

# DYNAMIC RESOLUTION SWITCHING FOR LIVE STREAMING

*Xin Xiong, Yixu Chen, Hai Wei, Yongjun Wu, Sriram Sethuraman*

Amazon Prime Video, Seattle, USA

## ABSTRACT

Conventional adaptive bitrate (ABR) streaming systems typically rely on static bitrate ladders to optimize Quality of Experience (QoE). While operationally simple, this “one-size-fits-all” approach neglects content-specific characteristics, often compromising streaming efficiency. Per-title optimization methods address this by predicting the rate-distortion convex hull directly from the source content, but their reliance on pre-encoding source analysis can limit their applicability to live streaming. Moreover, the objective video quality metrics (VQMs) they rely on are optimized for overall correlation with subjective scores rather than cross-over accuracy, often yielding inaccurate cross-over predictions and suboptimal ladder construction. To overcome both limitations, we introduce a Dynamic Resolution Switching (DRS) framework for live streaming that remains fully compatible with existing streaming protocols. Our approach augments static ladders with strategically selected representations guided by user bandwidth distributions and cross-over regions. The quality of these representations is then analyzed in real time to construct dynamic ladders. Central to this framework is a lightweight, bitstream-based VQM that ensures computational efficiency while maximizing the accuracy of subjective resolution cross-over prediction through training on Pairwise Comparison (PC) datasets. At each bitrate, the VQM evaluates all candidate representations to identify the resolution maximizing the quality score. This decision process, operating at a configurable granularity (e.g., per segment), drives the dynamic resolution switching mechanism specifically optimized for the metric. Experimental results validate the approach, demonstrating a significant performance gain (approximately 9% Bjøntegaard Delta rate reduction under the proposed VQM) while maintaining practical feasibility for live streaming.

**Index Terms**— Adaptive Bitrate Streaming, Video Quality Metric, Quality of Experience

## 1. INTRODUCTION

Video streaming has become an integral part of our daily lives. Due to the varying network bandwidth and device types among end-users, service providers widely implement adaptive bitrate (ABR) methods [1] to enhance the quality of experience (QoE). The cornerstone of an ABR implementation is the bitrate ladder, a set of bitrate-resolution pairs (representations) selected dynamically to accommodate varying bandwidth. The conventional approach to bitrate ladder construction involves the implementation of a static or “one-size-fits-all” ladder, utilizing predetermined bitrate-resolution pairs across all content. However, this methodology fails to account for content-specific characteristics, potentially yielding suboptimal results. To address this limitation, per-title optimized bitrate ladder construction has been proposed [2–4]. These methods exhaustively search the rate-distortion (RD) convex hull across pre-encoded copies to

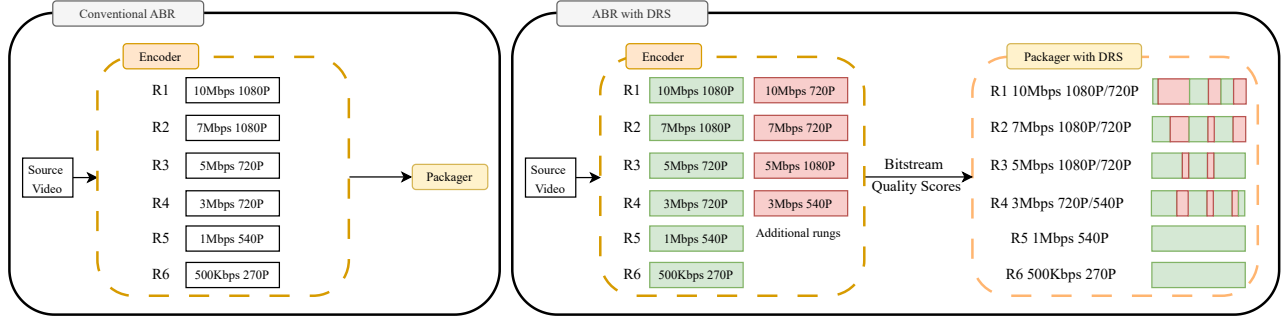
maximize QoE. However, the computational demands of these approaches make them impractical for live streaming applications.

Alternatively, some prior work has proposed to predict the convex hull directly from source content for ladder construction [5, 6]. For instance, [5] estimates resolution cross-over points, defined as the intersections of per-resolution RD curves, from content features in order to maximize QoE. While these approaches reduce the computational burden, they still require pre-encoding source analysis, limiting their applicability to live streaming. Moreover, they typically rely on objective VQMs (e.g., VMAF [7] and SSIM [8]) as proxies for ground-truth, primarily due to the labor-intensive nature of subjective studies. However, recent work indicates that VQMs, including VMAF and P1204.3 [9], often fail to accurately predict resolution cross-over points due to an inherent bias toward higher resolutions [10]. Additionally, conventional VQMs are trained on Absolute Category Rating (ACR) datasets. Cross-over points derived from ACR scores are suboptimal [11]; pairwise comparison (PC) datasets achieve better cross-over accuracy. Another common direction for content-adaptive encoding is variable bitrate (VBR) encoding [12], which saves bits during low-complexity scenes. However, VBR primarily optimizes bit allocation within a single representation rather than across representations, leaving the resolution selection problem unaddressed. Furthermore, while VBR reduces average bandwidth, peak CDN capacity must still be provisioned for the worst-case content complexity.

While prior cross-over-focused work [10, 11] has been primarily diagnostic, none has closed the loop from cross-over analysis to an operational real-time system. To bridge this gap, we propose a dynamic resolution switching (DRS) framework tailored for live streaming, transitioning cross-over insights from diagnosis to action. Furthermore, DRS achieves end-to-end bit savings from origin to edge to client and is complementary to both constant bitrate (CBR) and VBR encoding. Figure 1 illustrates the proposed DRS pipeline integrated within the ABR streaming system. In contrast to conventional ABR approaches, we augment the standard static ladder with a data-driven set of additional representations, to better utilize the unused capacity of the encoding server, especially for low-complexity codecs such as AVC [13]. During encoding, the encoder generates an expanded set of representations, each embedded with a computed quality score. The packager then dynamically filters these streams, selecting the resolution with the highest quality score for each bitrate. This filtering reduces the ladder size back to its standard static dimensions before delivery, thereby making our method compatible with existing live streaming protocols. The efficacy of this quality-driven resolution selection mechanism hinges on a VQM that balances computational efficiency with high accuracy in predicting subjective resolution cross-over points. Although standardized bitstream-based metrics like P1204.3 are computationally efficient, they often lack the accuracy required to reliably predict subjective resolution cross-over points. To address this limitation, we adopt the Encoder-Quantization-Motion (EQM) model [10] and

---

Xin’s work was done during an internship at Amazon Prime Video.



**Fig. 1:** The proposed dynamic resolution switching pipeline in live streaming. Compared to conventional ABR systems, our approach introduces additional encoded representations and utilizes a bitstream-based VQM to predict quality scores, which are then used for adaptive resolution selection at a configurable granularity (e.g., per segment) to maximize QoE.

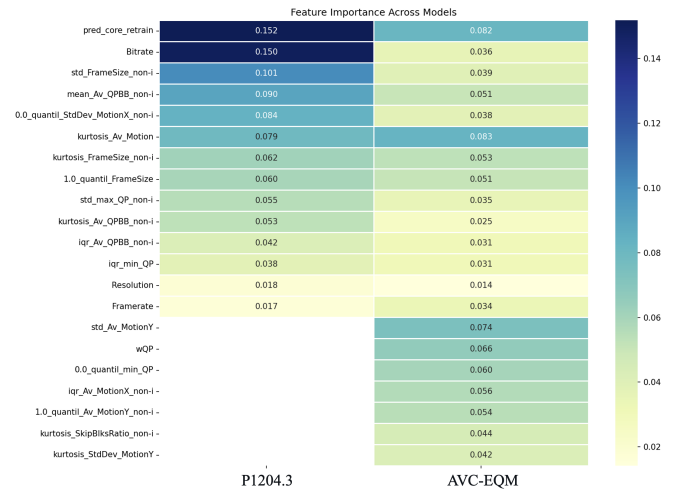
recalibrate it on the PC dataset from [11] for improved cross-over accuracy. The primary contributions of this work are summarized as follows:

- A DRS pipeline that is compatible with existing streaming protocols, which augments static bitrate ladders with strategically selected additional representations and dynamically selects the optimal resolution per segment to maximize QoE.
- AVC-EQM, an efficient bitstream-based VQM tailored for the DRS pipeline, with minimal overhead and a high correlation with subjective resolution cross-over points. We also introduce new evaluation measures that align cross-over assessment with the operational objective of DRS.
- Experiments on real-world sports videos demonstrate that our proposed DRS pipeline significantly improves QoE. Compared to a baseline static ladder, we achieve approximately 9% Bjøntegaard Delta rate (BD-rate) savings and a 0.15 gain in BD-AVC-EQM (on a 0–10 scale).

## 2. AVC-EQM MODEL

### 2.1. Feature selection

Although EQM [10] was initially developed for HEVC [14], the strict real-time constraints of live streaming often limit the feasibility of HEVC in high-density encoding scenarios. In contrast, AVC remains a more practical choice for our DRS pipeline due to its lower encoding complexity. In this work, we adapt the EQM model [10] from HEVC to AVC, and propose a specialized AVC-EQM. Given the inherent differences between the two codecs, such as hierarchical block partitioning structures and motion vector prediction mechanisms, we first identify the subset of EQM features in [10] that are applicable to AVC. We then employ Greedy Feature Selection (GFS) [15] on the Live Sport Cross-Over (LSCO) dataset [11], whose quality labels are generated via PC to ensure a better correlation with subjective resolution cross-over, in order to determine the optimal feature set for the AVC bitstream. The selected features are shown in Figure 2. The features that are exclusive to AVC-EQM exhibit significant importance, suggesting that they were likely overlooked by P1204.3, since its feature set was neither designed nor trained on datasets specifically targeting subjective resolution cross-over. To validate the effectiveness of AVC-EQM, we evaluate it from two perspectives: (i) its correlation with subjective resolution cross-over points, and (ii) its runtime complexity.



**Fig. 2:** Feature importance comparison (random forest residual model): retrained P1204.3 vs. AVC-EQM on the LSCO-AVC dataset. AVC-EQM leverages a broader set of significant features.

### 2.2. Benchmark on cross-over dataset

To assess the correlation between objective VQMs and subjective cross-over points, Zhu *et al.* [11] introduced the Resolution Cross-over Quality Loss (RCQL) to quantify the quality loss caused by cross-over prediction errors. RCQL employs the Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) [16] to interpolate the Rate-Distortion (RD) curves, identifying ground-truth subjective cross-over points at the curve intersections. While PCHIP is standard for calculating the BD-rate, it has limitations when handling noisy subjective quality labels recovered by Bradley-Terry or PWCMP models [17], which often exhibit non-monotonic behavior with respect to bitrate [18]. In such scenarios, PCHIP often yields unreliable results, potentially producing multiple ambiguous intersections because of the violation of monotonicity.

To address this limitation, we adopt a more robust alternative using a 4-parameter logistic function for fitting the ground-truth RD curve:

$$f(x) = \beta_2 + \frac{\beta_1 - \beta_2}{1 + \exp\left(-\frac{x - \beta_3}{|\beta_4|}\right)} \quad (1)$$

where  $\beta_3$  represents the horizontal inflection point. To ensure

**Table 1:** RCQL benchmark results on the LSCO-AVC dataset. P1204.3 and AVC-EQM are bitstream-based metrics, whereas the other VQMs are computed by upscaling decoded frames to the source resolution for full-referenced quality evaluation.

	Out-of-box								Retrained	
		PSNR <sub>y</sub>	SSIM	MS-SSIM	VMAF	VMAF 4k	VMAF 4k neg	P1204.3	P1204.3*	AVC-EQM
Correlation $\uparrow$	SROCC	0.728	0.815	0.812	0.826	0.822	0.824	0.895	0.912	<b>0.926</b>
	PLCC	0.684	0.741	0.781	0.832	0.825	0.827	0.895	0.906	<b>0.925</b>
$\Delta$ Bitrate (Kbps) $\downarrow$	1080p vs 720p	<b>862.66</b>	8153.23	4078.12	1193.89	1131.49	1117.85	1512.71	1404.45	1195.27
	720p vs 540p	<b>254.9</b>	4191.05	783.58	377.51	371.39	366.8	644.11	465.1	484.64
$RCQL_s$ (JOD $\times$ Kbps) $\downarrow$	1080p vs 720p	<b>178.49</b>	4825.25	1701.84	293.41	275.68	267.03	758.38	458.98	402.02
	720p vs 540p	<b>51.33</b>	3900.03	349.07	150.92	136.34	135.73	336.31	158.97	171.67
$RCQL_{avg}$ (JOD) $\downarrow$	1080p vs 720p	<b>0.1674</b>	0.5694	0.3299	0.1988	0.2022	0.2006	0.254	0.2854	0.2722
	720p vs 540p	<b>0.1564</b>	0.7573	0.3103	0.2037	0.188	0.1871	0.3493	0.2695	0.2873
Acc (%) $\uparrow$	1080p vs 720p	93.55	58.06	74.19	91.94	<b>95.16</b>	<b>95.16</b>	88.71	89.17	92.08
	720p vs 540p	<b>86.36</b>	62.12	80.3	83.33	<b>86.36</b>	<b>86.36</b>	80.3	84.17	79.17
QL (JOD) $\downarrow$	1080p vs 720p	0.214	0.383	0.361	0.658	0.768	0.768	0.287	0.172	<b>0.085</b>
	720p vs 540p	0.299	0.546	0.305	0.374	0.272	0.272	0.39	<b>0.159</b>	0.217

**Table 2:** RCQL benchmark results on the LSCO-HEVC dataset. EQM NR and EQM FR are adopted from [10]. HEVC-EQM denotes the EQM NR model retrained on the LSCO-HEVC dataset.

	Out-of-box									Retrained		
		PSNR <sub>y</sub>	SSIM	MS-SSIM	VMAF	VMAF 4k	VMAF 4k neg	P1204.3	EQM NR	EQM FR	P1204.3*	HEVC-EQM
Correlation $\uparrow$	SROCC	0.705	0.820	0.857	0.880	0.876	0.877	0.851	0.921	0.928	0.908	<b>0.943</b>
	PLCC	0.673	0.652	0.748	0.873	0.866	0.869	0.830	0.916	0.922	0.901	<b>0.943</b>
$\Delta$ Bitrate (Kbps) $\downarrow$	1080p vs 720p	<b>1277.47</b>	5479.27	2433.18	1515.0	1511.07	1377.85	1488.57	1595.54	2171.01	1519.58	1484.72
	720p vs 540p	560.69	2295.09	755.37	595.62	644.27	627.74	716.18	584.58	<b>544.54</b>	875.86	811.07
$RCQL_s$ (JOD $\times$ Kbps) $\downarrow$	1080p vs 720p	560.32	3581.81	1401.83	615.34	598.65	578.82	645.7	767.43	1440.03	<b>477.44</b>	556.65
	720p vs 540p	<b>112.27</b>	1843.95	255.25	115.58	131.22	123.82	184.58	131.76	120.34	312.88	216.91
$RCQL_{avg}$ (JOD) $\downarrow$	1080p vs 720p	0.2549	0.5956	0.3173	0.2596	0.2587	0.2515	<b>0.2182</b>	0.309	0.3858	0.2571	0.2508
	720p vs 540p	0.1193	0.4875	0.1942	0.0918	0.1039	<b>0.1012</b>	0.1995	0.1656	0.1441	0.2696	0.2646
Acc (%) $\uparrow$	1080p vs 720p	<b>87.5</b>	60.94	79.69	84.38	85.94	85.94	85.94	79.69	81.25	78.33	80.83
	720p vs 540p	80.3	68.18	80.3	<b>84.85</b>	<b>84.85</b>	<b>84.85</b>	78.79	75.76	81.82	76.25	76.25
QL (JOD) $\downarrow$	1080p vs 720p	0.249	0.517	0.29	0.294	0.321	0.321	0.39	0.341	0.365	0.18	<b>0.164</b>
	720p vs 540p	0.226	0.478	0.241	0.23	0.23	0.23	0.29	0.347	0.345	<b>0.176</b>	0.200

that the fitted curve models the saturation characteristic of RD curves, we constrain the optimization domain of  $\beta_3$  to the interval  $[r_{min}/2, r_{min}]$ , where  $r_{min}$  denotes the minimum bitrate present in each RD curve. Together with the inherent monotonicity of the logistic function, this enables more accurate cross-over determinations.

Leveraging the refined subjective cross-over points, we re-evaluated the performance of various VQMs. The results are summarized in Table 1 and Table 2. Following the RCQL evaluation in [11], three key metrics are reported: (1)  $\Delta$ Bitrate (Kbps) $\downarrow$  quantifies the absolute bitrate deviation between the subjective cross-over (determined from the fitted subjective RD curves) and the objective cross-over predicted by the VQM; (2)  $RCQL_s$  (JOD $\times$ Kbps) $\downarrow$  represents the cumulative quality loss, calculated as the integral area bounded by the two RD curves within the bitrate interval defined by the subjective and objective cross-over points; and (3)  $RCQL_{avg}$  (JOD) $\downarrow$  denotes the mean quality loss, calculated as the average perceptual loss in the cross-over divergence region.

While the RCQL metric introduced in [11] offers valuable insights into a VQM’s ability to locate exact resolution cross-over points, the primary objective of our DRS framework is slightly different. Rather than pinpointing the precise cross-over bitrate, our focus is on accurately identifying the optimal resolution at any given operating bitrate. To better align with this objective, we propose a complementary RCQL metric that quantifies the accuracy of rela-

tive quality ranking across resolutions. This methodology involves pairwise comparisons of quality scores between different resolutions (e.g., 720p vs. 1080p) at the same bitrate. For each pair, we evaluate the concordance between the objective VQM’s preference and the ground-truth subjective quality ranking. The resulting accuracy is denoted as  $Acc$  (%). In instances where the objective VQM incorrectly predicts the perceptual ranking (e.g., favoring 1080p when subjective scores indicate that 720p is superior), we quantify the magnitude of the subjective quality difference between the two resolutions. We then define  $QL$  (JOD) $\downarrow$  as the mean quality degradation averaged across all such incorrectly predicted events. The proposed complementary RCQL measures are particularly pertinent to the DRS framework, where decision-making relies directly on VQM scores. By quantifying the quality loss incurred by suboptimal resolution selection, they provide a practical evaluation aligned with operational DRS in live streaming.

Table 1 presents the RCQL benchmark on the LSCO-AVC dataset. P1204.3\* and AVC-EQM are trained and tested with 1000 runs of 5-fold cross-validation. The cross-validation procedure uses a content-based split, in which videos generated from the same source content are kept in the same group. For each RCQL metric, we first take the median (over the 1000 runs) for each source content and then average across all contents. In terms of overall correlation, AVC-EQM achieves the highest SROCC and PLCC, outperforming the other VQMs. While P1204.3\* shows a slight advantage on the

**Table 3:** Decoder side speed comparison on AVC bitstream. The proposed AVC-EQM is over 10 times faster than P1204.3.

Resolution	Bitrate(Mbps)	P1204.3(FPS)	EQM(FPS)
1080P	9	6.45	68.82
1080P	6	6.52	73.19
720P	2.2	8.99	99.13
720P	1.6	9.00	101.46
540P	1.2	10.30	111.84
540P	0.6	10.32	115.05

**Table 4:** Encoder side runtime complexity analysis. AVC-EQM introduces a small overhead of only 2% to 4%, making it suitable for live streaming applications.

Res.	Bit.	x264(s)	x264&EQM(s)	Overhead(%)
1080P	9	166.8	169.8	1.8
720P	1.6	74.3	76.8	3.3
540P	0.6	40.8	42.0	2.9

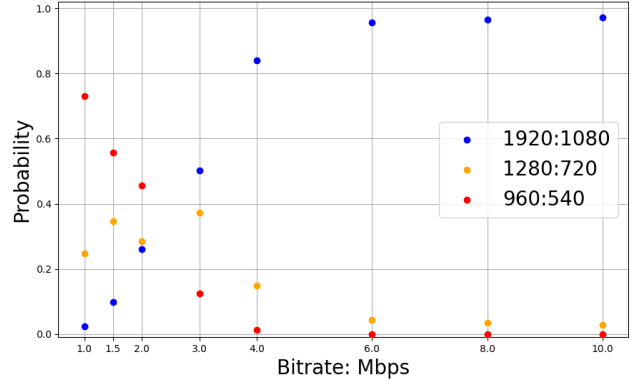
lower-resolution cross-over (720p vs. 540p), AVC-EQM is significantly better on the more critical high-resolution cross-over (1080p vs. 720p). Specifically, AVC-EQM reduces  $QL$  by approximately 50% relative to P1204.3\* (0.085 vs. 0.172). This indicates that AVC-EQM is particularly well suited to our proposed DRS pipeline. Furthermore, in light of the new subjective cross-over determination method and the new RCQL measures, we also re-benchmark the original HEVC bitstream-based EQM [10] on the LSCO-HEVC dataset (Table 2). Following a similar GFS process, the retrained HEVC-EQM generally surpasses the other bitstream-based metrics at identifying subjective cross-over points for 1080p vs. 720p.

### 2.3. Runtime complexity of AVC-EQM

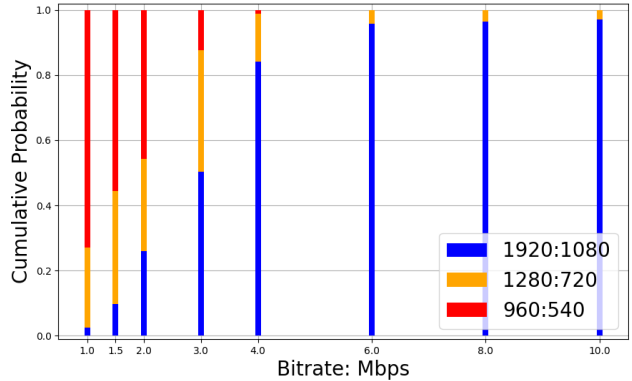
AVC-EQM can operate on both the encoder side (facilitating DRS in live streaming) and the decoder side (serving as a research prototype). We implement the decoder side AVC-EQM in libavcodec and the encoder side AVC-EQM in x264. For efficiency, the decoder side implementation supports frame-level multi-threading. Table 3 presents a speed comparison between P1204.3 and AVC-EQM (using 15 threads) for AVC bitstreams on an AWS EC2 c7a.48xlarge instance. Results are averaged over 15 source videos (60 fps, 8 seconds duration). Notably, for 1080p content, AVC-EQM is over 10 $\times$  faster than P1204.3. For the evaluation of encoder side integration, we use an 8-second 1080p SDR source video from the sports dataset on the same EC2 instance, with the x264 single-thread medium preset. As shown in Table 4, the additional computational overhead of AVC-EQM is small, approximately 2% to 4%, which validates its suitability for our proposed DRS pipeline.

## 3. EMPIRICAL EVALUATION

As the proposed DRS pipeline augments a static ladder with a small number of additional representations, we first describe how these representations are selected in subsection 3.1. We then evaluate the resulting performance in subsection 3.2.



(a) Probability of a resolution having the best score



(b) Probability stacked bar plot

**Fig. 3:** Statistical analysis of the resolution-quality relationship using AVC-EQM scores over 1800 GOPs of sports content. (a) Probability that each resolution yields the highest quality at a given bitrate. (b) Cumulative probability derived from (a), with resolutions distinguished by color.

### 3.1. GOP-level statistical analysis

This analysis uses a fixed Group of Pictures (GOP) size of 1 second, consistent with standard live streaming protocols. A 30-minute video sequence is encoded using an AVC codec across a diverse set of representations. Focusing on bitrates within the typical resolution switching range (see the *Rate* column in Table 5), the video is encoded at 540p, 720p, and 1080p. For each bitrate, we count the number of GOPs in which a specific resolution yields the highest quality as measured by AVC-EQM.

Figure 3(a) presents, for each bitrate, the probability that a given resolution achieves the highest VQM score. This probability is defined as the ratio of GOPs at which the resolution offers the best VQM score to the total number of GOPs (1800 for the 30-minute sequence). Figure 3(b) shows the corresponding cumulative probabilities. Taking 1080p as an example, at lower bitrates, only a few GOPs exhibit optimal quality at this resolution; however, as the bitrate increases, nearly all GOPs achieve best quality at 1080p. These findings suggest the design of augmented bitrate ladders for DRS: selecting the resolutions with a higher probability of optimality at each bitrate maximizes overall QoE as measured by the VQM. Given the strict constraints of live streaming, which limit the number of representations in the ladder, it is crucial to evaluate the marginal quality gain of each additional representation when optimizing the selection.

The optimization for representation selection can be formulated

as:

$$\begin{aligned} & \underset{\Theta}{\text{maximize}} && \sum_{r \in R} W_r \sum_{i=1}^N Q(i, r, \theta_{i,r}), \theta_{i,r} \in \Theta \\ & \text{subject to} && |\Theta| \leq K \end{aligned} \quad (2)$$

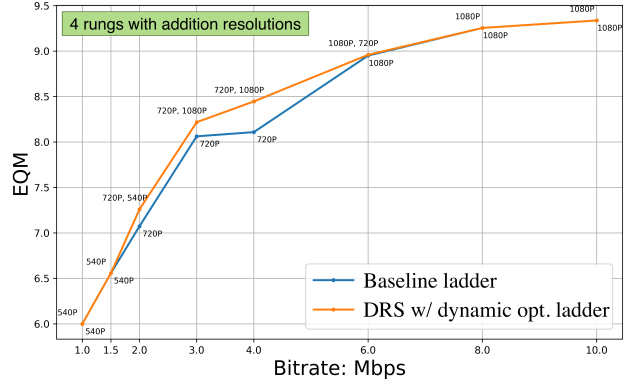
where  $R$  is the set of bitrates in the ladder,  $N$  is the total number of GOPs,  $\theta_{i,r}$  is the representation that offers the best score for the  $i$ -th GOP at bitrate  $r$ , and  $Q(\cdot)$  denotes the quality as measured by the VQM.  $W_r$  is the weight for the corresponding bitrate, which can be derived from the real-world network bandwidth distribution.  $K$  is the maximum number of representations in a ladder. The proposed formulation enables the construction of an augmented ladder that extends the static ladder for the DRS framework while optimizing for QoE.

**Table 5:** Enhancement of the baseline ladder for sports content using AVC-EQM scores. The *content optimized* ladder selects the resolution with the highest probability of achieving the best AVC-EQM score at each bitrate (see Figure 3), while the *dynamic optimized* ladder extends the baseline by adding 4 additional representations, which serve as representation candidates in the DRS pipeline and are optimized via Equation 2.

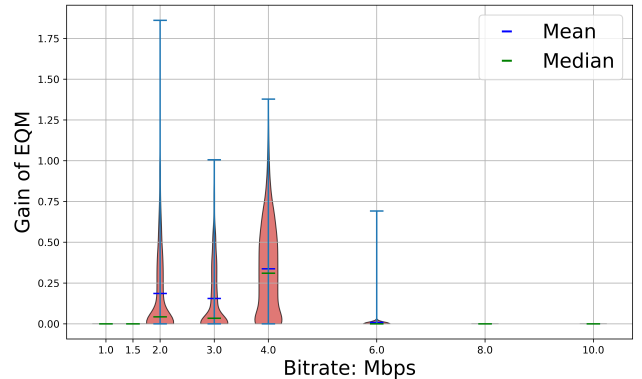
Rate	Baseline ladder	Content Opt.	Dynamic Opt.
1	540P	540P	540P
1.5	540P	540P	540P
2	720P	540P	540P, 720P
3	720P	1080P	720P, 1080P
4	720P	1080P	720P, 1080P
6	1080P	1080P	720P, 1080P
8	1080P	1080P	1080P
10	1080P	1080P	1080P

### 3.2. AVC-EQM based dynamic resolution switching

Table 5 presents the optimized bitrate ladder for sports content based on AVC-EQM scores. We evaluate the content optimized ladder and the dynamic optimized augmented ladder on a 30-minute test sequence that is distinct from the one used for representation selection. Figure 4 (a) shows the rate-quality curves comparing the baseline against our proposed DRS pipeline with the dynamic optimized ladder, using AVC-EQM (on a 0–10 scale) as the VQM. The results demonstrate that the proposed DRS framework achieves a significant performance gain over the baseline static ladder. Figure 4 (b) shows the distribution of AVC-EQM score improvements across GOPs, highlighting that some GOPs exhibit notable quality enhancement. Furthermore, substantial gains are observed particularly at 3 and 4 Mbps, demonstrating the efficacy of our method in this critical resolution cross-over region. While our evaluation focuses on sports content, the proposed DRS pipeline is designed to be generic: it selects the resolution with the highest VQM score at each bitrate, regardless of the content type. In addition, AVC-EQM and the ladder statistics can be re-estimated for other content using suitable subjective data, without modifying the DRS pipeline. Table 6 reports the results on two 30-minute test sequences from different sports. On average, our method achieves a BD-rate saving of 8.97% and a BD-quality gain of 0.152 (on a 10-point scale). These results, together with the minimal overhead of AVC-EQM, validate DRS as a practical and effective solution for QoE optimization in live streaming.



(a) RD Curves, BD-Rate: -9.367, BD-Quality: 0.153



(b) Distribution of EQM gains

**Fig. 4:** Performance evaluation of DRS using AVC-EQM scores on sports content. (a) RD curves for the baseline static ladder (blue) and DRS with the dynamic optimized ladder (orange). (b) Distribution of quality score gains across all GOPs; mean and median are marked by blue and green lines.

**Table 6:** Performance comparison between the content optimized ladder and our proposed DRS framework using an augmented ladder (4 additional representations) on two 30-minute sports sequences.

		Content Opt.	Dynamic Opt.
Sports 1	BD-Rate	-5.097	<b>-9.367</b>
	BD-Quality	0.082	<b>0.153</b>
Sports 2	BD-Rate	-5.274	<b>-8.567</b>
	BD-Quality	0.091	<b>0.151</b>

## 4. CONCLUSION

In this work, we proposed a dynamic resolution switching pipeline tailored for live streaming. The core of the method is selecting the optimal resolution, as measured by an efficient VQM, at any given bitrate during encoding. We first demonstrated the efficacy of the proposed AVC-EQM, validating its high correlation with subjective quality, its accuracy in estimating resolution cross-over points, and its runtime efficiency. We then extended the traditional static bitrate ladder by incorporating a statistical analysis of user-side bitrate distributions to maximize QoE, presenting a principled approach for selecting the additional representations. Empirical evaluation on real-world sports video sequences confirms significant quality improvements, demonstrating the overall effectiveness and practicality of the proposed method.

## 5. REFERENCES

- [1] Abdelhak Bentaleb, Bayan Taani, Ali C. Begen, Christian Timmerer, and Roger Zimmermann, “A survey on bitrate adaptation schemes for streaming media over http,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 562–585, 2019.
- [2] Jan De Cock, Zhi Li, Megha Manohara, and Anne Aaron, “Complexity-based consistent-quality encoding in the cloud,” in *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2016, pp. 1484–1488.
- [3] Vignesh V Menon, Jingwen Zhu, Prajit T Rajendran, Hadi Amirpour, Patrick Le Callet, and Christian Timmerer, “Just noticeable difference-aware per-scene bitrate-laddering for adaptive video streaming,” in *2023 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2023, pp. 1673–1678.
- [4] Ioannis Katsavounidis, “Iterative techniques for encoding video content,” Feb. 9 2021, US Patent 10,917,644.
- [5] Ahmed Telili, Wassim Hamidouche, Hadi Amirpour, Sid Ahmed Fezza, Christian Timmerer, and Luce Morin, “Convex hull prediction methods for bitrate ladder construction: Design, evaluation, and comparison,” *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 21, no. 7, pp. 1–23, 2025.
- [6] Jinhai Yang, Mengxi Guo, Shijie Zhao, Junlin Li, and Li Zhang, “Optimal transcoding resolution prediction for efficient per-title bitrate ladder estimation,” in *2024 Data Compression Conference (DCC)*, 2024, pp. 597–597.
- [7] Zhi Li, Anne Aaron, Ioannis Katsavounidis, Anush Moorthy, and Megha Manohara, “Toward a practical perceptual video quality metric,” *The Netflix Tech Blog*, vol. 6, no. 2, 2016.
- [8] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [9] Rakesh Rao Ramachandra Rao, Steve Göring, Peter List, Werner Robitza, Bernhard Feiten, Ulf Wüstenhagen, and Alexander Raake, “Bitstream-based model standard for 4k/uhd: Itu-t p. 1204.3—model details, evaluation, analysis and open source implementation,” in *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2020, pp. 1–6.
- [10] Yixu Chen, Zaixi Shang, Hai Wei, Yongjun Wu, and Sriram Sethuraman, “Encoder-quantization-motion-based video quality metrics,” in *2024 Picture Coding Symposium (PCS)*, 2024, pp. 1–5.
- [11] Jingwen Zhu, Yixu Chen, Hai Wei, Sriram Sethuraman, and Yongjun Wu, “Video quality assessment for resolution cross-over in live sports,” in *2025 IEEE International Conference on Multimedia and Expo (ICME)*, 2025, pp. 1–6.
- [12] Renata Teixeira, Zhi Li, Reenal Mahajan, and Wei Wei, “Smarter live streaming at scale: Rolling out vbr for all netflix live events,” *The Netflix Tech Blog*, 2026.
- [13] T. Wiegand, G.J. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, 2003.
- [14] Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, and Thomas Wiegand, “Overview of the high efficiency video coding (HEVC) standard,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, 2012.
- [15] Isabelle Guyon and André Elisseeff, “An introduction to variable and feature selection,” *Journal of machine learning research*, vol. 3, no. Mar, pp. 1157–1182, 2003.
- [16] Frederick N Fritsch and Judy Butland, “A method for constructing local monotone piecewise cubic interpolants,” *SIAM journal on scientific and statistical computing*, vol. 5, no. 2, pp. 300–304, 1984.
- [17] Maria Perez-Ortiz and Rafal K Mantiuk, “A practical guide and software for analysing pairwise comparison experiments,” *arXiv preprint arXiv:1712.03686*, 2017.
- [18] Fabian Brand, Christian Herglotz, and Andre Kaup, “On interpolation of subjective rate-distortion curves for video coder comparison,” in *2023 15th International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2023, pp. 189–192.