Link Adaptation Using Joint-Thompson Sampling

Vignatha Vinjam* Manjunath Kolavennu* Myna Vajha* Karthik Periyapattana Narayanaprasad†
* Department of EE, IIT Hyderabad † Department of AI, IIT Hyderabad
ee23resch14002@iith.ac.in, ee25mtech02002@iith.ac.in, mynav@ee.iith.ac.in, pnkarthik@ai .iith.ac.in

***Abstract*—The choice of Modulation and Coding (MCS) type for a particular channel condition is made through link adaptation (LA) algorithms that operate at the MAC layer. These algorithms rely on the ACK/NACK statistics and the channel quality index (CQI) feedback. Several existing works model LA as a multi-armed bandit (MAB) problem across cellular and Wi-Fi links. In the MAB formulation, each available MCS is a Bernoulli arm parameterized by its transmission success probability, and the goal is to design a selection strategy that accrues maximum reward. Several popular MAB algorithms, such as upper confidence bound (UCB) and Thompson Sampling (TS), have been proposed in the literature. Using the fact that MCS success probabilities are ordered, we propose the Joint-Thompson Sampling (Joint-TS) algorithm. Unlike classical TS, which assumes independent Beta distributions for each arm, Joint-TS utilizes a multivariate ordered Beta distribution as the prior to preserve the inherent monotonicity of success probabilities. Our simulation results show that while existing MAB algorithms fail in specific scenarios, Joint-TS delivers competitive throughput with robust, consistent performance in all scenarios.**

***Index Terms*—Thompson Sampling, multi-armed bandits, link adaptation, Wi-Fi, 5G, 6G, multivariate ordered Beta distribution**

I. Introduction

Link adaptation (LA) is an essential MAC layer algorithm that selects link parameters such as Modulation and Coding Scheme (MCS), power control, number of layers, and precoders for dynamic wireless channels. LA focusing on MCS selection is widely known as Adaptive Modulation and Coding (AMC) or as Adaptive Data Rate (ADR) in WiFi. The goal of MCS selection is to maximize expected throughput under an error rate constraint for a given channel state—the Signal-to-Interference-plus-Noise Ratio (SINR). Mathematically,

$$R_{\max}(c) \quad = \max_{i:\ \mathrm{BLER}(i,c)\,\leq\, p_{\mathrm{target}}} s_i \cdot (1-\mathrm{BLER}(i,c)),$$

where s_i is the spectral efficiency corresponding to MCS i, determined as product of number of bits per modulation symbol times the rate of forward error correction (FEC) code used (cf. Table I), and $\mathrm{BLER}(i,c)$ is the Block Error Rate at channel state (or effective SINR) c upon using MCS i. Effective SINR [1] maps the sub-carrier SINR vector to a single scalar, enabling the use of Additive White Gaussian Noise (AWGN) BLER tables. While the BLER target, p_{target}, is typically set to 0.1 [2], SINR-dependent BLER targets result in better performance of ARQ schemes at the MAC layer. [3] If the throughput-maximizing MCS always satisfies this BLER constraint, then the condition $\mathrm{BLER}(i,c) \leq p_{\mathrm{target}}$ can be ignored. We observe this holds true for the 5G NR waveform simulations conducted over the AWGN channel using the MATLAB 5G toolbox.

Myna Vajha is supported by ANRF-ECRG and IIT Hyderabad Seed Grant.

MCS (i)	M_i	F_i	s_i
0	2	120	0.2344
1	2	157	0.3066
2	2	193	0.3770
3	2	251	0.4902
⋮	⋮	⋮	⋮
27	6	910	5.3320
28	6	948	5.5547

TABLE I: 5G NR MCS configurations [4, 5.1.3.1 Table 1], where M_i is bits/symbol, F_i is FEC code rate ($\times 1024$), and $s_i = M_i \cdot F_i/1024$ is the spectral efficiency.

While uplink MCS selection is straightforward using the SINR estimates, the downlink selection is a complex task since the gNB lacks direct channel knowledge. As a remedy, a quantized MCS suggestion, called Channel Quality Index (CQI), can be requested from the UE. See [4, 5.2.2.1] for a list of 5G NR CQI options.

a) Prior Work: Link adaptation in wireless systems has traditionally relied on two core mechanisms: Inner Loop Link Adaptation (ILLA) and Outer Loop Link Adaptation (OLLA) (see [5], [6]). ILLA maps CQI or estimated SINR to an MCS using predefined lookup tables, while OLLA works in conjunction by adding an SINR offset adjusted via Acknowledgment (ACK)/ Negative Acknowledgment (NACK) feedback. The ratio of upward to downward offset adjustments is determined by the target BLER, ensuring the expected BLER remains constrained. Although OLLA is effective, it suffers from slow responses to fast-changing channels [7], sensitivity to step sizes, and a lack of per-CQI state information. An optimization scheme for initial OLLA offset selection is proposed in [8]; however, this requires periodic intervention at each cell-site.

To address these limitations, recent literature models LA as a multi-armed bandit (MAB) problem. MAB framework models each MCS as an arm generating Bernoulli rewards, allowing the algorithm to learn from history of rewards and arm selections in an online fashion. This facilitates principled exploration of uncertain arms while ensuring performance is maximized, as more reward samples are observed. MAB-based solutions can outperform OLLA in highly dynamic or unpredictable channels, requiring no manual parameter tuning and adapting quickly based on observed feedback. In [9], the authors formulate rate adaptation in 802.11 as an MAB problem and propose a KL-UCB based algorithm called G-

ORS, which assumes unimodality of the expected throughput. ϵ-greedy approaches [10] balance random exploration with the greedy exploitation of the empirically best-performing arm. MTS [11] and BayesLA [12] have similar approaches based on Thompson Sampling (TS). Here, independent Beta priors are maintained for the success probability of each MCS, at every CQI. In [13], a TS-based alternative to G-ORS is proposed. In [14] and [15], a version of the TS algorithm that operates on channel SINR values is proposed, wherein SINR estimates are mapped to MCS values using a lookup table along similar lines as in OLLA.

b) Our Contributions: We propose Joint-TS, a version of classical TS that exploits the monotonicity property of success probabilities across MCS values (i.e., lower MCS yields higher success probability). Unlike the independent Beta distributions employed in classical TS, Joint-TS utilizes a multivariate ordered Beta (MOB) prior [16]. The advantages of employing MOB distribution are: (a) the samples satisfy the monotonicity constraint, and (b) the posterior remains an MOB distribution. To address the challenge of sampling from the MOB distribution, we propose a method based on Gibbs sampling, a popular MCMC technique.

Using the pyitpp library [17], we demonstrate that while state-of-the-art unimodal algorithms (like UTS [13]) perform well with perfect CQI, their performance degrades significantly when CQI feedback is unavailable. Joint-TS maintains stable throughput in these deprived conditions. Furthermore, we highlight the limitations of Latent-TS (LTS) [14], which relies heavily on a pre-existing lookup table. While LTS performs well ideally, it is vulnerable in high Doppler environments and practical setups with inaccurate tables. Joint-TS operates independently of such tables while matching or exceeding LTS performance in high-mobility states.

The idea of exploiting the monotonicity property of the success probabilities appears in [18], wherein the authors propose Constrained TS (CoTS), an algorithm in which sampling from an MOB prior is achieved using a sequential inverse transform sampling technique. However, their technique results in samples that are not, in fact, jointly MOB distributed; for a more detailed discussion, see Sec. IV-B. Additionally, while [18] focuses on regret analysis for CoTS, our focus is on the performance of Joint-TS for Doppler scenarios.

Notations: For any positive integer K, let $[K] := \{1, 2, \cdots, K\}$. Given a vector $\underline{\Theta} = (\Theta_1, \ldots, \Theta_K)$ of length K, $\underline{\Theta}_{\sim i}$ denotes the vector $(\Theta_1, \cdots, \Theta_{i-1}, \Theta_{i+1}, \cdots, \Theta_K)$ of all elements in vector $\underline{\Theta}$ except Θ_i. We employ the following abbreviations: (a) i.i.d: independent and identically distributed, (b) PDF: probability density function, (c) CDF: cumulative distribution function, and (d) TTI: transmission time interval.

II. The MAB Problem Formulation

The MCS selection problem in link adaptation can be formulated as a multi-armed bandit (MAB) problem [19]. While we assume the channel state (SINR) to be fixed but unknown in our problem description, this formulation extends to time-varying channels (i.e., Doppler scenarios), as outlined in Sec. V. Let $[K]$ denote the set of K supported MCS levels, each representing a distinct arm. At every discrete time step $t = 1, 2, \ldots, T$, the transmitter selects an MCS arm $I_t \in [K]$ and obtains the reward $R_{I_t,t} = s_{I_t} X_{I_t,t}$. Here, s_{I_t} is the known spectral efficiency of MCS I_t, and $X_{I_t,t} \sim \mathrm{Bernoulli}(\theta_{I_t})$ represents the transmission outcome (1 for success, 0 for failure), governed by the unknown success probability θ_{I_t}. The spectral efficiencies are known and satisfy the relation $s_1 < s_2 < \cdots < s_K$, while the success probabilities are unknown but are known to be monotonically decreasing, satisfying $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_K$.

The goal of the transmitter is to maximize the expected cumulative *reward* (throughput) over the time horizon T. This requires balancing the exploration of untested MCS values with the exploitation of the empirically best-performing ones via sequential interaction with the environment.

We propose Joint-TS, a version of the classical TS algorithm, that exploits and retains the monotonicity property of the success probabilities throughout the entire time horizon T.

III. The Joint-Thompson Sampling Algorithm

In the classical TS algorithm [19], the unknown success probabilities are treated as random variables $\Theta_1, \ldots, \Theta_K$, with an initial prior specified by the product Beta distribution: $\Theta_i \sim \mathrm{Beta}(\alpha_i, \beta_i)$ independent of $\underline{\Theta}_{\sim i}$ for each $i \in [K]$. Here, $\alpha_1, \ldots, \alpha_K$ and $\beta_1, \ldots, \beta_K$ are fixed constants. At time step t, the TS algorithm selects arm $I_t = \mathrm{argmax}_{i \in [K]} s_i \Theta_i$. Upon obtaining reward $R_t = r$ from arm I_t, the parameters $(\alpha_{I_t}, \beta_{I_t})$ of arm I_t are updated as in (3) (with $I_t = i$) while leaving the parameters of other arms unchanged. The independence of $\Theta_1, \ldots, \Theta_K$ assumed at the start of TS algorithm clearly fails to capture the monotonicity in the success probabilities of the arms.

To preserve the monotonicity relation $\Theta_1 \geq \Theta_2 \geq \cdots \geq \Theta_K$ at every time step, we propose the Joint-TS algorithm in which $\Theta_1, \ldots, \Theta_K$ are sampled from a Multivariate Ordered Beta (MOB) distribution, introduced in [16]. Sampling from the MOB distribution inherently guarantees that the resulting samples satisfy the desired monotonicity constraint. The pseudocode for Joint-TS is presented in Algorithm 1, and its key facets are described below.

A. Multivariate Ordered Beta (MOB) Distribution

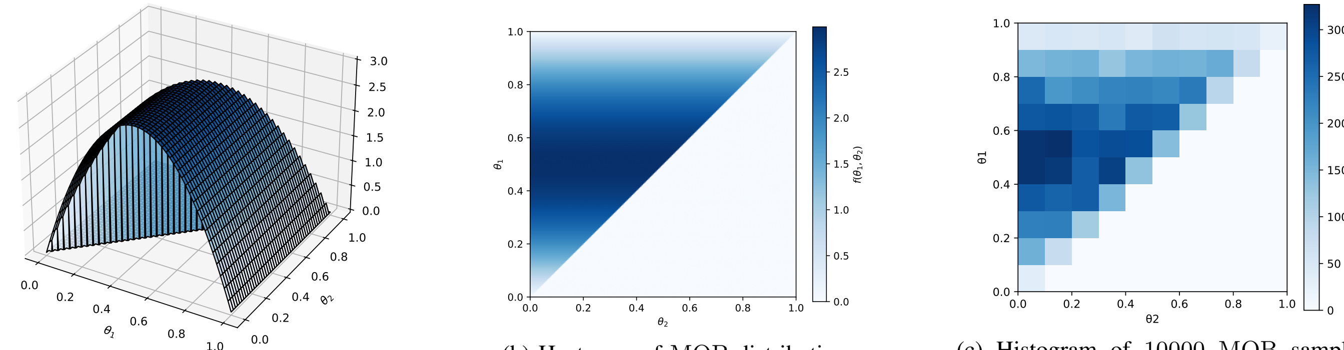
Fix $\underline{\alpha} = (\alpha_1, \cdots, \alpha_K)$ and $\underline{\beta} = (\beta_1, \cdots, \beta_K)$ with $\alpha_i, \beta_i > 0$ for all $i \in [K]$. We say that $\underline{\Theta} = (\Theta_1, \ldots, \Theta_K)$ follows MOB distribution with parameters $(\underline{\alpha}, \underline{\beta})$, denoted by $\mathrm{MOB}(\underline{\alpha}, \underline{\beta})$, if the joint PDF of $\Theta_1, \ldots, \Theta_K$ is given by

$$f_{\underline{\Theta}}(\underline{\theta}) = B(\underline{\alpha}, \underline{\beta})^{-1} \prod_{i=1}^{K} \theta_i^{\alpha_i - 1} (1 - \theta_i)^{\beta_i - 1}, \quad \underline{\theta} \in S, \tag{1}$$

where $\underline{\theta} = (\theta_1, \cdots, \theta_K)$, and

$$S := \{\underline{\theta} \in [0, 1]^K \mid \theta_1 \geq \theta_2 \geq \cdots \geq \theta_K\} \tag{2}$$

is the set of vectors that satisfy the monotonicity constraint, and $B(\underline{\alpha}, \underline{\beta})$ is the normalization constant.



(a) Joint PDF of the MOB distribution.

(b) Heat map of MOB distribution.

(c) Histogram of 10000 MOB samples generated via Gibbs sampling approach.

Fig. 1: Illustrations of the PDF of $\mathrm{MOB}((2,1),(2,1))$ distribution and the histogram samples collected using Gibbs sampling. The PDF is nonzero in the triangle $0 \leq \theta_2 \leq \theta_1 \leq 1$ and shows the effect of the ordering constraint.

B. Parameter Updates in the Joint-TS Algorithm

Let $X_{i,t}$ denote the reward from arm i at time t. Then,

$$P(X_{i,t} = x \mid \Theta_i = \theta) \;=\; \theta^x(1-\theta)^{1-x}, \quad x \in \{0,1\}.$$

On playing arm i and observing reward $X_{i,t} = r \in \{0,1\}$, the posterior on the parameters $\underline{\Theta} = (\Theta_1, \cdots, \Theta_K)$ is given by:

$$\begin{aligned} f_{\underline{\Theta}|X_{i,t}}(\underline{\theta} \mid r) &\propto f_{\Theta}(\underline{\theta}) \cdot P(X_{i,t} = r \mid \Theta_i = \theta) \\ &\propto \left(\prod_{j=1}^{K} \theta_j^{\alpha_j - 1}(1-\theta_j)^{\beta_j - 1} \right) \theta_i^r (1-\theta_i)^{1-r}, \quad \underline{\theta} \in S \end{aligned}$$

Clearly, (3) represents an MOB distribution with parameters

$$(\alpha_i, \beta_i) \leftarrow (\alpha_i + r,\;\; \beta_i + 1 - r), \quad (\alpha_j, \beta_j) \leftarrow (\alpha_j, \beta_j)\; \forall j \neq i. \tag{3}$$

Though the parameters of arms $j \neq i$ remain unchanged, the reward from arm i implicitly acts as feedback for the other arms due to the monotonicity constraint.

Unlike classical TS, which samples independent Beta distributions, Joint-TS, as seen in Algorithm 1, samples the vector $(\Theta_1, \ldots, \Theta_K)$ jointly at each step to inherently satisfy $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_K$. This implicit correlation makes direct sampling from the MOB distribution a non-trivial task, motivating our Gibbs sampling approach outlined below.

Algorithm 1 Joint-Thompson Sampling

```
Initialize: For all i ∈ [K], choose α_i > 0, β_i > 0.
for t = 1, 2, ..., T do
    Sample (Θ_1, ..., Θ_K) ~ MOB(α, β) (Algorithm 2).
    Select arm I_t = arg max_i Θ_i s_i
    Observe reward R_t = r from arm I_t
    (α_{I_t}, β_{I_t}) ← (α_{I_t} + r, β_{I_t} + 1 − r)
end for
```

Remark 1: If $\alpha_i = a$ and $\beta_i = b$ for all $i \in [K]$, for some $a, b > 0$, MOB samples can be generated by sorting K i.i.d $\mathrm{Beta}(a,b)$ samples. This idea does not extend to general MOB distributions where α_i and β_i can take any values > 0.

IV. Sampling From MOB Distribution

We use the Gibbs sampling technique [20] to sample from MOB distribution. Gibbs Sampling is an MCMC technique to sample from complex joint distributions that are difficult to directly sample from, but whose conditional distributions are easy to sample from. In what follows, we first define a *restricted Beta distribution*, and subsequently demonstrate that if a random vector is MOB distributed, then the conditional distributions follow restricted Beta distribution.

Definition 1: A random variable Y is said to follow *restricted Beta distribution* with parameters (α, β, ℓ, u), denoted by $Y \sim \mathrm{ResBeta}(\alpha, \beta, \ell, u)$, if the PDF of Y is given by

$$f_Y(y) = \begin{cases} \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta,\ell,u)}, & y \in [\ell, u], \\ 0, & \text{otherwise}, \end{cases} \tag{4}$$

where it is assumed that $\alpha, \beta > 0$, $0 \leq \ell < u \leq 1$, and $B(\alpha, \beta, \ell, u)$ is the normalization factor.
If $\ell = 0$ and $u = 1$, then $\mathrm{ResBeta}(\alpha, \beta, 0, 1) = \mathrm{Beta}(\alpha, \beta)$.

A. Sampling from Restricted Beta Distribution

Let $Y \sim \mathrm{ResBeta}(a, b, \ell, u)$ and let $X \sim \mathrm{Beta}(a, b)$. It follows that the CDF of Y is given by

$$F_Y(y) \;=\; \frac{F_X(y) - F_X(\ell)}{F_X(u) - F_X(\ell)}, \qquad y \in [\ell, u].$$

The well-known inverse transform sampling method may be used to sample from restricted Beta distribution.

B. Conditional Distribution $f_{\Theta_i|\underline{\Theta}_{\sim i}}$ under MOB

We now proceed to show that if a random vector is MOB distributed, then the conditional PDF of each random variable in the vector, conditioned over the remaining random variables, follows restricted Beta distribution.

Lemma 1: If $\underline{\Theta} \sim \mathrm{MOB}(\underline{\alpha}, \underline{\beta})$, then

$$\Theta_i \mid \{\underline{\Theta}_{\sim i} = \underline{\theta}_{\sim i}\} \sim \mathrm{ResBeta}(\alpha_i, \beta_i, \theta_{i+1}, \theta_{i-1}),$$

where by convention, $\theta_0 = 1$, $\theta_{K+1} = 0$.
Proof: Follows directly from the joint PDF in (1). $\square$

C. Gibbs Sampling from MOB Distribution

Fix $N \in \mathbb{N}$ and $(\underline{\alpha}, \underline{\beta})$ with $\alpha_i > 0$ and $\beta_i > 0$ for all $i \in [K]$. Let $\underline{\Theta}_1, \underline{\Theta}_2, \ldots$ be sampled as per the Gibbs sampling procedure outlined in Algorithm 2.

Algorithm 2 Gibbs Sampling from $\mathrm{MOB}(\underline{\alpha}, \underline{\beta})$ Distribution

Input: $(\underline{\alpha}, \underline{\beta})$ with $\alpha_i > 0$ and $\beta_i > 0$ for all $i \in [K]$, number of iterations $N \in \mathbb{N}$.
Generate $U_1, \ldots, U_K \overset{\text{i.i.d.}}{\sim} \mathrm{Uniform}([0,1])$, and order the samples samples in descending order. Let the ordered samples be $\Theta_{1,1}, \cdots, \Theta_{1,K}$, with $\Theta_{1,1} \geq \Theta_{1,2} \geq \cdots \geq \Theta_{1,K}$. Let $\underline{\Theta}_1 = (\Theta_{1,1}, \cdots, \Theta_{1,K})$.
for $t = 2, \ldots, N$ **do**
    for $i = 1, \ldots, K$ **do**
        Sample $\Theta_{t,i} \sim \mathrm{ResBeta}(\alpha_i, \beta_i, \Theta_{t-1,i+1}, \Theta_{t,i-1})$ for all $i \in [K]$; by convention, $\Theta_{t,0} = 1$ and $\Theta_{t,K+1} = 0$.
    end for
    Set $\underline{\Theta}_t = (\Theta_{t,1}, \ldots, \Theta_{t,K})$.
end for
Output: $\underline{\Theta}_N = (\Theta_{N,1}, \ldots, \Theta_{N,K})$.

Lemma 2: The sequence of random vectors $\underline{\Theta}_1, \underline{\Theta}_2, \cdots$ obtained via Algorithm 2 forms an S-valued Markov chain (where S is as defined in (2)). Furthermore,

$$\lim_{t \to \infty} \|\mathcal{L}(\underline{\Theta}_t \mid \underline{\Theta}_1 = \underline{\theta}) - \mathrm{MOB}(\underline{\alpha}, \underline{\beta})\|_{\mathrm{TV}} = 0 \qquad \forall\, \underline{\theta} \in S,$$

where $\mathcal{L}(\underline{\Theta}_t \mid \underline{\Theta}_1 = \underline{\theta})$ denotes the conditional law of $\underline{\Theta}_t$, conditioned on starting the Gibbs sampling in state $\underline{\Theta}_1 = \underline{\theta}$, and $\|\cdot\|_{\mathrm{TV}}$ denotes the total variation distance.
We skip the proof of this Lemma. It follows from the fact that the Markov chain $\{\underline{\Theta}_1, \underline{\Theta}_2, \ldots\}$ can be shown to be positive Harris-recurrent [21], [22, Chapter 13].

Lemma 2 assures that for N sufficiently large, the joint PDF of the samples output by Algorithm 2 is close to the desired $\mathrm{MOB}(\underline{\alpha}, \underline{\beta})$ distribution in total variation. In Joint-TS, we run Algorithm 2 with $N = 1000$ at each time step. For the case with $K = 2$ and $\underline{\alpha} = (2, 1)$, $\underline{\beta} = (2, 1)$, it can be seen in Fig. 1c that the 2D histogram of the samples generated via Algorithm 2 matches with the PDF of MOB distribution.

Remark 2: The sequential inverse transform sampling method in [18] suggests sampling $\Theta_1 \sim \mathrm{Beta}(\alpha_1, \beta_1)$ and $\Theta_i \mid \underline{\Theta}_{\sim i} \sim \mathrm{ResBeta}(\alpha_2, \beta_2, 0, \Theta_{i-1})$ for all $i > 1$. This fundamentally differs from the true conditional distributions of the MOB (see Sec. IV-B). For example, for $K = 2$, $\underline{\alpha} = (1000, 1000)$ and $\underline{\beta} = (1000, 1)$, the MOB joint PDF is concentrated around $(1, 1)$, whereas the method in [18] falsely concentrates Θ_1 around 0.5.

V. Simulation Setup and Results

The simulations are carried out using the Python itpp repository developed by Saxena et. al [17] that provides libraries to simulate the fading and Doppler models. We consider $K = 29$ MCS options available in 5G NR (Table I) and evaluate Joint-TS against OLLA, Modified Thompson Sampling (MTS) [11], Unimodal Thompson Sampling (UTS) [13], and Latent-TS (LTS) [14]. Note that LTS requires a pre-existing mapping table to correlate latent states with MCS success probabilities. While we assume this table is perfectly accurate in our simulations, relying on static lookup tables in real-world deployments can cause sub-optimal performance when actual channel conditions deviate from the modeled mapping.

a) Static Channel Simulations: We initialize $\alpha_i = \beta_i = 1$ for all $i \in [K]$ to run the TS and Joint-TS algorithms at three different SINR values: -4dB, 10dB, 20dB. Averaged over 20 independent trials, stable conditions allow OLLA, UTS, and LTS to exhibit the highest aggregate throughput. Joint-TS demonstrates competitive performance (e.g., achieving 2.54 bps/Hz at 10 dB SINR compared to LTS's 2.68 bps/Hz).

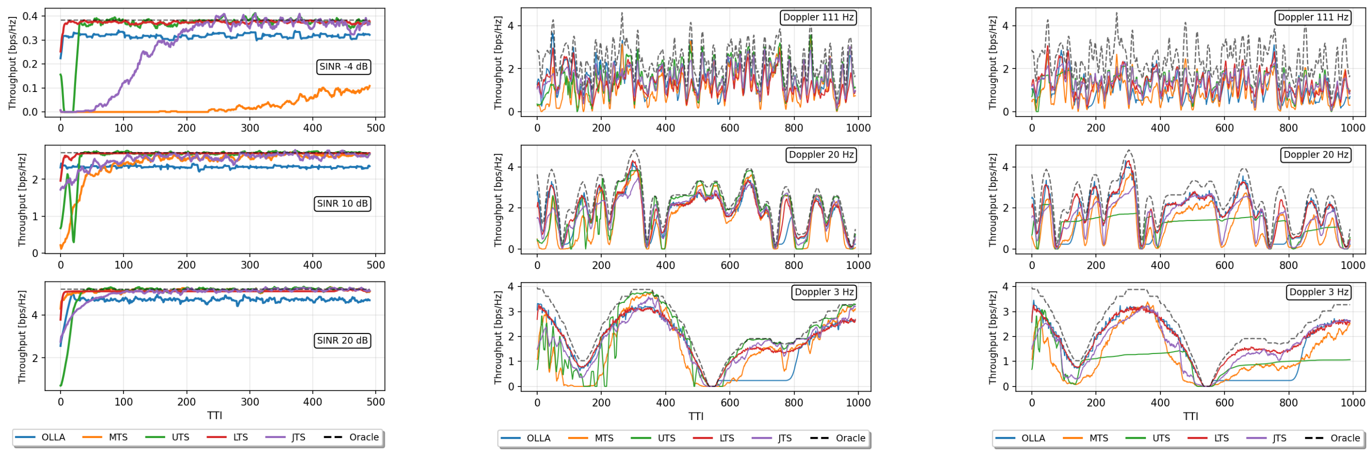
b) Doppler Channel: We present simulations for the Doppler setting in Fig. 2b and 2c that correspond to scenarios with perfect CQI and CQI-less settings. We assume unconstrained optimization to find the optimal arm for both these simulations. Three cases of Doppler shifts, $F_d = 3, 20, 111$ Hz, are considered at average SINR 10dB. This corresponds to coherence times of $\approx 60, 9, 1.6$ ms respectively, using the relation $T_{\text{coh}} = \frac{9}{16\pi F_d}$. We have assumed a single-path channel model here.

CQI setting	Doppler Shift (Hz)	Throughput (in bps/Hz)				
		OLLA	MTS	UTS	LTS	Joint-TS
Perfect CQI	3	1.64	1.53	1.89	1.89	1.80
	20	1.81	1.56	1.89	1.87	1.80
	111	1.61	1.24	1.56	1.29	1.56
No CQI	3	1.60	1.24	1.04	1.88	1.63
	20	1.72	1.13	1.21	1.87	1.49
	111	1.06	0.96	1.34	1.29	1.32

TABLE II: Average throughput across the 1000 TTIs at average SINR of 10db.

Perfect CQI setting: Assuming error and delay free CQI is available every TTI (500 μs), we run 20 parallel MAB instances. Note that for SINRs optimally mapped to $18, 19, 20$ are grouped to CQI mapped to MCS 18, as in [4, 5.2.2.1], due to MCS quantization. Therefore, even though the CQI is assumed available, each MAB may observe a range of SINRs over which single MCS need not be optimal. While UTS and LTS achieve the highest throughput at low/mid Doppler shifts, LTS degrades significantly at high Doppler shifts, whereas UTS and Joint-TS successfully maintain high throughput.

CQI-less settings: When CQI feedback is unavailable, a single MAB instance of a smoothed MTS/UTS/Joint-TS variant is used. The parameter update is modified to: $(\alpha_{I_t}, \beta_{I_t}) \leftarrow (\alpha_{I_t}, \beta_{I_t})\, e^{-\frac{\Delta t}{w}} + (r,\, 1-r)$, where Δt denotes the time interval since arm I_t was last explored, and w is a smoothing parameter [15]. A higher value of w corresponds to longer memory and slower adaptation to change, while a lower value of w results in faster response. Fig. 2c shows results for $w = 50$. The parameter w can be adapted based on Doppler estimates to achieve optimal performance. A detailed analysis of this is left for future work. We use the smoothening method proposed in [14] for LTS. In this setting, UTS struggles without CQI, while LTS maintains high throughput in low-mid Doppler settings.



(a) Static channel setting

(b) Perfect CQI at average SINR 10dB

(c) Without CQI at average SINR 10dB

Fig. 2: Throughput of OLLA, MTS, UTS, LTS, Joint-TS algorithms averaged over 20 samples. Oracle in the last 2 figures implies correspond to best possible throughput for the channel conditions. It also indicates the effect of Doppler on the SINRs.

Joint-TS demonstrates consistently robust performance across all conditions.

VI. Conclusion

Joint-TS exploits the monotonicity property of success probabilities of MCS values by generating samples from MOB distribution, thereby resulting in faster convergence in low to mid-range SINRs and throughput gains in both static and Doppler settings with and without CQI. Future work will investigate the performance of Joint-TS in scenarios with periodic (infrequent) CQI feedback and under conditions of noisy or imperfect CQI estimation.

References

[1] K. Brueninghaus, D. Astely, T. Salzer, S. Visuri, A. Alexiou, S. Karger, and G.-A. Seraji, "Link Performance Models for System Level Simulations of Broadband Radio Access Systems," in *IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, vol. 4, 2005, pp. 2306–2311 Vol. 4.

[2] P. Wu and N. Jindal, "Coding Versus ARQ in Fading Channels: How Reliable Should the PHY Be?" in *IEEE Global Communications Conference (GLOBECOM)*, 2009, pp. 1–6.

[3] S. Park, R. C. Daniels, and R. W. Heath, "Optimizing the Target Error Rate for Link Adaptation," in *IEEE Global Communications Conference (GLOBECOM)*, 2015, pp. 1–6.

[4] 3rd Generation Partnership Project (3GPP), "NR; Physical layer procedures for data," 3GPP, Technical Specification TS 38.214, 2018, release 15.

[5] D. Paranchych and M. Yavuz, "A Method for Outer Loop Rate Control in High Data Rate Wireless Networks," in *Proceedings IEEE 56th Vehicular Technology Conference (VTC)*, vol. 3, 2002, pp. 1701–1705 vol.3.

[6] K. I. Pedersen, G. Monghal, I. Z. Kovacs, T. E. Kolding, A. Pokhariyal, F. Frederiksen, and P. Mogensen, "Frequency Domain Scheduling for OFDMA with Limited and Noisy Channel Feedback," in *IEEE 66th Vehicular Technology Conference (VTC)*, 2007, pp. 1792–1796.

[7] V. Buenestado, J. M. Ruiz-Avilés, M. Toril, S. Luna-Ramírez, and A. Mendo, "Analysis of Throughput Performance Statistics for Benchmarking LTE Networks," *IEEE Communications Letters*, vol. 18, no. 9, pp. 1607–1610, 2014.

[8] A. Durán, M. Toril, F. Ruiz, and A. Mendo, "Self-Optimization Algorithm for Outer Loop Link Adaptation in LTE," *IEEE Communications Letters*, vol. 19, no. 11, pp. 2005–2008, 2015.

[9] R. Combes, J. Ok, A. Proutiere, D. Yun, and Y. Yi, "Optimal Rate Sampling in 802.11 Systems: Theory, Design, and Implementation," *IEEE Transactions on Mobile Computing*, vol. 18, no. 5, pp. 1145–1158, 2019.

[10] V. Saxena, J. Jaldén, J. E. Gonzalez, M. Bengtsson, H. Tullberg, and I. Stoica, "Contextual Multi-Armed Bandits for Link Adaptation in Cellular Networks," in *Proceedings of the 2019 Workshop on Network Meets AI & ML*, ser. NetAI'19. New York, NY, USA: Association for Computing Machinery, 2019, p. 44–49.

[11] H. Gupta, A. Eryilmaz, and R. Srikant, "Low-Complexity, Low-Regret Link Rate Selection in Rapidly-Varying Wireless Channels," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, 2018, pp. 540–548.

[12] V. Saxena and J. Jaldén, "Bayesian Link Adaptation under a BLER Target," in *IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2020, pp. 1–5.

[13] S. Paladino, F. Trovò, M. Restelli, and N. Gatti, "Unimodal thompson sampling for graph-structured arms," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, ser. AAAI'17. AAAI Press, 2017, p. 2457–2463.

[14] V. Saxena, H. Tullberg, and J. Jaldén, "Reinforcement Learning for Efficient and Tuning-Free Link Adaptation," *IEEE Transactions on Wireless Communications*, vol. 21, no. 2, pp. 768–780, 2022.

[15] A. Krotov, A. Kiryanov, and E. Khorov, "Rate Control With Spatial Reuse for Wi-Fi 6 Dense Deployments," *IEEE Access*, vol. 8, pp. 168 898–168 909, 2020.

[16] M. Al-Saidi, A. Kuznetsov, and M. Nediak, "On Ordered Beta Distribution and the Generalized Incomplete Beta Function," *Method. Comput. Appl. Prob.*, vol. 27, no. 1, Dec. 2024.

[17] V. Saxena, "py-itpp," https://github.com/vidits-kth/py-itpp, accessed: 2025-09-13.

[18] H. Gupta, A. Eryilmaz, and R. Srikant, "Link Rate Selection using Constrained Thompson Sampling," in *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*, 2019, pp. 739–747.

[19] W. R. Thompson, "On the Likelihood That One Unknown Probability Exceeds Another in View of the Evidence of Two Samples," *Biometrika*, vol. 25, no. 3/4, pp. 285–294, 1933.

[20] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 721–741, 1984.

[21] L. Tierney, "Markov Chains for Exploring Posterior Distributions," *the Annals of Statistics*, pp. 1701–1728, 1994.

[22] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*. Springer Science & Business Media, 2012.