Tensor-Based Modulation for Massive Random Access

Maxime Guillaud

Joint work with Alexis Decurninge and Ingmar Land

Huawei Technologies Paris

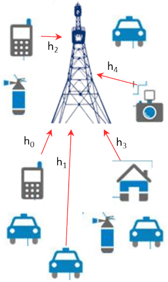
Workshop on Tensor Theory and Methods – Paris, November 23-24, 2022

HUAWEI

Outline of talk

Massive Random Access Scenario

- Multipoint-to-Point, K users
- **Random user activation** (K_a active), unknown to the receiver, no coordination
- Regime of interest:
  - Massively many users ($K \gg K_a$)
  - Small payload
  - High degree of contention with no packet loss



Potential fields of applications:

- Cellular systems: low-latency communications
- V2V: safety-related transmissions
- Aeronautics: ADS-B, radiotelephony

Non-Coherent Point-to-Point

- Block-fading channel with blocksize T

$$\mathbf{y} = \mathbf{s}h + \mathbf{w} \quad \in \quad \mathbb{C}^T$$

- Channel state $h \in \mathbb{C}$ is unknown to both transmitter and receiver

Classical approaches to design the transmitted sequence $\mathbf{s}$:

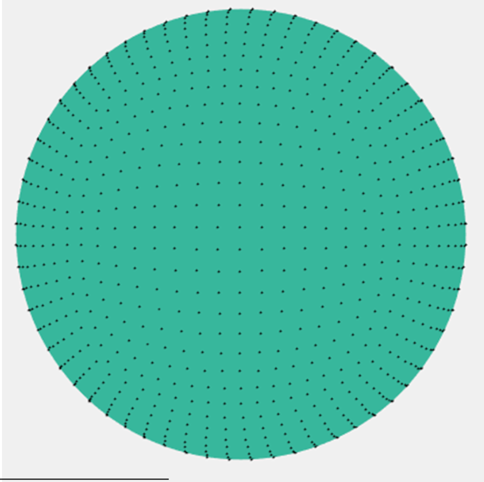
Pilot-based 1 reference symbol and $T-1$ information-bearing scalar symbols (QAM...): $\mathbf{s} = \begin{bmatrix} 1 \\ * \end{bmatrix}$

Grassmannian $\mathbf{s}$ is chosen from a vector modulation that remains identifiable under scalar (h) uncertainty [1]

[1] Lizhong Zheng and D. N. C. Tse. "Communication on the Grassmann Manifold: A Geometric Approach to the Noncoherent Multiple-Antenna Channel". In: **IEEE Transactions on Information Theory** 48.2 (Feb. 2002), pp. 359–383. ISSN: 0018-9448. DOI: 10.1109/18.978730.

Example of Grassmannian Modulation

Example of a vector constellation in $\mathbb{R}^3$ (on a sphere due to power constraints) from [2] – 768 possible states, encoding 9.58 bits



[2] Khac-Hoang Ngo et al. "Cube-Split: A Structured Grassmannian Constellation for Non-Coherent SIMO Communications". In: **IEEE Transactions on Wireless Communications** 19.20 (Mar. 2020).

Non-Coherent **Multiple-Access**, SIMO

- N receive antennas, channel state $\mathbf{h}_k \in \mathbb{C}^N$ for user k

$$\mathbf{Y} = \sum_{k=1}^{K} \mathbf{s}_k \mathbf{h}_k^T + \mathbf{W} \quad \in \quad \mathbb{C}^{T \times N}$$

Pilot-based: channel estimation (LS) with K **linearly independent sequences**, then coherent data modulation

Grassmannian: Difficult joint design of multiple constellations [3]

[3] Khac-Hoang Ngo et al. "Noncoherent MIMO Multiple-Access Channels: A Joint Constellation Design". In: **Proc. IEEE Information Theory Workshop (ITW) 2020**.

Non-Coherent **Random Access**

- The set $\mathcal{A}$ of active users is unknown

$$\mathbf{Y} = \sum_{k \in \mathcal{A}} \mathbf{s}_k \mathbf{h}_k^T + \mathbf{W}$$

- **Activity detection** needed to estimate $\mathcal{A}$
- Joint activity detection and CSI estimation typically requires complex algorithms

Unsourced Random Access

- Analyzed for the Gaussian MAC by Polyanskyi [4]
- All users employ the **same codebook** (no user-specific preamble or pilot sequence)
- Decoding is done up to a permutation of the active users
- Benefit: lower decoder complexity (especially for $K \gg K_{\mathrm{a}}$)
- User ID can still be included in the payload ($\lceil \log_2 K \rceil$ bits): allows an **unsourced** design at the physical layer to behave like a **sourced** scheme

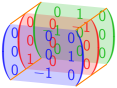
[4] Y. Polyanskiy. "A perspective on massive random-access". In: **IEEE Int. Symp. Inf. Theory (ISIT)**. 2017.

Our Contribution: **Tensor-Based Modulation**

- An **unsourced PHY-layer design** for massive random access
- Suitable for **fading channels**, non-coherent
- Suitable for **multiple antennas** at Rx (SIMO)
- Based on a **low-rank tensor** construction

Tensor Notations

Tensors are merely multi-dimensional data structures that generalize matrices to $d \geq 2$ dimensions



- Outer product and vectorization operator $\mathcal{V}$ (matrix case): for $\mathbf{a} \in \mathbb{C}^{\mathbf{m}}, \mathbf{b} \in \mathbb{C}^{\mathbf{n}}$, $\otimes =$ Kronecker product,

$$\mathcal{V}(\underbrace{\mathbf{a}\mathbf{b}^{\mathsf{T}}}_{\in\ \mathbb{C}^{m\times n}}) = \underbrace{\mathbf{a} \otimes \mathbf{b}}_{\in\ \mathbb{C}^{mn}}$$

- Generalization to tensors of order $d \geq 2$:

$$\mathcal{V}(\mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \cdots \circ \mathbf{a}^{(d)}) = \mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)} \otimes \cdots \otimes \mathbf{a}^{(d)}$$

- $\mathcal{V}$ is an isomorphism between the space of $(T_1, \ldots, T_d)$-dimensional tensors and the space of $(\prod_{i=1}^{d} T_i)$-dimensional vectors (with the respective sums)
- $\mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \cdots \circ \mathbf{a}^{(d)}$ is a rank-1 tensor

Random Access Fading Channel Model

- $K_{\mathrm{a}} = |\mathcal{A}|$ active transmitters
- SIMO channel from user k to Rx denoted by $\mathbf{h}_k \in \mathbb{C}^N$
- Block-fading with blocklength T
- Block synchronicity across the users
- User k transmits $\mathbf{s}_k \in \mathbb{C}^T$
- Signal received by the N antennas over the T channel accesses: $\mathbf{Y} \in \mathbb{C}^{T \times N}$, corrupted by additive noise $\mathbf{W} \in \mathbb{C}^{T \times N}$

$$\mathbf{Y} = \sum_{k \in \mathcal{A}} \mathbf{s}_k \mathbf{h}_k^T + \mathbf{W} \quad \xrightarrow{\mathcal{V}} \quad \mathbf{y} = \sum_{k \in \mathcal{A}} \mathbf{s}_k \otimes \mathbf{h}_k + \mathbf{w}$$

Tensor-Based Modulation

- Assume that the blocklength factorizes as $T = \prod_{i=1}^{d} T_i$
- Let the transmitted symbol $\mathbf{s}_k$ be a **rank-1 tensor** of dimension $(T_1, \ldots, T_d)$:

$$\mathbf{s}_k = \mathbf{x}_{1,k} \otimes \cdots \otimes \mathbf{x}_{d,k}$$

- $\mathbf{x}_{i,k}$ is from a finite **vector constellation sub-codebook** $\mathcal{C}_i \subset \mathbb{C}^{T_i}$, $i = 1 \ldots d$, suitable for P-t-P non-coherent communications
- At the receiver:

$$\mathbf{y} = \sum_{k \in \mathcal{A}} \underbrace{\mathbf{x}_{1,k} \otimes \cdots \otimes \mathbf{x}_{d,k} \otimes \mathbf{h}_k}_{\substack{\text{rank-1 tensor of order } d+1 \\ \text{and dimension} (T_1, \ldots, T_d, N)}} + \mathbf{w}$$

CP Decomposition, Unicity

- The decomposition of an arbitrary tensor into a **minimal number of rank-1 components** is known as the canonical polyadic decomposition (CPD)

$$\sum_{k\in\mathcal{A}} \mathbf{x}_{1,k} \otimes \cdots \otimes \mathbf{x}_{d,k} \otimes \mathbf{h}_k \Longrightarrow \begin{cases} \mathbf{x}_{1,1} \otimes \cdots \otimes \mathbf{x}_{d,1} \otimes \mathbf{h}_1 \\ \vdots \\ \mathbf{x}_{1,K_\mathrm{a}} \otimes \cdots \otimes \mathbf{x}_{d,K_\mathrm{a}} \otimes \mathbf{h}_{K_\mathrm{a}} \end{cases}$$

- Tensors ($d > 2$) have specific, non-trivial conditions for CPD unicity
- Unicity can only be achieved
  - up to a permutation between the K_a terms, and
  - up to d scalars per rank-1 term

Unique Decomposibility Conditions from [5]

- For $(T_1, \ldots, T_d, N)$-tensors with general choices of the $\mathbf{x}_{i,k}, \mathbf{h}_k$, the CPD is **almost surely unique** for tensors of rank $K_\mathrm{a} < \overline{R}$,

$$\overline{R} = \begin{cases} R^1 - 1 & \text{for } T_1 \geq R^1 \\ R^2 - 1 & \text{for } N \geq R^2 \\ R^0 & \text{otherwise} \end{cases}$$

where

$$R^0 = \left\lceil \frac{N \prod_{i=1}^{d} T_i}{N + \sum_{i=1}^{d} (T_i - 1)} \right\rceil \quad \text{(expected generic rank)},$$

$$R^1 = 2 - N + N \prod_{i=2}^{d} T_i - \sum_{i=2}^{d} (T_i - 1), \quad \text{and}$$

$$R^2 = 1 + T - \sum_{i=1}^{d} (T_i - 1)$$

- Much more **favorable rank scaling** than for the matrix case!

[5] L. Chiantini, Giorgio Ottaviani, and Nick Vannieuwenhoven. "An algorithm for generic and low-rank specific identifiability of complex tensors". In: **SIAM Journal on Matrix Analysis and Applications** 35.4 (2014), pp. 1265–1287.

Practical Multi-User Decoding Approach

- Approximate CPD for **user separation** (K_a known)

$$\left\{\hat{\mathbf{z}}_{i,k}, \hat{\mathbf{h}}_k\right\} = \underset{\substack{\mathbf{z}_{i,k}\in\mathbb{C}^{T_i} \\ \mathbf{h}_k\in\mathbb{C}^N}}{\operatorname{argmin}} \left\| \mathbf{y} - \sum_{k=1}^{K_\mathrm{a}} \mathbf{z}_{1,k} \otimes \cdots \otimes \mathbf{z}_{d,k} \otimes \mathbf{h}_k \right\|_2^2$$

  - Solvable with e.g. Gauss-Newton approach [6]
  - Yields acceptable results despite the lack of Eckhart-Young – like result
  - Optimization over a continuous space: does not account for the discrete nature of the constellation

- Single-user non-coherent demapping, **per sub-constellation** (low complexity)

$$\hat{\mathbf{x}}_{i,k} = \underset{\mathbf{x}_{i,k}\in\mathcal{C}_i}{\operatorname{argmax}} \frac{\left|\mathbf{x}_{i,k}^H \hat{\mathbf{z}}_{i,k}\right|}{\|\hat{\mathbf{z}}_{i,k}\|\|\mathbf{x}_{i,k}\|} \quad \forall i, k$$

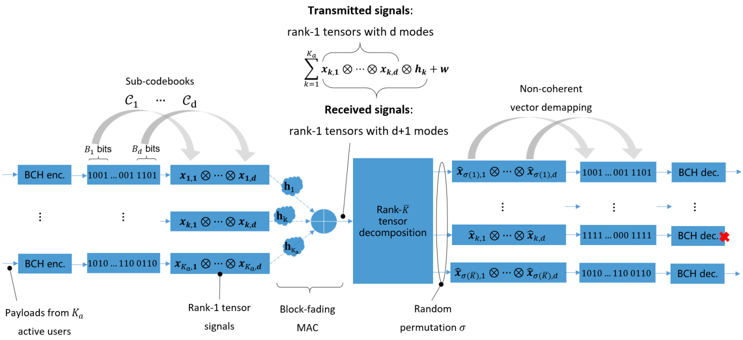
---

[6] L. Sorber, M. Van Barel, and L. De Lathauwer. “Optimization-based algorithms for tensor decompositions: Canonical polyadic decomposition, decomposition in rank-(Lr;Lr;1) terms, and a new generalization”. In: **SIAM J. Optim** 23.2 (2013), pp. 695–720.

Random Activation

- K_a **is unknown** to the receiver
- Try to compute the CPD anyway, and limit the number of rank-1 terms to an arbitrary limit $\bar{K}$
  - this generates at least $\bar{K} - K_\mathrm{a}$ spurious messages
- Add a **binary channel code** and use it as a CRC in an attempt to remove the $\bar{K} - K_\mathrm{a}$ "ghost" users
- Leverage the binary code for error **correction**: BCH

Tensor-Based Modulation: Recap



Achievable Degrees of Freedom (DoF)

- DoF characterizes the scaling of capacity at high SNR
- Based on the identifiability results and interpretation of $\mathbf{x}_{1,k}$ as a Grassmannian variable in dimension T_i
- At infinite SNR, each rank-1 component is equivalent to the parallel, noise-free transmission of d Grassmannian variables
  - Per-user complex DoF: $\sum_{i=1}^{d}(T_i - 1)$
  - Sum-DoF: $\mathrm{D}_{\mathrm{TBM}}(K_{\mathrm{a}}) = K_{\mathrm{a}} \sum_{i=1}^{d}(T_i - 1)$
- Tensor identifiability results indicate that at most $K_{\mathrm{a}} = \overline{R} - 1$ can be identifiable, hence the maximum Sum-DoF:

$$\mathrm{D}_{\mathrm{TBM}}(\overline{R} - 1) = (\overline{R} - 1) \sum_{i=1}^{d}(T_i - 1)$$

DoF Bounds

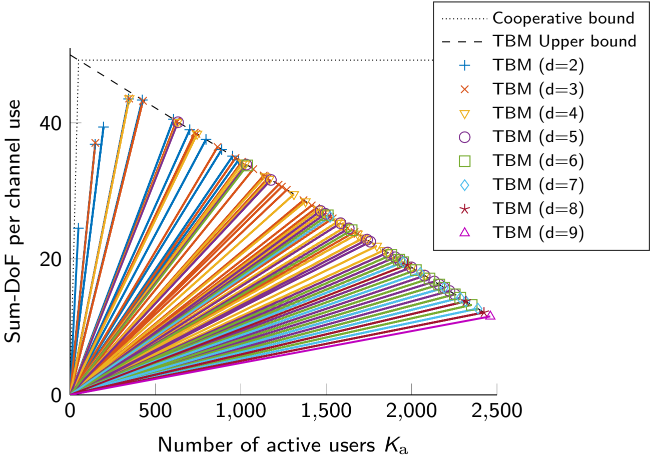
- DoF upper-bound based on the expression of $\overline{R}$ (independent of tensor size factorization):

$$\mathrm{D}_{\mathrm{TBM}}(K_{\mathrm{a}}) < N(T - K_{\mathrm{a}})$$

- Cooperative DoF upper-bound (point-to-point $N \times K_{\mathrm{a}}$ non-coherent channel):

$$\mathrm{D}_{\mathrm{coop}}(K_{\mathrm{a}}) = M^*(T - M^*) \quad \text{where} \quad M^* = \min(K_{\mathrm{a}}, N, \lfloor T/2 \rfloor)$$

Achievable Degrees of Freedom (DoF)



Achievable sum-DoF per channel use $(\mathrm{D_{TBM}}(K_\mathrm{a})/T)$ vs. K_a for different tensor sizes (d and T_i), for $T = 3200$ and $N = 50$ antennas. Markers denote the case $K_\mathrm{a} = \overline{R} - 1$, while the slope of the lines going through the origin represents the per-user DoF.

Further Practical Implementation Considerations

- Which **constellation for the $\mathcal{C}_i$**?
  - Grassmannian vector constellation design to deal with the scalar indeterminacy on each $\mathbf{x}_{i,k}$ (non-coherent)
  - We use our own low-complexity approach [7]
- Extension to **soft-demapping**
  - Compute coded bits log-likelihood ratios based on $\hat{\mathbf{x}}_{i,k}$ and $\mathcal{C}_i$
  - Need to know how "noisy" the output of the CPD behaves

[7] Khac-Hoang Ngo et al. "Cube-Split: A Structured Grassmannian Constellation for Non-Coherent SIMO Communications". In: **IEEE Transactions on Wireless Communications** 19.20 (Mar. 2020).

Simulation Performance Metrics

- Unsourced setting: Denote $\mathcal{L}$ the set of transmitted messages, and $\hat{\mathcal{L}}$ the messages output by the MU receiver
  - Message error rate (K_a unknown):

$$\mathsf{MER} = \min\left(\underbrace{\frac{|\mathcal{L} \setminus \hat{\mathcal{L}}|}{|\mathcal{L}|}}_{\substack{\text{average ratio}\\ \text{of missed messages}}} + \underbrace{\frac{|\hat{\mathcal{L}} \setminus \mathcal{L}|}{|\hat{\mathcal{L}}|}}_{\substack{\text{average ratio}\\ \text{of phantom messages}}}, 1\right)$$

  - Per-User Probability of Error (K_a assumed known):

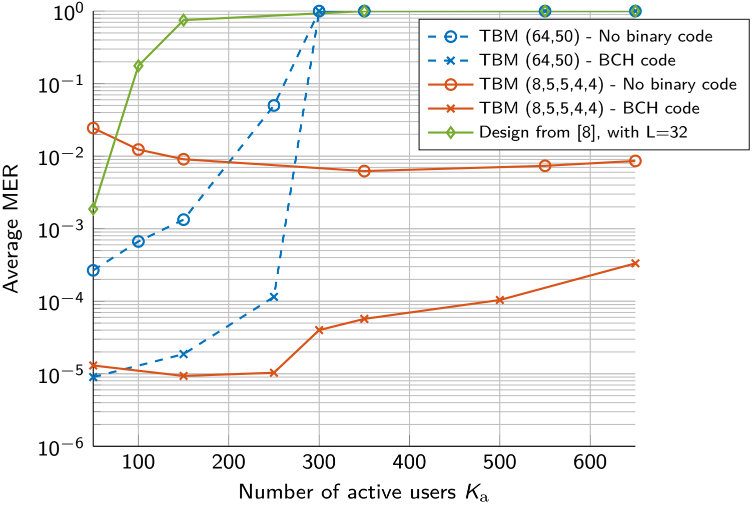
$$\mathsf{PUPE} = \frac{|\mathcal{L} \setminus \hat{\mathcal{L}}|}{|\mathcal{L}|}$$

- Sourced setting:
  - Packet error rate (PER) computed after user re-identification with decoded user ID

Simulation Parameters

- Blocksize $T = 3200$
- $N = 1$ or 50 antennas (with i.i.d. Rayleigh fading)
- 96 bits payload + 14 bits redundancy (BCH) = 110 bits
- Two tensor dimensions:
  - $(8, 5, 5, 4, 4)$ with $B_1, B_2, B_3, B_4, B_5 = 37, 21, 21, 16, 15$ bits in $\mathcal{C}_1, \ldots, \mathcal{C}_5$
  - $(64, 50)$ with $B_1, B_2 = 62, 48$ bits in $\mathcal{C}_1, \mathcal{C}_2$
- $\bar{K} = K_{\mathrm{a}}$ assumed to facilitate comparison with SotA
- A slightly improved receiver capable of interference cancelation (exploiting the binary code)

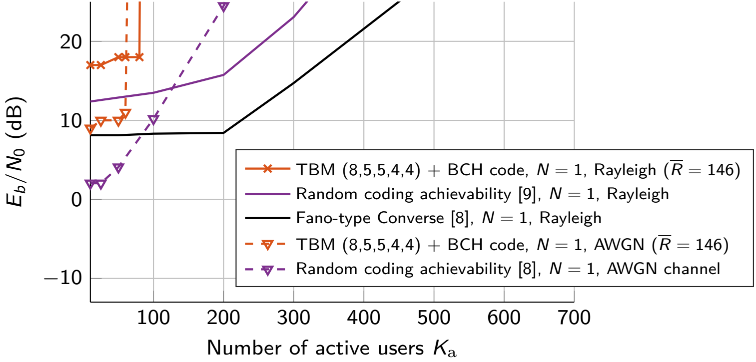
Message Error Rate, $N = 50$ antennas, $E_b/N_0 = 0$ dB



[8] A. Fengler et al. **Massive MIMO unsourced random access**. 2019. URL: https://arxiv.org/abs/1901.00828.

Minimum Achievable E_b/N_0, $N = 1$ antenna

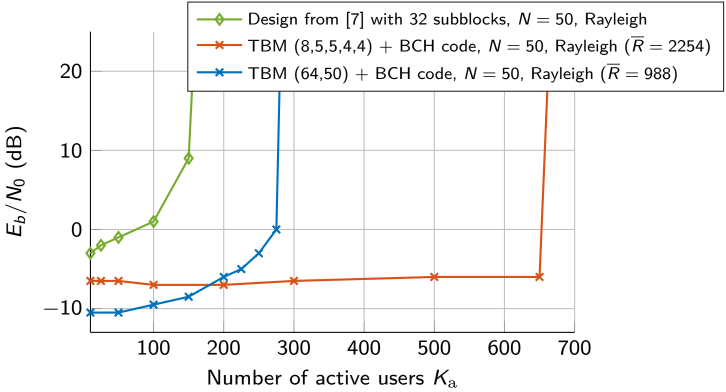
Y-axis: required E_b/N_0 to achieve PUPE ≤ 0.1



[9] S. Kowshik et al. **Energy efficient coded random access for the wireless uplink**. 2019. URL: https://arxiv.org/abs/1907.09448.

Minimum Achievable E_b/N_0, $N = 50$ antennas

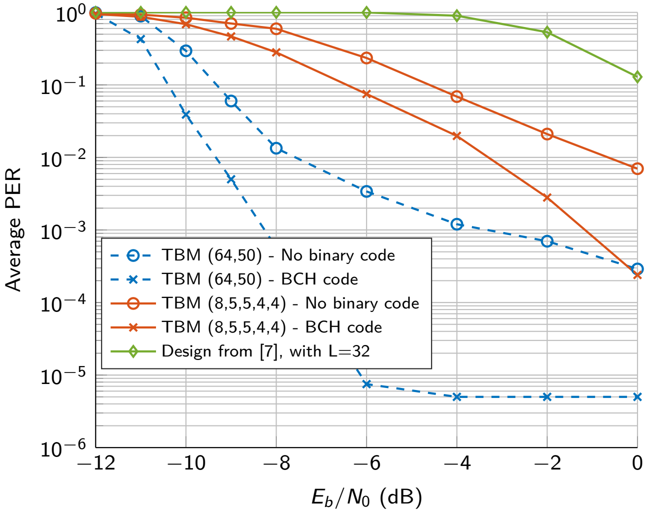
Y-axis: required E_b/N_0 to achieve PUPE ≤ 0.1



- Increasing N allows to support more users, lower required E_b/N_0
- K_a up to 650 active users, with 96-bit payloads, over $T = 3200$ channel uses $\longrightarrow$ **19.5 bits/cu!**

Sourced Setup, $N = 50$ antennas

PER vs. E_b/N_0 for $T = 3200$, $K_\mathrm{a} = 100$ and $K = 8192$



Possible Extensions

- Application to OFDM systems
  - Small (within cyclic prefix) synchronization errors can break the flat-fading assumption
  - A well-chosen mapping of $\mathbf{s}_k$ to the time-frequency grid can preserve the rank-1 tensor property of each user
- Totally asynchronous multiple access
- Fast time-varying channel

Conclusion

- Summary
  - Low-rank tensor structure allows user separation without relying on the discrete nature of the constellation
  - Designed for fading channels
  - Benefits from Rx diversity
  - No assumption about the fading distribution
- Outlook
  - Applications: stand-alone, or provide multi-reception to classical MAC schemes (ALOHA with multi-reception)
  - Theoretical understanding of the limits of low-rank tensor detection still lacking

Details in A. Decurninge, I. Land, and M. Guillaud. "Tensor-Based Modulation for Unsourced Massive Random Access". In: **IEEE Wireless Communications Letters** 10.3 (Mar. 2021). DOI: 10.1109/LWC.2020.3037523