

Neural-Network-Aided Supermode Basis Tracking for Feedforward Equalization in Strongly Coupled Multicore Fibers

Abdullah Oran*

Electrical and Electronics Engineering Department, Abdullah Gul University, Kayseri, 38039, Turkey

Received April 02, 2026; accepted October 10, 2026; published October 17, 2026

Abstract— Environmental perturbations make the supermode basis of strongly coupled multicore fibers drift in time, so that a receiver relying on a fixed basis suffers from inter-supermode crosstalk. We propose a feedforward receiver in which a small multilayer perceptron refines short sequences of pilot-based least-squares estimates of the channel unitary, and the refined estimate rotates the received core signals back into the supermode domain once per frame. In simulations of a six-core SC-MCF carrying 16-QAM, the tracker reduces the basis-estimation error by 2.3 dB with respect to pilot-only estimation, gains about 0.5 dB in required signal-to-noise ratio, and extends the reach at the hard-decision FEC limit from 1480 to 1635 km. It remains within 0.5 dB of an adaptive multiple-input multiple-output equalizer while avoiding its symbol-rate update loop.

Space-division multiplexing (SDM) increases the capacity of a single fiber by transmitting independent signals over several spatial channels [1]. In strongly coupled multicore fibers (SC-MCFs) the cores are placed so close together that light propagates in the collective eigenmodes of the core array, known as supermodes [2]. Random coupling along such fibers keeps the spatial-mode dispersion low, which makes them attractive for long-haul links [3]. The receiver must undo the mixing with multiple-input multiple-output (MIMO) digital signal processing (DSP), i.e., algorithms that compute every output stream from all received core signals [4–5].

Temperature changes, bending, and vibration perturb the core coupling, so the transformation between the transmitted supermodes and the received core signals is not fixed but drifts in time. A receiver that assumes a fixed (nominal or once-calibrated) supermode basis then suffers from inter-supermode crosstalk and a rising bit error rate (BER), defined as the fraction of erroneously detected bits. Adaptive MIMO equalizers based on the least-mean-squares (LMS) algorithm track this drift, but they update N^2 coefficients at the symbol rate inside a decision-feedback loop, which is difficult to parallelize in hardware [5].

Machine learning (ML) is increasingly used to model and compensate optical-channel impairments [6–8]. Here we use it in a narrow, well-defined role: a small multilayer perceptron (MLP) exploits the temporal correlation of the drift to denoise pilot-based least-squares (LS) estimates of

the channel unitary. The refined estimate is applied as a feedforward rotation that is updated only once per frame.

Let $\mathbf{a}(z, t) \in \mathbb{C}^N$ denote the complex field amplitudes in the N cores at position z and time t . In coupled-mode theory, they evolve as

$$\frac{d\mathbf{a}(z, t)}{dz} = j \mathbf{H}(z, t) \mathbf{a}(z, t), \quad (1)$$

where the Hermitian matrix $\mathbf{H}(z, t)$ (lossless propagation) contains the propagation-constant mismatches on its diagonal and the core-to-core coupling coefficients off the diagonal. We write it as

$$\mathbf{H}(z, t) = \mathbf{D} + \kappa \mathbf{C} + \sigma_\kappa \mathbf{G}(z, t), \quad (2)$$

where $\mathbf{D}$ is a diagonal matrix of small deterministic core detunings, $\mathbf{C}$ is the adjacency matrix of the core layout, κ is the nominal coupling coefficient and $\sigma_\kappa \mathbf{G}(z, t)$ is a random real-symmetric perturbation of the self- and nearest-neighbour terms. The supermodes of the unperturbed fiber are the columns of the unitary matrix $\mathbf{V}_0$ in the eigen-decomposition $\mathbf{H}_0 = \mathbf{D} + \kappa \mathbf{C} = \mathbf{V}_0 \mathbf{\Lambda}_0 \mathbf{V}_0^\dagger$, where $\mathbf{\Lambda}_0$ holds the supermode propagation constants and $(\cdot)^\dagger$ denotes the conjugate transpose. The fiber is divided into K sections of length Δz , each with its own perturbation $\mathbf{G}_i(t)$. To model a slow environmental drift, every perturbation follows a first-order Gauss–Markov process in time,

$$\mathbf{G}_i(t+1) = \rho \mathbf{G}_i(t) + \sqrt{1-\rho^2} \mathbf{W}_i(t), \quad (3)$$

where t is the frame index, $\mathbf{W}_i(t)$ is an independent random matrix with the same statistics as $\mathbf{G}_i$, and $1-\rho$ sets the drift rate. The fiber transfer matrix is $\mathbf{T}(t) = e^{j\mathbf{H}_K(t)\Delta z} \dots e^{j\mathbf{H}_1(t)\Delta z}$. If the supermodes are launched at the transmitter, the received core-domain samples after coherent detection and analog-to-digital conversion (ADC) are

$$\mathbf{y}[k] = \mathbf{U}_t \mathbf{s}[k] + \mathbf{n}[k], \quad \mathbf{U}_t = \mathbf{T}(t) \mathbf{V}_0, \quad (4)$$

where $\mathbf{s}[k]$ is the vector of supermode symbols at time index k and $\mathbf{n}[k]$ is circularly symmetric Gaussian noise that models amplified spontaneous emission (ASE). Model (4) is frequency-flat; it holds when the accumulated

* E-mail: abdullah.oran@agu.edu.tr



spatial-mode dispersion is small compared with the symbol period. Without perturbations $\mathbf{U}_t = \mathbf{V}_0 e^{j\Lambda_0 K \Delta z}$, so projection onto $\mathbf{V}_0$ followed by one complex gain per mode would separate the supermodes. The perturbations make $\mathbf{U}_t$ deviate from this nominal form and change with t .

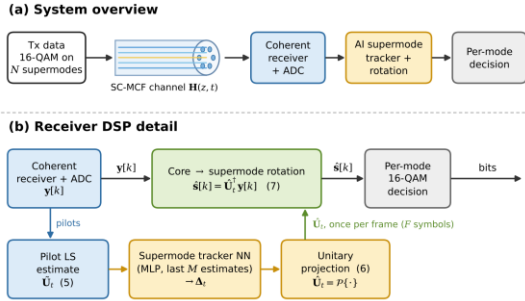


Fig. 1. AI-assisted supermode tracking receiver: (a) system overview; (b) receiver DSP detail.

Figure 1 shows the proposed receiver. Every frame of F symbols starts with P pilot symbols per mode, collected in $\mathbf{S} \in \mathbb{C}^{N \times P}$. The corresponding received samples $\mathbf{Y}_t$ give the LS estimate

$$\tilde{\mathbf{U}}_t = \mathbf{Y}_t \mathbf{S}^\dagger (\mathbf{S} \mathbf{S}^\dagger)^{-1}. \quad (5)$$

The MLP f_θ receives the relative rotations $\tilde{\mathbf{U}}_t^\dagger \tilde{\mathbf{U}}_{t-m} - \mathbf{I}$, $m = 1, \dots, M-1$, and the logarithm of the pilot residual power as a noise-level feature. Expressing the inputs relative to the newest estimate makes the network independent of the absolute basis orientation. Its output is a correction Δ_t , and the refined estimate is

$$\hat{\mathbf{U}}_t = \mathcal{P}\{\tilde{\mathbf{U}}_t (\mathbf{I} + \Delta_t)\}, \quad (6)$$

where $\mathcal{P}\{\mathbf{A}\} = \mathbf{Q}_1 \mathbf{Q}_2^\dagger$, obtained from the singular value decomposition $\mathbf{A} = \mathbf{Q}_1 \mathbf{\Sigma} \mathbf{Q}_2^\dagger$, is the closest unitary matrix in the Frobenius norm [9]. The data symbols are then recovered by a single rotation followed by per-mode decisions:

$$\hat{\mathbf{s}}[k] = \hat{\mathbf{U}}_t^\dagger \mathbf{y}[k]. \quad (7)$$

The MLP is trained offline to minimize the mean squared error between Δ_t and $\tilde{\mathbf{U}}_t^{-1} \mathbf{U}_t - \mathbf{I}$, using 80 000 frames from 200 simulated trajectories with random distance (200–2200 km), signal-to-noise ratio (SNR) offset (± 6 dB) and drift rate ($3 \times 10^{-4} \leq 1 - \rho \leq 3 \times 10^{-3}$), the Adam optimizer and early stopping. All results use independent channel realizations. Estimation quality is measured by the basis error $\varepsilon_t = \|\hat{\mathbf{U}}_t - \mathbf{U}_t\|_F^2 / N$.

Table 1 lists the simulation parameters; the ring detunings lift the degeneracy of the ring supermodes, and the SNR decreases with distance as ASE accumulates. Four receivers are compared: (i) a nominal fixed-basis receiver that projects onto $\mathbf{V}_0$ and applies one pilot-estimated gain per mode; (ii) pilot LS tracking, i.e., (6) with $\Delta_t = \mathbf{0}$; (iii) the proposed MLP tracker; and (iv) a one-tap $N \times N$ adaptive MIMO equalizer trained on the same pilots and updated by decision-directed (DD) LMS at every symbol,

with the best step size from $\{2, 5, 10\} \times 10^{-3}$ at each point. Perfect channel state information (CSI), $\hat{\mathbf{U}}_t = \mathbf{U}_t$, gives the lower bound.

Table 1. Simulation parameters

Parameter	Value
Cores / supermodes, N	6 (ring), detuning $\leq 0.15\kappa$
Perturbation	$\sigma_\kappa = 0.1\kappa$ ($\sigma_\kappa^2 = 0.01$), 10 km sections
Modulation	16-ary quadrature amplitude modulation (16-QAM)
Frame / pilots	$F = 1024$, $P = 16$ (1.6% overhead)
Drift rate	$1 - \rho = 10^{-3}$ per frame (unless stated)
SNR per mode	28 dB – $10 \log_{10}(L/100 \text{ km})$
MLP	505–128–128–72, ReLU, $M = 8$
Monte-Carlo	5 runs $\times$ 200 frames (2.4×10^7 bits/point)

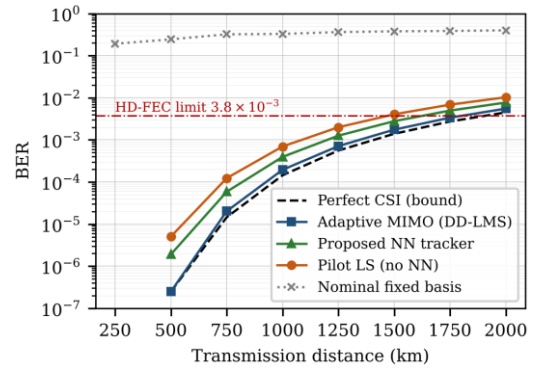


Fig. 2. BER versus transmission distance.

Figure 2 shows the BER versus distance. The nominal fixed-basis receiver fails at all distances ($\text{BER} > 0.19$) because the perturbations already redistribute about 30% of the power between supermodes after 250 km. At 1000 km the BER is 7.0×10^{-4} with pilot LS, 3.9×10^{-4} with the proposed tracker, 1.9×10^{-4} with adaptive MIMO and 1.4×10^{-4} with perfect CSI. At the hard-decision forward-error-correction (HD-FEC) limit of 3.8×10^{-3} , the reach grows from 1480 km (pilot LS) to 1635 km with the MLP, compared with 1810 km for adaptive MIMO and 1915 km for perfect CSI.

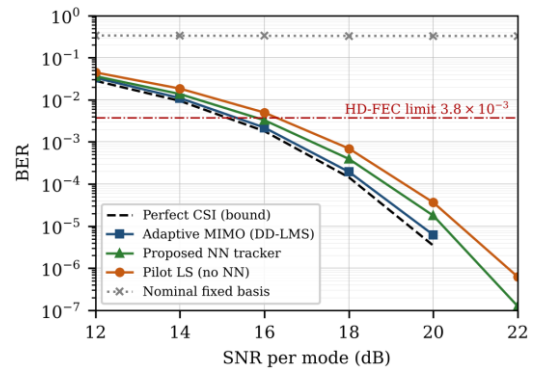


Fig. 3. BER versus SNR per mode at 1000 km.

Figure 3 gives the BER versus SNR at 1000 km. To reach a BER of 10^{-3} , the receivers need 16.5 dB (perfect CSI), 16.6 dB (adaptive MIMO), 17.1 dB (proposed) and 17.6 dB (pilot LS). The MLP thus recovers about half of the 1.2 dB penalty of pilot-only tracking without extra pilots; the remaining 0.5 dB gap to adaptive MIMO is the price of updating only once per frame.

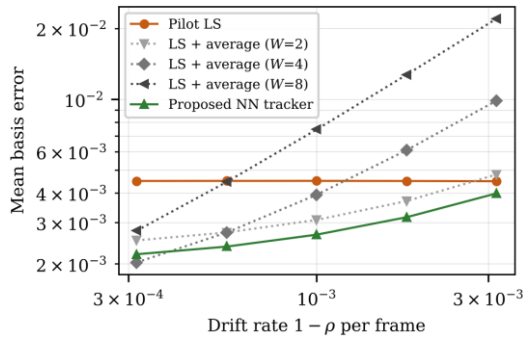


Fig. 4. Mean basis error at 1000 km versus drift rate (5 runs). LS + average: mean of the last W LS estimates.

Static processing fails because the drift is fast relative to the frame rate: at 1000 km, a basis calibrated once at $t = 0$ exceeds $\varepsilon_t = 10^{-1}$ within about 25 frames. Pilot LS keeps the error near 4.4×10^{-3} , and the MLP lowers it to about 2.6×10^{-3} . Fig. 4 compares the MLP with averaging the last W LS estimates. A fixed window helps only near one drift rate: long windows lag behind fast drift, short windows leave noise at slow drift. The MLP adapts to the drift rate without knowing it and reaches the lowest error at all but the slowest drift, where the best window ($W = 4$) is 0.35 dB better.

Per received symbol vector, adaptive MIMO needs $2N^2 + N = 78$ complex multiplications for $N = 6$, since it filters with N^2 coefficients and updates all of them at every symbol. The proposed receiver applies only the rotation (7), i.e., $N^2 = 36$ multiplications, at the symbol rate. The LS estimate, the MLP ($\approx 9 \times 10^4$ real multiplications) and the unitary projection are computed once per frame, adding about 25 complex multiplications per symbol for $F = 1024$. The total of about 61 is 22% lower. More important is the absence of a symbol-rate feedback loop: the rotation can be applied to blocks of samples in parallel, and no step size has to be tuned. Both receivers still scale as $\mathcal{O}(N^2)$ per symbol.

In conclusion, an MLP that refines pilot-based LS estimates of the drifting supermode basis enables a feedforward SC-MCF receiver. Relative to pilot-only tracking, it lowers the basis error by 2.3 dB at the nominal drift rate, saves about 0.5 dB of SNR and extends the HD-FEC reach by about 10%. It stays within 0.5 dB of adaptive MIMO while needing fewer multiplications and no symbol-rate update. Future work will extend the approach to frequency-selective channels with non-negligible spatial-mode dispersion, e.g., one rotation per subcarrier in

a frequency-domain equalizer [5],[10], and validate it experimentally.

References

- [1] D.J. Richardson, J.M. Fini, L.E. Nelson, *Nat. Photonics* **7**(5), 354 (2013).
- [2] C. Xia *et al.*, *Opt. Expr.* **19**(17), 16653 (2011).
- [3] T. Hayashi, Y. Tamura, T. Hasegawa, T. Taru, *J. Lightwave Technol.* **35**(3), 450 (2017).
- [4] R. Ryf *et al.*, *J. Lightwave Technol.* **30**(4), 521 (2012).
- [5] S.Ö. Arik, D. Askarov, J.M. Kahn, *J. Lightwave Technol.* **32**(10), 1841 (2014).
- [6] D. Zibar, M. Piels, R. Jones, C.G. Schäffer, *J. Lightwave Technol.* **34**(6), 1442 (2016).
- [7] B. Karanov *et al.*, *J. Lightwave Technol.* **36**(20), 4843 (2018).
- [8] T. Koike-Akino *et al.*, *J. Lightwave Technol.* **38**(11), 3059 (2020).
- [9] N.J. Higham, *SIAM J. Sci. Stat. Comput.* **7**(4), 1160 (1986).
- [10] K.-P. Ho, J.M. Kahn, *J. Lightwave Technol.* **29**(21), 3119 (2011).