

Predictors Elimination Technique for HEVC

JVEC-C0023

Mikhail Korman, Oleg Prosekov

24th Meeting: Geneva, CH, May/June 2016



Outline

Introduction

Proposed solution

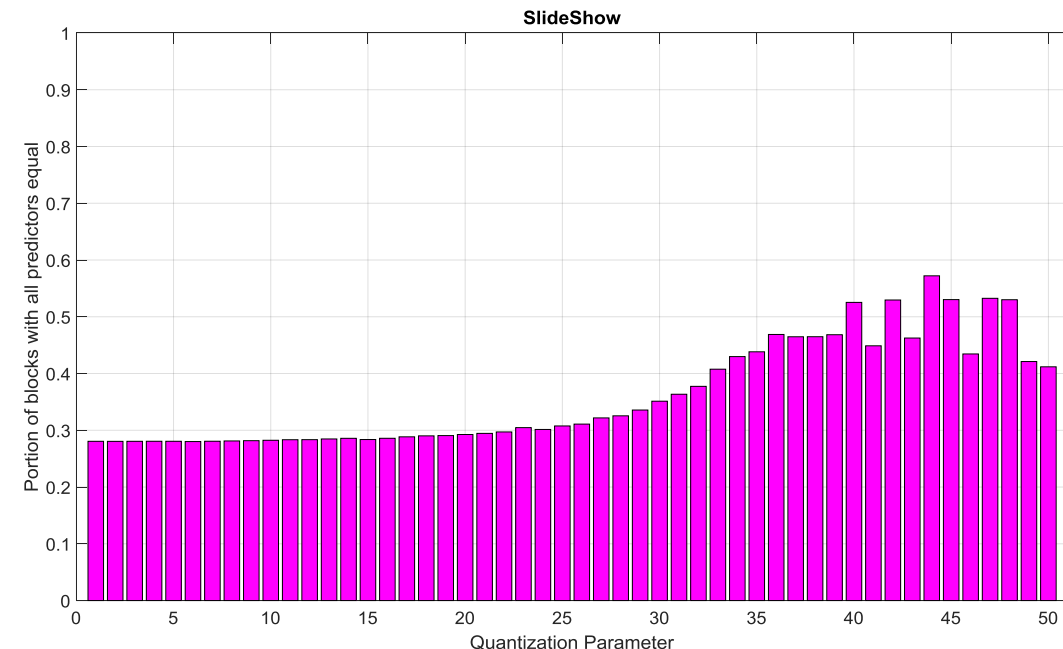
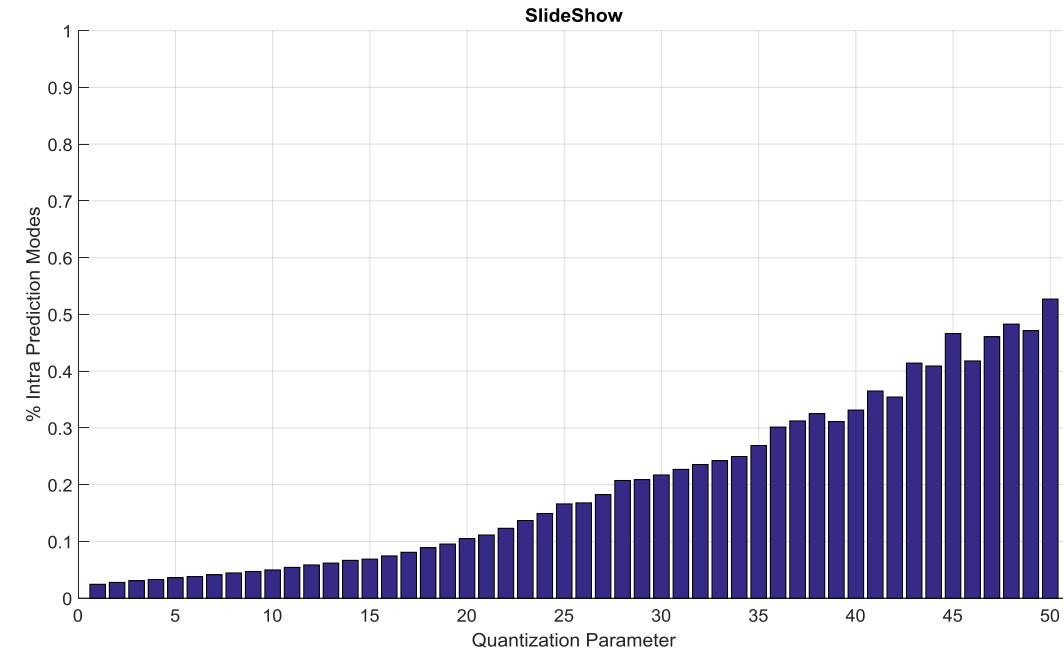
Results

Further research directions

Introduction

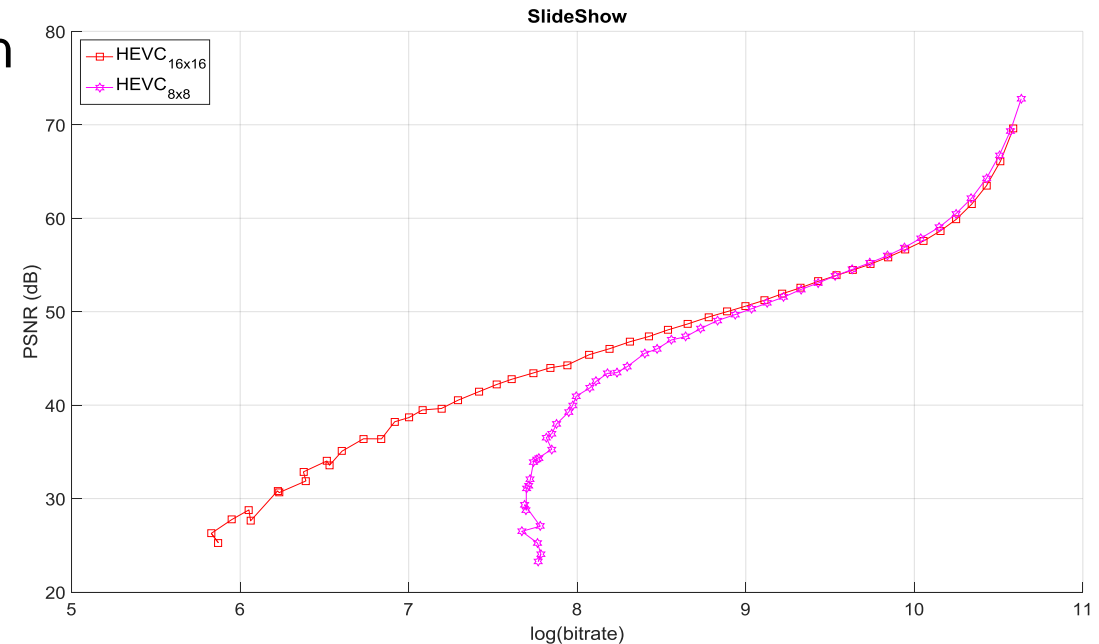
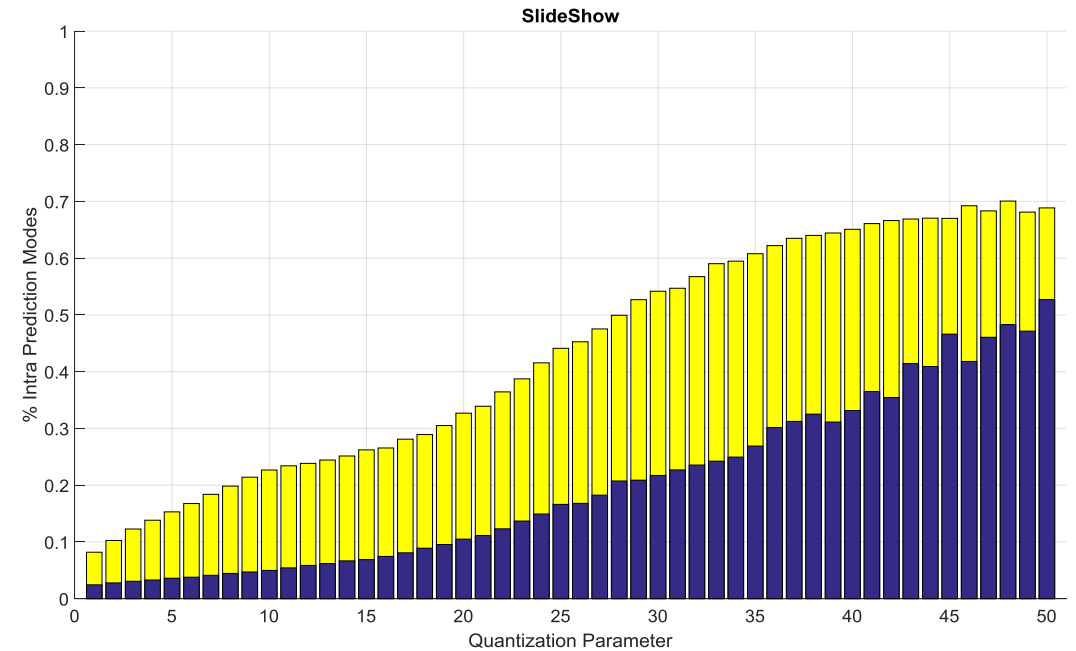
Problem statement

- Up to 70 % of bitrate represents Intra Coding modes in HEVC coded bitstream for Intra frames
- Multiple predictors may represent exactly the same or similar predictor blocks, and HEVC requires identifying one out of 35 modes in the bitstream
 - Quantization applied to coefficients allows treating similar (in some sense close) predictors as equivalent because the final residual block is being quantized
- **The main idea is to utilize predictors similarity for reducing the Intra modes representation in the bitstream**
- Efficiency of Intra modes coding is currently provided via MPM (Most Probable Mode) prediction; thus the way to take predictors similarity into account should not negatively impact MPM efficiency

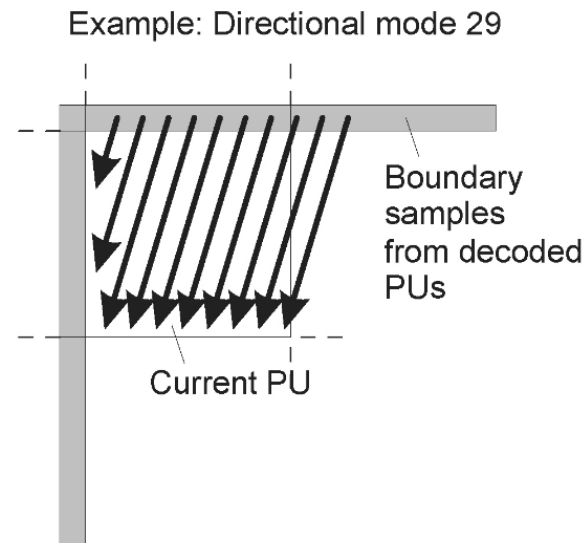
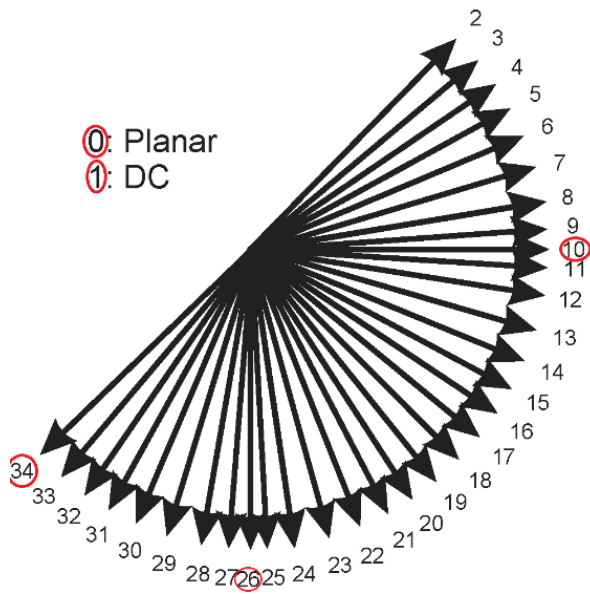


Intra Prediction: CU size

- HEVC defines CTU size in the range from 64x64 down to 8x8
 - Use of $N \times N$ shape leads to PU of size 4x4
 - HM (16.6+SCM5.2) reference code disables CU size of 8x8 due to reasons related to inter-frame prediction
- For the sake of simplicity and ease of implementation on the current stage the following assumptions were made for this contribution:
 - $2N \times 2N$ partition shape is disabled
 - 8x8 Coding Unit size and only such is enabled for Intra frames
 - CU for 4:2:0 source consists of 4 Luminance and 2 Chrominance PB's of size 4x4



Intra Prediction Modes



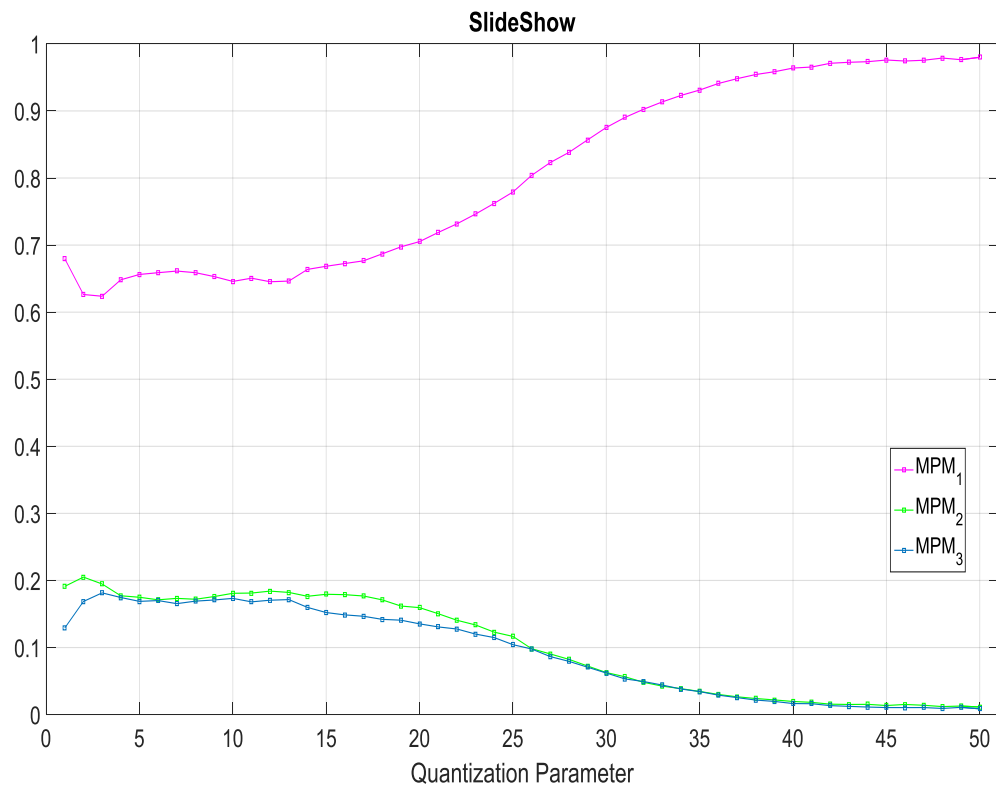
- 5 Chroma modes allowed (for both U and V Chroma PB)
 - One derived from Luma prediction – one bin goes to the bit stream
 - Planar, DC
 - Vertical, Horizontal

} 2 bits more to represent in the bit stream

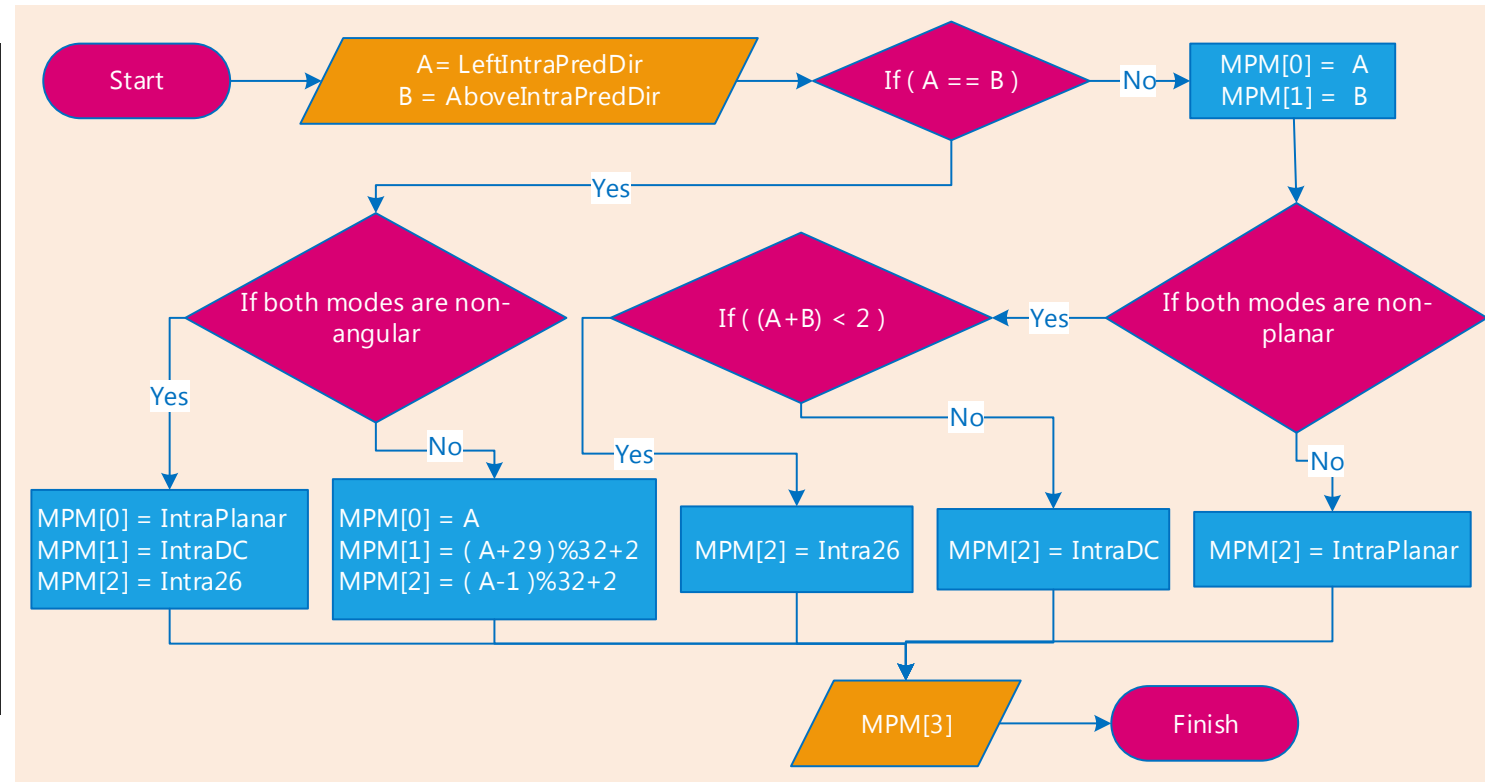
- 33 angular modes with different directional orientations are defined
- 2 non-angular modes: Planar and DC derived from all boundary samples
- Red circles indicate “preferred” modes used by MPM for Luma or Chroma
- Luma mode coding price:
 - 1 bin plus up to 2 bits code used in bit stream for MPM selection
 - Otherwise 5 bit fixed length code used in bit stream for non-MPMs modes

MPM coding

- MPM hit rate example is shown below



- MPM coding diagram shown below



Proposed solution

Aka Predictors Elimination Technique (PET)

Predictor Elimination Technic simplified (PET-1)

- The predictors P_a and P_b constructed from samples $P[i, j]$ are close if the following inequality is satisfied:

$$\|P_a - P_b\|_1 = \sum_{i,j} |P_a[i, j] - P_b[i, j]| < f_Q$$

- Groups of close predictors are formed

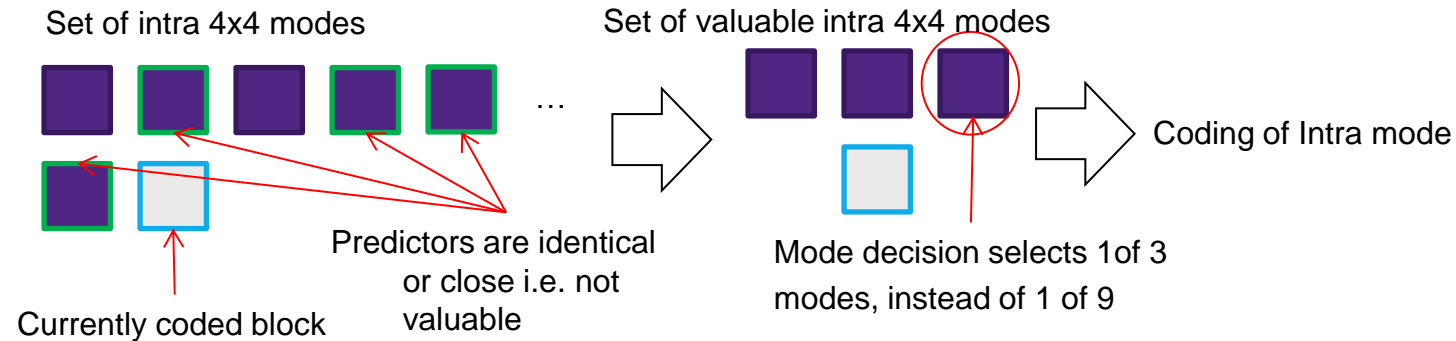
$$G' = \left\{ i_k \in G : \forall k \neq m: \|P_{i_k} - P_{i_m}\|_1 < f_Q \right\}, \text{ where } G \text{ is the set of all allowed predictor modes}$$

- Close predictors are replaced by a predictor representing the whole group it belongs to

$$RP(G') = \operatorname{argmin}_{i \in G'} (\|P_i - \bar{P}\|_1),$$

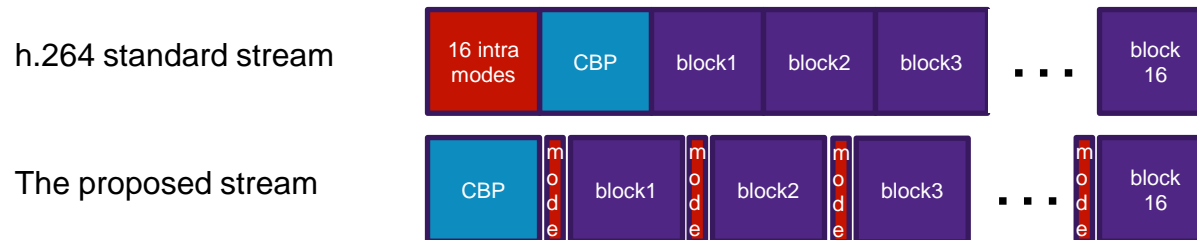
where \bar{P} is the averaged predictor over modes from G'

PET requires bit stream reordering



Intra modes info precedes block data in the macro-block header of standard h.264 coded stream

Mode information for each block requires all predictor blocks to be decoded first in order to let decoder know the length of prediction mode code.



PET-1 implementation for H.264

- PET was initially implemented for H.264
- Achieved Rate-Distortion gain is shown on the example of Launch sequence

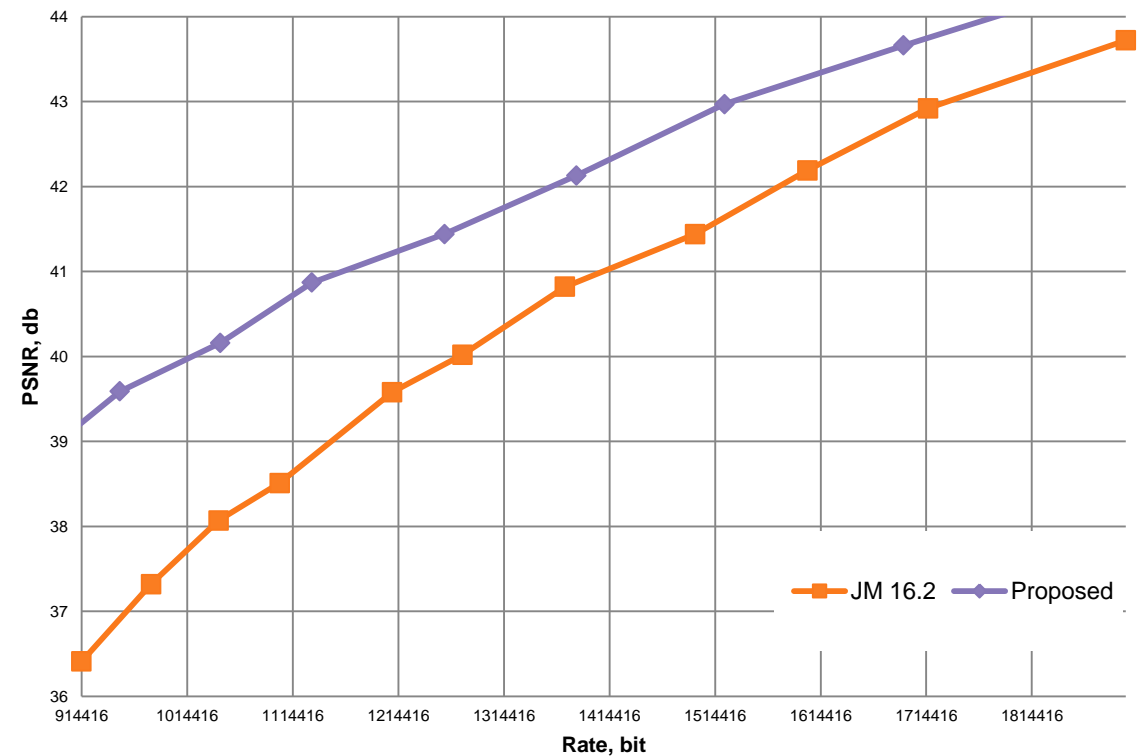


Compare the quality with the same rate (Stream "Launch"):

JM 16.2 Intra 4x4, PSNR = 38.51;

Rate-distortion curves (Q=20,21,...,30)

Intra 4x4 coding, stream "Launch"



Predictor Elimination Technic for HEVC (PET-2)

- The predictors P_a and P_b constructed from samples $P[i, j]$ are close if the following inequality is satisfied:

$$\|P_a - P_b\|_\infty = \max_{i,j} |P_a[i, j] - P_b[i, j]| < f_Q$$

- Groups of close predictors are formed

$$\{G_{non-angular}'\} = \begin{cases} \{0, 1\} & \text{if } \|P_0 - P_1\|_\infty < f_Q \\ \{0\}, \{1\} & \text{otherwise} \end{cases}$$

$$G_{angular}' = \left\{ i_k \in G_{angular} : \forall i_k, i_m \in \{p, \dots, p+q\}, k \neq m: \|P_{i_k} - P_{i_m}\|_\infty < f_Q \right\}$$

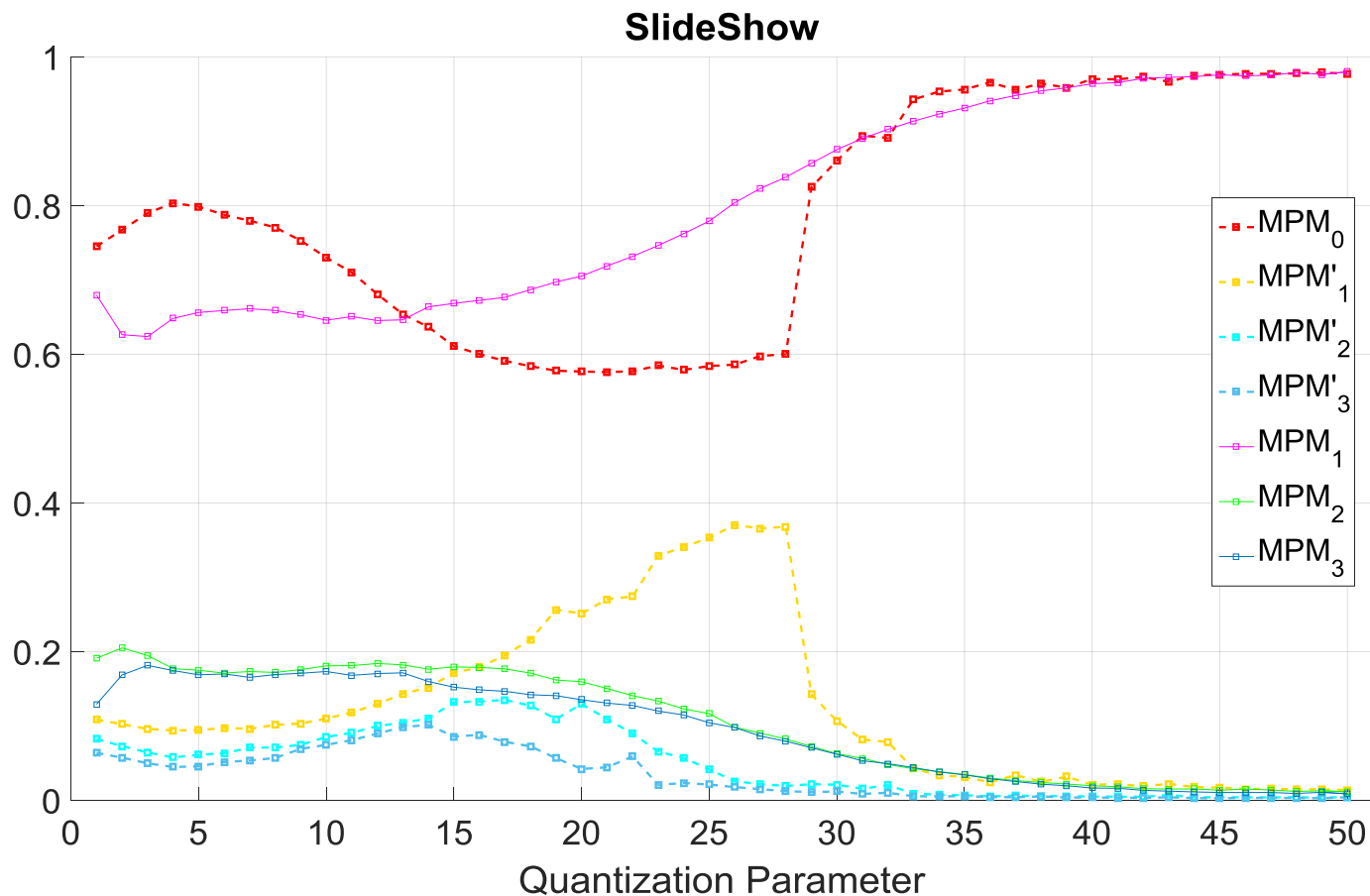
In case all 33 angular and both non-angular modes are grouped, grouping of angular and non-angular modes into a group combining all 35 modes is considered using the same norm and similarity threshold f_Q

- Close predictors are replaced by a predictor representing the whole group it belongs to
- Chroma modes are grouped into one group only when all predictors in both components are similar

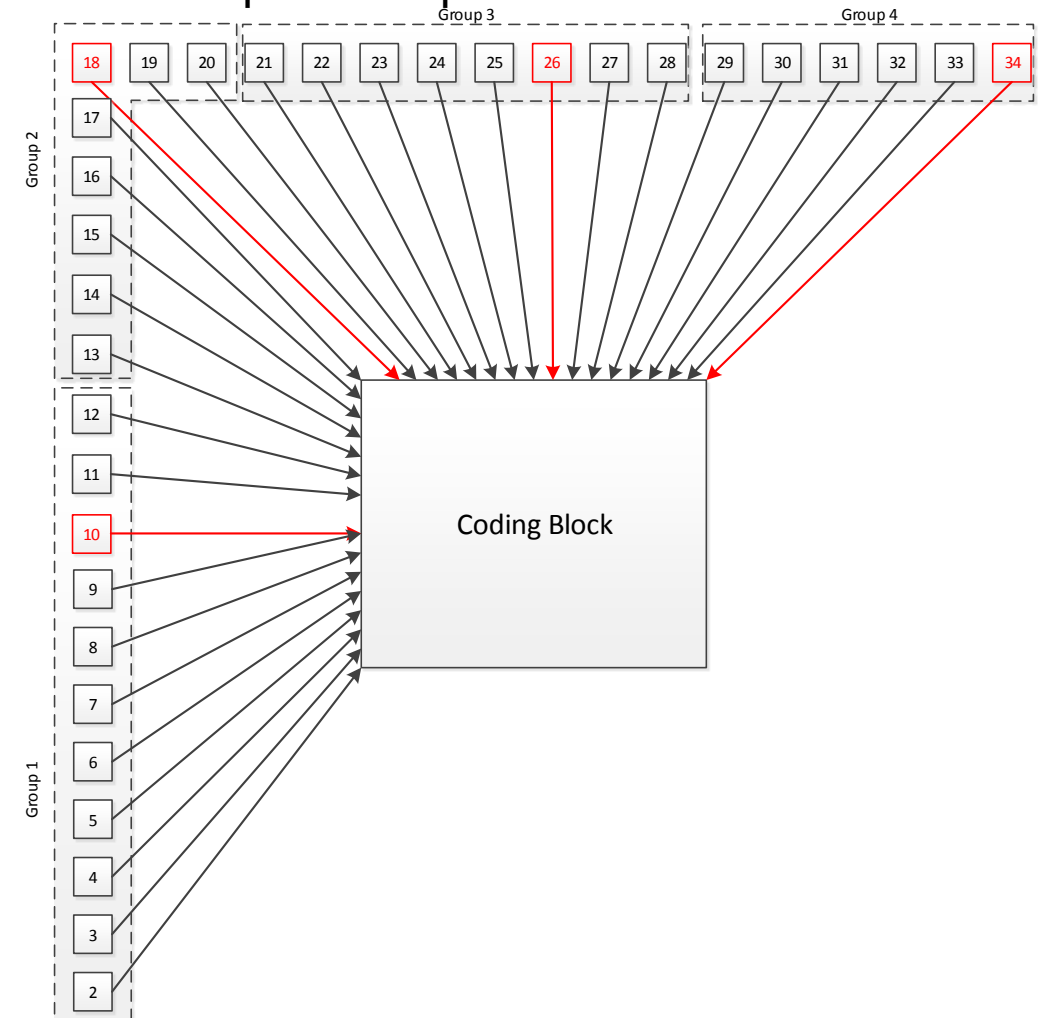
$$RP(G') = \begin{cases} 0, & \text{if } 0 \in G' \\ 1, & \text{otherwise if } 1 \in G' \\ 26, & \text{otherwise if } 26 \in G' \\ 10, & \text{otherwise if } 10 \in G' \\ 34, & \text{otherwise if } 34 \in G' \\ \operatorname{argmin}_{i \in G'} (\|P_i - \bar{P}\|_\infty), & \text{otherwise} \end{cases}$$

Grouping for Angular Predictors and MPM statistics

- For the sequence depicted the first MPM of HEVC usually does not require indication in the bit stream (shown as MPM_0) due to all modes grouped



- Proposed angular modes grouping requires less comparison operations



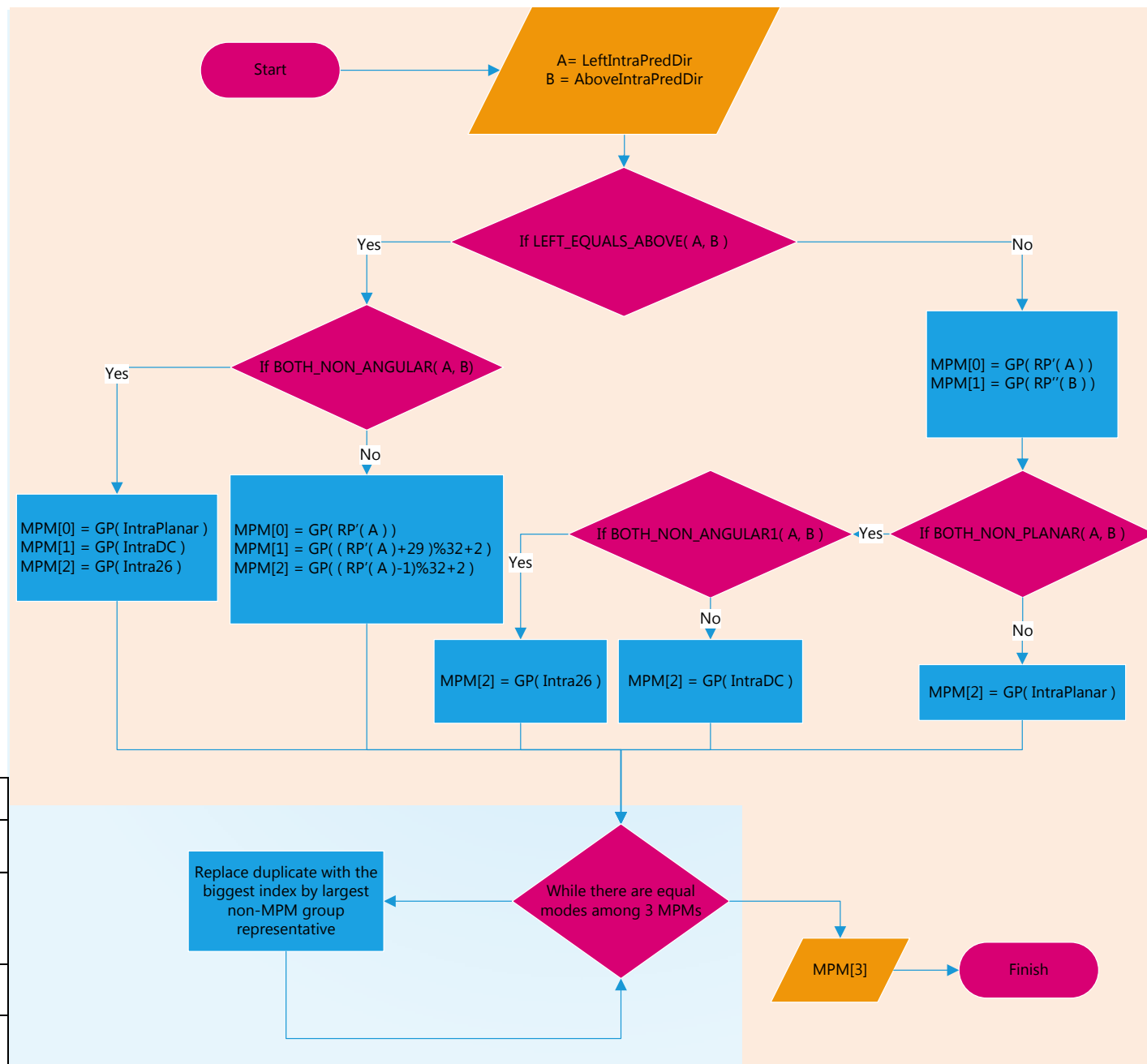
Modified MPM for PET-2

- New steps in MPM array construction are shown on blue background
- Modes re-numeration used in the scheme:

$$GP(x) = i: x \in G_i,$$

$RP'(x)$ and $RP''(x)$ denote representative predictors for groups in PB's to the left and above to the current PB being processed

Function	HEVC condition	Modified condition
LEFT_EQUALS_ABOVE(A, B)	A==B	$\exists i: RP'(A), RP''(B) \in G_i$
BOTH_NON_ANGULAR(A, B)	(A<2) && (B<2)	$\{2,...,35\} \cap GP(RP'(A)) = \emptyset$ $\{2,...,35\} \cap GP(RP''(B)) = \emptyset$
BOTH_NON_ANGULAR1(A, B)	(A+B) <2	$RP(GP(RP'(A))) + RP(GP(RP''(B))) < 2$
BOTH_NON_PLANAR	A && B	$RP(GP(RP'(A))) \&\& RP(GP(RP''(B)))$



Bit stream re-ordering for PET-2

- Intra Mode data placement in HEVC bit stream:

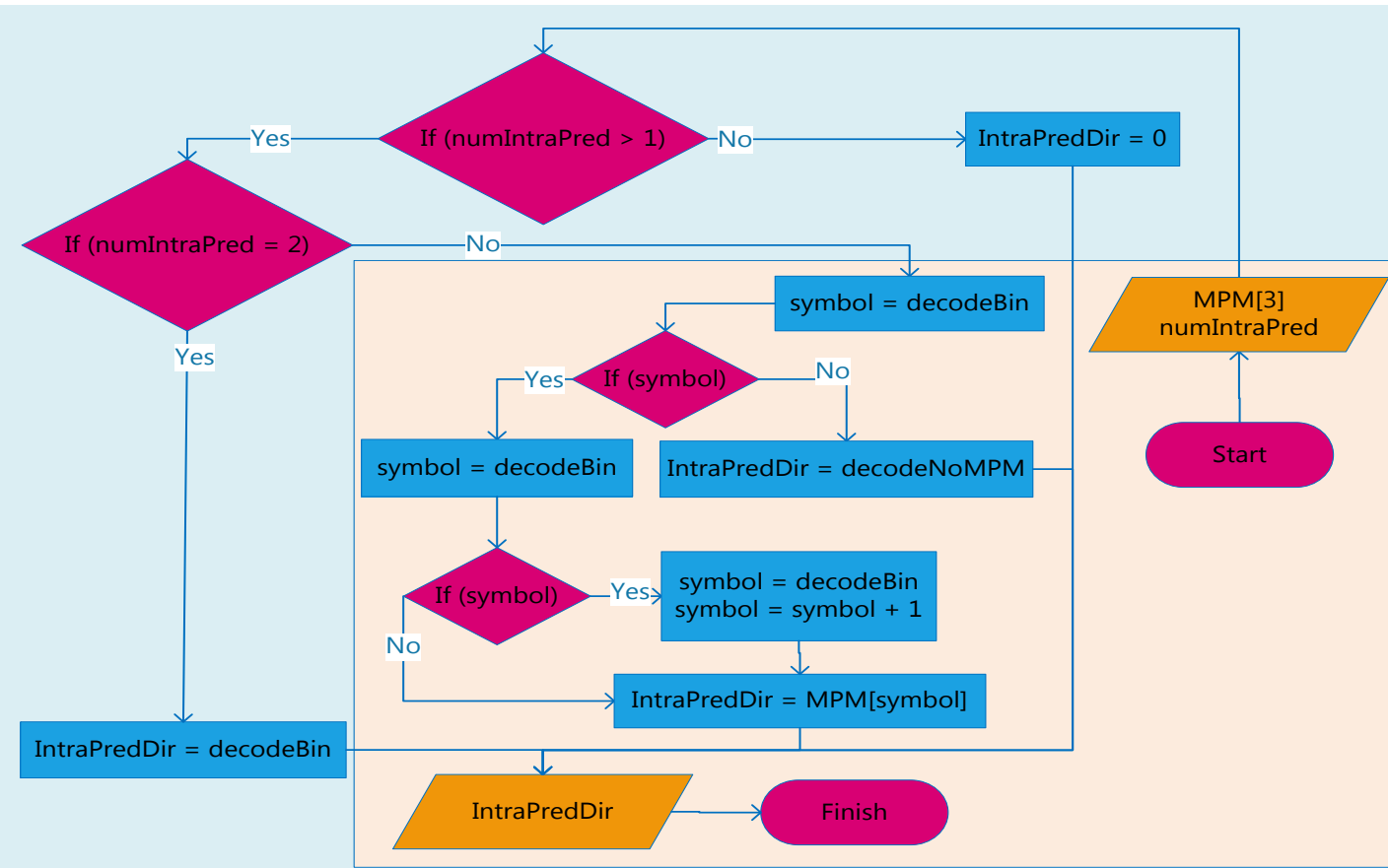
Mode for LumaPU0	Mode for LumaPU1	Mode for LumaPU2	Mode for LumaPU3	Mode for Chroma	Luma Residual PU0	Luma Residual PU1	Luma Residual PU2	Luma Residual PU3	Chroma Residual U	Chroma Residual V
------------------	------------------	------------------	------------------	-----------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------

- Proposed Intra Mode data placement for 4:2:0 input:

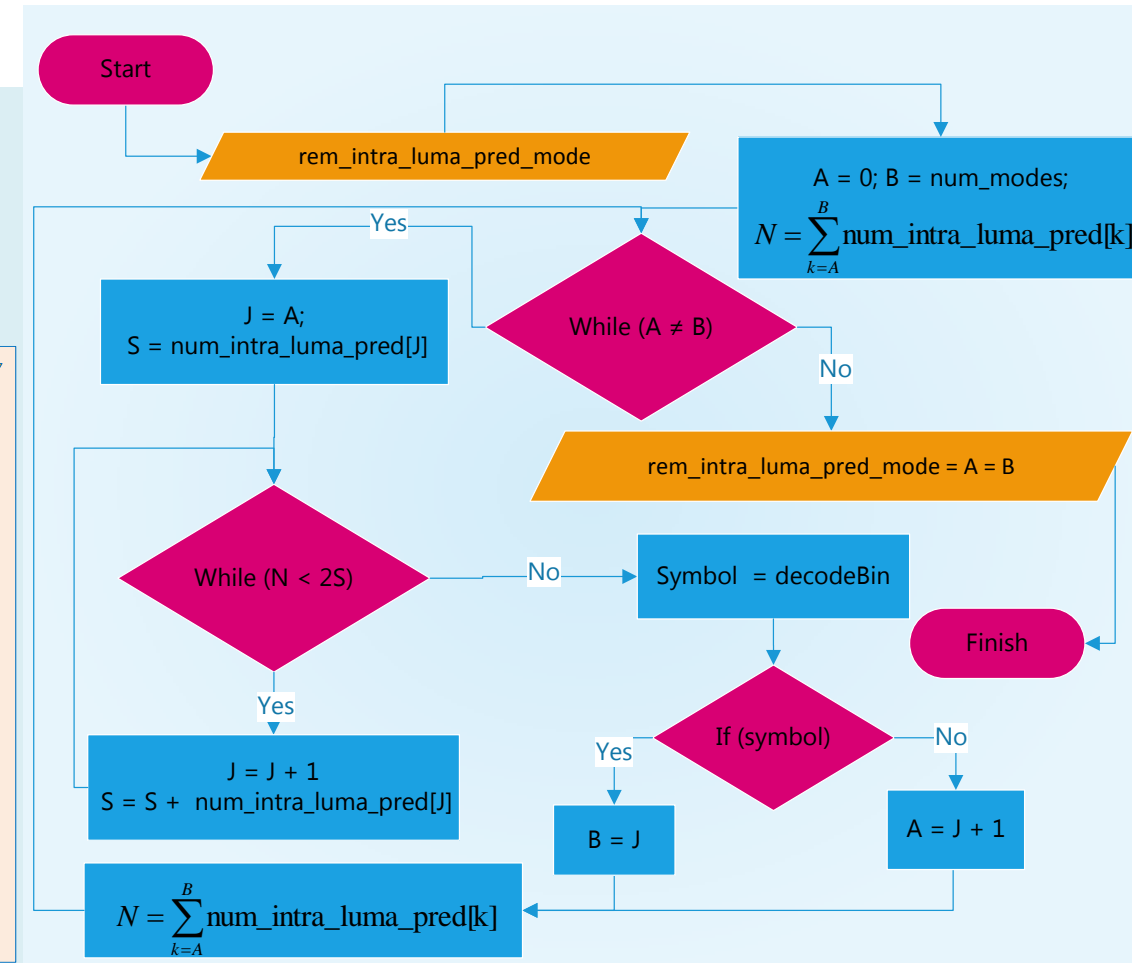
Mode for LumaPU0	Luma Residual PU0	Mode for LumaPU1	Luma Residual PU1	Mode for LumaPU2	Luma Residual PU2	Mode for LumaPU3	Luma Residual PU3	Mode for Chroma	Chroma Residual U	Chroma Residual V
------------------	-------------------	------------------	-------------------	------------------	-------------------	------------------	-------------------	-----------------	-------------------	-------------------

Modes information coding

- Shortcuts added to modes coding scheme
 - Modifications shown on blue background



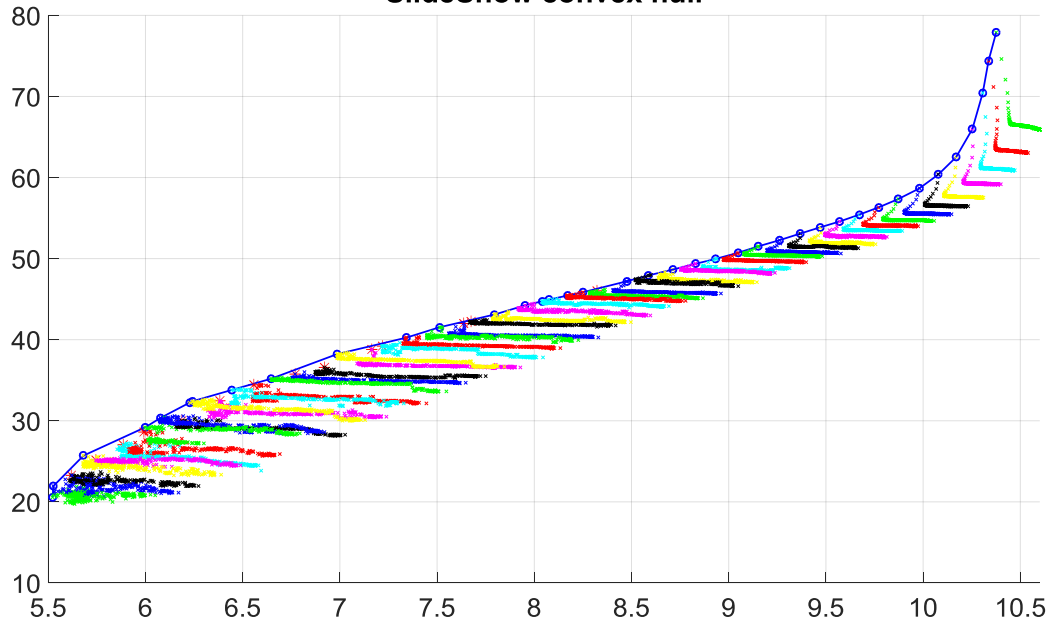
- Shannon-Fano coding of non-MPM modes



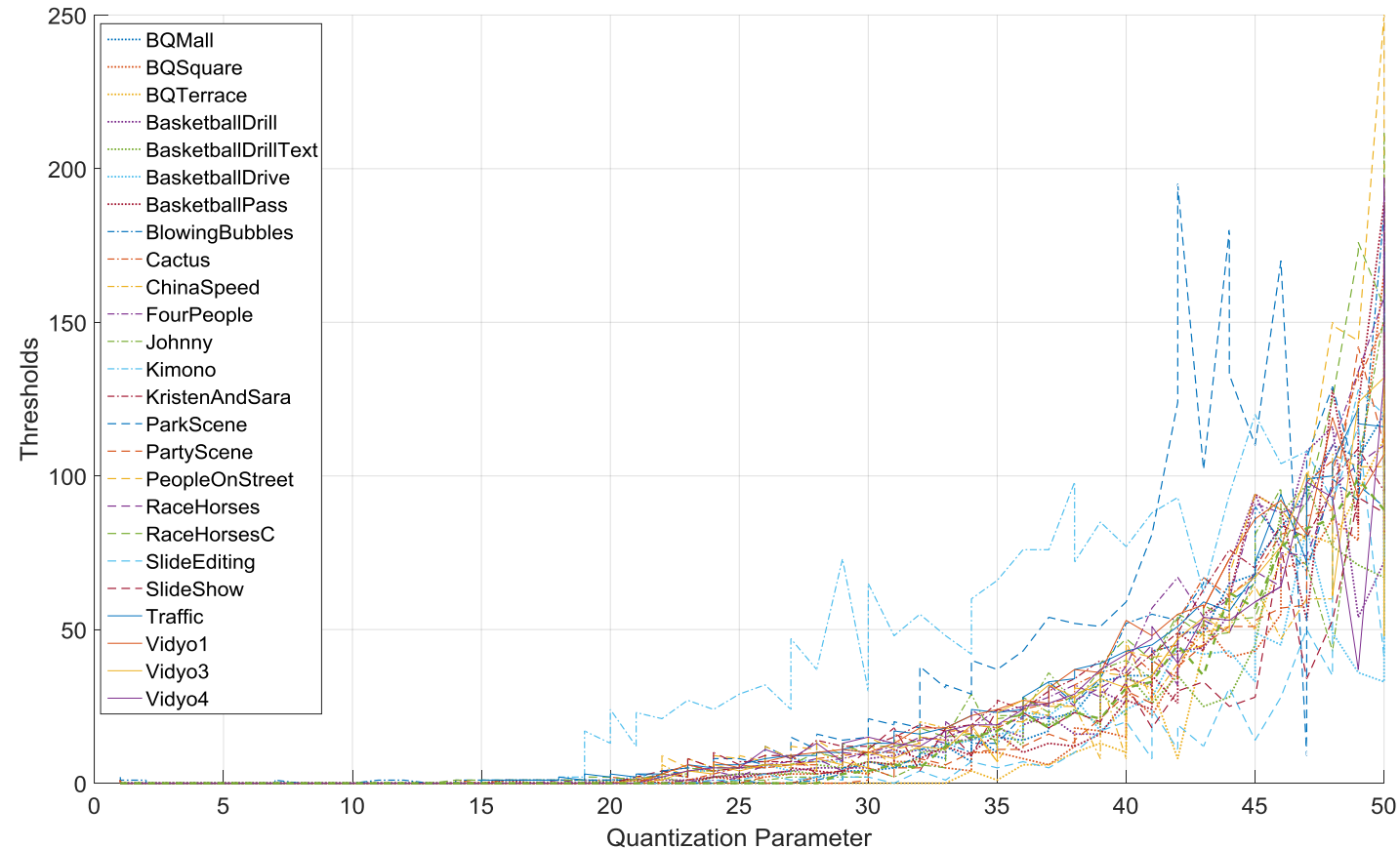
Sequence thresholds selection

- Rate distortion points are put onto a rate-distortion plane
 - MSE domain is used for distortion
- Convex hull of the resulting set defines optimal thresholds for a given sequence

SlideShow convex hull

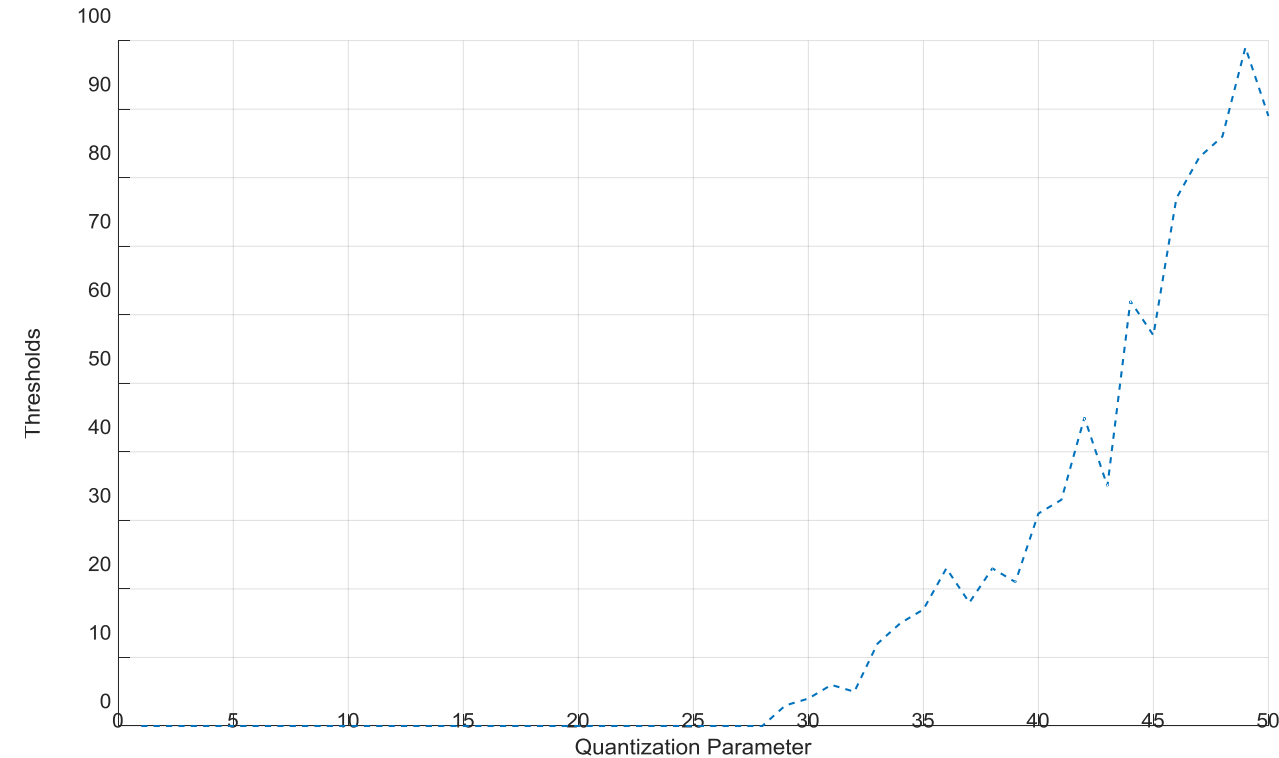
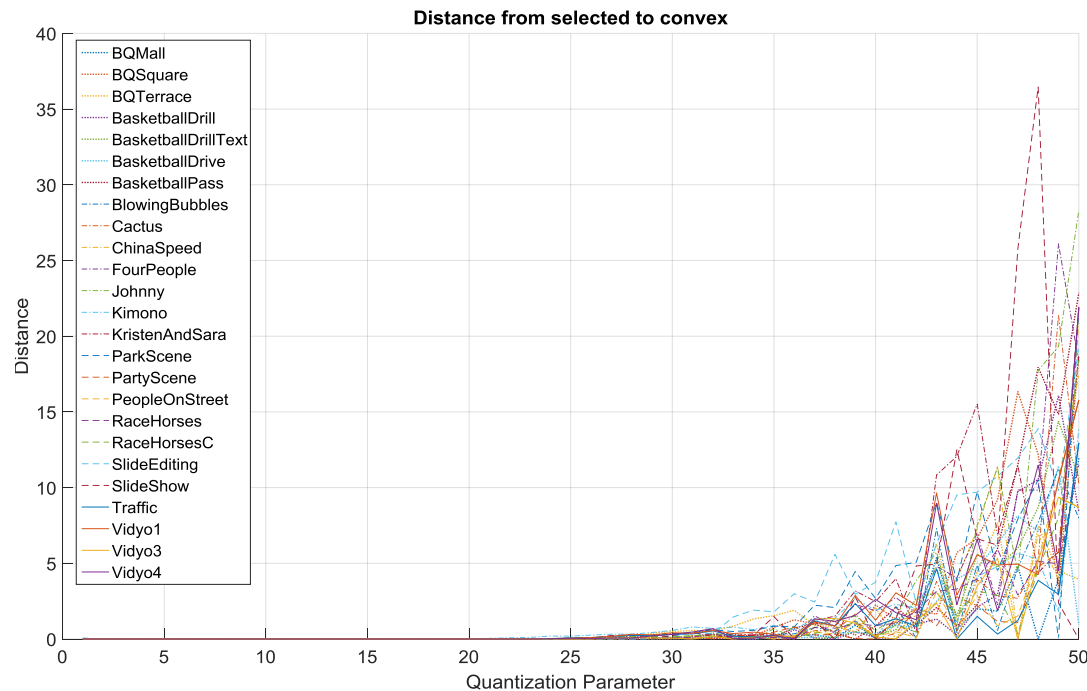


- Thresholds f_Q in range 0..255 are tried for each Q and each test sequence and optimal results are depicted below



Sequence independent thresholds selection

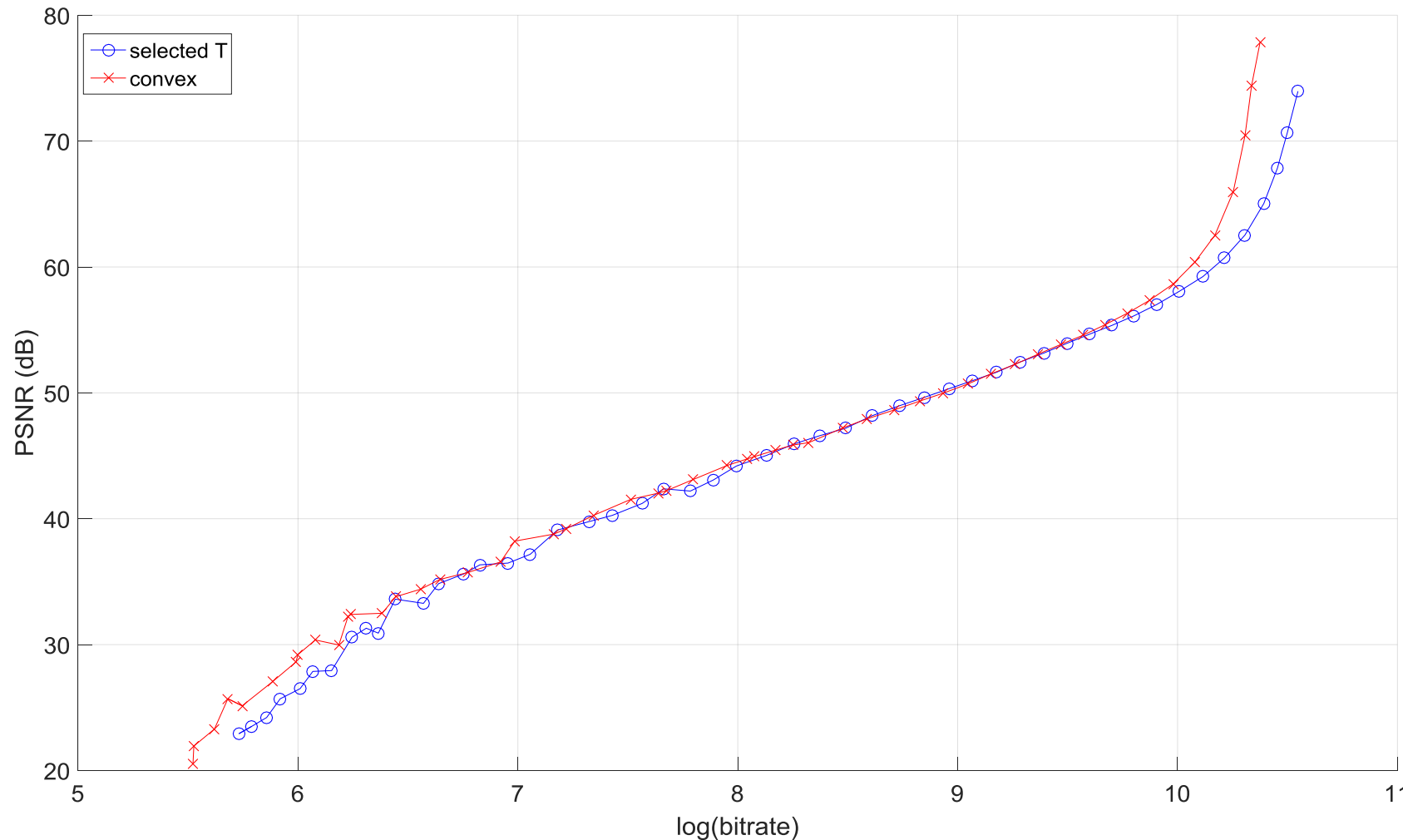
- Sequence independent thresholds are selected with help of least squares method
 - For each Q the distance to convex hulls optimized for individual sequences is being minimized



- $f_Q(1:28) = 0$
- $f_Q(29:50) = \{3, 4, 6, 5, 12, 15, 17, 23, 18, 23, 21, 31, 33, 45, 35, 62, 57, 77, 83, 86, 99, 89\}$

Sequence independent thresholds vs sequence thresholds

- Impact of sequence independent thresholds on SlideShow sequence depicted below indicates room for further improvements possible



RESULTS

Modifications to HM reference

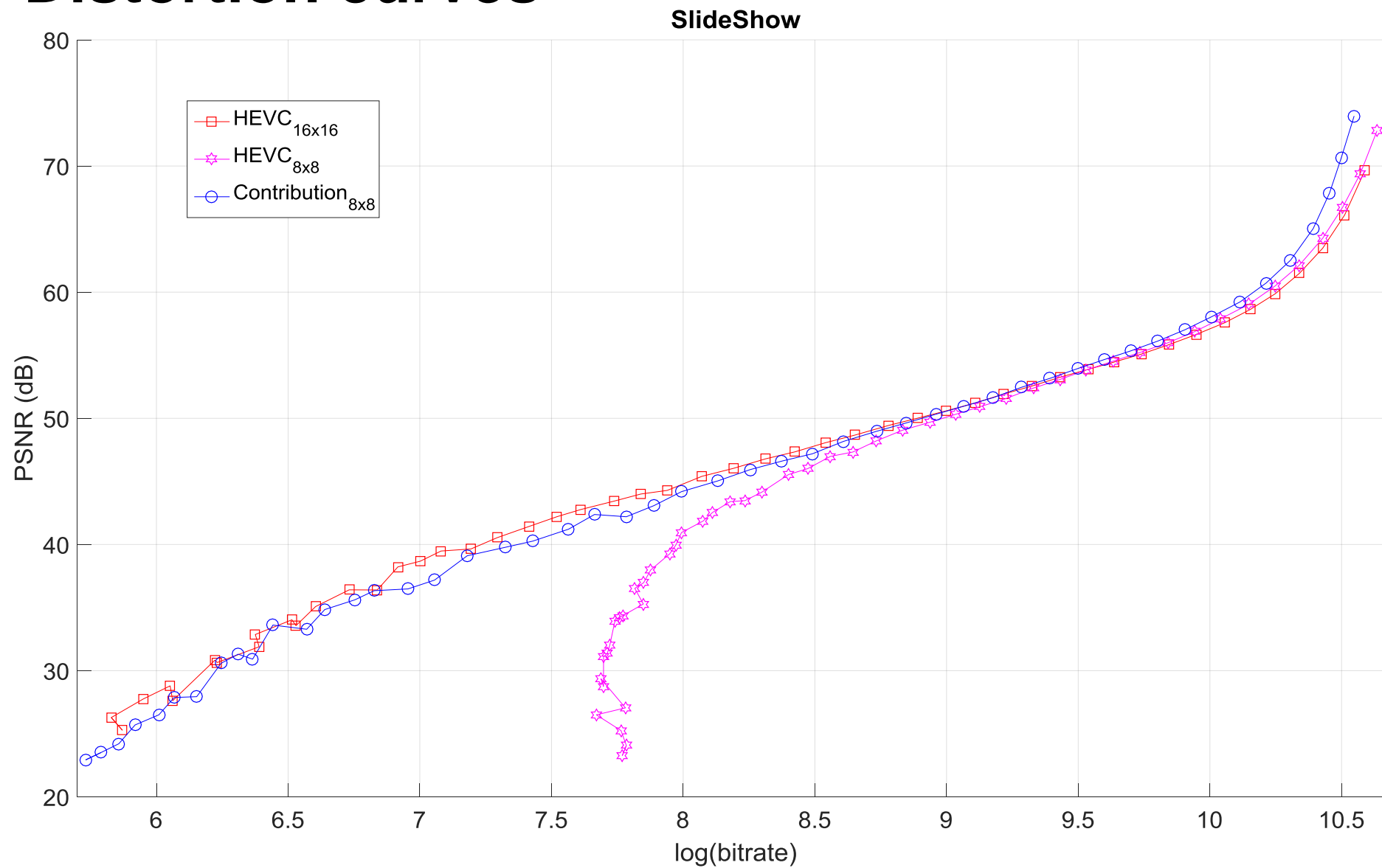
- We use HM-16.6+SCM5.2
- Set `bFlag` to `true` in `calcRdCost`
 - Estimate and pre-estimate steps both floating point cost values
- Rate estimation function modifications – requires floating point RD-cost values
 - Current Rate estimation function for i -th prediction unit \widehat{R}_i^{PU}
$$\widehat{R}_i^{PU} = \left\lfloor \{r_{i-1}^{PU}\} + \widetilde{R}_i^{PU} \right\rfloor,$$
 - where fractional part of bits collected by $(i - 1)$ -th PU r_{i-1}^{PU} is hidden in arithmetic coder state
 - where integer and fractional part estimator `m_fracBits` from `TEncBinCABACCounter` class is denoted as \widetilde{R}_i^{PU}
 - Proposed Rate estimation function

$$\widehat{R}_i^{PU} = \left\lfloor \widetilde{R}_i^{PU} \right\rfloor + \{ \widetilde{R}_i^{PU} \}$$

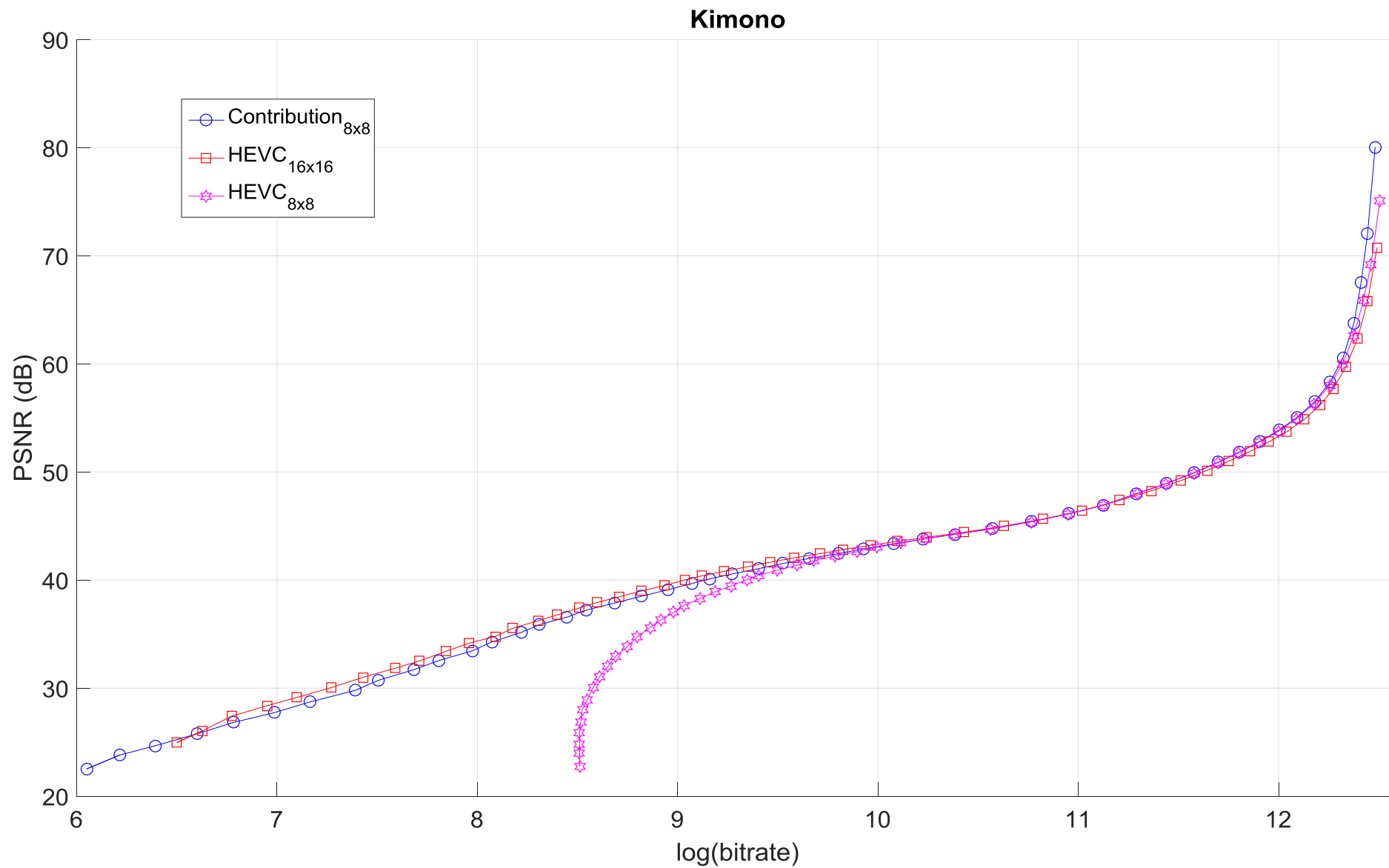
BD-rate results

Sequence name	% of bitrate saving vs CU size 8x8	average PSNR difference vs CU size 8x8	% of bitrate saving vs CU size 16x16	average PSNR difference vs CU size 16x16	% of bitrate saving vs integer RD-cost	average PSNR difference vs integer RD-cost	% of bitrate saving vs float RD-cost	average PSNR difference vs float RD-cost
BQMall	-9,724163854	0,722076047	2,533703265	0,035421287	-0,136926404	0,092672777	-0,060069594	0,007099948
BQSquare	-5,220651905	0,610371645	-3,683784128	0,470020732	-0,85896723	0,080462738	-0,061778216	0,002189962
BQTerrace	-9,2750297	0,531513635	4,951605404	-0,320208165	-0,708541143	0,061846179	-0,002190754	0,001144217
BasketballDrill	-14,75078301	0,801481375	6,061578429	-0,172919617	-0,791872024	0,08511662	-0,153569907	0,016160076
BasketballDrillText	-13,49574137	0,806507158	4,765093298	-0,131663049	-0,580088912	0,058237429	0,123574902	-0,000339352
BasketballDrive	-24,41627756	0,474097498	14,61405249	-0,416831175	3,598849867	0,059446407	0,1119441	-0,000427038
BasketballPass	-12,96827381	0,700501708	4,440966693	-0,199542856	-1,031651107	0,099113682	-0,429280634	0,03193402
BlowingBubbles	-9,189989003	0,46944702	-1,072025613	0,23285155	-1,610921218	0,084489934	-0,049014697	0,002734033
Cactus	-14,37132728	0,706396934	6,680158284	-0,038507966	1,558130272	0,07822568	0,057668054	-0,001334787
ChinaSpeed	-11,79594088	1,149437583	0,85854345	-0,172333796	-1,362367954	0,113942031	-0,191996924	0,023725253
FourPeople	-19,00718623	1,250801128	6,657654674	-0,32140283	0,153384472	0,112513986	-0,136302927	0,004284742
Johnny	-25,03883606	1,255308037	9,563594581	-0,375012036	0,953065585	0,099084775	-0,139503105	-0,006459255
Kimono	-28,77518709	0,932268171	6,479387324	0,19165455	4,332191544	0,097952023	-0,044971308	-0,002994573
KristenAndSara	-23,5114247	1,372723215	7,968134921	-0,364169271	0,392046232	0,111358007	-0,117938228	0,00961824
ParkScene	-21,31823708	0,653847948	11,57859648	-0,356896489	1,285535792	0,093769237	0,125660007	-0,001816718
PartyScene	-7,025192177	0,409654978	-2,82135326	0,272329955	-1,693882296	0,073242941	0,011466755	-0,002401764
PeopleOnStreet	-11,87801126	0,912766578	4,894650549	-0,0902157	0,926729155	0,103977427	0,001511001	0,002562629
RaceHorses	-7,146083169	0,591985959	-0,099524819	0,138103884	-0,660422492	0,084118324	-0,054099588	0,010934265
RaceHorsesC	-10,83070048	0,65926298	2,773937302	-0,091925323	-0,781220706	0,09848195	-0,257662188	0,021092375
SlideEditing	-6,114292682	0,794870406	-3,687825599	0,40934888	-0,725760464	0,075894011	-0,080935292	0,010421221
SlideShow	-40,59629937	1,91262327	2,970280773	-0,172582827	-0,797449029	0,125118265	-0,194155572	0,023489963
Traffic	-17,03355619	0,984579892	6,099084241	-0,121139421	0,7498769	0,107690638	-0,043598149	0,006804123
Vidyo1	-25,38779962	1,279729787	8,293107461	-0,30297432	1,167726144	0,102903012	-0,06086272	-0,007131115
Vidyo3	-23,37260801	1,305150008	9,454754268	-0,611533455	-0,593458354	0,120469943	-0,350830544	0,029132823
Vidyo4	-27,80267809	1,195824516	11,5131471	-0,634864918	0,484187656	0,094401846	-0,118459304	-0,000013007
MEAN	-16,801850823	0,8993290990	4,8715007027	-0,1257996950	0,1307277714	0,0925811945	-0,0846157933	0,0072164113

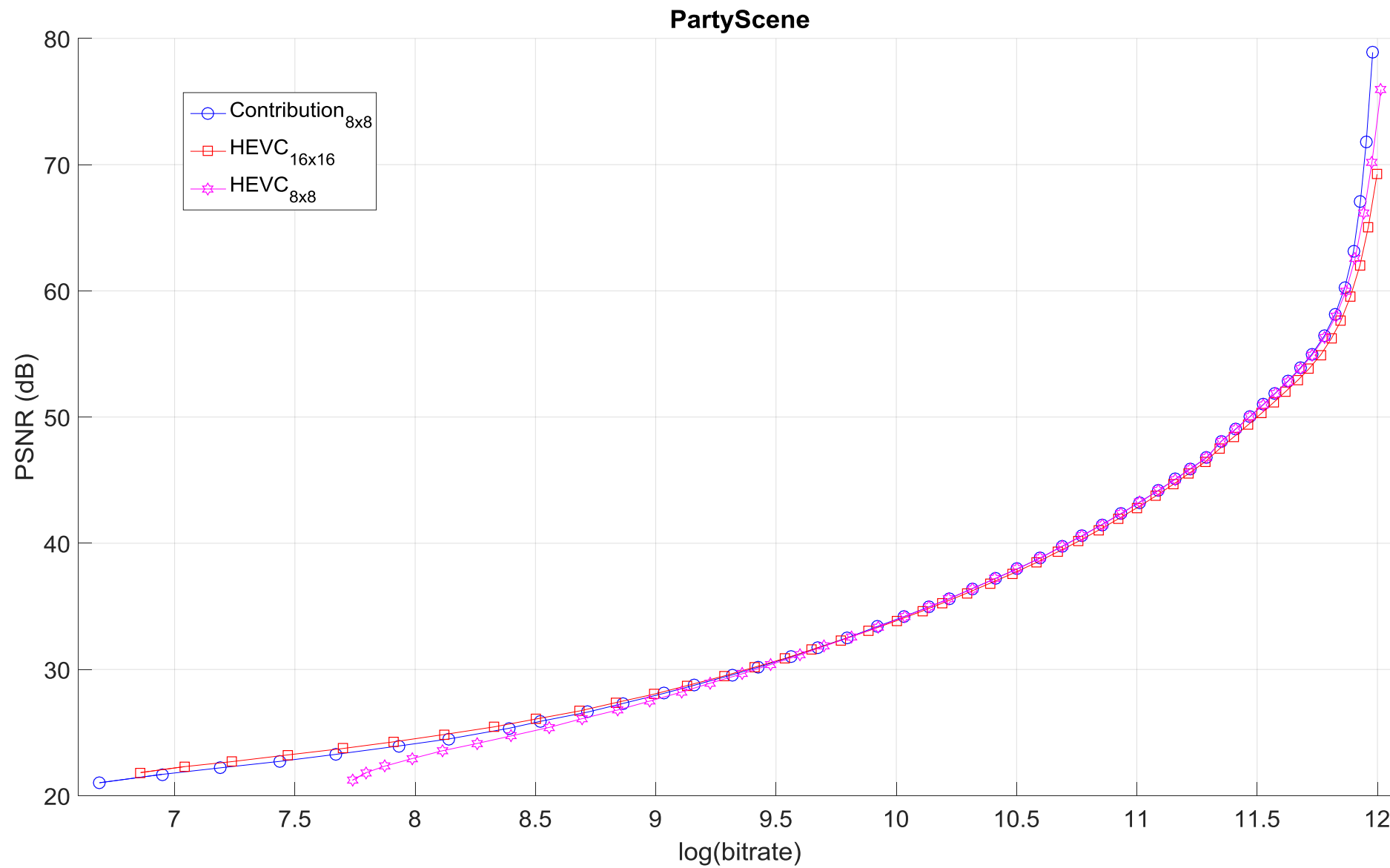
Rate-Distortion curves



Rate-Distortion curves



Rate-Distortion curves



Further research directions

Proposal status

1. This is not yet a final proposal but a direction for further investigations
2. Investigations planned
 1. Support 4:2:2 and 4:4:4
 2. Support PB's of bigger sizes and CTU hierarchy
 3. Chroma modes coding can be further improved
 4. Further MPM modifications may bring extra gain
 5. Speedup options
 6. Other quality metrics

