communication rates has been investigated in [11]. Channel power optimization following the Gaussian noise model and alternation of high and low rate channels or polynomial spreading of channels with multiple discrete rates was investigated in [12]. Integer programming methods of estimating rate assignments combined with routing have been proposed [13], [14]. An exponentially complex method of power, rate, and spectral width assignment was described in [15]. Extensive investigations have been performed of the performance of networks with variable transceiver data rates and truncated rate allocations [16], [17].

This paper provides an efficient method of optimizing discrete channel rates in order to approach the bound obtained with continuous channel rates, taking advantage of optimal solutions to power optimization problems. In Section II, existing convex capacity and margin objectives are combined to form useful intermediate objectives for discrete-rate optimization. These objectives also have application in safely optimizing newly provisioned channels on fibers with existing traffic. An unbiased capacity metric is formulated, as the existing point-to-point capacity metric does not generalize to mesh networks in the same manner as the margin optimization problem. In Section III, two main heuristics are proposed to efficiently solve the discrete-rate optimization problem. Several higher-complexity solutions to problems of this class are also addressed. Section IV presents the results of discrete-rate optimization for point-to-point links with various discrete rate steps, and for a mesh network using the new mesh capacity metric.

II. Convex Sub-Problems

A communications modem will typically implement a discrete set of communication rates. These may be 50, 100, 200 Gb/s, as determined by the use of BPSK, QPSK, and 16-QAM modulation formats, or may be more finely quantized via advanced coding techniques [1]–[3]. Each of the implemented discrete rates will have a corresponding SNR requirement. Changing the data rate from one rate to the next higher rate changes the required SNR. This change in SNR is referred to as an SNR step, while the change in data rate is referred to as a rate step. A given data rate step, at different baud rates, corresponds to different SNR steps. Smaller rate steps correspond to smaller SNR steps, and smaller SNR requirement changes are more likely to have optimal power allocations providing feasible margins. Hence discrete-rate optimization, supporting data rates above and below the continuous optimum, performs best with small SNR steps between rates.

In most applications of interest, the optimal discrete rate choices are only those that are adjacent to the continuous-rate optimum. Let us examine why this is true.

At the continuous-rate capacity-optimizing power allocation, the nonlinear noise is 3 dB below the amplifier noise level for almost all channels [9]. Uniformly increasing channel powers will lead to a decrease in the SNR. Achieving rates above the continuous-rate optimum generally requires a combination of raising the power of the channel in question above that of the continuous-rate capacity-optimizing power allocation, and lowering the powers of neighboring channels. The amount by which the SNR of a particular channel can be increased in this way is limited by the proportion of the nonlinear noise that is due to inter-channel nonlinear components.

Given the optimum continuous channel rates and the corresponding power allocation, in all reasonable scenarios the search for the optimal discrete rate set need only look at the two discrete rates, which are the greatest discrete rate less than the continuous optimum and the smallest discrete rate greater than the continuous-rate optimum. These are referred to as the truncated and upgraded discrete rates. This holds when receiver coding is typical, such as following the Shannon capacity curve minus some coding gap [18], and the coding gap at higher rates is no lower than at lower rates. If the set of discrete rates is coarse, as in current systems, the level of mediation provided by inter-channel nonlinear effects cannot support the large jump in SNR required to upgrade by more than one step. With rate separations that are dramatically non-uniform, one can create counter-examples. In systems with very low spectral efficiencies, even small uniform rate steps correspond to significant SNR steps, so rate quantization can be considered coarse. Systems with uniform data rate steps need only look at the upgraded and truncated rates.

A. Fixed Implementation Noise

Implementation noise refers to power-independent SNR contributions. Even with effective analog gain control, limitations on precision in digital-to-analog and analog-to-digital conversion are important sources of implementation noise. Fixed-point precision in digital signal processing is another such source. Equation (1) shows how the power-independent SNR due to implementation noise $SNR_{\text{implementation}}$ combines with the optical channel SNR SNR_{optical} to produce the overall system SNR SNR_{total}, which is a function of fiber input signal powers P.

$$SNR_{\text{total}}(P) = \frac{1}{SNR_{\text{implementation}}^{-1} + SNR_{\text{optical}}^{-1}(P)} \quad (1)$$

Capacity and margin optimizations using the Gaussian noise model retain convex or concave forms with the inclusion of implementation noise.

When operating at low SNR levels, such as when carrying small numbers of bits per symbol, implementation noise has a negligible impact on the overall SNR. With high numbers of bits per symbol, high SNR levels are required for error-free communication. As the optical SNR increases, the SNR due to implementation noise becomes significant and eventually limits the achievable overall SNR.

As shown in Fig. 2, the SNR difference required to achieve a given rate step is strongly dependent on the implementation SNR. This figure shows the SNR steps for a system without implementation noise, and for systems with a range of implementation noise levels. 25 Gb/s rate steps are considered with SNRs calculated for a 50 Gbaud dual-polarization coherent system following the Shannon capacity limit minus a 3.5 dB combined margin and coding gap. Systems with higher implementation noise levels have a narrower range of data rates with low SNR steps between consecutive data rates, and higher minimum SNR steps. Both of these effects reduce the efficiency of discrete rate assignment.

B. Mesh Capacity Optimization

Capacity optimization of a point-to-point link with a continuous set of communication rates following the Shannon capacity curve with an SNR gap has been demonstrated to be concave for high SNR [9]. In point-to-point links, capacity optimization is a proxy for the true underlying system metric of economic utility. In mesh networks, the same capacity optimization

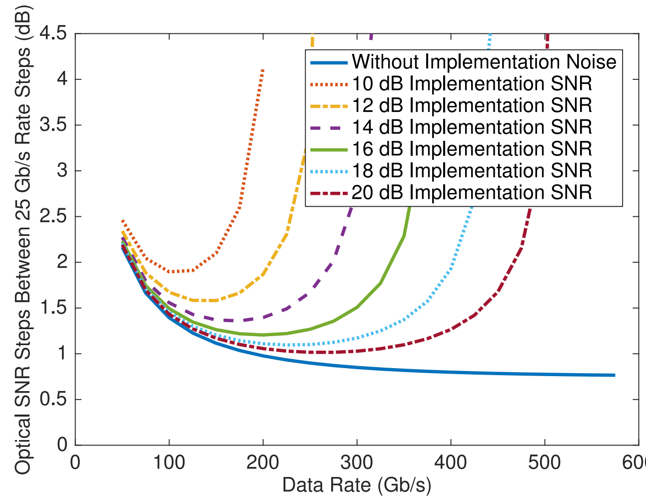

Fig. 2. Optical SNR steps between successive communication rates are shown for 25 Gb/s rate steps in a 50 Gbaud dual-polarization coherent system following the Shannon capacity limit minus a 3.5 dB combined margin and coding gap.

objective fails to optimize the underlying economic utility, since carrying a given data rate over a longer distance provides more value. A large portion of this utility comes from regenerators being avoided on long-range channels.

The effect of this mismatch between the total capacity metric and the economic utility is apparent in the powers that are assigned to channels with different total path lengths that share a section of fiber. For systems without implementation noise, the derivative of capacity with respect to log-SNR is monotone increasing. This means that a given SNR increase gives the greatest capacity return when a high-rate channel is improved. On any given section of fiber in a mesh network, optimizing with the total capacity metric will assign higher powers to channels with short total path lengths than to channels with longer total path lengths. Over this section of fiber, the local SNR is higher for the channels with short path lengths than the channels with long ones due to the difference in allocated powers. Given that the channels with longer total path lengths cross many sections of fiber, their total optical SNR is further reduced as compared to the short-range channels. Short-range channels also have nonlinear impact on fewer other wavelengths, compounding the bias under the total capacity metric. In systems with implementation noise, capacity optimization disfavors channels with extremely short lengths, but short or medium-range channel lengths are favored over long-range ones.

The distance-weighted mesh optimization metric proposed here is

$$A = \sum_{i=1}^{L} a_i C_i d_i^{\xi}, \quad (2)$$

where C_i is the capacity of optical path i in a mesh with L paths under optimization, d_i is the distance covered by the optical path, and a_i and ξ are weighting parameters. The values of a_i and ξ will vary based on the specific network in question. The weights a_i are typically one, but certain high-priority traffic can receive extra weight. The value of ξ depends on the economic trade-off between distance and rate for a particular network. For $\xi \approx 1$ the bias towards short-range channels is removed. For higher values of ξ, the penalty on short-range channels is found to increase without significant benefit for long-range channels. This is due to the challenge of further increasing rates on very long-range channels because of the large number of other channels impacted by their long paths.

The metric (2) is a concave function in logarithmic power variables where the individual channel capacities are concave functions as a_i, d_i, and ξ are constants. For systems with continuous coding following the Shannon capacity curve, concavity is obtained in the region where the channel SNRs satisfy $\log(SNR + 1) \approx \log(SNR)$ using logarithmic power variables [9].

C. Protected Channel Margin Optimization

The minimum-margin optimization problem, considered in [9], seeks to optimize the full set of system channels. In a live system, such green-field deployments are rare. Typically a number of new channels are provisioned in an existing system. The goal of optimizing the minimum margin of a set $\mathcal{A}$ of new channels while maintaining strict operational margins for the set of existing channels $\mathcal{B}$ is slightly different than the objective considered in [9].

This objective is

$$\begin{aligned} &\text{maximize } \min_{n\in\mathcal{A}}(M_n(\boldsymbol{x})) \\ &\text{subject to } M_m(\boldsymbol{x}) \geq 1 \qquad \forall m \in \mathcal{B}, \end{aligned} \tag{3}$$

where $M_n(\boldsymbol{x})$ is the margin in channel n as a function of power vector $\boldsymbol{x}$ containing the powers input for each channel to the link.

In convex form, using $e^{\boldsymbol{y}} = \boldsymbol{x}$ and inverse margins M_n^{-1}, this is best represented as the problem

$$\begin{aligned} &\text{minimize } e^s \\ &\text{subject to } M_n^{-1}(e^{\boldsymbol{y}}) \leq e^s \qquad \forall n \in \mathcal{A} \\ &\qquad\qquad M_m^{-1}(e^{\boldsymbol{y}}) \leq 1 \qquad \forall m \in \mathcal{B}. \end{aligned} \tag{4}$$

The optimization problem (4) is readily solvable using the methods of [9]. In these constrained convex optimizations, a feasible starting point simplifies the required algorithm. A feasible start can be obtained by solving a regular minimum-margin problem with the desired SNR requirements on channels $\mathcal{B}$ and negligible SNR requirements on channels $\mathcal{A}$.

A concave mixed capacity and margin objective can be formulated similarly, as shown in (5), using the mesh capacity metric of (2). This supports capacity optimization of new channels on a network with existing channels that have margin requirements to maintain.

$$\begin{aligned} &\text{maximize } \sum_{i\in\mathcal{A}} a_i C_i d_i^{\xi} \\ &\text{subject to } M_m^{-1}(e^{\boldsymbol{y}}) \leq 1 \qquad \forall m \in \mathcal{B} \end{aligned} \tag{5}$$

For the objective of (5), channels in set $\mathcal{A}$ are optimized for capacity with continuous rates, while channels in set $\mathcal{B}$ are maintained with strict operational margins. For point-to-point links, typically $a_i = 1$ and $\xi = 0$, leaving the traditional sum-of-capacities metric. The optimization complexity of this objective is closer to that of a minimum-margin problem than a pure capacity objective due to the fixed margin inequalities for channels in set $\mathcal{B}$.

In the process of discrete-rate optimization, a large number of minimum-margin problems will be solved. Taking advantage of sparsity patterns can significantly reduce the computational complexity of minimum-margin optimization. For high-baud-rate channels, individual four-wave mixing components are generally very small, but there are many such components. Ignoring the four-wave mixing terms reduces the model accuracy, but if only the XPM terms are included in the discrete Gaussian noise model evaluation, sparsity patterns are introduced in the nonlinear derivatives. Taking advantage of these sparsity patterns can significantly reduce computation time, and can change the mesh minimum-margin optimization scaling to be closer to linear in network size.

III. Discrete Rate Optimization

Given that real modems implement a discrete set of communication rates, we find that the desired objective of maximizing the total discrete rate is a different optimization problem than maximizing the continuous-rate capacity. In this section, we formulate the optimization problem of maximizing the total discrete rate. This will use the unbiased mesh metric, but for point-to-point links the distance and channel weighting factors drop out.

Given a set of rates assigned to channels, and a mapping from rates to margined SNR requirements, the feasibility of the system is determined by solving a margin optimization problem [9] or (4). A negative (dB) margin result indicates the infeasibility of an assigned rate set. Convexity of the optimization objective is very valuable here, as it means that there is one optimal power allocation for a given rate assignment, and that there is a sharp distinction between feasible and infeasible rate sets as determined by the margins of their optimal power allocations.

The problem of maximizing the discrete communication rate of a point-to-point link combines discrete rate variables with continuous logarithmic power variables. Let $M_n(r_n, e^{\boldsymbol{y}})$ be the margin of channel n given its assigned discrete rate r_n of discrete rate vector $\boldsymbol{r}$ and power vector $e^{\boldsymbol{y}}$. The discrete-rate objective, given discrete rate set $\mathcal{D}$, optimization channel set $\mathcal{A}$, and potentially empty margin channel set $\mathcal{B}$ is

$$\begin{aligned} &\text{maximize } \sum_{n\in\mathcal{A}} a_n r_n d_n^{\xi} \\ &\text{subject to } M_n^{-1}(r_n, e^{\boldsymbol{y}}) \leq 1 \qquad \forall n \in \mathcal{A}. \\ &\qquad\qquad M_m^{-1}(e^{\boldsymbol{y}}) \leq 1 \qquad \forall m \in \mathcal{B} \\ &\qquad\qquad \boldsymbol{r} \in \mathcal{D}^N \end{aligned} \tag{6}$$

Constants a_n, d_n, and ξ follow from the distance-weighted mesh metric (2). For point-to-point links, typically $a_i = 1$ and $\xi = 0$, leaving just the discrete channel rate term r_n.

If the continuous-rate capacity objective is convex, then (6) is a mixed-integer convex problem, which is NP-hard [19].

Very general methods, such as branch-and-bound and other cutting-plane heuristics, can eventually find good or asymptotically optimal solutions to integer convex problems [19]–[21]. These methods will generally be very slow. Insight into the structure of the problem can lead to more efficient heuristic solutions. For this problem, the simple convex upper bound obtained with continuous rates, and the fact that rate truncation is strictly feasible, are simplifying factors. Typical heuristics for mixed-integer convex problems are diving methods and the feasibility pump [21].

The building block that forms the foundation for most solution methods of (6) is the margin optimization problem obtained for a fixed rate assignment [9] or (4). For a point-to-point link, solving this problem using a theoretical channel model takes on the order of a second using a desktop computer, and is open to dramatic performance improvements from parallelization. Cutting plane-based methods augment the basic margin problem with a sequence of additional domain inequalities, which increases the problem scale and difficulty. Given the short time required to solve the point-to-point optimization, there remains the possibility of solving a large number of these convex margin problems in the process of solving the desired discrete-rate objective.

Computation time issues are aggravated when considering margin optimization of mesh networks, optimization with additional complexities, or live system optimization. The near quadratic scaling of computational complexity with network size can make each minimum-margin problem significant for mesh networks [9]. On a live system, if receiver measurements are incorporated into the optimization, the adjustment, stabilization, and measurement time needs to be included in the time required for each margin optimization iteration. These computational complexity or time constraints have led to this paper's focus on solution methods to the discrete-rate maximization problem (6) that aim to minimize the number of minimum-margin problems solved.

The time required to solve the minimum-margin problem depends upon the number of active channels rather than the subset being assigned discrete rates. The number of active channels will depend upon the size of the network under optimization. The size of a given network also limits the total number of channels to which discrete rates can be assigned. For the green-field deployments examined here, scaling of network size thus scales both the size of the discrete-rate optimization problem and the complexity of the minimum-margin problem. To separate the margin optimization complexity from the discrete-rate optimization complexity, we examine complexity in terms of the number of calls to solve the minimum-margin problem.

Given a fixed system and Y channels with unassigned rates, a linear-call method will solve $\mathcal{O}(Y)$ margin optimization problems. In Section III-A, we examine such linear-call methods. The overall search space for the optimal discrete channel rate allocation is exponential, and naively $\mathcal{O}(2^Y)$, due to inter-channel nonlinear interactions. Raising the rate and corresponding SNR requirement of one channel makes it harder to raise the rate of its neighbors. Any linear-call method will therefore only obtain accurate margin information, via solving a minimum-margin optimization, for a very small subset of possible solutions. On the order of one dimension of the $|\mathcal{A}|$-dimensional solution space

```
Solve continuous-rate capacity optimization
Truncate continuous-rate solution to discrete rates
Solve minimum-margin problem, obtaining dual variables
Put all channels to be optimized in the set untested
while untested is not empty do
    Remove channel with loosest dual variable from untested
    Upgrade this channel
    Solve minimum-margin problem, obtaining dual variables
    if margin solution is infeasible then
        Revert channel upgrade
        Restore previous minimum-margin problem results
    end if
end while
```

Fig. 3. Sequential linear-call dual variable-based discrete-rate optimization algorithm.

for discrete rate vector $\boldsymbol{r}$ must be fixed for each margin problem solved.

For mesh networks, it is desirable to be more computationally efficient than linear-call methods. In Section III-B, we examine logarithmic-call methods for solving the discrete-rate optimization problem, requiring $\mathcal{O}(\log(Y))$ margin optimization calls.

A. Linear-Call Methods

An efficient diving method is enabled by solving the margin optimization problem in a manner that provides dual variables [10]. The dual variables of [9] provide a measure of how much each channel is limiting the minimum margin optimization. Channels with looser dual variables can potentially support higher SNR requirements and correspondingly higher rates. At the minimum-margin operating point, the loosest dual variable corresponds to the channel with the greatest marginal return in SNR for an increase in channel power, while maintaining the margin requirements of the remaining channels. In the limit of small discrete rate steps, this marginal indication of the channel to which a higher rate should be assigned becomes exact, but truncation of the continuous-rate optimum also converges towards the continuous-rate bound in such a scenario. With larger discrete rate steps and corresponding SNR jumps, the marginal indication provided by the dual variables is not exact for the discrete jump in SNR to the next rate, but remains very precise.

Starting with the truncated-rate lower bound rate allocation, solving a minimum-margin problem can provide dual variables that will indicate the best candidate channels for upgrading to the next rate, in order. Solving a minimum-margin problem with the first candidate channel upgraded in rate will determine feasibility and provide a new set of dual variables, if feasible. If infeasible, the next candidate channel can be upgraded instead and checked for feasibility. This forms the sequential algorithm of Fig. 3 for determining a feasible quantized rate pattern that is close to optimal, via $\mathcal{O}(Y)$ margin optimization solutions. It is ideal in the limit of small discrete rate steps, and is otherwise very good given accurate margin optimization solutions and dual variables.

The dual variable-based sequential method extends poorly to mesh networks. The dual variable indication of the channel that would provide the greatest increase in marginal capacity is on the sum of capacities metric. If the desired metric incorporates distance, such as the objective in (2), then the dual variables are poor indicators.

Next, let us consider the feasibility pump, which has varying complexity, but can be implemented with linear-call complexity. Mixed-integer convex problems can be relaxed to convex problems by allowing all integer valued variables to be continuous. The optimal solution to the relaxed problem is almost always infeasible with the integral constraints. The feasibility pump is a heuristic method to obtain a feasible solution to the mixed-integer problem [20]. The feasible domain of the discrete-rate objective (6) includes all discrete rate assignments strictly smaller than the continuous-rate solution, including assigning all channels a rate of zero, such that simple truncation of the continuous-rate convex solution is feasible. In order to apply the feasibility pump usefully, additional constraints must be added to reduce the feasible domain. One such constraint is obtained by estimating the desired minimum feasible discrete-rate capacity $C^* < C$, and requiring that all solutions have total rate at least C^*. Obtaining a good solution requires an accurate C^* just below the greatest feasible integer solution, which requires solving for a sequence of C^* values. The feasibility pump requires solving a series of relaxed convex problems, but with the additional constraints, the problem complexity is higher than the minimum-margin problem. The feasibility pump is good at finding feasible solutions, but maintaining numerical stability without greatly increasing computation time becomes a challenge as the domain gets reduced pushing towards optimal integral solutions.

These establish linear-call methods, which require a large number of minimum-margin calls even for point-to-point links. In order to further reduce complexity we look at methods that solve only a number of minimum-margin problems proportional to the logarithm of the number of rates assigned.

B. Logarithmic-Call Method

For point-to-point links with a consistent rate step, upgrades to different channels are fully interchangeable and the number of feasible upgrades directly corresponds to the desired capacity metric. If we assume access to an oracle that will provide the optimal pattern of channel rate upgrades for a given upgrade count, the problem of discrete rate assignment can be reduced to a binary search (bisection) to find the maximum number of feasible upgrades. This gives a complexity that is logarithmic in the number of minimum-margin problems solved during the search. Since such an oracle does not exist, determining the pattern of channel upgrades is an assignment problem. This separates the discrete-rate problem into a binary feasibility search across assignment problems. Due to the cubic Kerr nonlinearity, the resulting assignment problems are cubic and NP hard [22]. The advantage of partitioning the discrete-rate problem in this manner is that near-optimal solutions to the assignment problem do not require solving a minimum-margin problem.

In order to obtain the binary vector of rate upgrades, the proposed solution method for the assignment problem is a repeated comparison of all binary swaps. For each upgrade in turn, the channel margin is evaluated for the upgrade in its original position and after swapping to each open non-upgraded channel. The channel with the greatest margin is selected. This is a greedy placement algorithm requiring a quadratic number of iterations.

Input: set $\mathcal{A}$ of channels to optimize
Output: set $\mathcal{A}^+$ of upgraded channels
1: Solve continuous-rate capacity optimization
2: Truncate continuous-rate solution to discrete rates
3: $lower \leftarrow 0$
4: $upper \leftarrow |\mathcal{A}|$
5: **while** infeasible or $(upper\text{-}lower) > 0$ **do**
6:     *upgradeCount* $\leftarrow$ ceil$((upper + lower)/2)$
7:     Spread *upgradeCount* upgrades uniformly over set $\mathcal{A}$
8:     Let $\mathcal{A}^+$ denote channels with upgrades
9:     Let $\mathcal{A}^-$ denote channels without upgrades
10:     **for** 1:*numRepetitions* ≈ 3 **do**
11:         **for each** element α of $\mathcal{A}^+$ **do**
12:             Remove α from $\mathcal{A}^+$ and insert into $\mathcal{A}^-$
13:             **for each** element β of $\mathcal{A}^-$ **do**
14:                 Insert β into $\mathcal{A}^+$
15:                 Estimate powers given upgrade allocation $\mathcal{A}^+$
16:                 Measure margin of β given estimated powers
17:                 Return β to $\mathcal{A}^-$
18:             **end for**
19:             Let $\beta^* \in \mathcal{A}^-$ be channel with greatest margin
20:             Insert β^* into $\mathcal{A}^+$ in place of α
21:         **end for**
22:     **end for**
23:     Solve minimum-margin problem given upgrade set $\mathcal{A}^+$
24:     **if** margin solution is feasible **then**
25:         $lower \leftarrow$ *upgradeCount*
26:     **else**
27:         $upper \leftarrow$ *upgradeCount*-1
28:     **end if**
29: **end while**

Fig. 4. Bisection-based logarithmic-call discrete-rate optimization algorithm.

When this solution method for the assignment problem is combined with a bisection search for the maximum number of feasible upgrades, the result is the algorithm of Fig. 4. This algorithm accounts for the nonlinear interactions between neighboring channels from channel powers adapting to upgraded or downgraded rates. The optimal placement for a given upgrade depends on the arrangement of all other potential upgrades, so the swapping process is repeated *numRepetitions* ≈ 3 times as the placement of a given upgrade improves with the improvement of the placement of other upgrades. No noticeable improvement in quantization performance was observed for *numRepetitions* > 3.

In order to evaluate the margin for a proposed upgrade vector, such as when comparing upgrade placement swaps, an ideal solution is obtained by solving a minimum-margin problem. This fails to result in a logarithmic-call method. The algorithm of Fig. 4 makes a call on line 15 to estimate the channel powers given a particular pattern of rate upgrades. Good solutions to the assignment problem can be obtained with very roughly estimated power allocations. Further, in order for the proposed algorithm to provide an efficient solution method, a power estimation method that is dramatically faster than solving a minimum-margin problem is required. Good results are obtained using estimated channel power assignments as simple as multiplying the required SNR by $3/2$ of the ASE, which corresponds to the assumption of fixed nonlinear noise equal to half of the ASE level. This estimate is sufficient for point-to-point links, but produces poor results for more complex mesh networks.

For this paper, a more accurate, but still fast power estimation procedure is used. The initial step to solving the discrete-rate capacity problem is to solve it for continuous rates, and thereby obtain the continuous-rate-optimizing power allocation. For a given channel, decreasing the rate and corresponding SNR requirement from the continuous-rate-optimum pushes the power assignment towards the linear regime, where nonlinearity is negligible. Increasing the data rate and corresponding SNR requirement above the continuous-rate optimum is difficult due to the presence of neighboring channels. Small single-channel SNR increases are possible for large single-channel power increases, but many neighboring channels are negatively impacted by increasing the power of a single channel. Without other neighboring channels being downgraded in rate and power, upgraded data rates will never be feasible. Even if other channels are downgraded, there remains significant nonlinearity, which necessitates a more-than-linear increase in channel power to achieve a higher SNR. For the results of this paper, (7) and (8) describe the estimated power P used for upgraded and downgraded channels in solving the upgrade assignment problem. The estimated powers are functions of P_{capacity}, the power assigned from continuous-rate capacity optimization, SNR_{capacity}, the channel SNR delivered in the continuous-rate solution, and the SNRs of the upgraded and downgraded discrete rates above and below the continuous-rate solution.

$$P_{\text{upgrade}} = P_{\text{capacity}} \times \left(\frac{SNR_{\text{upgrade}}}{SNR_{\text{capacity}}}\right)^2 \tag{7}$$

$$P_{\text{downgrade}} = P_{\text{capacity}} \times \left(\frac{SNR_{\text{downgrade}}}{SNR_{\text{capacity}}}\right)^{1.2} \tag{8}$$

Alternative approaches to the problem of power estimation were considered and found inferior. The set of derivatives of channel SNRs with respect to channel powers can be used to provide a local linear approximation to the channel SNR vector as a function of channel powers. Inverting this linear equation gives an approximate linear expression for the channel power allocation as a function of required SNRs for a given upgrade pattern. The power variation for downgraded channels can be approximately linear, but upgraded channels are distinctly nonlinear, so such a linear approximation provides a poor estimate of channel powers. Various second-order approximations exist, making use of both the set of derivative vectors and Hessian matrices of channel SNR with respect to channel power. One such second-order method can be obtained by formulating the minimum-margin problem in differentiable form using barrier functions [9], and then taking a single second-order Newton step. The first Newton step uses the nonlinearity derivatives of the starting point, the capacity-optimizing power allocation, which can be cached. For each different pattern of required SNRs for given upgrade patterns, new margin objective derivatives are formed using the fixed cached nonlinearity derivatives. Subsequent Newton steps cannot be cached due to the unique direction of the first step, and thus have time complexity proportional to the full margin optimization solution. For a system with N channels, the second-order estimate is $\mathcal{O}(N^3)$ times slower than that of (7) and (8), while providing less effective power estimates due to increased variance as compared to the true minimum-margin-optimizing power allocation.

The ultimate test of the power estimation method proposed in (7) and (8) is to compare the resulting discrete rate to that obtained with the ideal channel powers obtained from solving the minimum-margin problem. Due to the large amount of computation required for $\mathcal{O}(N^2 \log_2(N))$ minimum-margin calls, this is presented in Fig. 5(b) for a few points selected from the portion of the staircase offering the greatest possible benefit from improved discrete-rate optimization.

In addition to reducing the computational complexity of model-based optimization, achieving good performance with roughly estimated power allocations has significant benefits for live system optimization. Live system power optimization methods will typically not have access to optimization dual variables, as measurement and sample noise will lead to the use of different optimization methods than for theoretical models. Testing a large number of potential discrete-rate allocations is unreasonable on a live system. The rough power estimation used with this method supports the determination of the discrete rate pattern with estimated powers on an approximate theoretical channel model. This is much faster and minimizes system reconfigurations required. While absolute powers and margins will have modeling errors, the pattern determination is robust to such errors, as it depends upon comparing relative margins of different upgrade locations. With the final trial pattern of upgrades, a feasibility determination is then performed using the live system.

For a mesh network with S fiber sections and N channels per fiber, a quadratic-time swapping solution for the allocation problem scales faster in S than solving the minimum-margin optimization problem. Despite the negligible time required for point-to-point links, for sufficiently large meshes, the time required for the allocation problem could become dominant. For each upgrade, checking only the best N channels, as determined by maintaining information from position swap checks on upgrades with prior indices, reduces the search from $\mathcal{O}(S^2 N^2)$ to $\mathcal{O}(SN^2)$ which will be dominated by the minimum-margin solution time.

For mesh networks incorporating path length into the metric (2), or systems combining various discrete rate steps, the number of channel upgrades only partially correlates with the optimization metric. Given that channel rate upgrades remain a binary vector, assigning the upgrade pattern can be once again solved with a binary assignment problem. The equivalence in value of rate upgrades on different channels for point-to-point links is generally lost in mesh networks due to varying path lengths. When using a swapping method for arranging upgrades, the metric of the channel with the greatest margin is invalid. The inclusion of path length in the capacity metric leads to the selection metric for best upgrade channel being a function of channel

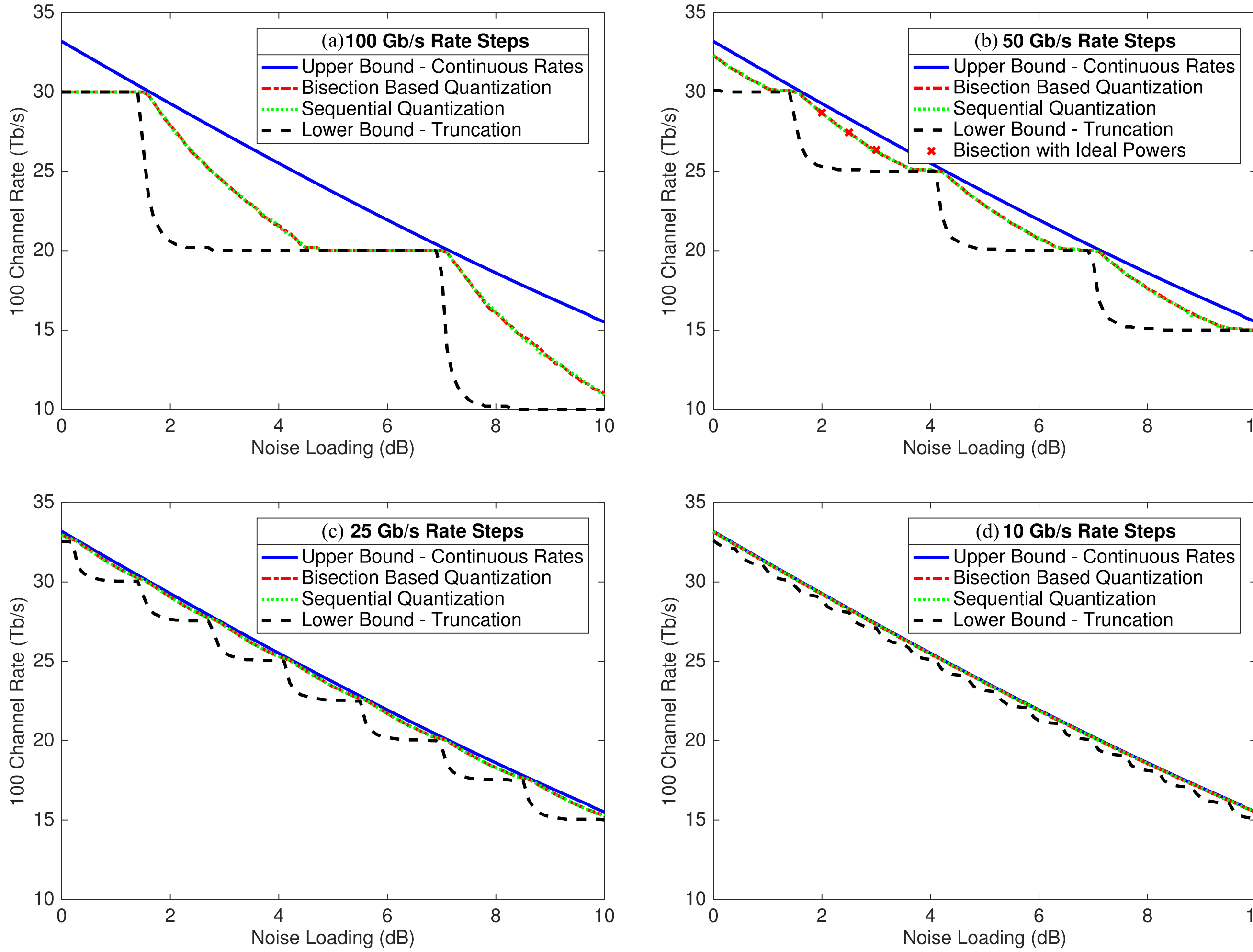

Fig. 5. Fiber capacity optimization with discrete rates on a point-to-point link. The convex continuous-rate upper bound, and the lower bound obtained by discrete truncation of the upper bound, provide loose limits on the possible performance of discrete-rate quantization. Sequential discrete-rate optimization, requiring $\mathcal{O}(N)$ margin optimization solutions, and a bisection-based method requiring $\mathcal{O}(\log(N))$ margin optimization solutions are shown. Channels are 50 GHz wide with ideal rectangular frequency spectra. A 10 dB range of operating conditions are shown in order to model discrete-rate optimization across a wide range of channel rates, encompassing the SNR range of multiple discrete rates. Fiber parameters are: 20 spans of fiber, each 100 km long with 21 dB of loss and dispersion $D = 17\ \mathrm{ps} \cdot \mathrm{nm}^{-1}\ \mathrm{km}^{-1}$. Amplifiers with 6 dB noise figures are assumed. Modulation and coding with a 1.5 dB gap to the Shannon capacity is assumed. (a) 100 Gb/s per channel rate granularity. The lower bound averages a rate of 78.5% of the continuous-rate bound across the range of conditions shown, while both methods of quantization achieve an average of 89.3%. (b) 50 Gb/s per channel rate granularity. The lower bound averages a rate of 90.0% of the continuous-rate bound, while both methods of quantization achieve an average of 96.5%. (c) 25 Gb/s per channel rate granularity. The lower bound averages a rate of 94.8% of the continuous-rate bound, while both methods of quantization achieve an average of 99.1%. (d) 10 Gb/s per channel rate granularity. The lower bound averages a rate of 97.9% of the continuous-rate bound, while the bisection-based quantization method achieves an average of 99.8% and the sequential method achieves an average of 99.9%.

margin and path length $f(M_i, d_i^{\xi})$. For the mesh network optimization results of this paper, the upgrade assignment metric of the form (9) is used, with θ tuned based on the distance exponent ξ. Assignment metric (9) is linear in d_i^{ξ}, similar to (2). For $\xi = 0$, where distance is removed from the mesh capacity metric, this metric for upgrade assignment reduces to the metric for a point-to-point link:

$$f(M_i, d_i^{\xi}) = d_i^{\xi} M_i^{\theta}. \tag{9}$$

Exponent θ of (9) is naturally much larger than ξ due to assignment feasibility depending very strongly on the margin achieved by patterns of grouped upgrades. In determining the best channel for an upgrade, margin differences between placing the upgrade in candidate channel A versus channel B will have a very significant impact on the feasibility of the proposed upgrade allocation. In comparing the margin between two candidate upgrade channels, a 0.2 dB (approximately 5%) difference in margin will have a very significant impact on the feasibility of a proposed upgrade allocation. The distance component of metric (9) can vary by a factor of 20 or more between network paths, and the exponent θ must keep this from dominating over the relative margin between two candidate upgrade channels. For the results of this paper, $\theta = \frac{120}{\ln(10)} \approx 52$ was chosen following tuning.

The distance-adjusted metric removes the invariance of channel index in assigning rate upgrades as different channels have different path lengths, and a fixed rate upgrade can have different weights in different channels. The removal of this invariance has the result of an increased sensitivity to the accuracy of power estimates during upgrade assignment.

C. General Integer Program Solvers

Packages for solving general nonlinear integer programs exist [23]. The nonlinear discrete capacity optimization problem is a poor fit for general integer program solvers. The nonlinear interaction is cubic in channel powers, preventing the use of many integer program packages that support only quadratic nonlinearities. The continuous power variables are a key component of the discrete-rate objective (6). With integer power variables, the formerly convex power optimization problem becomes a hard nonlinear integer program itself. General integer program packages are also limited to application on theoretical system models without the possibility of feedback from a live system operating over real-world channels. While general integer solvers may eventually obtain solutions that are asymptotically optimal, the need for timely solutions on large mesh networks will generally limit the general solver to linear-iteration solutions that pin on the order of one dimension per iteration. Specific efficient algorithms with that complexity have been discussed in Section III-A.

IV. Results

A. Discrete Rate Optimization of Point-to-Point Links

Discrete-rate optimization provides notable increases in fiber data rate over truncated rate assignments for large discrete-rate steps. For very small rate steps, discrete-rate optimization becomes asymptotically optimal much faster than truncation.

In order to examine the performance of quantization as the continuous-rate allocation crosses the discrete rate thresholds, a range of ASE noise loading is used corresponding to noise figures from 6 to 16 dB for the simulated optical amplifiers. The optimal power allocation, and hence total nonlinear noise, adapts to the ASE noise level, serving as equivalent to an increase in system length. Increasing the ASE level rather than adding additional spans allows a continuous range of SNR values to be evaluated. The simulated system features 20 spans of single-mode fiber, each 100 km long with 21 dB of loss and dispersion coefficient $D = 17\ \text{ps} \cdot \text{nm}^{-1}\ \text{km}^{-1}$. Modulation and coding with a 1.5 dB gap to the Shannon capacity is assumed, and simulations are performed for systems hosting 100 channels with a 50 Gbaud symbol rate.

With 100 Gb/s rate steps on 50 Gbaud channels, rate quantization increases the average lower bound achieved rate by 11.1%, averaged over one full staircase step with noise loading range 1.4 dB to 6.9 dB, as shown in Fig. 5(a). At 50 Gb/s rate steps on 50 Gbaud channels, rate quantization increases the average lower bound achieved rate by 7.2% for the same range of noise levels, with results shown in Fig. 5(b). With 25 Gb/s rate steps on 50 Gbaud channels, the absolute capacity gains are smaller, but the SNR step between channels is small enough that rate quantization achieves better than 99% of the continuous-rate upper bound averaged across both the same 1.4 dB to 6.9 dB range of noise loading, and the full range shown in Fig. 5(c). If the trend in the average truncated rate capacity indicated by Fig. 5 continues, a discrete rate set size of 5 Gb/s or below on 50 Gbaud channels would be required to achieve 99% of the continuous-rate upper bound.

The staircase formed by the truncated discrete-rate lower bound arises from the continuous-rate capacity allocation crossing the threshold for a particular discrete rate. The staircase has some curvature to the troughs, due to the capacity-maximizing continuous-rate allocation not being spectrally flat and instead curving up at the edges of the spectrum band, where channels have fewer interfering neighbors [9]. For the 100 Gb/s rate steps of Fig. 5(a), the optimized discrete-rate allocations are only able to improve upon truncation for about half of the width of a truncated rate staircase step. Where the optimized and truncated discrete rates are the same, the SNR increase required to achieve the next higher rate is beyond that which is achievable for any channel given the portion of inter-channel nonlinear interaction, and the neighboring channel restrictions. This is also slightly visible for the 50 Gb/s rate steps of Fig. 5(b). The two finer discrete rate steps, 25 Gb/s and 10 Gb/s, allow rate interpolation over the full width of the truncated rate staircase steps.

The threshold rate step size for supporting rate interpolation across the full noise range, and hence near-optimal discrete rate assignments, depends upon the ratio of SNRs of consecutive data rates. The ratio of SNRs of consecutive rates given a fixed rate step varies with baud rate, coding gap, implementation noise level, and total data rate. Fig. 2 shows the SNR increases required for a 50 Gbaud system with 25 Gb/s rate steps with varying levels of implementation noise. For the simulations of this paper, the inter-channel nonlinear interaction is able to mediate about a 1 dB optical SNR step between consecutive rates. This is what is provided for 150 Gb/s or higher data rates at 50 Gbaud with 25 Gb/s rate steps, without implementation noise. Implementation noise raises the size of steps in optical SNR between consecutive rates, and correspondingly reduces the efficiency of discrete-rate optimization for a fixed data rate step.

Fig. 5(b) includes additional trial points where bisection-based quantization is performed with ideal powers during the upgrade assignment problem solution. Ideal powers in this case are obtained by solving a convex minimum-margin problem. The repeated check of all $O(N^2)$ binary swaps for each potential upgrade count requires significant computation to identify ideal channel powers. Noise loading levels in the middle of the rate staircase have been selected where there is the most potential for improvement from improved quantization with ideal channel powers in swapping margin evaluation. The greatest gap to capacity is found at the larger 100 Gb/s rate step at the point where no channels are upgraded, but it is trivial to place a single upgrade at the edge of the channel band and verify infeasibility. The greatest potential impairment from rough power estimation is found where the SNR jump to upgraded rates is small enough to be mediated by the inter-channel nonlinear component, and where there are already many upgrades, such that the pattern of upgrade placement becomes important. For the three trial noise levels, the bisection-based quantization with

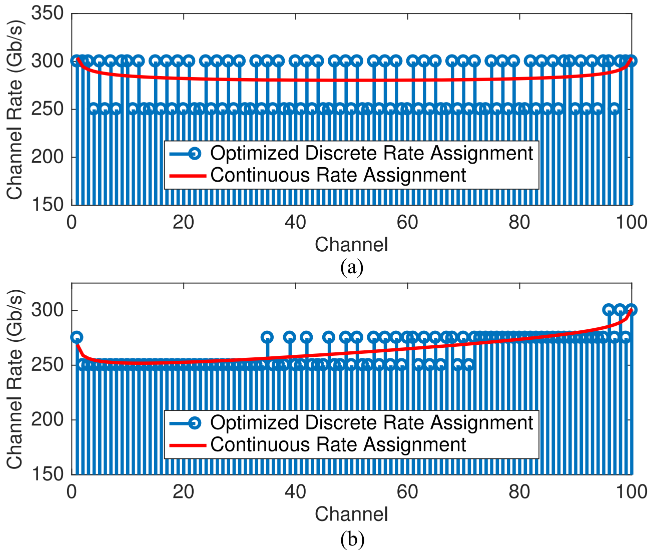

Fig. 6. Sample optimized discrete rate allocations. (a) 50 Gb/s rate steps, simulation parameters as in Fig. 5 with 2.5 dB of excess noise. 47 channels are upgraded, achieving an optimized rate of 27.45 Tb/s out of a 28.314 Tb/s bound. (b) 25 Gb/s rate steps, simulation parameters as in Fig. 5 with 2.5 dB of excess noise and a 2 dB noise ramp across the channel band. 30 channels are upgraded, achieving an optimized rate of 26.225 Tb/s out of a 26.431 Tb/s bound.

TABLE I
MEAN LOST CAPACITY (MLC) PER TRANSCEIVER FOR THE POINT-TO-POINT SCENARIOS OF FIG. 5

	Step Size			
Allocation Method	100 Gb/s	50 Gb/s	25 Gb/s	10 Gb/s
Truncation (Gb/s)	48.2	24.5	12.3	4.95
Bisection (Gb/s)	25.6	7.95	2.07	0.46
Sequential (Gb/s)	25.5	7.97	2.05	0.34

estimated powers is found to upgrade 67, 47 and 26 channels. With ideal powers, the bisection-based quantization method is able to place 68, 47 and 27 feasible upgrades. From this we conclude that the very coarse power estimate presented in (7) and (8) results in very close to ideal performance for the greedy binary swap solution to the nonlinear upgrade assignment problem. The rate allocation resulting in 47 upgraded channels is illustrated in Fig. 6(a).

The mean lost capacity (MLC) per transceiver was introduced as a discrete-rate performance metric in [16]. For the point-to-point scenarios of Fig. 5, the MLC per transceiver is reported in Table I. The MLC is averaged over the 1.4 dB to 6.9 dB range for one full staircase step at a 100 Gb/s rate step, as shown in Fig. 5(a). With truncation, the observed results for MLC are half of the discrete step size, as in [16]. The bisection-based or sequential discrete-rate optimizations significantly improve the MLC over the truncated rate lower bound. With 100 Gb/s rate steps on 50 Gbaud channels, the optimized rate MLC is half of that with truncation. With 50 Gb/s rate steps, the optimized rate MLC is 33% of that with truncation. With 25 Gb/s rate steps, the optimized rate MLC is six times smaller than that with

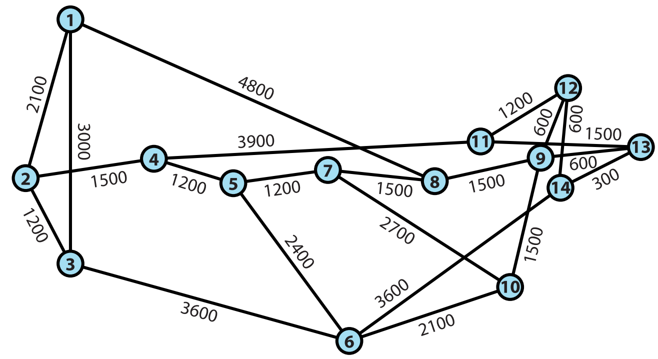

Fig. 7. The NSFNET network used as a reference network in mesh simulations, with link lengths in km [9], [24].

truncation. With 10 Gb/s rate steps, the optimized rate MLC is 10 to 15 times smaller than that with truncation.

Fig. 6 illustrates the optimized discrete rate assignments for two scenarios. Fig. 6(a) presents a basic flat-noise scenario with 50 Gb/s rate steps. The varying density of upgrades placed derives from the number of interfering neighbors, which is reduced at the edges of the channel band, and the number of upgrades placed nearby. As 47 channels are upgraded, the upgrade pattern cannot be exactly symmetric on 100 channels. Fig. 6(b) presents a scenario with 25 Gb/s rate steps and 2 dB of noise tilt across the channel band. The noise tilt dominates in determining upgrade placement, but the two edges of the band also benefit from discrete-rate optimization.

B. Mesh Discrete Rate Optimization With Distance Metric

Using a distance-adjusted capacity metric (2), unbiased discrete-rate optimization on mesh networks is feasible. Mesh optimization in this paper has been performed exclusively with the logarithmic-call bisection-based method. For mesh networks, the difference in computation time between linear and logarithmic-calls to margin optimization become very apparent. Further, with the desired distance-adjusted capacity metric, the margin optimization dual variables used in the successful sequential method no longer align with the desired metric.

Fig. 8 shows the results of discrete-rate optimization on the 14-node NSFNET mesh network of Fig. 7 using the distance-times-data-rate mesh metric of (2). Optical connections are randomly populated in the mesh network with a uniform probability of connecting all pairs of nodes. Congestion routing [24], with the congestion metric augmented by the expected nonlinear interference from neighboring channels, is used to perform routing and wavelength assignment for the channels. The network is filled until blocking, resulting in an average of 72% utilization of wavelengths in all fibers, with two fibers per link for bi-directional communication. The uniform connection demand among all pairs of nodes fills the long mid-continent fibers more than the short regional ones. Data rates are optimized following the conclusion of routing and wavelength assignment. With optimized continuous rates the average data rate per

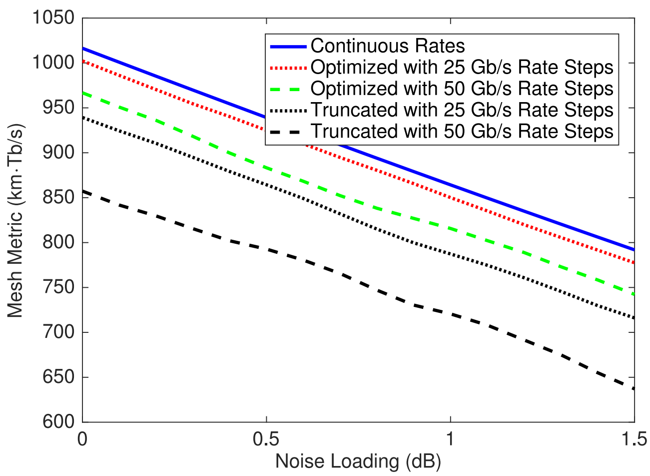

Fig. 8. Mesh capacity optimization using the distance-adjusted mesh metric of (2) with $a = 1$ and $\xi = 1$. An implementation noise SNR of 16 dB, and 2 dB of reliability margin are used in calculating the required SNR for channel data rates. Routing and wavelength assignment is done by congestion routing [24] with the congestion metric adjusted for number of nonlinearly interfering neighbors to incorporate wavelength assignment. Fiber parameters are as in Fig. 5, except for path lengths determined by the NSFNET mesh link lengths [9], [24] . The NSFNET mesh network diagram is shown in Fig. 7.

connection in this network is about 200 Gb/s, which corresponds to dual-polarization QPSK at 50 Gbaud.

The lower bound of Fig. 8 remains the discrete rate assignment with discrete rate strictly smaller than the continuous-rate optimum allocation. The truncated rate staircase of the point-to-point links of Fig. 5 vanishes in this mesh scenario due to the diversity of link lengths. The lower bounds with 50 Gb/s and 25 Gb/s discrete rate steps achieve an average of 83.5% and 91.6% of the continuous-rate upper bound, respectively. Discrete-rate optimization with 50 Gb/s rate steps averages 94.3% of the continuous-rate upper bound, for a 13% increase in distance-times-capacity metric over the lower bound. Discrete-rate optimization with 25 Gb/s rate steps achieves 98.4% of the continuous-rate upper bound, for a 7.5% increase in distance-times-capacity metric over the lower bound. While optimizing the mesh network discrete rates for the $\xi = 1$ distance-times-capacity metric, the uniform load metric ($\xi = 0$) is also improved, with increases of 10.4% for 50 Gb/s rate steps and 6.0% for 25 Gb/s rates steps.

While the point-to-point results cover a scenario with spectrally flat noise and no implementation noise, the mesh network results cover a much more complex discrete-rate optimization scenario. An implementation noise SNR of 16 dB is used, increasing the SNR steps. The pattern of channel assignments, along with some vacancies, lead to varying nonlinear noise levels and unusual optimal power allocations. In the point-to-point results of Section IV-A, averages of 67.5% and 83.2% of the fiber capacity lost due to truncation is recovered from optimization for 50 and 25 Gb/s rate steps, respectively. For the mesh network results of Fig. 8, averages of 65.8% and 81.1% of the mesh metric lost due to truncation is recovered from optimization for 50 and 25 Gb/s rate steps, respectively. The consistency of optimization performance in recovering capacity lost due to truncation, comparing the simple point-to-point and complex mesh application scenarios, supports the general value of optimization.

V. Discussion

Performing discrete-rate capacity optimization in an efficient manner depends upon being able to obtain information about the quality of various arrangements of rate upgrades without solving a minimum-margin problem for each potential arrangement. The sequential optimization of Section III-A obtains information about the quality of placement from the margin optimization dual variables. For the bisection method, the upgrade assignment problem obtains information about many potential patterns via roughly estimated power allocations.

In the limit of small discrete rate steps, truncating the continuous-rate assignment for each channel will approach capacity-achieving. This is efficient and simple from the modem standpoint, but pushes complexity external to the modem to source data streams of the precise rates available. If the precise rates cannot be delivered, and transmitted data is padded, then the capacity gain from the finely discretized rate is lost. Discrete-rate optimization was shown in Section IV-A to converge towards the continuous-rate upper bound much more rapidly than what is achieved via truncated rates and decreasing the discrete rate step. This is also achieved with less external complexity, as standardized rates can be used as discrete rate steps.

The proposed optimization methods are heuristics for the NP-hard discrete-rate problem (6). As heuristics, they are suboptimal solution methods, but valuable gains in system capacity have been demonstrated in Section IV. The dual-variable-based sequential method and the logarithmic-call solution method for solving upgrade assignment problems are entirely independent approaches to the discrete-rate optimization problem. The only part that is shared is the minimum-margin problem solver, which has been demonstrated to be convex and optimal for the Gaussian noise model regime under investigation [9]. While there are many approaches to discrete rate assignment that provide inferior performance, closer to the lower bound which is naive truncation, the very similar performance of these two best independent heuristics suggests a sharp decrease in the number of feasible discrete rate allocations near the rates obtained by these methods. The quadratic-call method obtained by solving the upgrade assignment method with ideal channel powers supports this hypothesis by obtaining solutions that support zero to one more rate upgrades at the cost of dramatically higher computational complexity.

The results of Fig. 5 cover a scenario with a sequence of identical spans, and avoiding SRS or non-flat amplifier noise. With non-identical spans, and SRS or other complexities, performing the continuous-rate optimization becomes more difficult and may incorporate different power levels for different span lengths. The same power estimation method discussed in Section III-B extends to such more complex scenarios as it adjusts the continuous-rate solution which absorbs the complexity. If average powers differ per span, or for mesh networks where channel neighbors vary along their path, SNR levels must be evaluated by the standard method of summing reciprocals for each distinct section. Non-flat noise spectra are well supported by these algorithms, which is supported by the results of Fig. 6(b) and the mesh results of Section IV-B with varying nonlinear interference, and SNR steps, per channel due to link path diversity.

In mesh networks, the method of routing and spectrum assignment has a significant impact on the overall network performance [15], [25]. The first stage of discrete-rate optimization, optimizing the continuous-rate channel capacities or mesh metric, obtains the solution that optimally utilizes the state of the network following routing and wavelength assignment. The second stage of discrete-rate optimization then attempts to approach this upper bound with discrete rates. The network performance obtained will be significantly impacted by the routing and wavelength assignment method, but the relative performance gain of discrete-rate optimization as compared to truncation will be preserved.

The mesh network results presented use $\xi = 1$, as this removes the bias in power assignments and resulting link SNRs during continuous-rate capacity optimization. Similar improvements in the capacity metric are obtained for a range of values of ξ, including $\xi = 0$. Without a good assignment metric such as that of (9), optimization of discrete rates performs poorly for $\xi \neq 0$. The case $\xi = 0$ is special, as the ordering of channel upgrades depends only on the channel margins. Thus, the results for $\xi = 1$ present a much more challenging optimization scenario than the point-to-point links or a mesh using the metric with $\xi = 0$.

The continuous-rate upper bound is convex and optimally solvable, but does not give a very tight bound on the observed feasible discrete rate allocations with large rate steps. Tighter convex bounds are obtained by solving mixed capacity / margin objectives for all potential discrete rate upgrade patterns for a certain subset of channels. Due to the exponential number of optimizations required, for 80- or 100-channel point-to-point links, only a small portion of the channels can be fixed in such trials, and little improvement in the upper bound can be obtained.

VI. Conclusion

An efficient heuristic for solving the discrete-rate capacity optimization problem is presented. This algorithm minimizes the number of margin optimization calls required for model-based optimization, or minimizes the number of system configuration trials required if live system measurements are used in determining discrete rate assignment feasibility. For 50 Gbaud channels, 50 Gb/s rate steps with truncation naively average 90.0% of the continuous-rate upper bound across the 10 dB range of noise levels tested. The proposed heuristic is able to increase the average utilization to 96.5% of the continuous-rate upper bound. In a mesh network using the unbiased metric of distance times capacity on each channel, discrete-rate optimization is able to increase the average capacity metric over truncated rates by 13% with 50 Gb/s rate steps and 7.5% with 25 Gb/s rate steps.

References

[1] G.-H. Gho, L. Klak, and J. M. Kahn, "Rate-adaptive coding for optical fiber transmission systems," *J. Lightw. Technol.*, vol. 29, no. 2, pp. 222–233, Jan. 2011.

[2] M. Karlsson and E. Agrell, "Multidimensional modulation and coding in optical transport," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 876–884, Feb. 2017.

[3] K. Roberts, Q. Zhuge, I. Monga, S. Gareau, and C. Laperle, "Beyond 100 Gb/s: Capacity, flexibility, and network optimization [invited]," *J. Opt. Commun. Netw.*, vol. 9, no. 4, pp. C12–C24, 2017.

[4] *Flex Ethernet Implementation Agreement* (IA # OIF-FLEXE-01.0), Mar. 2016.

[5] J. Campello, "Practical bit loading for DMT," in *Proc. IEEE Int. Conf. Commun.*, 1999, vol. 2, pp. 801–805.

[6] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3857–3879, Dec. 2012.

[7] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN model of non-linear propagation and its applications," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 2014.

[8] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Opt. Express*, vol. 22, no. 13, pp. 16 335–16 362, 2014.

[9] I. Roberts, J. M. Kahn, and D. Boertjes, "Convex channel power optimization in nonlinear WDM systems using Gaussian noise model," *J. Lightw. Technol.*, vol. 34, no. 13, pp. 3212–3222, Jul. 2016.

[10] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[11] H. Nakashima, Y. Akiyama, T. Hoshida, T. Oyama, T. Tanimura, and J. C. Rasmussen, "Launch power optimization co-operated with digital nonlinear compensator for elastic optical network," in *Proc. Opto-Electron. Commun. Conf.*, 2015, pp. 1–3.

[12] D. J. Ives, P. Bayvel, and S. J. Savory, "Adapting transmitter power and modulation format to improve optical network performance utilizing the Gaussian noise model of nonlinear impairments," *J. Lightw. Technol.*, vol. 32, no. 21, pp. 3485–3494, Nov. 2014.

[13] D. J. Ives, P. Bayvel, and S. J. Savory, "Routing, modulation, spectrum and launch power assignment to maximize the traffic throughput of a nonlinear optical mesh network," *Photon. Netw. Commun.*, vol. 29, no. 3, pp. 244–256, 2015.

[14] A. Pagès, J. Perelló, S. Spadaro, and J. Comellas, "Optimal route, spectrum, and modulation level assignment in split-spectrum-enabled dynamic elastic optical networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 6, no. 2, pp. 114–126, 2014.

[15] L. Yan, E. Agrell, H. Wymeersch, and M. Brandt-Pearce, "Resource allocation for flexible-grid optical networks with nonlinear channel model," *J. Opt. Commun. Netw.*, vol. 7, no. 11, pp. B101–B108, 2015.

[16] D. J. Ives, A. Alvarado, and S. J. Savory, "Throughput gains from adaptive transceivers in nonlinear elastic optical networks," *Photon. Netw. Commun.*, vol. 35, pp. 1280–1289, Mar. 2017.

[17] D. A. A. Mello, A. N. Barreto, T. C. de Lima, T. F. Portela, L. Beygi, and J. M. Kahn, "Optical networking with variable-code-rate transceivers," *J. Lightw. Technol.*, vol. 32, no. 2, pp. 257–266, Jan. 2014.

[18] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and J. G. David. Forney, "MMSE decision-feedback equalizers and coding-part II: Coding results," *IEEE Trans. Commun.*, vol. 43, no. 10, pp. 2595–2604, Oct. 1995.

[19] R. Hemmecke, M. Köppe, J. Lee, and R. Weismantel, "Nonlinear integer programming," in *50 Years of Integer Programming 19582008: The Early Years and State-of-the-Art Surveys*, M. Jünger, T. Liebling, D. Naddef, G. Nemhauser, W. Pulleyblank, G. Reinelt, G. Rinaldi, and L. Wolsey, Eds. New York, NY, USA: Springer-Verlag, 1999.

[20] M. Fischetti, F. Glover, and A. Lodi, "The feasibility pump," *Math. Program.*, vol. 104, no. 1, pp. 91–104, 2005.

[21] P. Bonami and J. P. Gonçalves, "Heuristics for convex mixed integer nonlinear programs," *Comput. Optim. Appl.*, vol. 51, no. 2, pp. 729–747, 2012.

[22] P. M. Pardalos, L. S. Pitsoulis, and T. D. Mavridou, "Nonlinear assignment problems," in *Large Scale Computations in Air Pollution Modelling. NATO Science Series* (Series 2: Environmental Security), vol. 57., Z. Zlatev, J. Brandt, P. J. H. Builtjes, G. Carmichael, I. Dimov, J. Dongarra, H. van Dop, K. Georgiev, H. Hass, and R. S. Jose, Eds. Dordrecht, The Netherlands: Springer, 1999, pp. 261–273.

[23] IBM Corporation, Armonk, NY, *IBM ILOG CPLEX software*. (2017). [Online]. Available: www.ibm.com/us-en/marketplace/ibm-ilog-cplex, Accessed on: Aug. 18, 2017.

[24] S. J. Savory, “Congestion aware routing in nonlinear elastic optical networks,” *IEEE Photon. Technol. Lett.*, vol. 26, no. 10, pp. 1057–1060, May 2014.

[25] D. J. Ives, P. Bayvel, and S. J. Savory, “Routing, modulation, spectrum and launch power assignment to maximize the traffic throughput of a nonlinear optical mesh network,” *Photon. Netw. Commun.*, vol. 29, pp. 244–256, Mar. 2015.

Ian Roberts received the Bachelor’s degree from the University of Ottawa, Ottawa, ON, Canada, and the Master’s degree from Stanford University, Stanford, CA, USA, both in electrical engineering, and is currently working toward the Ph.D. degree at Stanford University. Before his graduate studies, he worked at Ciena on DSP hardware development for WaveLogic 3 Extreme. His interests include optical communication, optimization, and compensation of nonlinear propagation. Mr. Roberts is a student member of OSA.

Joseph M. Kahn (F’00) received the A.B., M.A., and Ph.D. degrees in physics from the University of California (U.C.), Berkeley, CA, USA, in 1981, 1983, and 1986, respectively. From 1987 to 1990, he was at AT&T Bell Laboratories, Crawford Hill Laboratory, Holmdel, NJ, USA. He demonstrated multi-gigabit-per-second coherent optical fiber transmission systems, setting world records for receiver sensitivity. From 1990 to 2003, he was a faculty member in the Department of Electrical Engineering and Computer Sciences, U.C., performing research on optical and wireless communications. Since 2003, he has been a Professor of electrical engineering at Stanford University, Stanford, CA, USA, where he heads the Optical Communications Group. In 2000, he helped found StrataLight Communications, where he served as Chief Scientist from 2000 to 2003. StrataLight was acquired by Opnext, Inc. in 2009. His current research interests include fiber-based imaging, spatial multiplexing, rate-adaptive and spectrally efficient modulation and coding methods, coherent detection and associated digital signal processing algorithms, digital compensation of fiber nonlinearity, and free-space systems. Prof. Kahn served as a Technical Editor of the IEEE Personal Communications Magazine from 1993 to 2000. From 2009 to 2014, he was an Associate Editor of the IEEE/OSA JOURNAL OF OPTICAL COMMUNICATIONS AND NETWORKING. He received the National Science Foundation Presidential Young Investigator Award in 1991.