# Enhancing 1+1 Fault Tolerance Protection for Elastic Optical Networks: A Deep Reinforcement Learning Approach

Patricia Morales, Jorge Bermúdez Cedeño, Jesús Chaffe, Hermann Pempelfort, Nicolás Jara
*Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile*
Corresponding author: patricia.morales@usm.cl

***Abstract*—Dedicated 1+1 protection remains the fastest fault-tolerance mechanism for elastic optical networks, but its deterministic nature often leads to spectrum over-allocation and fragmentation under dynamic traffic and failure conditions. This work benchmarks two reproducible 1+1 heuristics against a deep reinforcement learning approach that preserves the primary route selection rule while learning to intelligently choose the protection path. The proposed controller observes spectral occupancy, fragmentation, and demand descriptors to minimize long-term blocking while maintaining strict link-disjoint 1+1 protection. Extensive simulations across five representative network topologies under multiple failure scenarios show that DRL-1+1-KD consistently reduces the blocking probability by up to two orders of magnitude compared to Classic 1+1 and achieves additional gains over enhanced deterministic heuristics, particularly under severe failures. These results indicate that learning-assisted backup selection improves spectrum efficiency and resilience while preserving operational simplicity.**

***Index Terms*—Fault Tolerance, Elastic Optical Networks, Protection, Deep Reinforcement Learning**

## I. Introduction

Elastic Optical Networks (EONs) enable fine-grained spectrum allocation via flexible frequency slots, enabling efficient support of heterogeneous bandwidth demands [1]–[3]. However, this flexibility also increases the impact of failures, as a single disrupted link may affect a large volume of high-capacity traffic. Consequently, survivability and spectrum efficiency have become central design objectives in modern optical backbone networks [4]–[7].

Dedicated 1+1 protection remains the fastest recovery mechanism, since traffic is transmitted simultaneously over a working and a backup path [3], [7]. Despite its robustness, deterministic implementations typically duplicate spectrum resources independently of instantaneous network conditions, leading to inefficient utilization and aggravated spectrum fragmentation under dynamic traffic loads or persistent failures. The simplest deployment, referred to as Classic 1+1, precomputes two shortest link-disjoint paths per source–destination pair and reuses them indiscriminately, ignoring current failures and spectrum availability [3].

To partially mitigate these limitations, enhanced deterministic heuristics have been proposed that evaluate multiple link-disjoint candidates and basic feasibility constraints at request arrival. While such schemes improve adaptability, their decisions remain rule-based and short-sighted, often favoring the same backup routes and progressively concentrating fragmentation on critical links, which limits long-term blocking reduction [6], [8].

In this work, we address this limitation by learning how to select the protection route while preserving the deterministic behavior of the primary path. We introduce DRL-1+1-KD (Deep Reinforcement Learning-based 1+1 protection with k-disjoint routes), a framework in which the working route is fixed using a shortest-path heuristic, and a policy learned via Proximal Policy Optimization selects the backup route based on the instantaneous spectrum state and demand characteristics. This design isolates the impact of learning to the protection decision, enabling transparent comparison with deterministic baselines and facilitating practical deployment.

The contributions of this paper are threefold: (i) the definition of two reproducible deterministic benchmarks for dedicated 1+1 protection in EONs; (ii) the design of a DRL-based controller that selects backup routes while enforcing strict link disjointness; and (iii) an extensive evaluation on five benchmark topologies under multiple failure scenarios, showing consistent reductions in blocking probability.

The remainder of the paper is organized as follows: Section II reviews related research, Section III details the proposed framework, Section IV presents numerical evaluations, and Section V summarizes conclusions and future work.

## II. Related Work

Survivability in optical backbone networks has traditionally relied on static protection schemes such as dedicated 1+1, 1:1 switching, and shared backup path protection. Numerous heuristic S-RMLSA (Survivable Routing, Modulation Level and Spectrum Assignment) formulations have been proposed to reduce the spectral overhead of these schemes by optimizing sets of link-disjoint paths, selecting modulation formats, or solving mixed-integer optimization problems under physical-layer constraints. While such approaches offer deterministic behavior and interpretability, their decisions are inherently myopic and cannot explicitly account for long-term spectrum fragmentation or evolving failure patterns without repeated re-optimization [3], [7].

Projects ANID FONDECYT 11251432 and 1250775, and USM PI_LII_24_15 are greatly acknowledged.

978-3-903176-78-2 © 2026 IFIP

More recently, machine-learning-based controllers have been explored for routing, modulation selection, and spectrum assignment in elastic optical networks. Supervised and unsupervised methods require labeled datasets and retraining when network conditions change, which limits their applicability to operational protection mechanisms [9]. Deep reinforcement learning, including DQN, PPO, and actor–critic variants, enables online policy learning from interaction and has shown promising results for dynamic routing and spectrum assignment in unprotected EON.

Only a limited number of studies have applied DRL to dedicated 1+1 protection. A double-agent DRL-based approach has been proposed in [10], where separate agents handle working and protection decisions for EON optimization. Existing approaches typically allow the agent to jointly select both primary and backup routes or relax strict link-disjointness constraints, reducing transparency and complicating operational validation [11], [12]. In contrast, the DRL-1+1-KD framework proposed in this work isolates the backup-route decision under strict link-disjoint protection, while preserving a deterministic primary-path heuristic. This design enables controlled comparison with established heuristics and facilitates integration into existing planning and control workflows.

## III. DRL-based Protection Proposal

To mitigate the combinatorial nature of the 1+1 S-RMLSA problem under dynamic arrivals and failures, we cast backup-route selection as a sequential decision process. The environment models the optical network, spectrum occupancy, traffic dynamics, and failure events so that both heuristic and learning-based controllers can be evaluated consistently.

### A. Deterministic Baselines

Two reproducible baselines are considered. Classic 1+1 precomputes the two shortest link-disjoint paths per source–destination pair and always reuses them as working and backup routes. H-1+1-KD extends this approach by evaluating a set of K = 5 precomputed link-disjoint candidates and, at request arrival, selecting the two shortest feasible paths that satisfy spectrum continuity and reach constraints. Both schemes employ identical modulation formats, slot widths, and duplication rules, providing fair and reproducible reference points.

### B. Environment Design

All controllers are evaluated in the same Gymnasium environment, instantiated across five benchmark topologies (NSFNet, Eurocore, UKNet, EONet, and ARPANET), each with 344 frequency-slot units per link. Traffic arrivals follow a Poisson process, and failures are injected on high-betweenness links to emulate severe disruptions and on moderate-centrality links to represent maintenance scenarios. The environment continuously tracks spectrum occupancy and physical-layer feasibility to validate working and protection paths.

TABLE I
Observation vector structure for DRL-1+1-KD (single-column width)

| Segment | Description |
|---|---|
| Source one-hot ($N$ bits) | Encodes the source node of the current demand. |
| Destination one-hot ($N$ bits) | Encodes the destination node. |
| Per-path spectrum ($344K$ bits) | For each of the $K$ link-disjoint candidates, binary availability of every FSU along the path. |
| Per-path demand class ($K \cdot \|\mathcal{D}\|$ bits) | For each candidate path, indicators for the admissible bandwidth requests in $\mathcal{D} = \{1,2,3,4,5,6,8,9,11,15,18,22,44\}$. |

### C. Observation and Action Space

The DRL agent observes the network state through a structured MultiBinary vector of length $2N + 344K + K \cdot |\mathcal{D}|$, encoding the source and destination nodes ($2N$), the spectrum availability along the K candidate protection paths ($344K$), and the requested bandwidth class ($K \cdot |\mathcal{D}|$). Table I summarizes the composition of this observation vector.

The action space is restricted to the protection decision: the primary route is always selected as the shortest feasible candidate, while the agent chooses one of the remaining $K - 1$ link-disjoint paths as backup. Infeasible actions violating spectrum continuity or reach constraints are penalized.

### D. Reward Function

At each decision step, the agent receives a positive reward ($+1$) when both working and protection paths are successfully established, a neutral reward (0) when only the working path is feasible, and a negative reward ($-1$) when the request is blocked. Although spectrum fragmentation is not explicitly included in the reward, its impact is implicitly captured through future blocking events, which dominate the long-term return. This choice avoids competing objectives while encouraging protection decisions that preserve spectrum availability over time.

### E. Operational Flow

For each incoming request, the primary route is fixed using the deterministic heuristic, and the DRL policy selects the protection path based on the current spectrum state and demand characteristics. This design isolates learning to the backup selection process, simplifying interpretability and deployment in operational controllers.

## IV. Numerical Examples

The numerical study contrasts the Classic 1+1 heuristic, the enhanced H-1+1-KD heuristic, and the DRL-1+1-KD agent under identical traffic, physical-layer, and failure assumptions. Simulations are executed on five benchmark topologies—NSFNet, Eurocore, UKNet, EONet, and ARPANET—each sharing the precomputed set of five link-disjoint paths per source–destination pair and experiencing Poisson arrivals with different loads. Failures are injected on the two highest betweenness links of each topology, and on a

TABLE II
PHYSICAL LAYER MODEL BASED ON MODULATION FORMAT AND MAXIMUM ACHIEVABLE REACH.

| Mod. | MAR [km] | 10 | 40 | 100 | 400 | 1000 |
|---|---|---|---|---|---|---|
| BPSK | 19900 | 1 | 2 | 5 | 18 | 44 |
| QPSK | 9900 | 1 | 1 | 3 | 9 | 22 |
| 8-QAM | 5400 | 1 | 1 | 2 | 6 | 15 |
| 16-QAM | 2700 | 1 | 1 | 2 | 5 | 11 |
| 32-QAM | 1300 | 1 | 1 | 1 | 4 | 9 |
| 64-QAM | 700 | 1 | 1 | 1 | 3 | 8 |

TABLE III
SIMULATION SETTINGS.

| Network Settings | |
|---|---|
| Topologies | NSFNet, Eurocore, UKNet, EONet, ARPANET |
| FSUs per Link | 344 (full C band) |
| Modulation Formats | BPSK, QPSK, 8-QAM, 16-QAM, 32-QAM, and 64-QAM |
| **Connection Settings** | |
| Traffic Model | Poisson process |
| Requests per episode | 1000 |
| Routes per pair | 5 precomputed using K-shortest paths |
| Assignment Mode | One-shot 1+1 protection |
| Failure Injection | Highest-betweenness links per topology |
| **DRL Agent Settings** | |
| Algorithm | PPO |
| Training Steps | $8 \times 10^6$ |
| Observation Space | MultiBinary of size $2N + K \times 344 + K \times \lvert\mathcal{D}\rvert$ |
| Neural Network | Multi Layer Perceptrons |
| Learning Rate | $3 \times 10^{-5}$ |
| Discount Factor | 0.99 |

moderate betweenness link to emulate maintenance windows or fiber cuts, while non-failure episodes capture nominal operation. In our experiments, the blocking probability serves as the main evaluation metric.

The DRL agent uses PPO with a shared policy-value network based on a Multi-Layer Perceptron. The agent training is performed in 8000 episodes of 1000 time steps each. The agent is trained independently for each topology, and policy transfer across networks is left for future work. Table II details the modulation reach constraints that bound the feasible spectrum allocations, while Table III consolidates the network, traffic, and DRL hyperparameters used in all experiments. To keep comparisons fair, both deterministic heuristics use the same modulation reach table, spectrum width assumptions, and failure set-up.

Detailed trends in Fig. 1 illustrate how the DRL policy reacts to both congestion and failures on NSFNet. Under nominal load the Classic 1+1 baseline reaches $1.9 \times 10^{-2}$ blocking at 13k Erlangs, whereas H-1+1-KD halves that value and DRL-1+1-KD further lowers it to $5.0 \times 10^{-3}$. Once link 7–8 is removed, Classic 1+1 saturates near $1.2 \times 10^{-1}$ blocking because it stubbornly reuses the same two paths, while H-1+1-KD and the agent exploit alternative candidates, keeping the probability below $8.7 \times 10^{-3}$ and $4.8 \times 10^{-3}$, respectively. Similar gaps appear for failures (4–6) and (8–11), confirming that the learned backup choices consistently preserve spectrum even in the most stressed scenarios.

Eurocore and UKNet amplify this advantage because their high centrality links create pronounced bottlenecks when they fail (Figs. 2 and 3). For Eurocore at 42k Erlangs, the DRL agent keeps blocking below $4.8 \times 10^{-3}$ without failures and below $3.6 \times 10^{-3}$ when link 7–9 fails, in contrast with the $3.45 \times 10^{-2}$ and $1.13 \times 10^{-1}$ exhibited by the Classic heuristic. Even the more adaptive H-1+1-KD remains at least $1.5\times$ worse than DRL across all Eurocore scenarios because it still tends to reuse the second-shortest path. On UKNet, failures on links 3–18 or 7–9 push Classic 1+1 above $1.1 \times 10^{-1}$ in blocking, while the agent keeps it in the low $10^{-2}$ range by distributing protection requests among the remaining candidates.

The meshed EON and ARPANET topologies confirm that these gains extend to larger networks (Figs. 4 and 5). On EON, the DRL policy stays below $1.1 \times 10^{-2}$ even at the highest load and under severe failures such as (3–8), whereas the aggregated heuristics fluctuate between $3 \times 10^{-2}$ and $1.0 \times 10^{-1}$. ARPANET shows the steepest contrast: removing link 6–8 drives Classic 1+1 beyond $2.3 \times 10^{-1}$ blocking, H-1+1-KD confines it to $3.1 \times 10^{-2}$, and DRL-1+1-KD halves it again to $2.9 \times 10^{-2}$ while also reducing the non-failure blocking by more than $60\%$.

Across all evaluated topologies and failure scenarios (Figs. 1–5), Classic 1+1 exhibits the highest blocking probability, H-1+1-KD achieves intermediate performance, and DRL-1+1-KD consistently attains the lowest blocking levels, often by approximately one order of magnitude.

Table IV presents the same data expressed as the percentage increase of each heuristic relative to the DRL-1+1-KD agent, where the percentage is calculated as:

$$\text{Percentage Points} = \frac{\text{Heuristic} - \text{DRL-1+1-KD}}{\text{DRL-1+1-KD}} \times 100 \quad (1)$$

Heuristic can be either Classic 1+1 or H-1+1-KD. The results underscore that Classic 1+1 can exceed the DRL-1+1-KD blocking rate by more than $10^5\%$ under severe outages, whereas H-1+1-KD typically remains within $80\%$–$140\%$. This consolidated view also highlights the steadier slope of DRL-1+1-KD: its blocking curves rise slowly with load, indicating that the policy is not merely reacting to failures but also proactively allocating spectrum to avoid future contention.

Across all examined loads, the DRL-1+1-KD policy reduces the blocking probability relative to both Classic 1+1 and H-1+1-KD, particularly once failures constrain the candidate set. By diversifying backup routes and more evenly distributing spectrum use, the agent increases resource availability for future requests. Classic 1+1, while fast and familiar, cannot react to link outages because it always reuses the same pair of paths, whereas H-1+1-KD filters candidates but still tends to saturate the second-shortest path and aggravate fragmentation on heavily used links. The DRL policy maintains the same

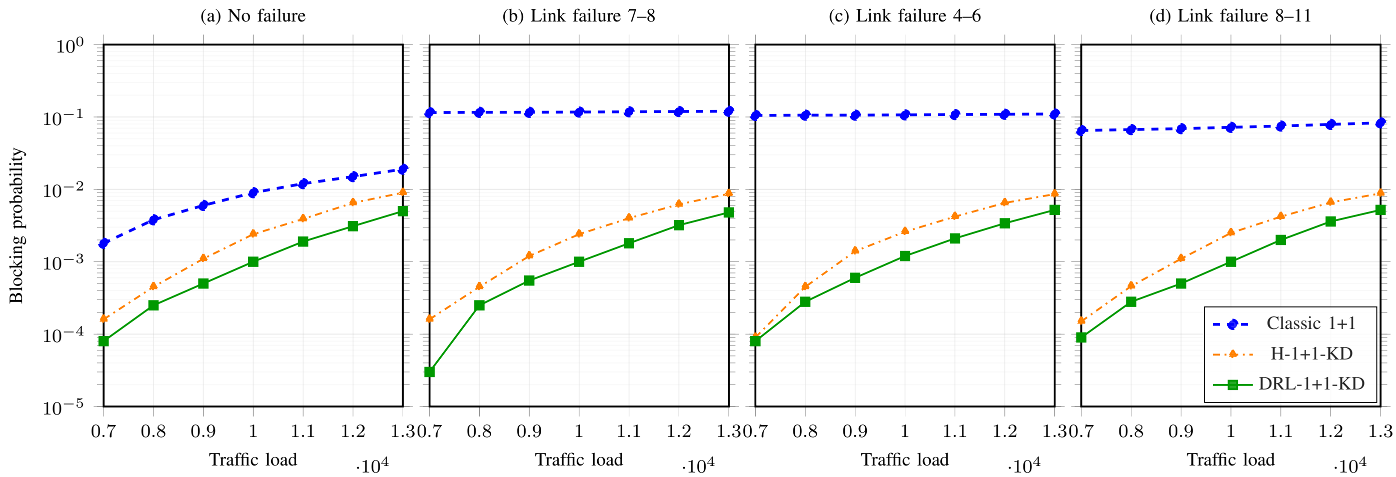


Fig. 1. Blocking probability comparison in the NSFNET topology under no-failure and single-link failure scenarios.

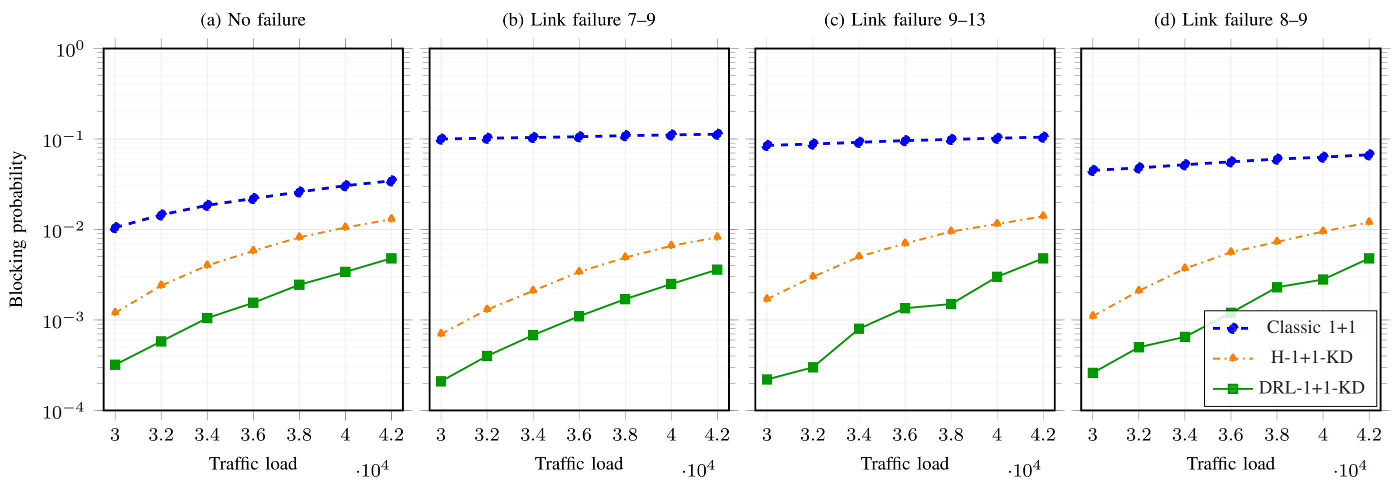


Fig. 2. Blocking probability comparison in the Eurocore topology under no-failure and single-link failure scenarios.

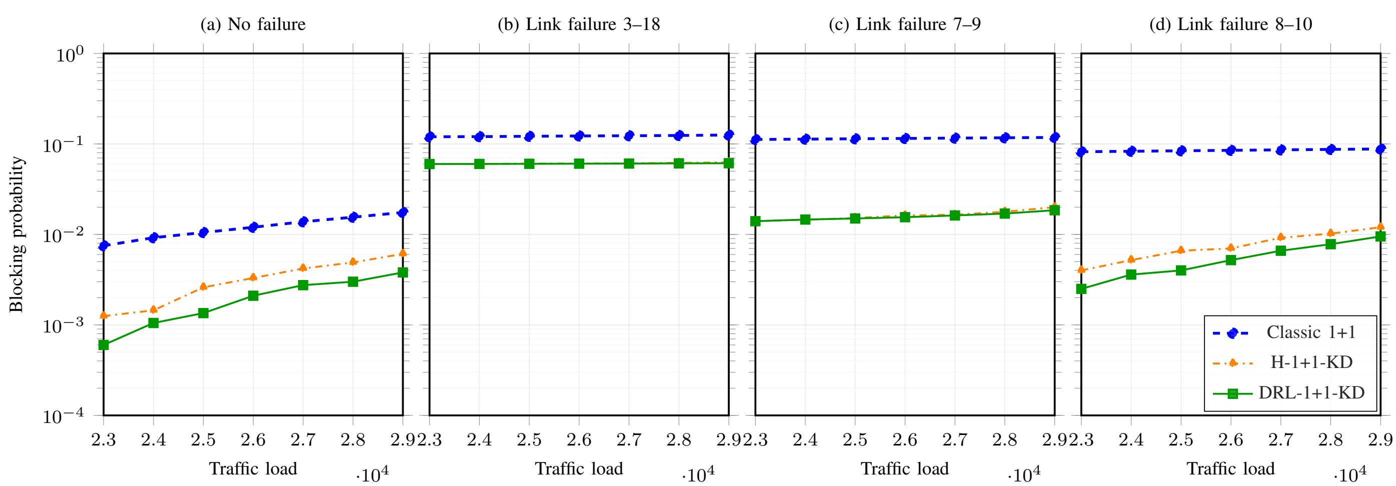


Fig. 3. Blocking probability comparison in the UKNET topology under no-failure and single-link failure scenarios.

instantaneous 1+1 protection behavior once allocations are accepted, ensuring comparable service continuity after failures.

For example, on NSFNet ($N = 14$, $K = 5$, $|\mathcal{D}| = 13$) the observation length becomes $2N + 344K + K|\mathcal{D}| = 1\,813$

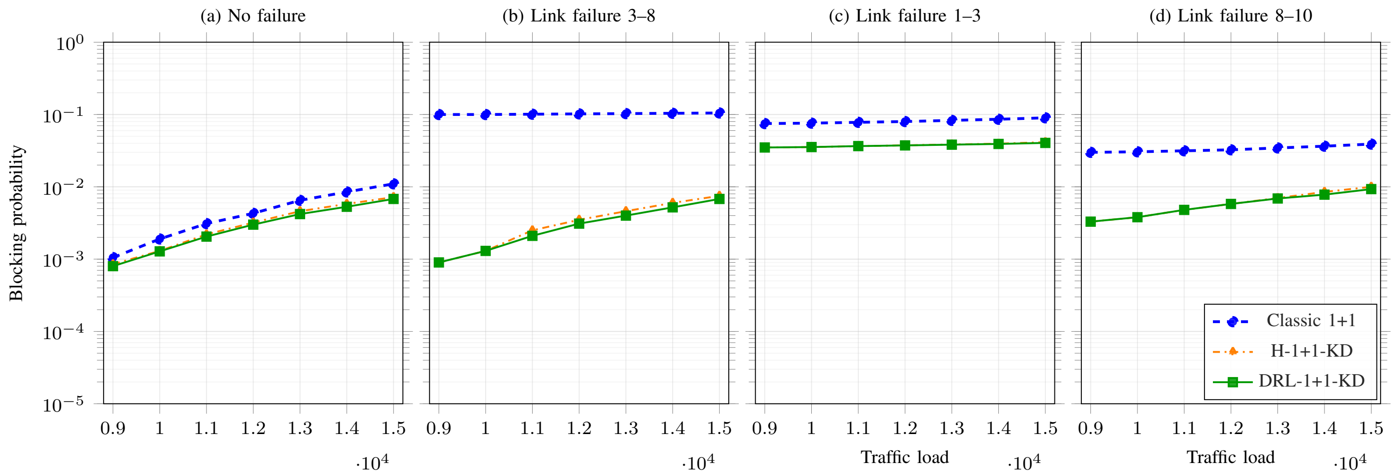


Fig. 4. Blocking probability comparison in the EON topology under no-failure and single-link failure scenarios.

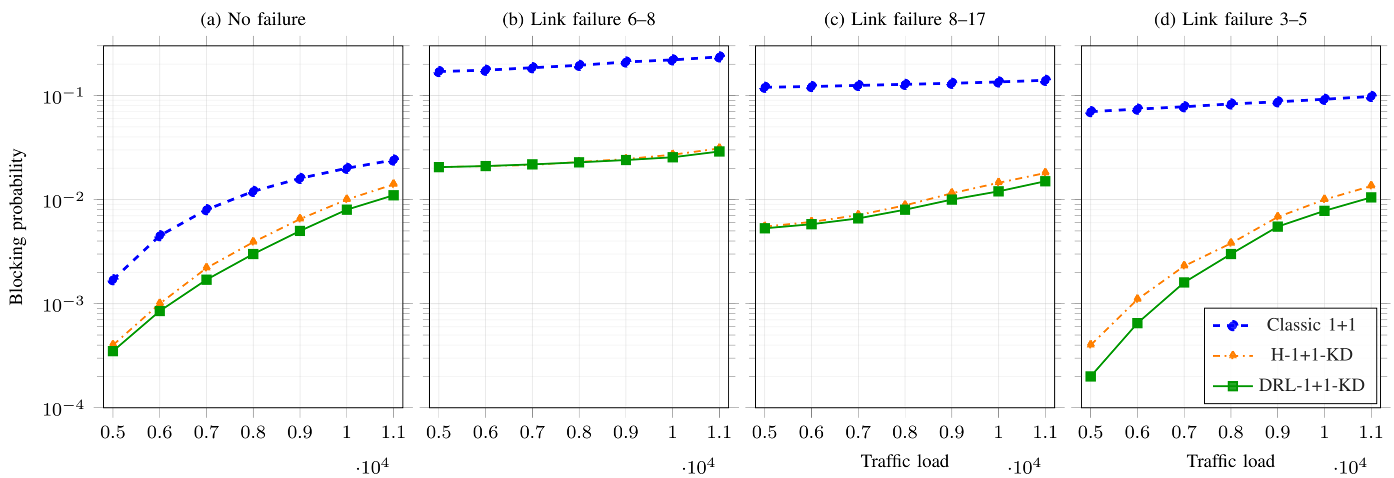


Fig. 5. Blocking probability comparison in ARPANET under no-failure and single-link failure scenarios.

MultiBinary inputs per decision step, while the action set contains four protection-route choices (one for each remaining link-disjoint candidate). The other topologies yield analogous observation sizes according to their number of nodes.

Simulation results on NSFNet, Eurocore, UKNet, EONet, and ARPANET, under injected failure scenarios, demonstrate that DRL-1+1-KD consistently outperforms Classic 1+1, achieving blocking probability reductions of up to 99.9% in NSFNet, 99.8% in Eurocore, 51% in UKNet, 88% in ARPANET, and 99% in EONet on the most heavily loaded links. When compared to H-1+1-KD, the proposed approach delivers additional gains of up to 81% in NSFNet, 70% in Eurocore, marginal improvements (<2%) in UKNet, up to 6.5% in ARPANET, and up to 16% in EONet, while maintaining a more balanced spectrum utilization. These results show that the DRL agent adapts effectively to diverse topological constraints, achieving consistent gains across heterogeneous networks and performance comparable to deterministic heuristics in highly constrained scenarios.

Across all experiments, three consistent trends are observed: (i) Classic 1+1 rapidly saturates under failures due to repeated reuse of fixed backup paths; (ii) H-1+1-KD partially alleviates this effect but tends to concentrate protection traffic on a small subset of candidate routes; and (iii) DRL-1+1-KD systematically distributes backup allocations across available candidates, reducing spectrum fragmentation and preserving future feasibility.

## V. Conclusions

This paper examined dedicated 1+1 protection in elastic optical networks by comparing two deterministic heuristics—Classic 1+1 and H-1+1-KD—with the proposed DRL-1+1-KD controller. By keeping the primary route selection rule fixed, the study isolated the impact of learning on the protection decision, enabling a transparent and reproducible comparison.

Results across five benchmark topologies and multiple failure scenarios show that learning-assisted backup selection consistently reduces blocking probability with respect to deterministic schemes, particularly in networks offering

TABLE IV
BLOCKING PROBABILITY INCREASE OF HEURISTIC SCHEMES RELATIVE TO THE DRL-1+1-KD AGENT (PERCENTAGE POINTS).

| **Topology** | **Scenario** | **Classic 1+1** | | | | | | | **H-1+1-KD** | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $L_1$ | $L_2$ | $L_3$ | $L_4$ | $L_5$ | $L_6$ | $L_7$ | $L_1$ | $L_2$ | $L_3$ | $L_4$ | $L_5$ | $L_6$ | $L_7$ |
| **NSFNet** (7k–13k) | | | | | | | | | | | | | | | |
| | No failure | 2150.0 | 1420.0 | 1100.0 | 800.0 | 531.6 | 383.9 | 280.0 | 100.0 | 80.0 | 120.0 | 140.0 | 105.3 | 109.7 | 80.0 |
| | Link 7–8 | 383233.3 | 46300.0 | 20990.9 | 11600.0 | 6455.6 | 3618.8 | 2400.0 | 433.3 | 80.0 | 118.2 | 140.0 | 122.2 | 93.7 | 81.2 |
| | Link 4–6 | 131150.0 | 37757.1 | 17566.7 | 8816.7 | 5042.9 | 3105.9 | 2015.4 | 12.5 | 60.7 | 133.3 | 116.7 | 100.0 | 91.2 | 65.4 |
| | Link 8–11 | 72122.2 | 23828.6 | 13700.0 | 7100.0 | 3650.0 | 2094.4 | 1496.2 | 66.7 | 64.3 | 120.0 | 150.0 | 110.0 | 83.3 | 69.2 |
| **Eurocore** (30k–42k) | | | | | | | | | | | | | | | |
| | No failure | 3181.2 | 2400.0 | 1661.9 | 1319.4 | 961.2 | 797.1 | 618.8 | 275.0 | 313.8 | 281.0 | 274.2 | 234.7 | 208.8 | 170.8 |
| | Link 7–9 | 47519.0 | 25400.0 | 15194.1 | 9536.4 | 6311.8 | 4340.0 | 3038.9 | 233.3 | 225.0 | 208.8 | 209.1 | 188.2 | 164.0 | 127.8 |
| | Link 9–13 | 38536.4 | 29233.3 | 11400.0 | 7011.1 | 6500.0 | 3300.0 | 2087.5 | 672.7 | 900.0 | 525.0 | 418.5 | 533.3 | 283.3 | 191.7 |
| | Link 8–9 | 17207.7 | 9500.0 | 7900.0 | 4566.7 | 2508.7 | 2150.0 | 1295.8 | 323.1 | 320.0 | 469.2 | 366.7 | 217.4 | 239.3 | 150.0 |
| **UKNet** (23k–29k) | | | | | | | | | | | | | | | |
| | No failure | 1150.0 | 776.2 | 677.8 | 471.4 | 401.8 | 416.7 | 360.5 | 108.3 | 38.1 | 92.6 | 57.1 | 52.7 | 63.3 | 60.5 |
| | Link 3–18 | 100.0 | 100.5 | 101.5 | 102.5 | 103.0 | 103.3 | 105.1 | 0.0 | 0.2 | 0.3 | 0.5 | 0.5 | 1.3 | 2.1 |
| | Link 7–9 | 700.0 | 674.0 | 660.0 | 641.9 | 616.0 | 588.2 | 537.8 | 0.0 | -0.7 | 1.3 | 4.5 | 1.9 | 4.7 | 8.1 |
| | Link 8–10 | 3180.0 | 2205.6 | 2000.0 | 1534.6 | 1203.0 | 1015.4 | 826.3 | 60.0 | 44.4 | 65.0 | 34.6 | 39.4 | 30.8 | 26.3 |
| **EON** (9k–15k) | | | | | | | | | | | | | | | |
| | No failure | 31.2 | 48.4 | 51.2 | 43.3 | 54.8 | 60.4 | 61.8 | 6.2 | 1.6 | 7.3 | 6.7 | 9.5 | 9.4 | 5.9 |
| | Link 3–8 | 11011.1 | 7592.3 | 4709.5 | 3190.3 | 2475.0 | 1900.0 | 1444.1 | 0.0 | 0.0 | 19.0 | 12.9 | 15.0 | 15.4 | 10.3 |
| | Link 1–3 | 114.3 | 114.1 | 113.1 | 113.9 | 116.7 | 119.4 | 122.2 | 0.0 | 0.0 | -0.3 | 0.3 | 0.5 | 0.8 | 2.5 |
| | Link 8–10 | 809.1 | 702.6 | 556.3 | 460.3 | 400.0 | 367.9 | 319.4 | 0.0 | 0.0 | 0.0 | 0.0 | 1.4 | 9.0 | 7.5 |
| **ARPANET** (5k–11k) | | | | | | | | | | | | | | | |
| | No failure | 385.7 | 429.4 | 370.6 | 300.0 | 220.0 | 150.0 | 118.2 | 14.3 | 17.6 | 29.4 | 30.0 | 30.0 | 25.0 | 27.3 |
| | Link 6–8 | 729.3 | 733.3 | 748.6 | 755.3 | 775.0 | 762.7 | 710.3 | 0.0 | 0.0 | -1.4 | 0.9 | 2.1 | 5.9 | 6.9 |
| | Link 8–17 | 2164.2 | 2003.4 | 1793.9 | 1500.0 | 1210.0 | 1025.0 | 833.3 | 3.8 | 5.2 | 7.6 | 10.0 | 15.0 | 20.8 | 20.0 |
| | Link 3–5 | 34900.0 | 11284.6 | 4775.0 | 2666.7 | 1481.8 | 1079.5 | 833.3 | 100.0 | 69.2 | 43.7 | 26.7 | 23.6 | 28.2 | 28.6 |

multiple disjoint alternatives and under severe disruptions. The advantage of the DRL policy lies in its ability to internalize the long-term spectral impact of protection choices, implicitly avoiding recurrent fragmentation of critical links and preserving feasibility for future requests.

In highly constrained topologies with limited path diversity, the DRL policy naturally converges toward heuristic behavior, indicating that learning cannot overcome fundamental structural limitations but does not degrade performance either. This behavior confirms that the proposed approach is robust across heterogeneous network conditions.

Overall, the results suggest that incorporating learning into the protection-path selection process can substantially improve spectrum efficiency and resilience without altering the operational simplicity of traditional 1+1 protection.

Future work will extend the framework to multi-band scenarios, investigate graph-based encoders for richer state representations, and explore adaptive policies that can switch between dedicated and shared protection modes.

## REFERENCES


[1] Y. Wang, Z. Su, N. Zhang, R. Xing, D. Liu, T. H. Luan, and X. Shen, "A survey on metaverse: Fundamentals, security, and privacy," *IEEE communications surveys & tutorials*, vol. 25, no. 1, pp. 319–352, 2022.

[2] Z. Zhang, S. Yin, S. Guo, Z. Lin, Y. Chen, and S. Huang, "Dynamic buffering cost-saving multi-path routing under differential delay constraint in eons," in *Asia Communications and Photonics Conference*, pp. Su2A–114, Optica Publishing Group, 2018.

[3] C. Zhao, X. Li, S. Qi, Z. Zhao, K. Qiao, D. Li, Y. Liu, J. Xin, L. Zhang, and S. Huang, "A survey on reliability evaluation of optical networks," *IEEE Communications Surveys & Tutorials*, vol. 28, pp. 3443–3477, 2026.

[4] Z. Luo, S. Yin, L. Zhao, Z. Wang, W. Zhang, L. Jiang, and S. Huang, "Survivable routing, spectrum, core and band assignment in multi-band space division multiplexing elastic optical networks," *Journal of Lightwave Technology*, vol. 40, no. 11, pp. 3442–3455, 2022.

[5] K. Assis, R. Almeida Jr, M. Reed, A. Santos, H. Dinarte, D. Chaves, H. Li, S. Yan, R. Nejabati, and D. Simeonidou, "Protection by diversity in elastic optical networks subject to single link failure," *Optical Fiber Technology*, vol. 75, p. 103208, 2023.

[6] R. Goscien, K. Walkowiak, M. Klinkowski, and J. Rak, "Protection in elastic optical networks," *IEEE Network*, vol. 29, no. 6, pp. 88–96, 2015.

[7] K. Walkowiak, M. Klinkowski, B. Rabiega, and R. Goścień, "Routing and spectrum allocation algorithms for elastic optical networks with dedicated path protection," *Optical Switching and Networking*, vol. 13, pp. 63–75, 2014.

[8] S. Bala, S. Gorai, P. Sarkar, and M. Chatterjee, "Blocking aware offline survivable path provisioning of connection requests in elastic optical networks," *International Journal of Parallel, Emergent and Distributed Systems*, vol. 39, no. 1, pp. 116–137, 2024.

[9] F. Musumeci, C. Rottondi, G. Corani, S. Shahkarami, F. Cugini, and M. Tornatore, "A tutorial on machine learning for failure management in optical networks," *Journal of Lightwave Technology*, vol. 37, no. 16, pp. 4125–4139, 2019.

[10] X. Luo, C. Shi, L. Wang, X. Chen, Y. Li, and T. Yang, "Leveraging double-agent-based deep reinforcement learning to global optimization of elastic optical networks with enhanced survivability," *Optics express*, vol. 27, no. 6, pp. 7896–7911, 2019.

[11] Z. Zhang, Z. Luo, W. Zhang, S. Yin, S. Huang, X. Wang, N. An, and Z. Luo, "A multi-agents survivable deep reinforced routing and spectrum assignment algorithm in elastic optical network," in *2022 IEEE 12th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, pp. 1–4, IEEE, 2022.

[12] C. Zhao, X. Li, J. Xin, Y. Liu, D. Li, and S. Huang, "Intelligent agent-based dynamic reliability evaluation for optical networks: A comprehensive framework and case study," *IEEE Communications Magazine*, 2024.