Photonic Network Communications (2024) 47:18–29
https://doi.org/10.1007/s11107-023-01008-x

ORIGINAL PAPER

Dual optimization for enhancing TRA in MCF-based SDM-EONs

Shrinivas Petale[1] · **Suresh Subramaniam**[1]

Received: 1 August 2023 / Accepted: 30 November 2023 / Published online: 3 January 2024

Abstract
A fine-grained, flexible frequency grid of elastic optical transmission and space division multiplexing (SDM) in conjunction with spectrally efficient modulations is an ideal solution for the impending capacity crunch. The routing, modulation, core, and spectrum assignment (RMCSA) problem is an important lightpath resource assignment problem in SDM elastic optical networks (SDM-EONs). Intercore Crosstalk (XT) degrades the quality of parallel transmissions on adjacent cores, and the RMCSA algorithm must satisfy XT requirements while optimizing network performance. There is an indirect trade-off between spectrum utilization and XT tolerance; while higher modulations are more spectrum-efficient, they are also less tolerant of XT because they allow fewer connections between neighboring cores on the overlapping spectrum. In the presence of XT, the Tridental Resource Assignment algorithm (TRA) has been shown to optimally assign resources in multicore fiber networks. Using the tridental coefficient (TC), this algorithm strikes a balance between spectrum utilization and XT. However, TRA is computationally expensive because TC is calculated for all the possible candidate resource sets for each lightpath request. In this paper, we demonstrate that acceptable or better TRA performance can be achieved while reducing computational overhead than traditional TRA. We achieve that using different stages of offline optimization, required only once. We demonstrate that by using tailored weights in TC, the performance of TRA can be enhanced further. We observe that acceptable TRA performance can be achieved with as low as 40% of the total computations and that adjusting weights reduces the bandwidth blocking of TRA, which is already performing better than the baseline algorithms, by approximately two orders of magnitude.

Keywords TRA · Tridental resource assignment algorithm · Dual optimization · SDM EON · Crosstalk · Multicore fibers

1 Introduction

Multicore fiber (MCF)-based space division multiplexed elastic optical networks (SDM-EONs) can provide a viable solution to bandwidth constraints in cloud services, 5G and 6G communications, high-resolution game streaming, and data center networks which rely mainly on optical networks [1]. SDM-EON enables parallel optical signal transmission over MCFs utilizing distance-adaptive multi-carrier transmission. However, the intercore crosstalk (XT) between these parallel transmissions on weakly coupled cores degrades quality of transmission (QoT) [2, 3].

The problem of route, modulation, core, and spectrum (RMCSA) assignment deals with selection of resources. In MCF, modulation, core and spectrum are the dominant parameters affecting the trade-off between spectrum utilization and XT accumulation in the network. The intensity and effect of XT vary according to the selected core, modulation, and spectrum. Higher modulation requires less spectrum to transmit information, but they are extremely sensitive to XT. However, lower modulations are less sensitive to XT interference but require more spectrum. On the other hand, the physical geometry of the core determines the number of adjacent cores and, consequently, the amount of XT accumulation. Choosing a core with a greater number of adjacent cores with active parallel transmissions increases the likelihood of XT accumulation. Due to spectrum continuity and contiguity constraints of optical communication, the current spectrum selection determines the availability of spectrum and spectrum options for future connection requests. A suboptimal choice of core, modulation, and spectrum can also result in fragmentation, QoS obstruction, detours, and other issues.

✉ Shrinivas Petale
srpetale@gmail.com

[1] The George Washington University, Washington, DC, USA

Prior works [4, 5] have addressed the RMCSA problem with the help of an RMCSA algorithm, the tridental resource assignment algorithm (TRA). TRA takes into account the effect of selection of core, modulation, and spectrum to balance the spectrum utilization versus XT accumulations trade-off. Further, an improved variant of TRA for transparent and translucent networks is proposed in [5]. It is demonstrated that TRA is extraordinarily effective at allocating resources to balance the trade-off. It does that by computing the tridental coefficient (TC) for all available resource combinations. The TC is the sum of three parameters: capacity loss [4], spectrum utilization, and spectrum location [5]. TRA relies on offline path priority calculation which helps it to select better routes and balance the load in the network. It then looks for the resource candidate that offers the lowest TC.

TRA searches through all possible combinations of resources, resulting in a greater computational burden in order to obtain the optimal resources. In addition, equal weight was assumed for all three parameters in the calculation of TC. We found that by changing the weights we can improve the performance of TRA. This improvement is then used to compensate the time and computational overload in search of the resource combination with the lowest TC. In this paper, we first demonstrate that it is possible to save computational time while maintaining an acceptable increase in bandwidth blocking by reducing the number of candidate resources. Then, we present a method for obtaining the weights to further enhance the efficacy of TRA. TRA outperforms both the baseline and literature-based RMCSA algorithms [4, 5]. With the proposed optimizations, we observe additional substantial enhancements to TRA's performance.

In our previous work [6], we presented two different ways to optimize the performance of TRA. We showed that the reduction of resources leads to saving computational time and still results in acceptable performance. We also showed that if we weight the tridental coefficient (TC) and then optimize the weights using bruteforce method, we can improve the performance of TRA. A preliminary version of this work was presented in [6]. This paper makes following additional contributions over [6]:

i. dual optimization of TRA for various scenarios (main highlight of the work),
ii. detailed analysis of blocking variation with respect to weights and step-size,
iii. percentage saving of resources per datarate instead of a constant resource utilization [6],
iv. detailed study and analysis of TRA for different topologies with different XT thresholds,
v. proposed variants of TRA to achieve equivalent performance with as little as half of the computations.

The paper is organized as follows. The network model and problem statement are introduced in Sect. 2. Important concepts useful to understand the proposed work is presented in Sect. 3. Details of simulation setup required to understand the results are given in Section 4. The detailed steps of proposed dual optimization of TRA are presented in Sect. 5. Section 6 presents simulation results and the last section concludes the work.

2 Network model and preliminaries

In this section, we provide more context of the network model used in the paper and details about the SDM-EON and its characteristics involved in its design. The abbreviations used in the paper are listed in Table 1. We assume SDM-EON with multicore fibers (MCFs) with weakly coupled cores packed within a single fiber. Each link comprises a single MCF in both directions, but the concepts can be extended to networks with multiple fibers per link. Coherent transceivers (TRXs) achieving flexible and adaptive data

Table 1 List of abbreviations

Abbreviation	Phrase
SDM	Space division multiplexing
RMCSA	Routing, modulation, core, and spectrum assignment
SDM-EON	Space division multiplexed elastic optical networks
XT	Intercore crosstalk
TRA	Tridental resource assignment algorithm
TC	Tridental coefficient
MCF	Multicore fiber
QoT	Quality of transmission
TRX	Transceivers
FS	Frequency slice
PM	Polarization multiplexed
QPSK	Quadrature phase shift keying
QAM	Quadrature amplitude modulation
SCh	Superchannel
XT_{μ}	Average XT between two adjacent cores after a single span of propagation
EVM	Error-vector magnitude
BER	Bit-error rate
RC	Resource candidates
SP	Shortest paths
CL	Capacity loss
BBP	Bandwidth blocking probability
DT	Generic German topology
EURO	European topology

transmission are used. These TRXs operate on a flexible grid with a granularity of 12.5 GHz, enabling the allocation of optical carriers in finer frequency slots whose entire bandwidth occupies 37.5 GHz which is three frequency slices (FSs) [7–10]. The flexibility in grid granularity allows for efficient spectrum utilization, making it possible to accommodate varying data rates and different modulations.

With the capability of adjustable bit rates and support for multiple modulations, the TRXs in SDM-EON offer adaptability to meet diverse application requirements. We use the TRX which operates at 28 GBaud. The set of modulations is denoted as D. Polarization multiplexed-quadrature phase shift keying (PM-QPSK), polarization multiplexed-8 quadrature amplitude modulation (PM-8QAM), PM-16QAM, PM-32QAM, and PM-64QAM ($|D| = 5$), are considered for their offered spectral efficiency and sensitivity to crosstalk. To accommodate high bit rate requests that exceed the capacity of a single TRX with a particular modulation, the SDM-EON employs superchannels (SChs). A superchannel aggregates multiple optical carriers, each occupying three frequency slots (37.5 GHz of spectrum), to form a higher-capacity transmission channel. To ensure spectral separation between SChs and minimize inter-channel interference, a 12.5 GHz guard band is utilized. We do not consider traffic grooming and regeneration [11]. There exists K shortest paths between each source-destination pair [12].

We examine 3-core and 7-core geometries in the network model [2, 13, 14]. XT due to parallel transmissions on overlapping spectrum only affects neighboring cores, and cores that are not adjacent remain unaffected. For instance, in a 3-core fiber, each core has two neighboring cores [15]. Similarly, in a seven-core fiber, the outer cores are surrounded by three adjacent cores, while the center core is encircled by six adjacent cores [15]. Moreover, we ensure spatial continuity, meaning that each lightpath is assigned to the same core throughout its path. Lightpath requests are received dynamically at a Poisson arrival rate, with an exponentially distributed mean holding time of one (arbitrary units). Additionally, the datarates of the lightpaths are uniformly distributed within a predefined range. The set of datarates is denoted as M and the ith datarate in Gbps is denoted as m_i.

The closest neighboring cores that contribute to the XT when the overlapping spectrum on these cores is occupied by another ongoing parallel transmission are called as litcores and the number is denoted by γ. An analytical model based on coupled-power theory is used to determine the mean XT (XT_μ) in MCF links as shown in (1). Here, path length (L), power-coupling coefficient (h), and the number of adjacent cores contributing to XT in a given core and spectrum band (γ) are used to calculate XT_μ. Maintaining a low XT level ensures high-quality transmission and minimizes the need for regeneration, which can impact network performance and add complexity. In the resource allocation phase, core continuity, spectrum continuity, and contiguity constraints are imposed to facilitate seamless transmission and switching between cores. An average XT of − 40 dB and − 25 dB between neighboring cores after a single span of propagation is maintained, respectively. These values allow the use of advanced modulations without requiring complex DSP algorithms and ensures efficient spectral utilization.

$$\mathrm{XT}_\mu = \frac{\gamma - \gamma \exp(-(\gamma+1)hL)}{1 + \gamma \exp(-(\gamma+1)hL)} \tag{1}$$

Transmission reach of a modulation is the maximum distance that can be traversed by a lightpath while maintaining the Quality of Transmission (QoT) without the requirement of regeneration. The maximum transmission reach ($L_{\max}$) can be calculated using (2), where L_{span}, EVM_0, and $\mathrm{EVM}_{\mathrm{span}}$ represent the span length, the error-vector magnitude (EVM) to achieve the target bit-error rate (BER), and the EVM of the span, respectively [7]. The EVM-obtained transmission reach guarantees a BER of 3.8×10^{-3} for h of 3.15×10^{-9} and span length of 50 km.

$$L_{\max} = L_{\mathrm{span}} \left\lfloor \frac{\mathrm{EVM}_0^2}{\mathrm{EVM}_{\mathrm{span}}^2} \right\rfloor. \tag{2}$$

The associated transmission reach model for − 40dB is shown in Table 2 and for -25dB is shown in Table 3 when a TRX of 28 GBaud and the above network model are used [7]. The transmission reach in the network depends on both the selected modulation and γ. Higher modulations offer greater spectral efficiency but are more sensitive to crosstalk, resulting in a shorter transmission reach compared to lower modulations for the same value of γ. Additionally, as γ increases, the transmission reach decreases due to the increased accumulation of XT.

To regulate the choice of modulation for a lightpath, a maximum value of γ is always considered by TRA. The modulations are searched from higher to lower to balance out the spectrum consumption. For dth modulation, the allowable number of γ litcores are calculated by checking the corresponding TR, denoted as T_d^γ, with the path length, L, such that the $L \leq T_d^\gamma$.

3 Definitions and important concepts

In this section, we discuss important concepts, calculations and definitions useful for understanding the proposed optimization.

3.1 Spectrum requirement and slice window

The standard equation to compute the number of FSs for ith datarate m_i and dth modulation is (3). Where n_i is the number of FSs required by TRX/carrier, δ is the spectrum width of an FS and η_d is the spectral efficiency of the dth modulation.

Table 2 Transmission reaches (in km) of modulations for different values of allowable litcore (γ) for $XT_{\mu} = -40$dB for 28GBaud transceiver (from Table I in [7])

γ (Litcores)	XT per span (dB)	Modulations ($\|D\| = 5$)				
		PM-QPSK ($d = 1$)	PM-8QAM ($d = 2$)	PM-16QAM ($d = 3$)	PM-32QAM ($d = 4$)	PM-64QAM ($d = 5$)
0	None	5200	2050	1100	550	250
1	−40.0	4650	1850	1000	500	250
2	−37.0	4200	1650	900	450	200
3	−35.2	3850	1500	800	400	200
4	−34.0	3550	1400	750	400	150
5	−33.0	3300	1300	700	350	150
6	−32.2	3050	1200	650	300	150

Table 3 Transmission reaches (in km) of modulations for different values of allowable litcore (γ) for $XT_{\mu} = -25$dB (restrictive) for 28GBaud transceiver (from Table I in [7])

γ (Litcores)	XT per span (dB)	Modulations ($\|D\| = 5$)				
		PM-QPSK (d=1)	PM-8QAM (d=2)	PM-16QAM (d=3)	PM-32QAM (d=4)	PM-64QAM (d=5)
0	None	5200	2050	1100	550	250
1	−25.0	1100	400	200	100	50
2	−22.0	600	200	100	50	0
3	−20.2	400	150	50	0	0
4	−19.0	300	100	50	0	0
5	−18.0	250	100	50	0	0
6	−17.2	200	50	0	0	0

$$\beta_d^{\mathrm{m}} = \left\lceil \frac{m_i}{(n_t\delta)\eta_d} \right\rceil \times n_t \tag{3}$$

The collection of spectrum slices required for a datarate on a lightpath is called as slice window. A connection can only be assigned an *available* slice window. Definition V.1 of [4] and Definition 6.1 of [5] define the available slice window precisely. An available slice window is essentially a window of spectrum slices where all the spectrum slices are free, can accommodate the current connection based on the XT constraint of the selected modulation, and does not violate the QoT constraint of ongoing connections on overlapping spectrum on adjacent cores according to their XT sensitivities.

3.2 Filtered modulations

Let D be the complete set of modulations. Let D^{m} be the filtered set of modulations obtained from D using β_d^{m} as per the definition of Def. 1.

Definition 1 (Filtered Modulations) The filtered set of modulations, denoted as D^{m}, is a set of filtered candidate modulations from the complete set of modulations, denoted as D, for mth datarate. The filtering is done based on the spectrum requirement, β_d^{m}, by forming the subsets of modulations by grouping modulations with same β_d^{m} in a subset and then selecting the lowest modulation in each subset as a candidate modulation in D^{m}.

We understand the filtered modulation with an example of 200 Gbps ($= m_5$) connection. The spectrum requirement for a 200 Gbps connection for PM-QPSK ($\eta_1 = 2.96$), PM-8QAM ($\eta_2 = 4.44$), PM-16QAM ($\eta_3 = 5.92$), PM-32QAM ($\eta_4 = 7.4$) and PM-64QAM ($\eta_5 = 8.88$) [7, 10] when $n_t = 3$ and $\delta = 12.5$ GHz is 6 ($= \beta_1^5$), 6, 3, 3, and 3 ($= \beta_5^5$), respectively. The set of modulations D reduces from { PM-QPSK, PM-8QAM, PM-16QAM, PM-32QAM, PM-64QAM } to D^m = { PM-QPSK, PM-16QAM }.

3.3 Resource candidates

The number of resource candidates (RCs) per core for a datarate m, denoted as N^{m}, is calculated using (4). Thus, the maximum number of RCs on shortest path for a datarate m is given by (5). Every slice window is separated from adjacent slice windows by a guardband, denoted as g_{b}. Remember that as the spectrum continuity is imposed, an RC is counted only once on all the links. In addition, the RC is considered as available when slice window on each link is available.

$$N_{\mathrm{d}}^{\mathrm{m}} = ((S + g_{\mathrm{b}}) - (\beta_d^{\mathrm{m}} + g_{\mathrm{b}}) + 1) \\ = (S - \beta_{\mathrm{d}}^{\mathrm{m}} + 1) \tag{4}$$

$$N_{\max}^{\mathrm{m}} = \sum_{d \in D^{\mathrm{m}}} N_d^{\mathrm{m}} C \\ = \left(|D^{\mathrm{m}}|(S+1) - \sum_{d \in D^{\mathrm{m}}} \beta_d^{\mathrm{m}} \right) C \tag{5}$$

We take the example of a 200 Gbps (= m_5) connection to understand the calculation of N_d^m and $N_{\max}^{\mathrm{m}}$. We obtained $\beta_1^5 = 6$, $\beta_2^5 = 6$, $\beta_3^5 = 3$, $\beta_4^5 = 3$, $\beta_5^5 = 3$, and $|D^{\mathrm{m}}| = 2$ in the example discussed in Section 3.2. For $S = 320$, $g_b = 1$, using (4), we get $N_1^5 = 315$, $N_2^5 = 315$, $N_3^5 = 315$, $N_4^5 = 318$, and $N_5^5 = 318$. For a seven-core fiber (C = 7), using (5), we get 4347 (= $(2 \times (320 + 1) - 21) \times 7$). All the calculated values resource candidates for different core types and datarates are given in Table 5.

3.4 Tridental resource assignment (TRA)

The RMCSA problem is complex, especially when impairments such as XT are taken into account; therefore, we decompose it into two distinct sub-problems: an offline planning component that selects path priorities and a dynamic MCSA provisioning algorithm.

3.4.1 Offline planning

Traffic forecasting can be used to predict the long-term average traffic intensities for various node-pairs, so we presume that the dynamic traffic loads on all routes (i.e., $s - d$ pairs) are known. The path priorities for a route are chosen so as to distribute the traffic load evenly across the multiple paths. We observe that these path priorities merely establish the search order for paths for incoming connection requests. The actual availability of a path is contingent upon the network state at the time a connection request is received and cannot be predicted in advance. Path prioritization is determined by path probabilities that optimize link traffic. Similar to our previous work in [16], K shortest paths (SPs) for each $s - d$ pair are precomputed, and a MILP is solved offline to determine the path probabilities. The SPs are ranked according to their path probabilities, from most likely to least likely. The objective function (z_{fs}) of the MILP is to minimize the sum of the average and maximal link loads, as shown in (6), where $\mathfrak{L}$ is number of links in the network, e is an arbitrary network link, F^{e} is the number of fibers on link e, C is the number of cores, R is the number of routes in the network, r is an arbitrary route, W_{r} is the load for route r; depends on the traffic pattern; $W_{\mathrm{r}} = 1, \forall\, r$ for uniform traffic, $u_{\mathrm{r}}^{\mathrm{e}}$ denotes the load distribution on link e on route r [5].

$$\text{Minimize } z_{fs} = \frac{1}{\mathfrak{L}} \sum_{e=1}^{L} \sum_{r=1}^{R} \frac{W_r u_{\mathrm{r}}^{\mathrm{e}}}{F^{\mathrm{e}} C} + \max_{\mathrm{e}} \sum_{r=1}^{R} \frac{W_{\mathrm{r}} u_{\mathrm{r}}^{\mathrm{e}}}{F^{\mathrm{e}} C} \tag{6}$$

3.4.2 Online provisioning

Capacity Loss The capacity loss (CL) represents the loss in capacity when a slice window is assigned with a modulation, core and spectrum. The CL for the nth slice window on the cth core using dth modulation on kth shortest path, r^k, is denoted as $\psi_{n,c}^{k,d}$ and is calculated using (7). It is the difference between the capacity before provisioning, denoted as $\upsilon_n^{k,d}$, and the capacity after provisioning, denoted as $\upsilon'^{k,d}_{n,c}$. The total CL of a slice window on a path is denoted as $\psi'(l_{\Delta(\mathrm{r,m})})$ and is calculated using (8), where Δ(r, m) denotes the incoming request which has datarate m and arrived on route r, p_r^k is path priority of r^k, $\psi_{n,c}^{r^k,d}$ is the CL on r^k, Z^k represents the collection of all shared paths of r^k; $Z^k = i_1, i_2, \ldots, i_z$, p_z is the path priority of zth shortest path, and $\psi_{n,c}^{i^{z,d}}$ is the CL on zth shortest path [5].

$$\psi_{n,c}^{k,d} = \upsilon_n^{k,d} - \upsilon'^{k,d}_{n,c}. \tag{7}$$

$$\psi'(l_{\Delta(\mathrm{r,m})}) = p_r^k \psi^{r^k,d}_{n,c} + \sum_{z=1}^{|Z^k|} p_z \psi^{i^{z,d}}_{n,c}. \tag{8}$$

Tridental Coefficient The capacity loss is a measurement of how much future capacity is lost due to the assignment of a candidate slice window with a specific core-MF pair to an incoming request. The selection of MF, core, and slice window impacts the quantity of required spectrum and fragmentation levels in addition to the capacity loss. We could penalize the MFs that require more spectrum and implement a first-fit policy for lower-indexed slice windows to reduce fragmentation. In order to account for these effects, the definition of Tridental Coefficient (TC) includes two additional factors in addition to CL.

The TC of $l_{\Delta(\mathrm{r,m})}$ is the sum of the normalized values of CL, the extent of the required slice window in terms of the number of FSs, and the beginning index of the slice window. It is represented as $\Psi(l_{\Delta(\mathrm{r,m})})$ and is provided in (9). The normalization is done using the respective maximum values, viz., maximum possible CL denoted by $\max \psi'(l_{\Delta(\mathrm{r,m})})$ [5], largest possible demandsize of datarate m denoted as β_1^m, and highest possible index of a slice window from (4). $\max \psi'(l_{\Delta(\mathrm{r,m})})$ is calculated using path priorities and the maximum CL, which is equal to C, on the primary route and shared paths.

$$\Psi(l_{\Delta(\mathrm{r,m})}) = \frac{\psi'(l_{\Delta(\mathrm{r,m})})}{\max \psi'(l_{\Delta(\mathrm{r,m})})} + \frac{\beta_d^{\mathrm{m}}}{\beta_1^{\mathrm{m}}} + \frac{n}{S - \beta_d^{\mathrm{m}} + 1}. \tag{9}$$

3.5 Weighted tridental coefficient

In traditional TRA, each parameter is weighted equally. However, each parameter in TC contributes differently in improving the performance of TRA. Thus, we weighted the three parameters in (9) with α, β and $(1-\alpha-\beta)$ as shown in (10) such that $0 \le \alpha, \beta, (\alpha+\beta) \le 1$. All the results shown in [5] are obtained when the weights are 1 in (9), which is similar to $\alpha = \beta = \frac{1}{3}$ in (10).

$$\Psi(l_{\Delta(\mathrm{r,m})}) = \alpha \times \frac{\psi'(l_{\Delta(\mathrm{r,m})})}{\max \psi'(l_{\Delta(\mathrm{r,m})})} + \beta \times \frac{\beta_d^{\mathrm{m}}}{\beta_1^{\mathrm{m}}} + (1-\alpha-\beta) \times \frac{n}{S-\beta_d^{\mathrm{m}}+1} \tag{10}$$

3.6 Optimization space

The value of TC changes with respect to the value of α and β. We can increment α and β starting from 0 while maintaining the constraint that $0 \le \alpha+\beta \le 1$. Assume that the step size is ϵ. The total number of combinations of α and β (and $(1-\alpha-\beta)$) is denoted as ζ and is shown in Eq. 11.

$$\zeta = \frac{1}{2}\left(\left\lfloor \frac{1}{\epsilon} \right\rfloor + 2\right)\left(\left\lfloor \frac{1}{\epsilon} \right\rfloor + 1\right) \tag{11}$$

Suppose the step size is set to 0.1, i.e., $\epsilon = 0.1$. This means that the set of $\{\alpha, \beta\}$ are $\{0, 0\}$, $\{0, 0.1\}, \cdots, \{0, 1\}$, $\{0.1, 0\}$, $\{0.1, 0.1\}, \cdots, \{0.1, 0.9\}$, $\{0.2, 0\}$, $\{0.2, 0.1\}, \cdots, \{0.2, 0.8\}$, and so on. Remember that when $\alpha = 0.2$, β can take a value from 0 to 0.8 to ensure that $\alpha + \beta \le 1$ holds true for all the combinations. Using (11), there will be $\zeta = 66$ $(= \frac{1}{2} \times 12 \times 11)$ such combinations.

4 Simulation Setup

In this section, we present the simulation setup to perform all the experiments in Sects. 5 and 6. We use two topologies, generic German (DT) and European (EURO) shown in Fig. 1. Each core in the MCF has 4 THz of C-band spectrum with 320 FSs ($S = 320$) each of size 12.5 GHz. Connections arrive according to a Poisson process and have a mean exponentially distributed holding time of one. We simulate 110,000 connection requests and utilize the first 10,000 connections to establish steady state in each iteration. Ten iterations are executed to get 95% confidence intervals. The datarates are in the range of 40–400 Gbps with 40 Gbps granularity. We present the results for three-core and seven-core fibers, however, the optimization approach is not limited by the number of cores or number of fibers per link. There exist three shortest paths between each source-destination pair ($K = 3$). According to priority-based path selection in Section 3.4.1, each source-destination pair has three shortest paths which are arranged from higher to lower path priority. We use a total of five MFs ($|D| = 5$), PM-QPSK ($\eta_1 = 2.96$),

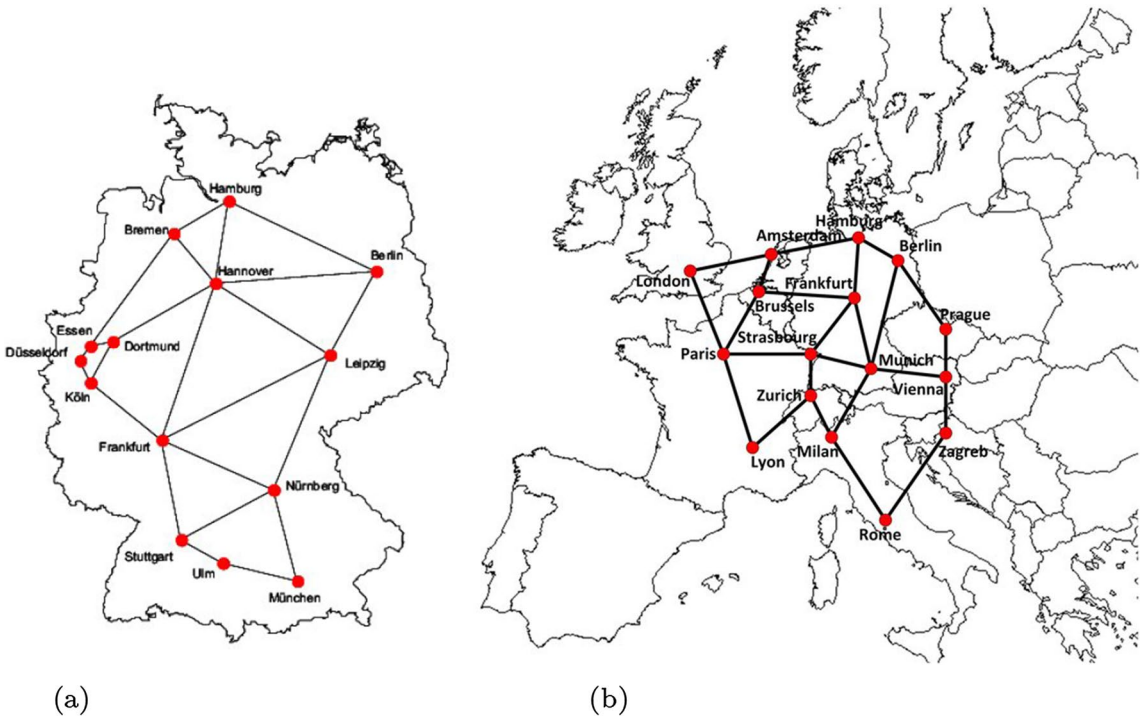


Fig. 1 Network topologies **a** DT (12 Nodes), **b** EURO (16 Nodes) [17]

PM-8QAM ($\eta_2 = 4.44$), PM-16QAM ($\eta_3 = 5.92$), PM-32QAM ($\eta_4 = 7.4$) and PM-64QAM ($\eta_5 = 8.88$) [7, 10]. We consider the average XT between two neighboring cores after a single span of propagation, denoted as XT_μ, of -40 dB and -25 dB. The transmission reaches for each modulation and number of occupied overlapping spectrum on adjacent cores (γ) from Tables 2 and 3 are used. The desired number of RCs for each datarate is uniformly picked from the complete set. It is done only once at the beginning of resource allocation stage. We use the results when $\epsilon = 0.01$ which gives $\zeta = 5151$. We set g_b to 1, i.e., one guardband to separate carriers.

5 Dual optimization for applied TRA

In this section, we outline two main stages of dual-optimization to improve the speed and performance of TRA as follows. The stages, which include the analysis of bandwidth blocking to find the amount of allowed resources and finding optimized weights, are discussed in detail as follows.

5.1 Reducing the computational complexity

TRA assigns the best RC to a connection request on the prioritized shortest path between source and destination. An RC is a combination of core, modulation, and slice window [5]. TRA calculates TC for each RC and chooses the RC with lowest TC. For a given datarate, TRA checks the modulation with highest XT tolerance if the spectrum requirement is same for two or more modulations. The spectrum requirement for each datarate and modulation pair for the network model in Sect. 2, spectrum efficiencies of {2.96 (η_1), 4.44, 5.92, 7.4, 8.88 (η_5)} [7, 10] and using (3) is given in Tab. 4.

Using the β_d^{m} value and D^{m} for each datarate, we can calculate the number of RCs per datarate using (5), and the results when $S = 320$ (C-band) are displayed in Table 5. It is evident that an average of 2172 RCs for $C = 3$ and 5069 RCs for $C = 7$ are always examined to check for the availability and calculate TC when a connection request arrives in the network. Each RC undergoes a QoT estimation procedure based on the XT constraints of the selected modulation and the modulation of ongoing connections on adjacent cores. Since the number of RCs per datarate varies, we use a network-level percentage of RCs. This means that, for any datarate, this percentage of total RCs, resulting into nearest lower integer value, are considered when calculating TC. When the percentage of RCs is decreased, bandwidth blocking increases. However, the increase is incremental before becoming substantially greater. The percentage of RCs per datarate which results into a permissible BBP reduction is subsequently used as the first of two parameters to improve TRA.

Table 4 Spectrum requirement for different datarates for modulations (β_d^m)

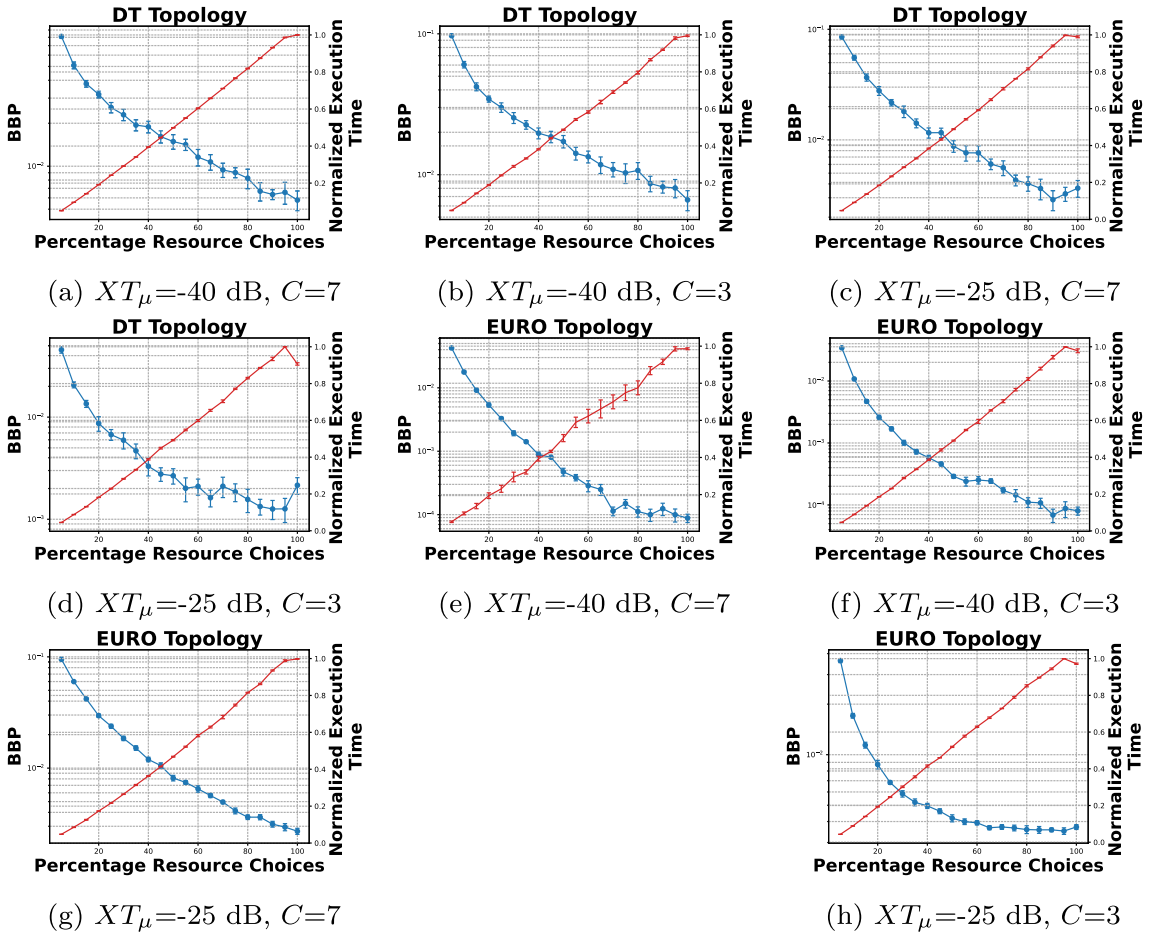
Datarate (Gbps)	Modulations ($\|D\| = 5$)				
	PM-QPSK ($d = 1$)	PM-8QAM ($d = 2$)	PM-16QAM ($d = 3$)	PM-32QAM ($d = 4$)	PM-64QAM ($d = 5$)
40	3	3	3	3	3
80	3	3	3	3	3
120	6	3	3	3	3
160	6	3	3	3	3
200	6	6	3	3	3
240	9	6	6	3	3
280	9	6	6	6	3
320	9	6	6	6	3
360	12	9	6	6	6
400	12	9	6	6	6

The time complexity of TRA algorithm is $O(K|D||B_d^m|\mathfrak{L}SC)$ [5], where K is the number of shortest paths between every pair of nodes, D is a set of all modulations, B_d^{m} is the set of all slice windows, $\mathfrak{L}$ is the maximum possible links per path, S is the number of FSs on a core, and C is the number of cores. The RCs cover the most time complexity, $O(|D||B_d^{\mathrm{m}}|SC)$. Lowering RCs speeds up TRA by reducing computational overhead.

The performance of TRA when percentage RCs is changed in various scenarios is shown in Fig. 2. The simulations are run for 10 seeds with 10% of actual connections for earlier convergence and results are presented with 95% confidence interval. The variation of bandwidth blocking probability (BBP) is on left y-axis whereas the normalized execution time is on right y-axis. It is clear that in all the scenarios the performance of TRA is acceptable above ≈50% of the RCs. It is important to notice that running the optimization experiment for lower traffic is intentional for early convergence. The BBP values are different from when the experiment would have been run for 100% traffic; however, variation in BBP with respect to the RCs is similar in both the cases. With the reduction in RCs, we save significant time and it is approximately equal to the percentage saving of RCs as can be seen in all the scenarios in Fig. 2. This gives the network operator the flexibility to choose the resources to decide the convergence of TRA even before looking at the actual performance with respect to the RCs. The average execution time is larger for 95% as compared to 100% because of the earlier convergence when more resources are available. For example, one of the factors is searching for resources on other shortest paths when the resources are not available on the first one with the highest path priority.

Table 5 Resource candidates ($N^m_{\max}$) for different datarates for various cores

Cores	Datarates (Gbps) ($\|M\| = 10$)									
	40, ($N^1_{\max}$)	80	120	160	200	240	280	320	360	400, ($N^{10}_{\max}$)
3	918	918	1872	1872	1863	2808	2799	2799	2772	2772
7	2142	2142	4368	4368	4347	6552	6531	6531	6468	6468



(a) XT_μ=-40 dB, C=7 (b) XT_μ=-40 dB, C=3 (c) XT_μ=-25 dB, C=7

(d) XT_μ=-25 dB, C=3 (e) XT_μ=-40 dB, C=7 (f) XT_μ=-40 dB, C=3

(g) XT_μ=-25 dB, C=7 (h) XT_μ=-25 dB, C=3

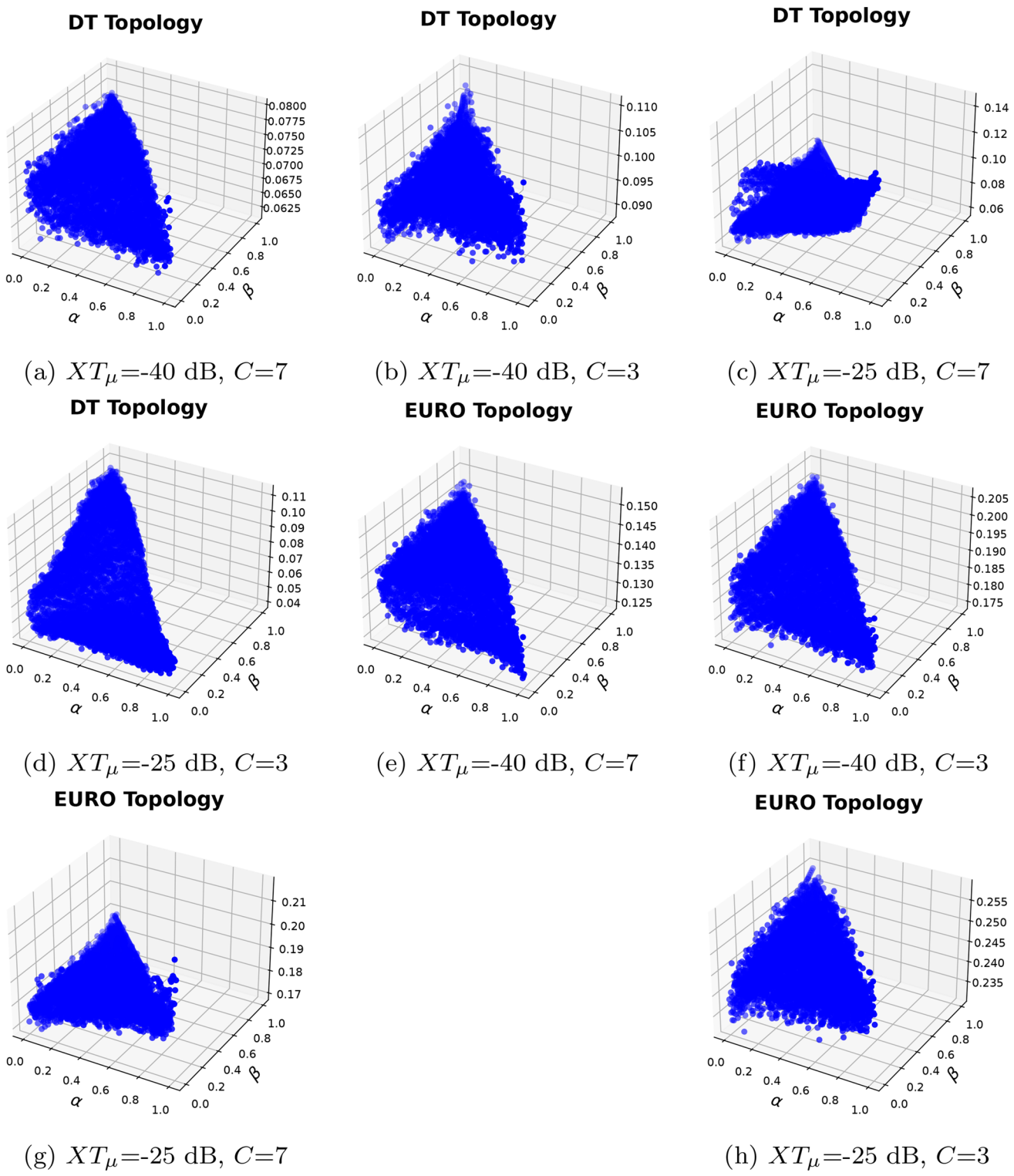
Fig. 2 Variation in BBP and time of execution of TRA for different percentage RCs in various scenarios of XT^μ and C for DT topology and EURO topology

5.2 Weight optimization

The weighted TC is given in (10) and α and β are used to weight the parameters. The optimal weights of α and β lead to better performance of TRA as compared to all the weights are equal. To enhance TRA, we tune α and β in two sub-stages. In the first sub-stage, we execute TRA for 10% of connection requests including the warmup calls for different sets of α and β with a gradual increase of ϵ. In the second sub-stage, we find the α and β with the lowest observed BBP. The variation of TRA for different α and β for various scenarios is plotted in Fig. 3. In most of the cases, the BBP is lowest for higher value of α and lower value of β.

6 Simulation results

We now present the results on performance of TRA and its dual-optimized variants in a variety of scenarios. The details of simulation setup are given in Sect. 4. We observed the performance of TRA for higher load and lower connection requests in Fig. 2 for various scenarios. We observed that lowering the candidate RCs per datarate upto 35% (and 45% in some cases) leads to slighter increase in BBP. However, the execution time reduces significantly and is equivalent to the percentage of candidate RCs for assignment. The time of execution per connection



(a) XT_{μ}=-40 dB, C=7 (b) XT_{μ}=-40 dB, C=3 (c) XT_{μ}=-25 dB, C=7

(d) XT_{μ}=-25 dB, C=3 (e) XT_{μ}=-40 dB, C=7 (f) XT_{μ}=-40 dB, C=3

(g) XT_{μ}=-25 dB, C=7 (h) XT_{μ}=-25 dB, C=3

Fig. 3 Variation in BBP of TRA for different pairs of α and β in various scenarios of XT^{μ} and C for DT topology and EURO topology

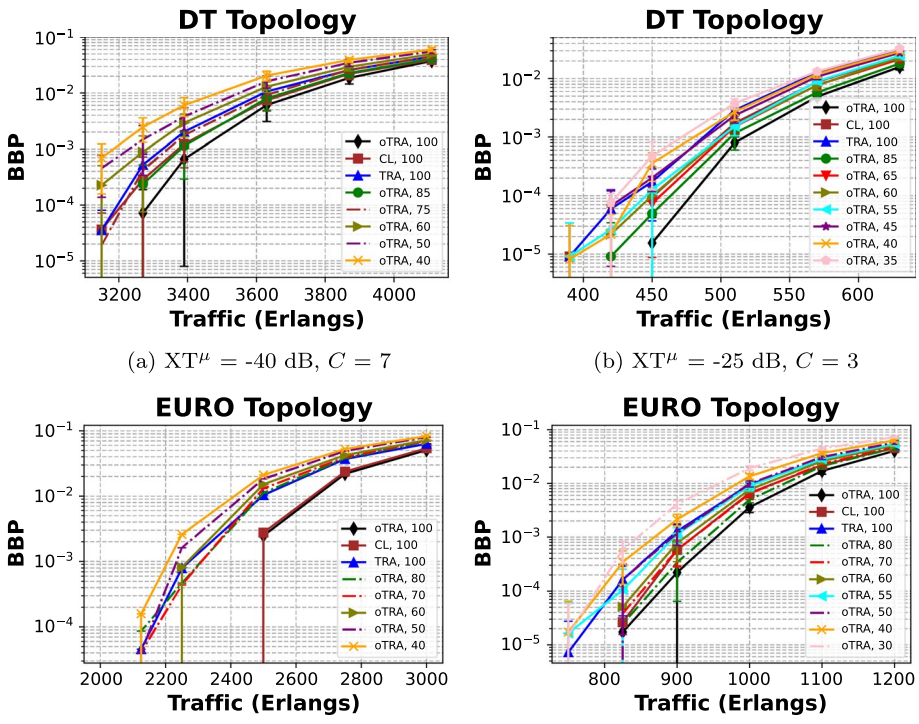
in TRA takes about 5–10 ms depending on the topology. The reduction in execution time is directly proportional to the reduction in RCs. For example, the execution time of oTRA at 60% RCs reduces to 59.84% and 61.76% of that of oTRA and TRA at 100% RCs, respectively, in DT topology. We captured the percentage of RCs for which the performance of TRA is acceptable. In the second stage we ran TRA for reduced load for ζ sets of α and β as shown

in Fig. 3. We fetched the values of α and β for which the BBP is lowest, denoted as B_l. Interestingly, there are pairs of α and β for which the BBP is higher than the BBP when $\alpha = \beta = \frac{1}{3}$ (i.e., traditional TRA), denoted as B. This means that traditional TRA has still a better performance than when α and β are set to values lead to worst case BBP, denoted as B_{m}. The observed B_l, B and B_{m} are shown in Table 6. Table 6 also shows the optimal set of α and β in various scenarios. We thus have reduced RCs which leads to increase in BBP and optimal weights which leads to decrement in BBP. We combine them such that we can get better or equal performance that TRA but with reduced computations. The traditional TRA uses equal weights for all the parameters which is equivalent to $\alpha = \beta = \frac{1}{3}$ and uses 100% of RCs. We denote it as 'TRA, 100'. When optimal values of α and β are used, the optimal TRA is obtained and we denote it as oTRA. TRA is a better version on CL algorithm which is nothing but TRA with $\alpha = 1$ and $\beta = 0$.

We finally compare the performance of TRA, CL and oTRA with various values of percentage candidate RCs for various scenarios in Fig. 4. 'oTRA, 100' works better than all the variants of TRA. The performance of oTRA degrades with reduction in percentage of RCs. In addition, CL works better than TRA in these scenarios where in CL α is set to 1 which doesn't require any optimization. We also observe that 75% of RCs in $\mathrm{XT}^{\mu} = -40$ dB and $C = 7$ in DT topology (Fig. 4a), 40% of RCs in $\mathrm{XT}^{\mu} = -25$ dB and $C = 3$ in DT topology (Fig. 4b), 60% of RCs in $\mathrm{XT}^{\mu} = -40$ dB and $C = 7$ in EURO topology (Fig. 4c), 50% of RCs in $\mathrm{XT}^{\mu} = -40$ dB and $C = 3$ in EURO topology (Fig. 4d), leads to an equivalent performance as that of TRA. Remember that percentage reduction in RCs is equivalent to percentage reduction in time as observed in Fig. 2. Thus with this dual optimization we can get equal or better results of TRA with as low as 40% of computations. We also observed that at higher loads, the difference in performance among all the variants of TRA is smaller. This happens due to extreme congestion and the limited number of candidates to select the best candidate resource from. Remember that the slice window is selected only when it is free, satisfies self-XT constraint and satisfies

Table 6 Sets of (α, β) corresponding to lowest BBP at $C = 3$ and $\mathrm{XT}_{\mu} = -40$ dB

Topology	XT^{μ} dB	C	Load (Erlang)	(α, β)	B_l	B	B_{m}
DT	− 40	7	6250	(0.85, 0.12)	0.0617305	0.076335	0.0796275
DT	− 25	3	750	(0.79, 0.01)	0.0410257	0.0600539	0.110998
EURO	− 40	7	5000	(0.98, 0.02)	0.124976	0.141398	0.152205
EURO	− 40	3	2000	(0.88, 0.06)	0.17495	0.192362	0.20527



(a) XT^{μ} = -40 dB, $C = 7$

(b) XT^{μ} = -25 dB, $C = 3$

(c) XT^{μ} = -40 dB, $C = 7$

(d) XT^{μ} = -40 dB, $C = 3$

Fig. 4 Variation in BBP of different versions of TRA for different loads in various scenarios of XT^{μ} and C for DT topology and EURO topology

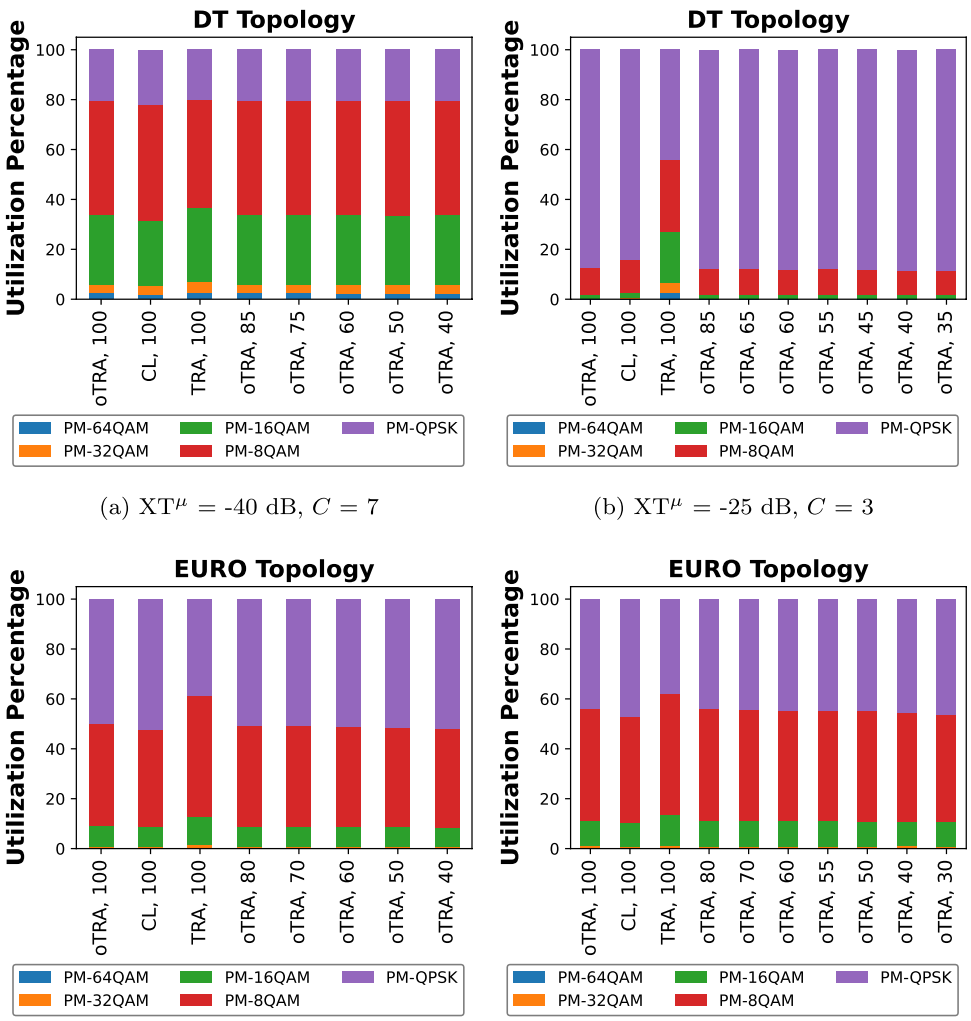
XT constraint imposed by ongoing connections on adjacent cores [5]. These constraints limit the number of RCs per connection. The limited number of choices lead to similar pattern of selection of resources by all the versions of TRA for higher load.

Finally we present the utilization of modulations for all the variants of TRA in Fig. 5. Using the utilization of modulations, we can understand the change in trade-off between spectrum utilization and XT accumulations which in turn affects the BBP. It is observed that the variants of oTRA, CL and TRA have different utilization of modulations. In addition, the variants of oTRA have almost similar utilization with more use of lower modulations (PM-QPSK in this case) and decrease in use of higher modulation (PM-8QAM in this case). Furthermore, we observe that for a particular load value, different versions of TRA have different performance. This happens due to a better pattern of selection of RCs, in other words modulations, by optimizing weights in the TC. The variation in selection of modulations as a part of RC, as shown in Fig. 5, leads to balancing the trade-off in a unique way. Please note that the graphs in Fig. 5 show the percentage of utilization of various modulations and not the actual number.

In future, the work can be expanded in two directions. First, machine learning can be used to replicate the strategy of selection of RCs by TRA and find RCs and optimal weights together without the need to re-run the network for each combination of weights and RCs. Second, a ML-aided standard template to optimize *any* RMCSA algorithm, similar to our previous work [15], can be used to obtain *guaranteed* improvements.

7 Conclusion

The tridental resource assignment (TRA) algorithm assigns route, modulation, core, and spectrum optimally in multi-core fiber-based space division multiplexed elastic optical



(a) XT^{μ} = -40 dB, $C = 7$

(b) XT^{μ} = -25 dB, $C = 3$

(c) XT^{μ} = -40 dB, $C = 7$

(d) XT^{μ} = -40 dB, $C = 3$

Fig. 5 Utilization of modulations of different versions of TRA for different loads in various scenarios of XT^{μ} and C for DT topology and EURO topology

networks. Using tridental coefficient (TC), TRA effectively balances spectrum utilization and intercore crosstalk (XT) accumulations at the expense of computational overhead. This study presents a dual-optimization strategy for optimizing TRA to obtain comparable or superior performance. We obtain a reduced number of resource options and optimized weights in TC, which expedites convergence with the desired blocking of bandwidth. The performance of the TRA, which was already superior to literature baseline algorithms, can be enhanced by considering only capacity loss. Every version of TRA, pre- and post-optimization, maintains the balance in trade-off between spectrum utilization and XT accumulations. With the help of these offline optimizations, acceptable performance of TRA is achieved with as low as 40% of the total amount of computation. Furthermore, the optimizations also leads to an improvement of approximately two orders of magnitude in the bandwidth blocking performance of TRA as compared to the baseline algorithms.

Acknowledgements This work was supported in part by NSF Grants CNS-1813617 and CNS-2210343.

Data availability The datasets generated during the current study are available from the corresponding author (srpetale@gmail.com) on reasonable request.

Declarations

Conflict of interest The authors declare no conflicts of interest.

References

1. Shariati, B., Klonidis, D., Tomkos, I., Marom, D., Blau, M., Ben-Ezra, S., Gerola, M., Siracusa, D., Macdonald, J., Psaila, N., et al.: Realizing spectrally-spatially flexible optical networks, pp. 4–9. IEEE Photon, Society Newsletter (2017)
2. Klinkowski, Miroslaw, Lechowicz, Piotr, Walkowiak, Krzysztof: Survey of resource allocation schemes and algorithms in spectrally-spatially flexible optical networking. Optical Switching and Networking **27**, 58–78 (2018)
3. Arpanaei, Farhad, Ardalani, Nahid, Beyranvand, Hamzeh, Shariati, Behnam: QoT-aware performance evaluation of spectrally-spatially flexible optical networks over FM-MCFs. Journal of Optical Communications and Networking **12**(8), 288–300 (2020)
4. Petale, Shrinivas, Zhao, Juzi, Subramaniam, Suresh: Tridental Resource Assignment Algorithm for Spectrally-Spatially Flexible Optical Networks. In ICC 2021-2021 IEEE International Conference on Communications (ICC), pages 1-7. IEEE, (2021)
5. Petale, Shrinivas, Zhao, Juzi, Subramaniam, Suresh: TRA: an efficient dynamic resource assignment algorithm for MCF-based SS-FONs. Journal of Optical Communications and Networking **14**(7), 511–523 (2022)
6. Petale, Shrinivas, Subramaniam, Suresh: Enhanced TRA Algorithm for Space Division Multiplexed Elastic Optical Networks. In IEEE International Conference on Advanced Networks and Telecommunications Systems, pages 1-6. IEEE, (2022)
7. Rottondi, Cristina, Martelli, Paolo, Boffi, Pierpaolo, Barletta, Luca, Tornatore, Massimo: Crosstalk-aware core and spectrum assignment in a multicore optical link with flexible grid. IEEE Transactions on Communications **67**(3), 2144–2156 (2018)
8. Walkowiak, Krzysztof, Klinkowski, Miroslaw, Lechowicz, Piotr: Dynamic routing in spectrally spatially flexible optical networks with back-to-back regeneration. Journal of Optical Communications and Networking **10**(5), 523–534 (2018)
9. Klinkowski, Miroslaw, Zalewski, Grzegorz: Dynamic crosstalk-aware lightpath provisioning in spectrally-spatially flexible optical networks. IEEE/OSA Journal of Optical Communications and Networking **11**(5), 213–225 (2019)
10. Chatterjee, Bijoy Chand, Wadud, Abdul, Ahmed, Imran, Oki, Eiji: Priority-Based Inter-Core and Inter-Mode Crosstalk-Avoided Resource Allocation for Spectrally-Spatially Elastic Optical Networks. IEEE/ACM Transactions on Networking, (2021)
11. Kmiecik, Wojciech, Walkowiak, Krzysztof: A performance study of dynamic routing algorithm for SDM translucent optical networks with assistive storage. Optical Switching and Networking **38**, 100572 (2020)
12. Oliveira, Helder MNS, da Fonseca, Nelson LS: Multipath routing, spectrum and core allocation in protected SDM elastic optical networks. In 2019 IEEE Global Communications Conference (GLOBECOM), pages 1-6. IEEE, (2019)
13. Agrawal, Anuj, Bhatia, Vimal, Prakash, Shashi: Core Arrangement Based Spectrum-Efficient Path Selection in Core-Continuity Constrained SS-FONs. In International IFIP Conference on Optical Network Design and Modeling, pages 570-583. Springer, (2019)
14. Brasileiro, Italo, Costa, Lucas, Drummond, Andr e: A survey on challenges of Spatial Division Multiplexing enabled elastic optical networks. Optical Switching and Networking, page 100584, (2020)
15. Petale, Shrinivas, Subramaniam, Suresh: An ML Approach for Crosstalk-Aware Modulation Format Selection in SDM-EONs. In 2022 International Conference on Optical Network Design and Modeling (ONDM), pages 1-6. IEEE, (2022)
16. Jingxin, Wu., Subramaniam, Suresh, Hasegawa, Hiroshi: Efficient dynamic routing and spectrum assignment for multifiber elastic optical networks. IEEE/OSA Journal of Optical Communications and Networking **11**(5), 190–201 (2019)
17. Goscien, Roza, Kucharzak, Michal: On the efficient optimization of unicast, anycast and multicast flows in survivable elastic optical networks. Optical Switching and Networking **31**, 114–126 (2019)