Learning-based Point Cloud Decoding With Independent and Scalable Reduced Complexity

Mohammadreza Ghafari[1,2], André F. R. Guarda[2], Nuno M. M. Rodrigues[2,3], and Fernando Pereira[1,2]

[1]*Instituto Superior Técnico – Universidade de Lisboa,* [2]*Instituto de Telecomunicações,* [3]*ESTG – Politécnico de Leiria*
PORTUGAL

IEEE International Conference on Image Processing, Abu Dhabi, UAE | October 27-30, 2024

instituto de telecomunicações

Outline

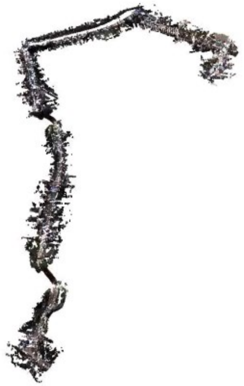
instituto de telecomunicações

1 Context and Objective

What is a Point Cloud?

- A point cloud (PC) is a 3D visual representation model consisting of an unordered set of 3D points, characterized by their coordinates (x, y, z).
- Target applications include virtual/augmented reality for entertainment and communication, autonomous driving, and among others.

https://sitesurvey.co.uk/



from Danillo Graziosi

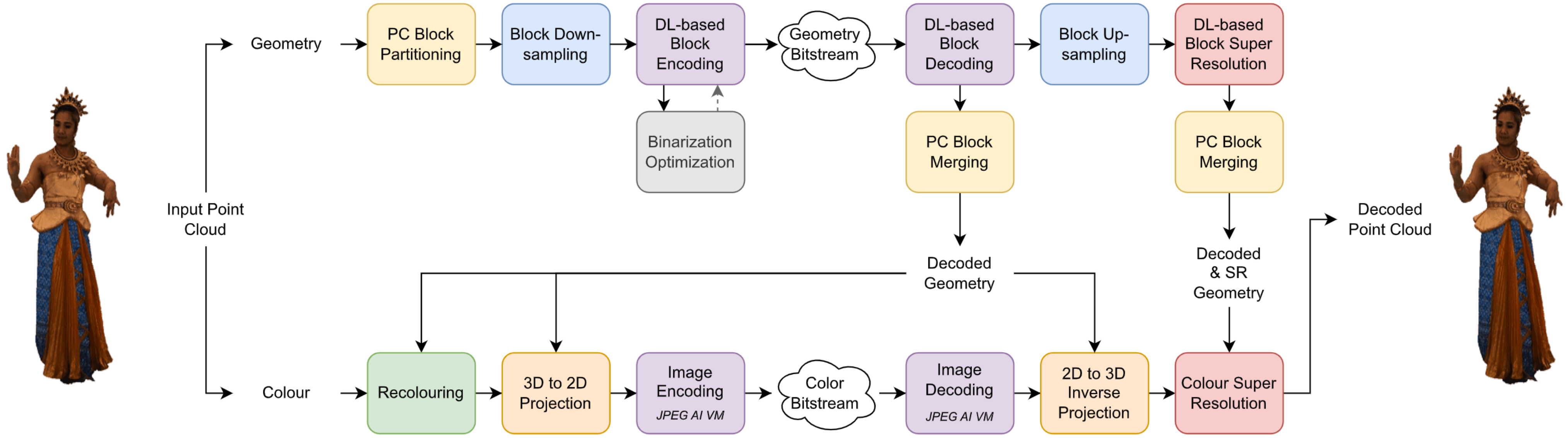
Why Point Cloud Coding is a Must?

- PCs may need millions of points to offer accurate representations, realism, and immersion, requiring very large amounts of data.
- Efficient PC coding solutions are essential for practical storage and transmission.

JPEG PCC Standard

instituto de telecomunicações

- ✶ JPEG recently finalized the JPEG Pleno Learning-based Point Cloud Coding (JPEG PCC) standard.
- ✶ JPEG PCC offers improved Rate-Distortion (RD) performance, notably outperforming the MPEG PC coding standards, particularly for geometry coding of solid static PCs.



Learning-based Coding Challenge: Model Complexity

- DL-based codecs have become state-of-the-art in many fields, including image and PC coding.

- Nevertheless, their high model complexity presents challenges for decoding in resource-constrained devices, such as smartphones.

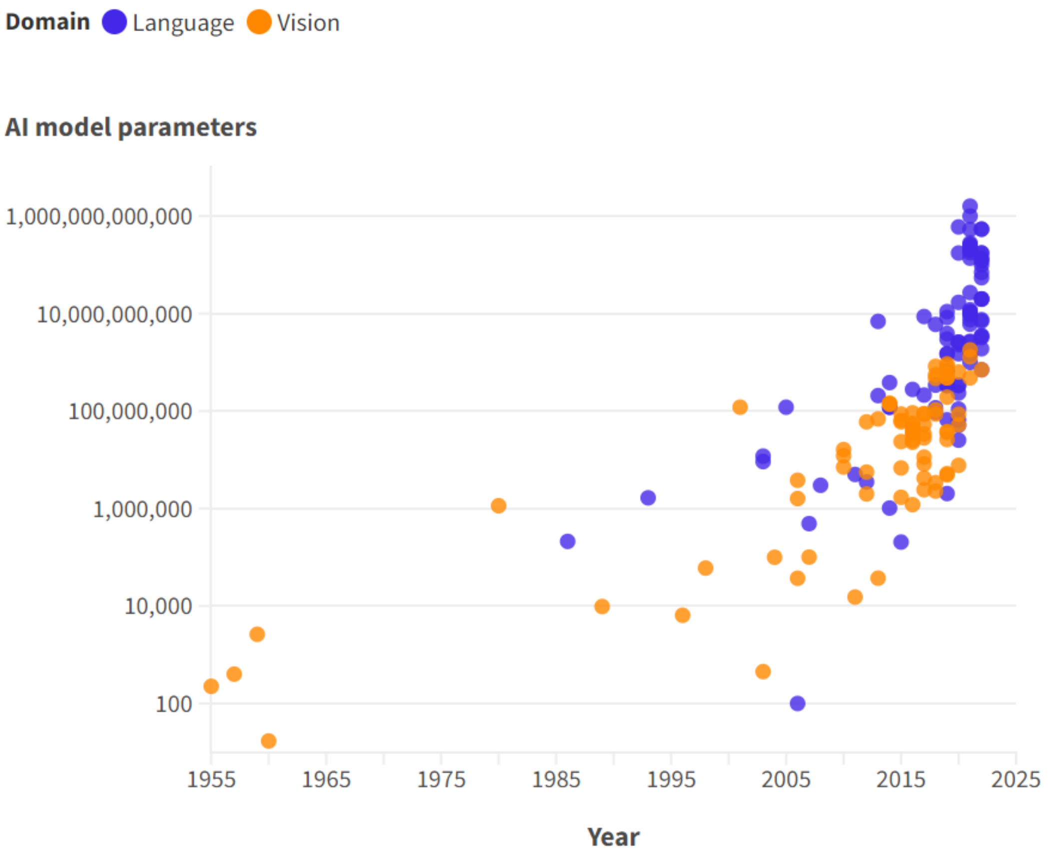
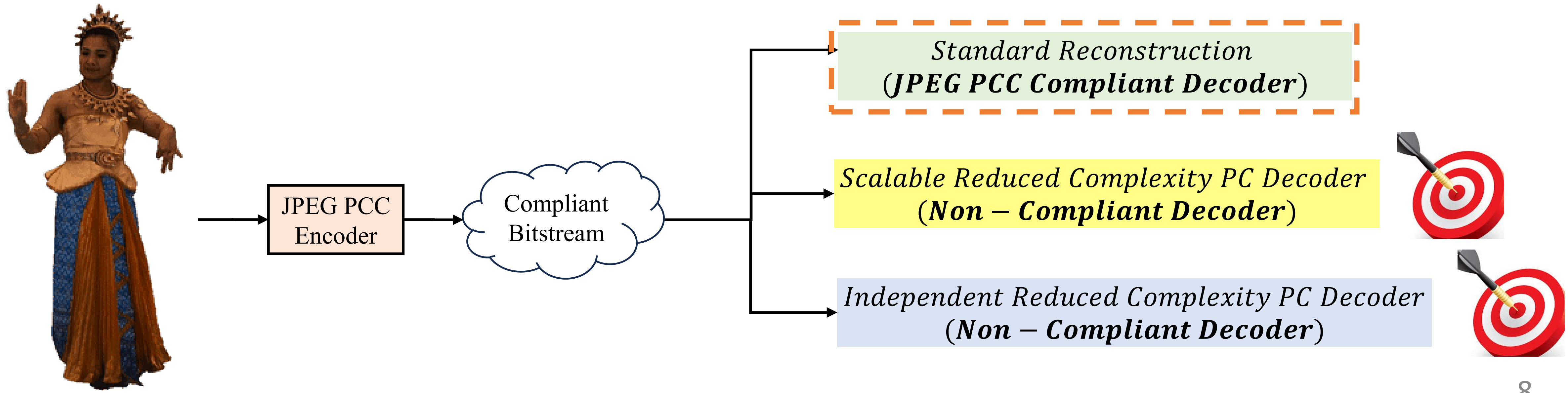


Image source: https://www.worldbank.org/

Objective

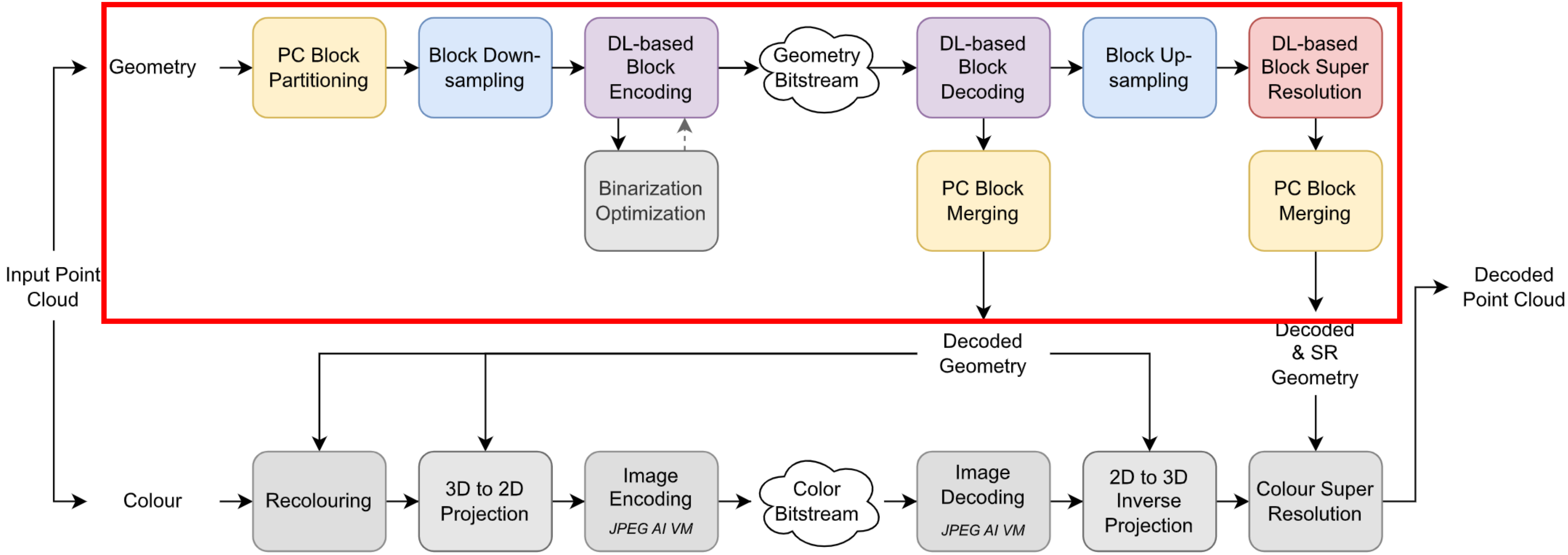
Propose two novel reduced complexity decoding solutions, one of them based on the adoption of scalability principles, to decode the same JPEG PCC compliant bitstream with minimal impact on RD performance.



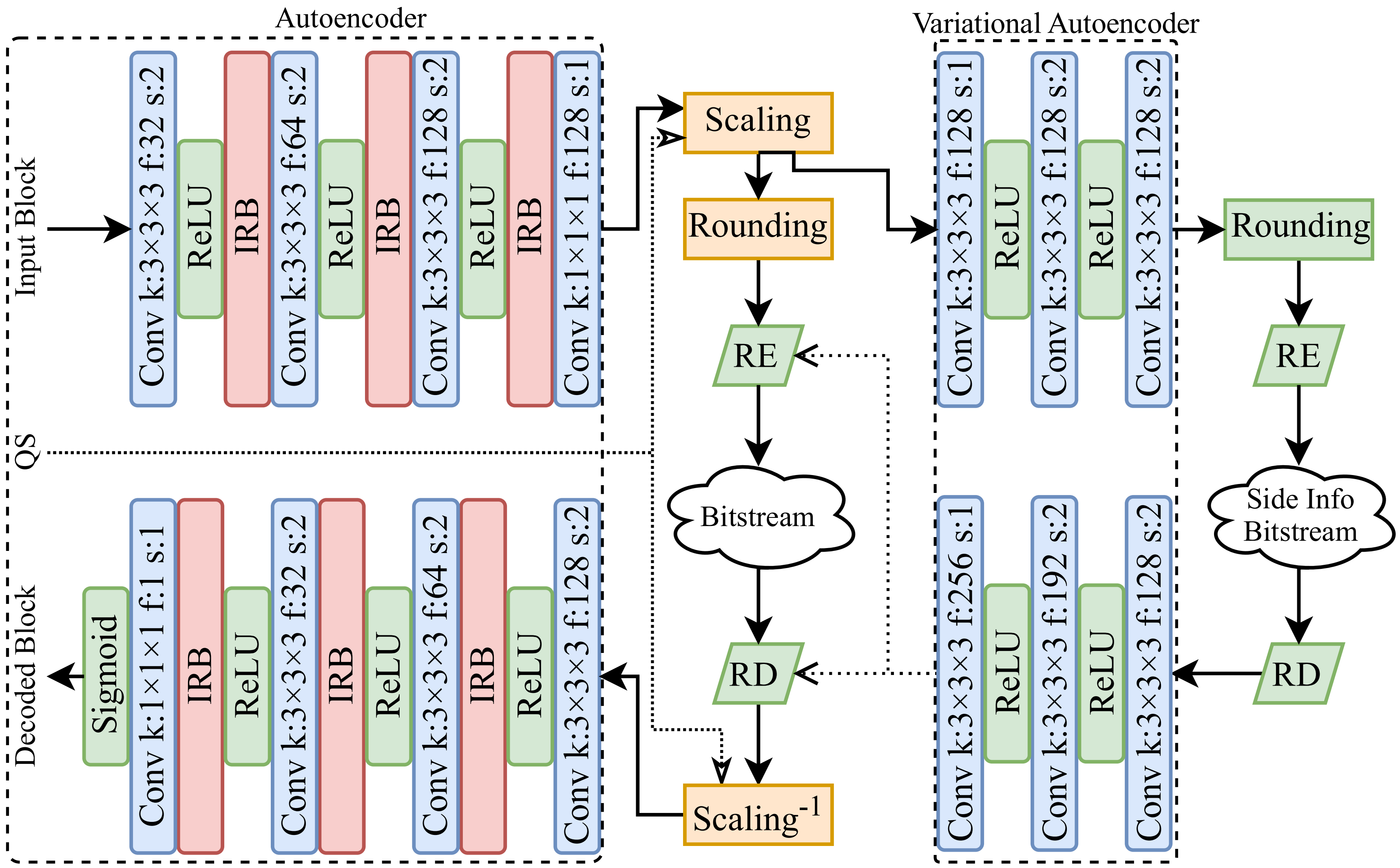
2

JPEG PCC Standard

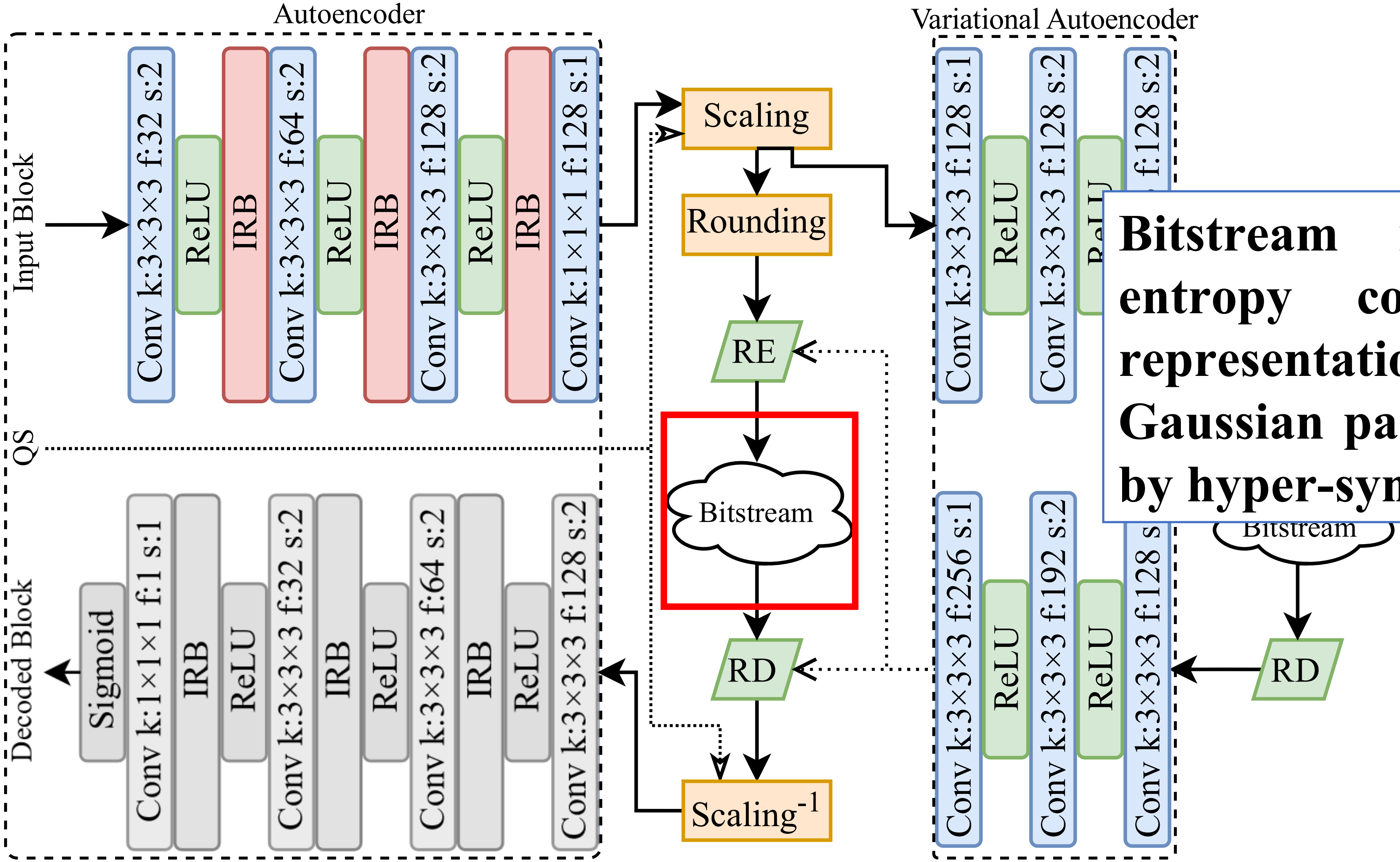
JPEG PCC Standard Architecture: Geometry-Only



JPEG PCC DL-based Coding Model



JPEG PCC DL-based Coding Model: Bitstream

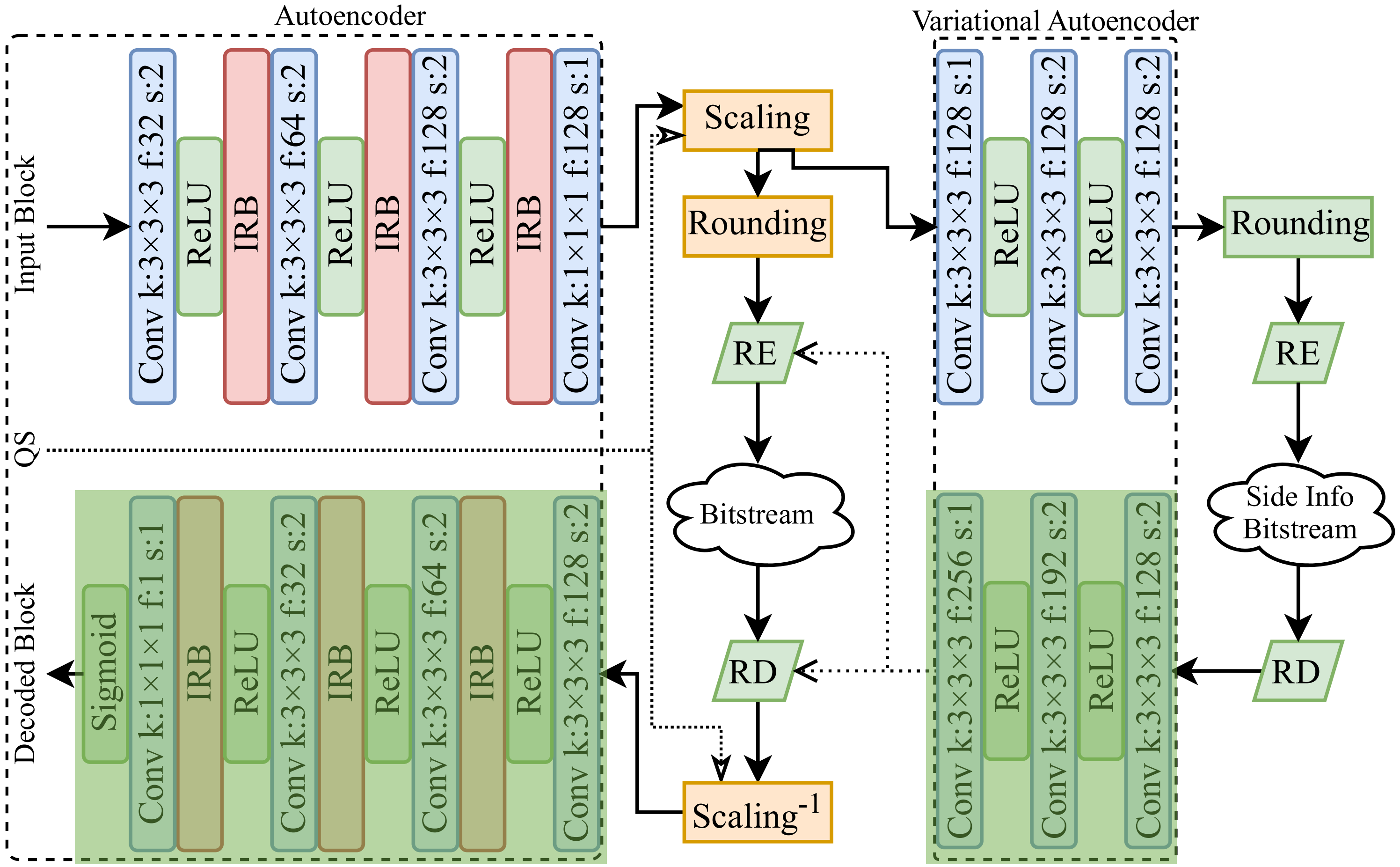


Bitstream is generated by entropy coding the latent representations conditioned by Gaussian parameters produced by hyper-synthesis transform.

RE: Range Encoder

RD: Range Decoder

JPEG PCC Decoding Model Complexity



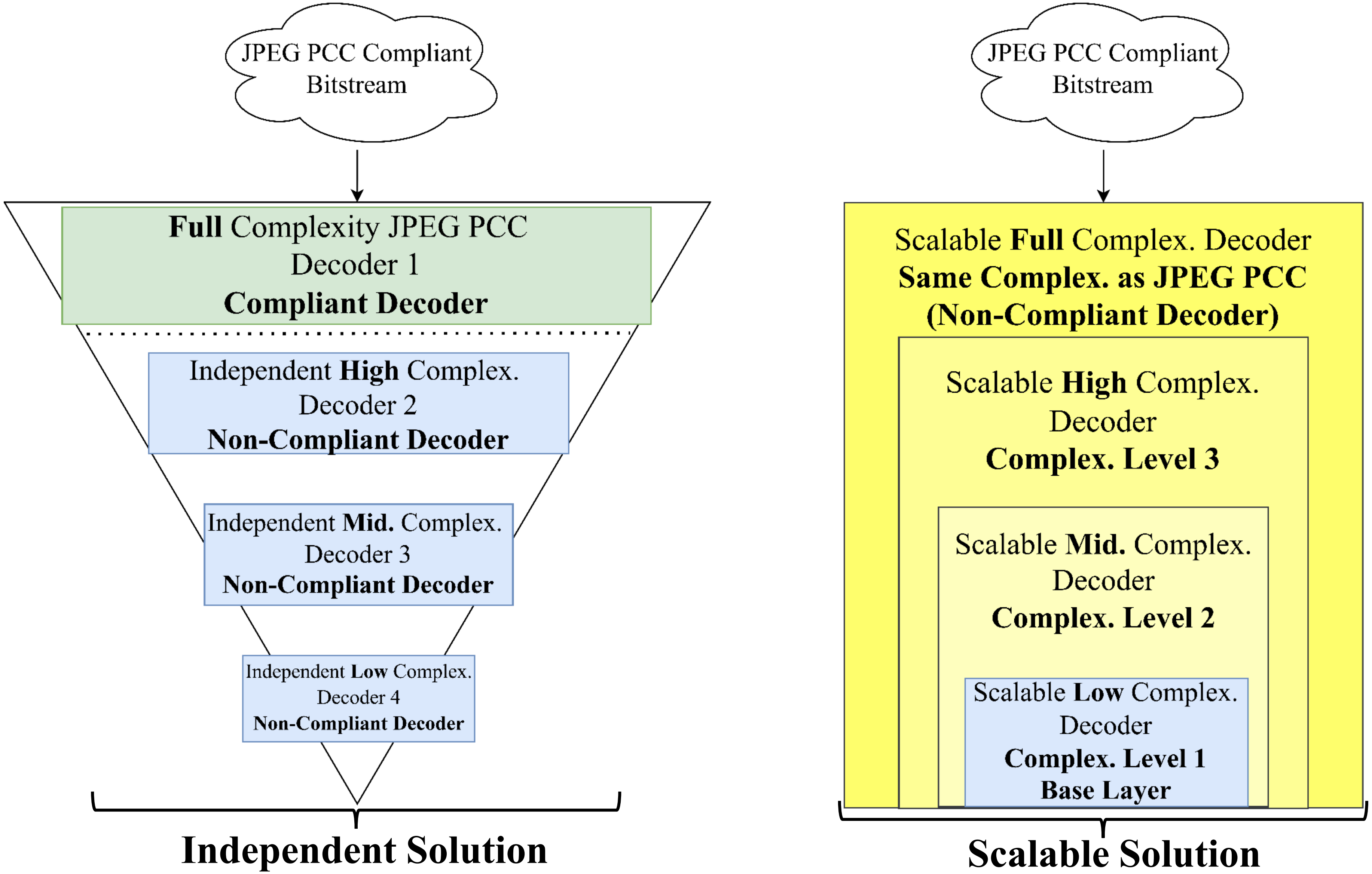
- ★ The decoding model complexity in terms of number of parameters is 2783089 (~2.8M)!
- ★ Devices with limited computational resources may struggle to deal with this level of complexity.

3 Proposed Reduced Complexity PC Decoding Solutions

instituto de telecomunicações

3.1 Lower Complexity Decoding Solutions: Definition

Proposed Reduced Complexity Decoding Solutions

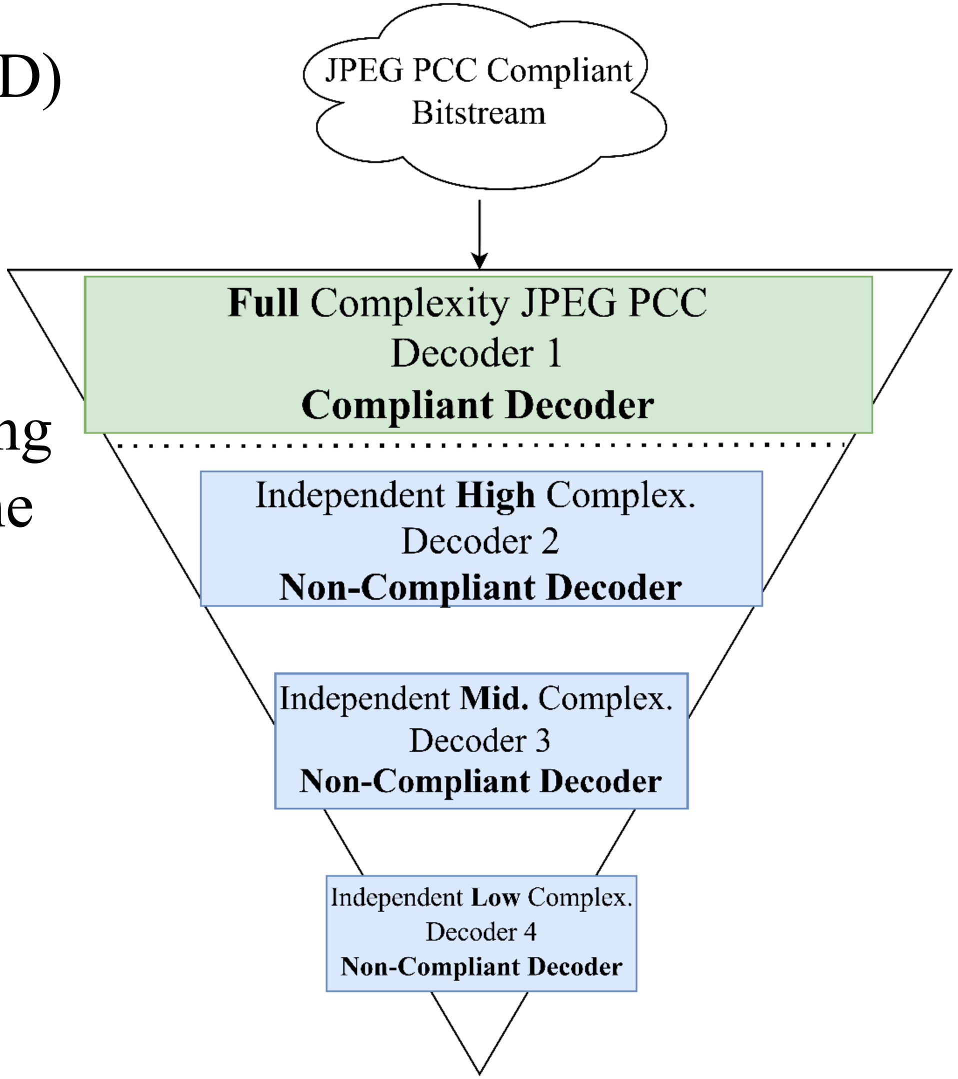


Independent Complexity JPEG PC Decoding Solution

- **I**ndependent **C**omplexity **JPEG** **PC** **D**ecoding (IC-JPCD) provides multiple independent reduced complexity decoder models.

- Independent reduced complexity is achieved by reducing a given percentage of the parameters associated with the filters in each layer of the JPEG PCC synthesis transform.

- Each independent model offers a trade-off between the complexity and RD performance.

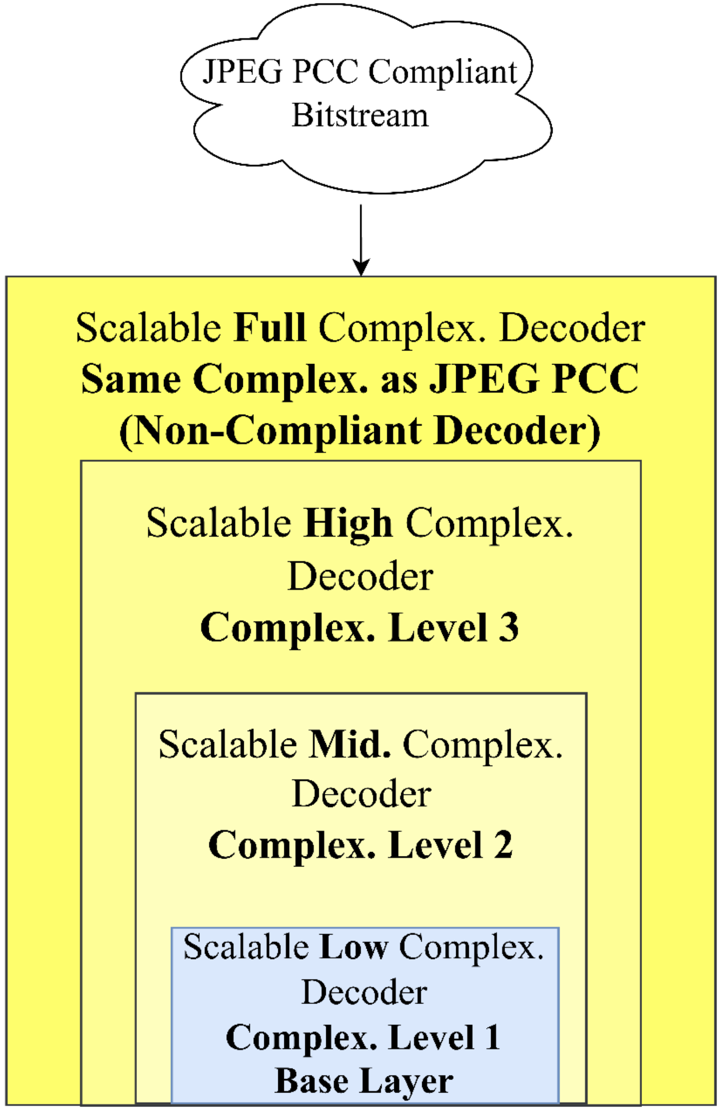


Scalable Complexity JPEG PC Decoding Solution

- Scalable **C**omplexity **J**PEG **PC D**ecoding (SC-JPCD) uses one single scalable decoder model, designed with a hierarchical structure of interdependent layers over a base layer.

- Scalability is achieved by progressively learning a set of additional parameters of the synthesis transform for each level, while reusing the parameters of the previous complexity level.

- Consumes less memory compared to the previous independent approach as it requires only one full model for all decoding complexity levels.



instituto de
telecomunicações

3.2 Lower Complexity Decoding Solutions: Design

Designing Varying Complexity PC Decoders: A Three Phases Process

The designing of the two strategies for the two reduced complexity decoding solutions, independent and scalable, includes three phases:

Regularization: Objective and Process

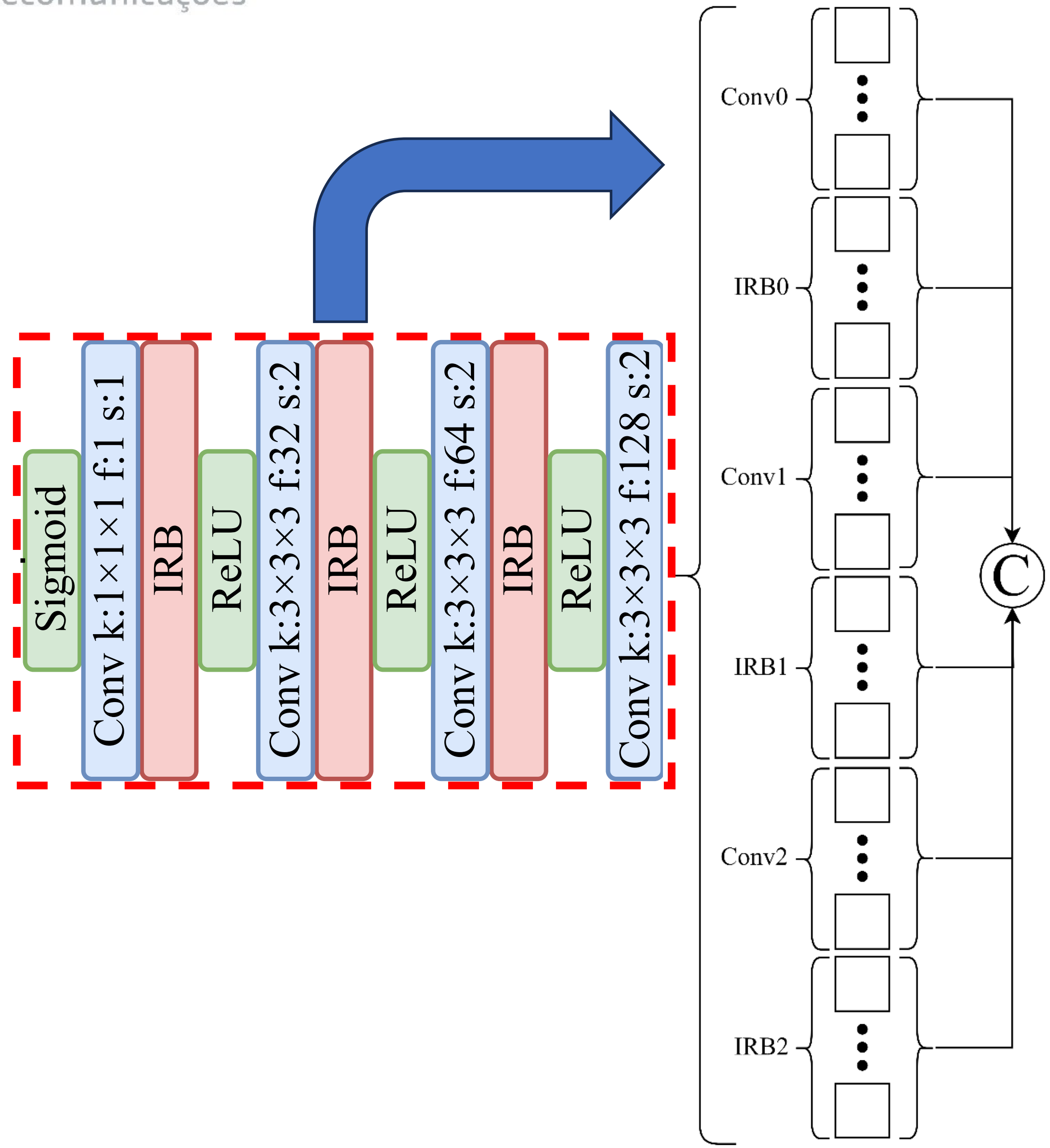
- The model parameters vary in their impact on RD performance, with some having more influence than others.
- The objective of the regularization step is to determine the filters in each layer that contribute the least to the RD performance by minimizing the norm of the parameters associated with those filters.
- The initial JPEG PCC models are trained using a new RD loss function integrating an L1 regularization term.
- $Loss = \underbrace{Distortion + \lambda * Coding\ Rate}_{\text{RD Loss function}} + \underbrace{\rho * \sum |w|}_{\text{L1 regularization}}$

instituto de
telecomunicações

Pruning: Objective and Process

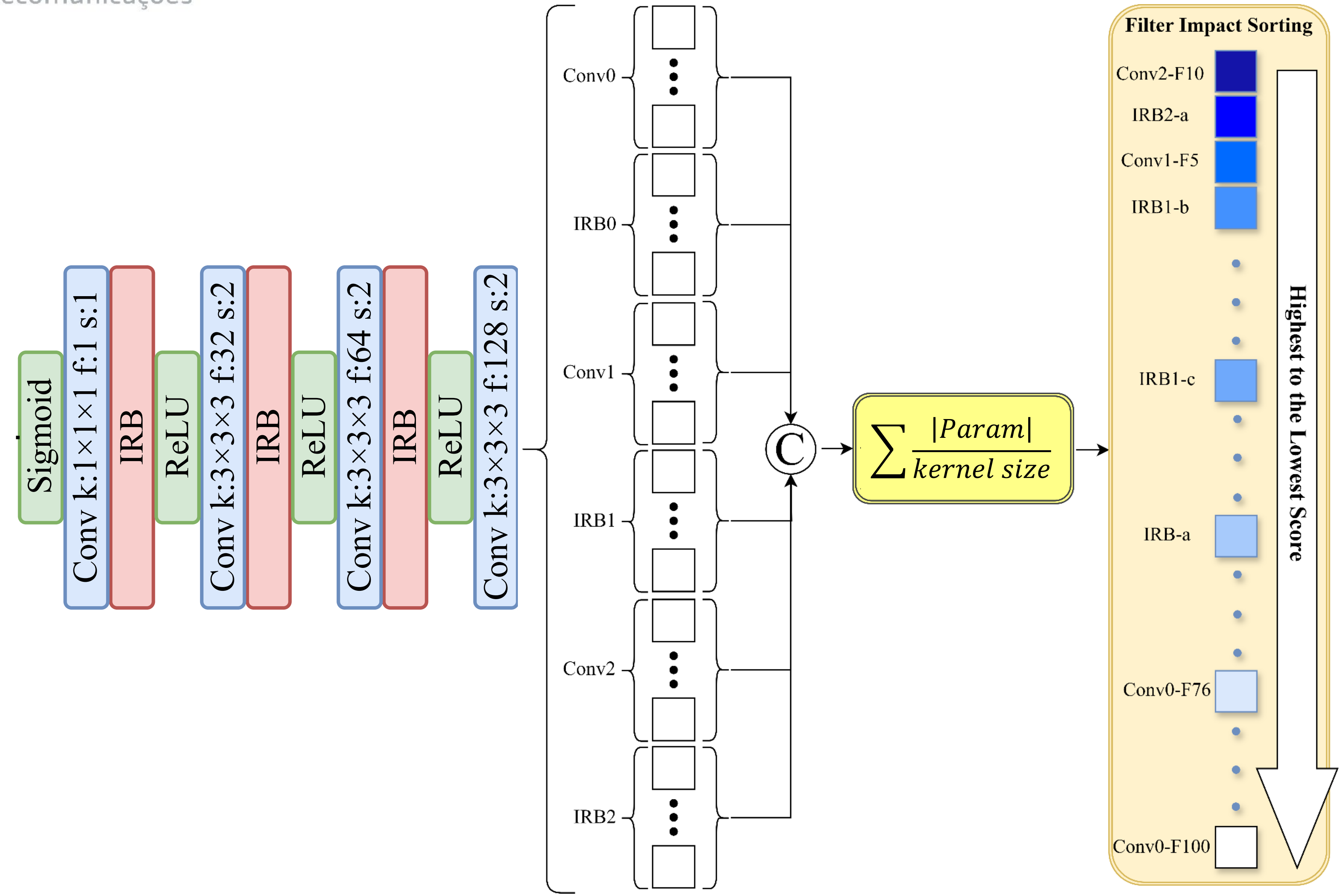
- The objective of the pruning is to remove the least important filters from each layer of synthesis transform.
- Global structured pruning is adopted, which means that individual parameters are not pruned, but rather the entire filters over all the layers of synthesis transform.
- Consequently, each layer may have a different number of pruned filters depending on their impact on the RD performance.

Pruning Procedure

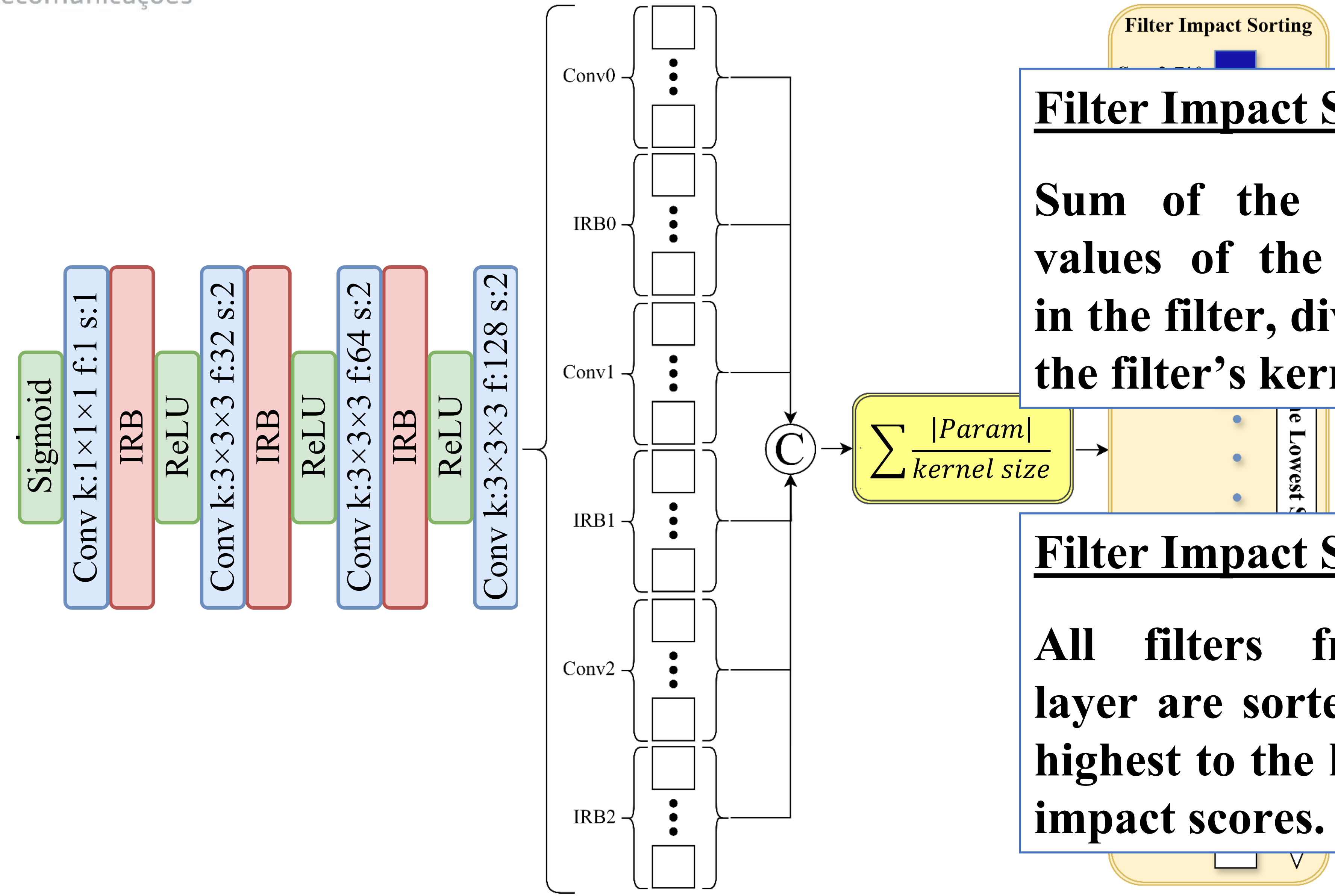


Filters are gathered from each layer of synthesis transform.

Pruning Procedure



Pruning Procedure



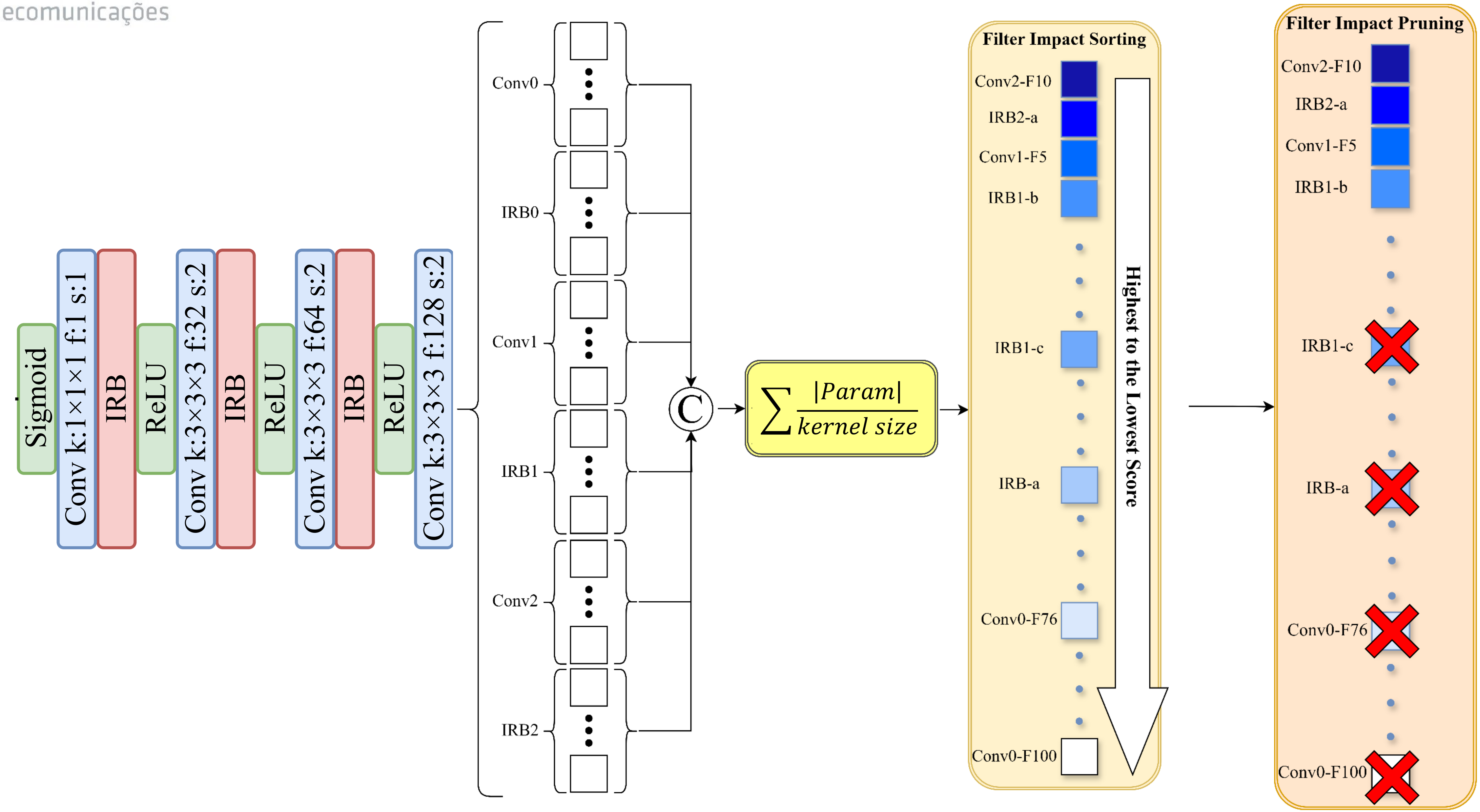
<u>Filter Impact Scoring:</u>

Sum of the absolute values of the weights in the filter, divided by the filter's kernel size.

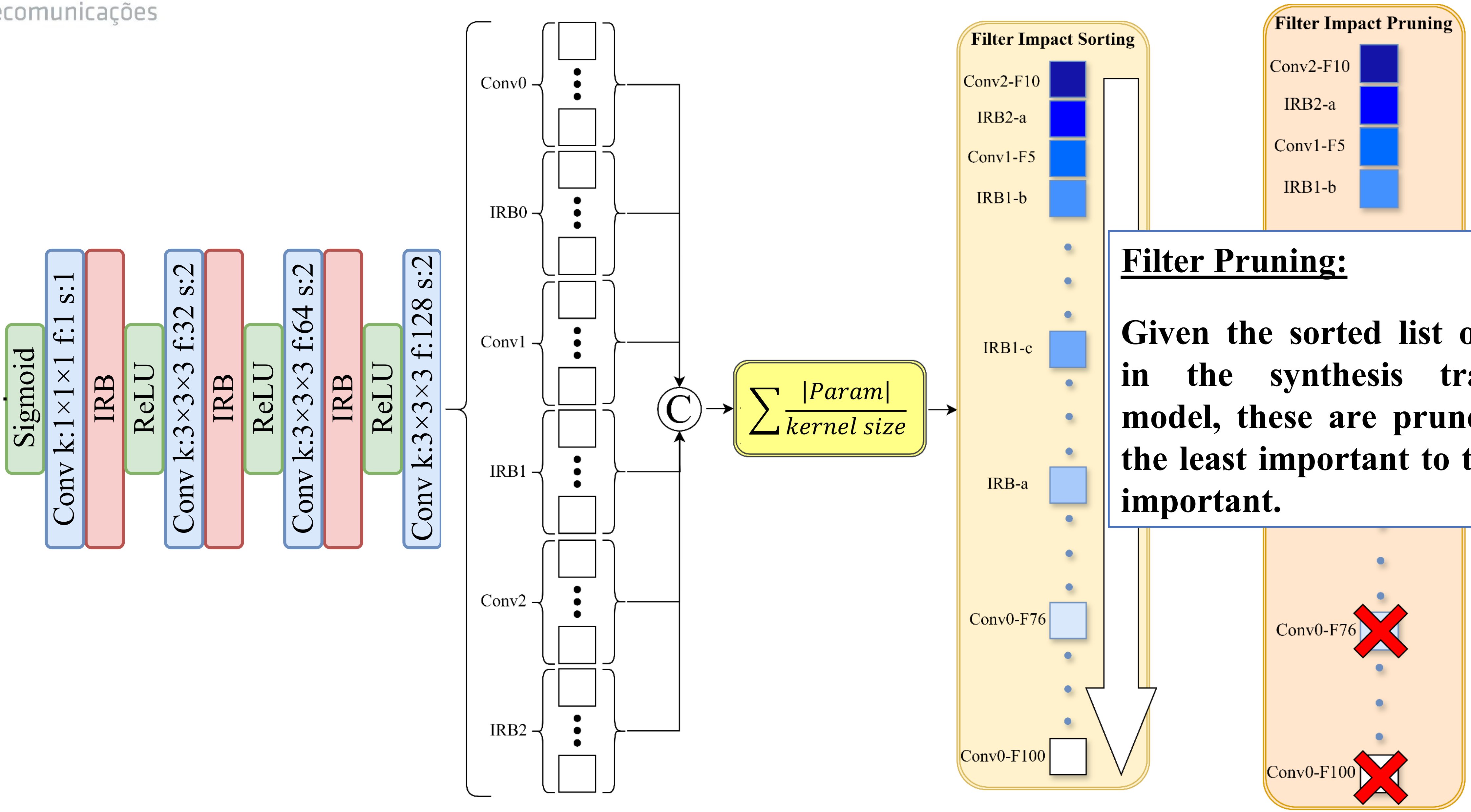
<u>Filter Impact Sorting:</u>

All filters from every layer are sorted from the highest to the lowest filter impact scores.

Pruning Procedure



Pruning Procedure



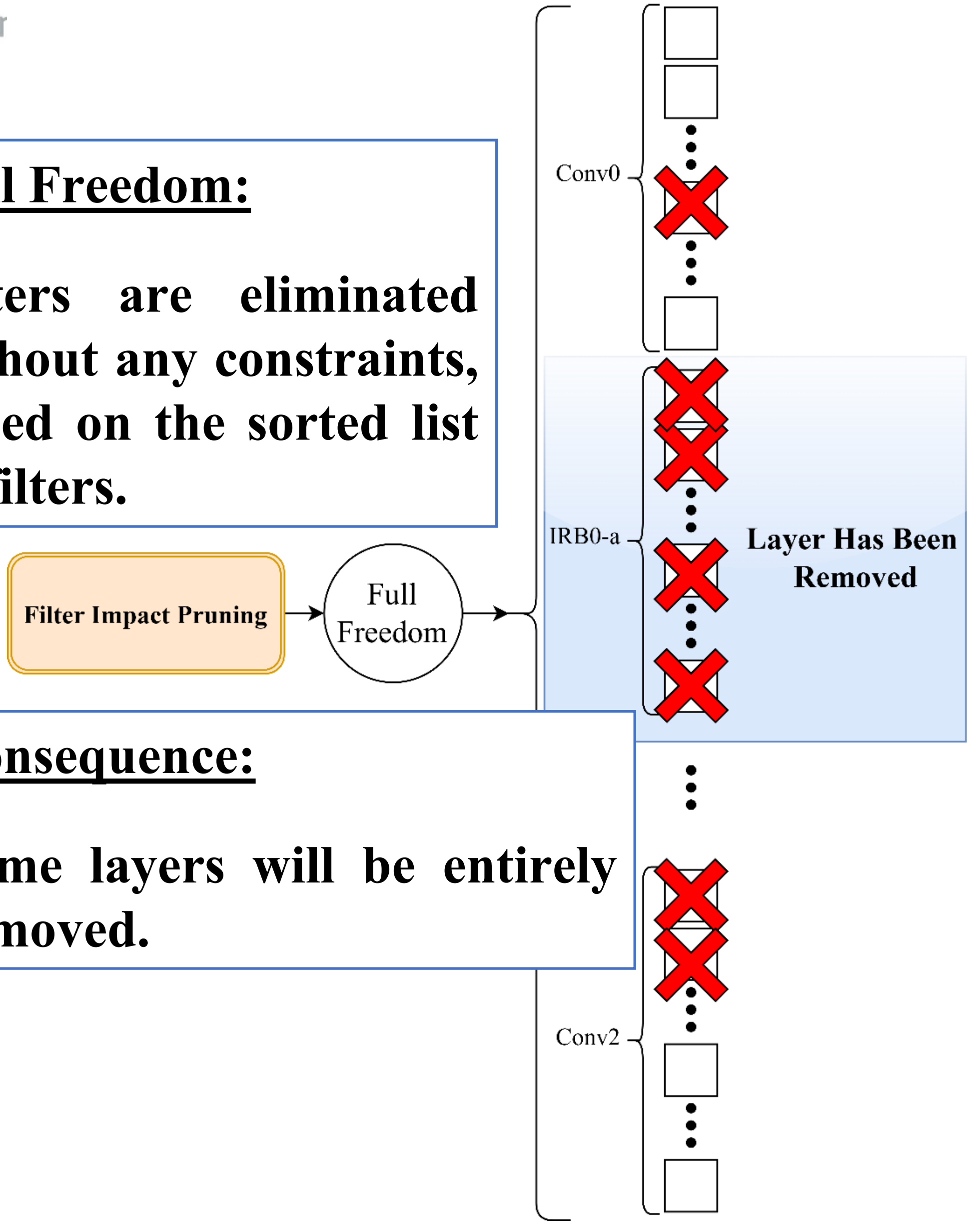
<u>Filter Pruning:</u>

Given the sorted list of filters in the synthesis transform model, these are pruned from the least important to the most important.

Pruning Strategies: Full Freedom vs Layer Restriction

Full Freedom:

Filters are eliminated without any constraints, based on the sorted list of filters.

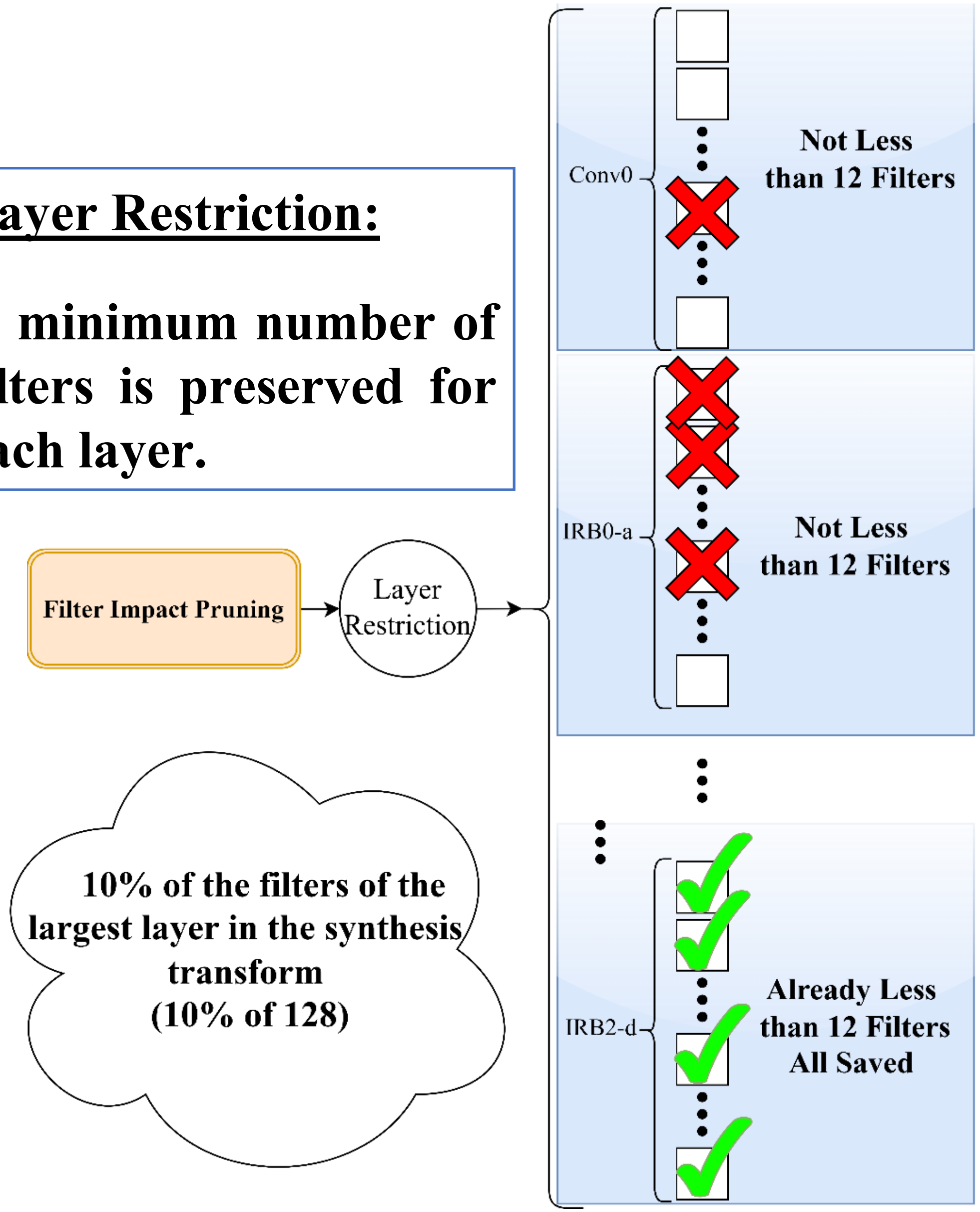


Consequence:

Some layers will be entirely removed.

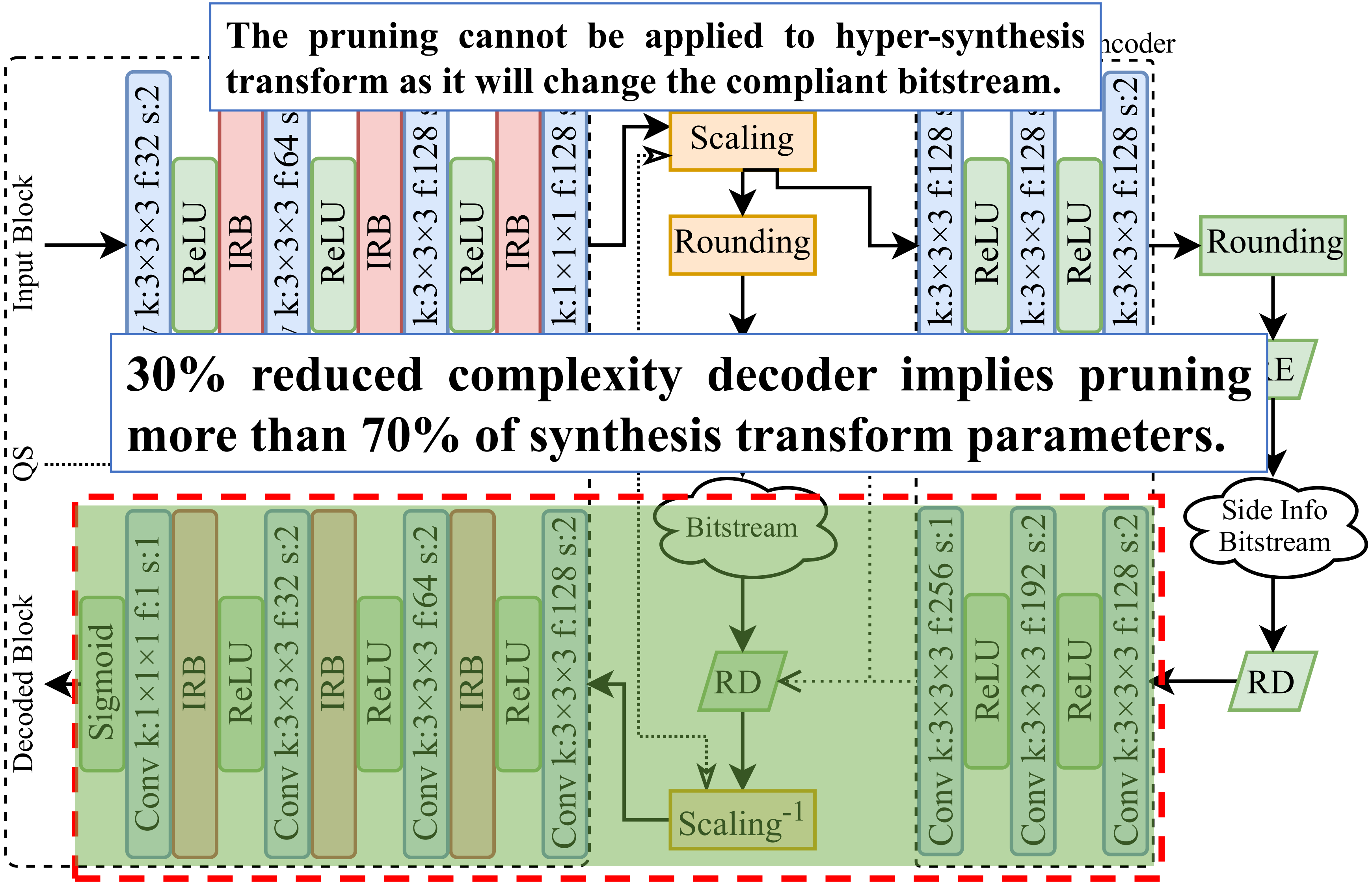
Layer Restriction:

A minimum number of filters is preserved for each layer.



Full Decoder Model vs Synthesis Transform Model

The pruning cannot be applied to hyper-synthesis transform as it will change the compliant bitstream.



30% reduced complexity decoder implies pruning more than 70% of synthesis transform parameters.

instituto de telecomunicações

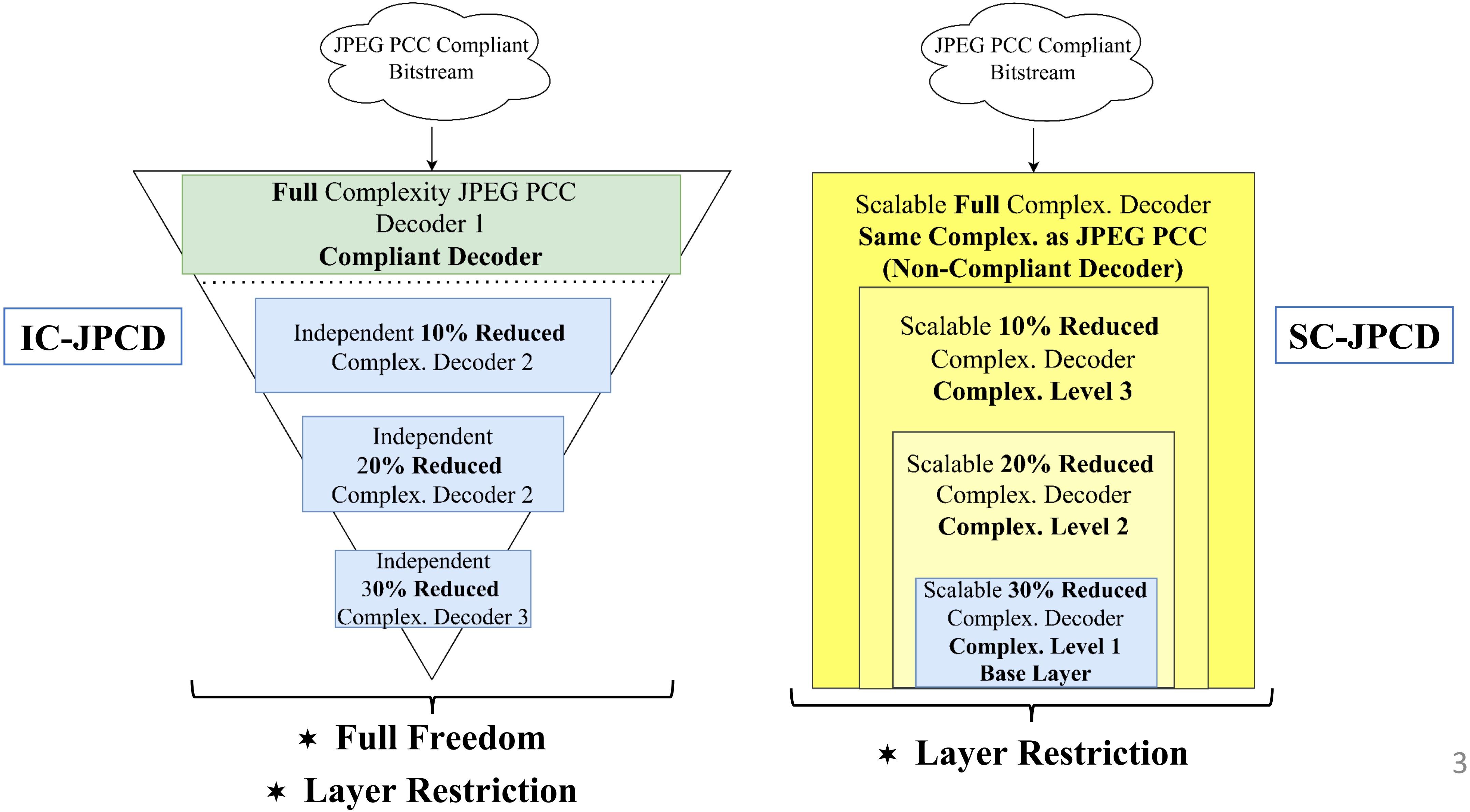
Training: Objective and Process

- The objective is to train the new pruned architecture (per target rate) to achieve the best possible RD performance.

- For each **<u>independent</u>** decoder layer:
  - Encoder part is initialized with JPEC PCC encoder model which is kept frozen.
  - Then, the pruned synthesis transform architecture is randomly initialized and trained.

- For each **<u>scalable</u>** decoder layer:
  - For scalable decoder base layer, the training approach is the same as for the independent solution.
  - All the synthesis transform filters from the previous scalable layer are initialized and kept frozen.
  - Then, the new filters (depending on the target percentage) are trained with random initialization.

Lower Complexity Non-Compliant Decoders



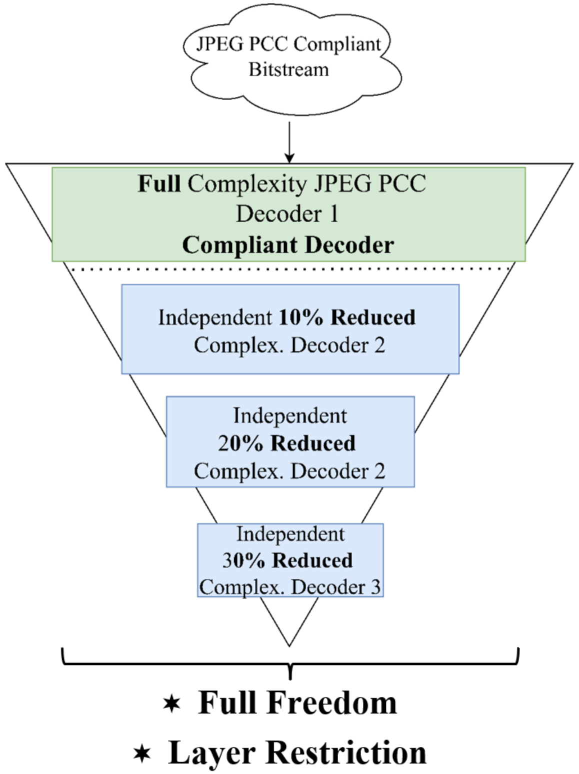
4 Performance Assessment

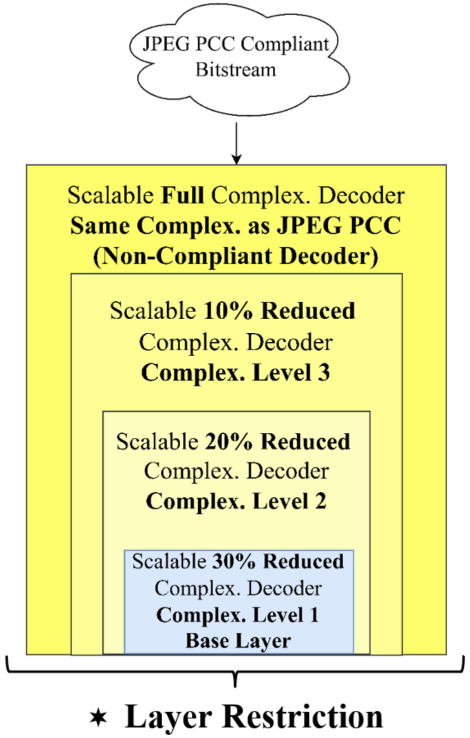
instituto de telecomunicações

JPEG Pleno PC Test Dataset

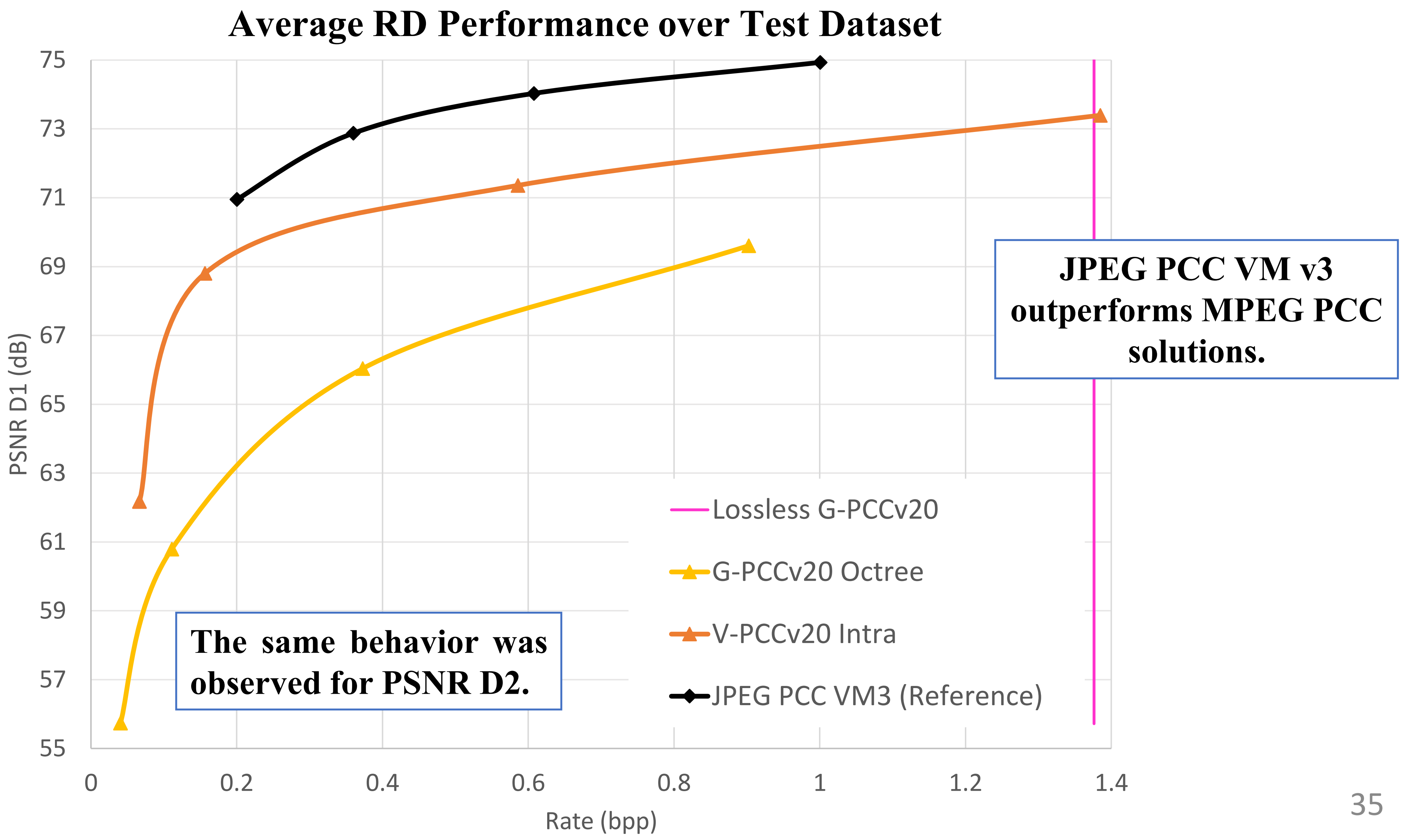
Decoding Solutions Under Comparison

- **Anchors:**
  - JPEG PCC VM v3
  - MPEG G-PCC Octree v20
  - MPEG V-PCC Intra v20
- **Independent Decoding Solution**
  - IC-JPCD: 30% (Layer Restriction, Full Freedom)
  - IC-JPCD: 20% (Layer Restriction, Full Freedom)
  - IC-JPCD: 10% (Layer Restriction, Full Freedom)
- **Scalable Decoding Solution**
  - SC-JPCD: 30% (Layer Restriction)
  - SC-JPCD: 20% (Layer Restriction)
  - SC-JPCD: 10% (Layer Restriction)

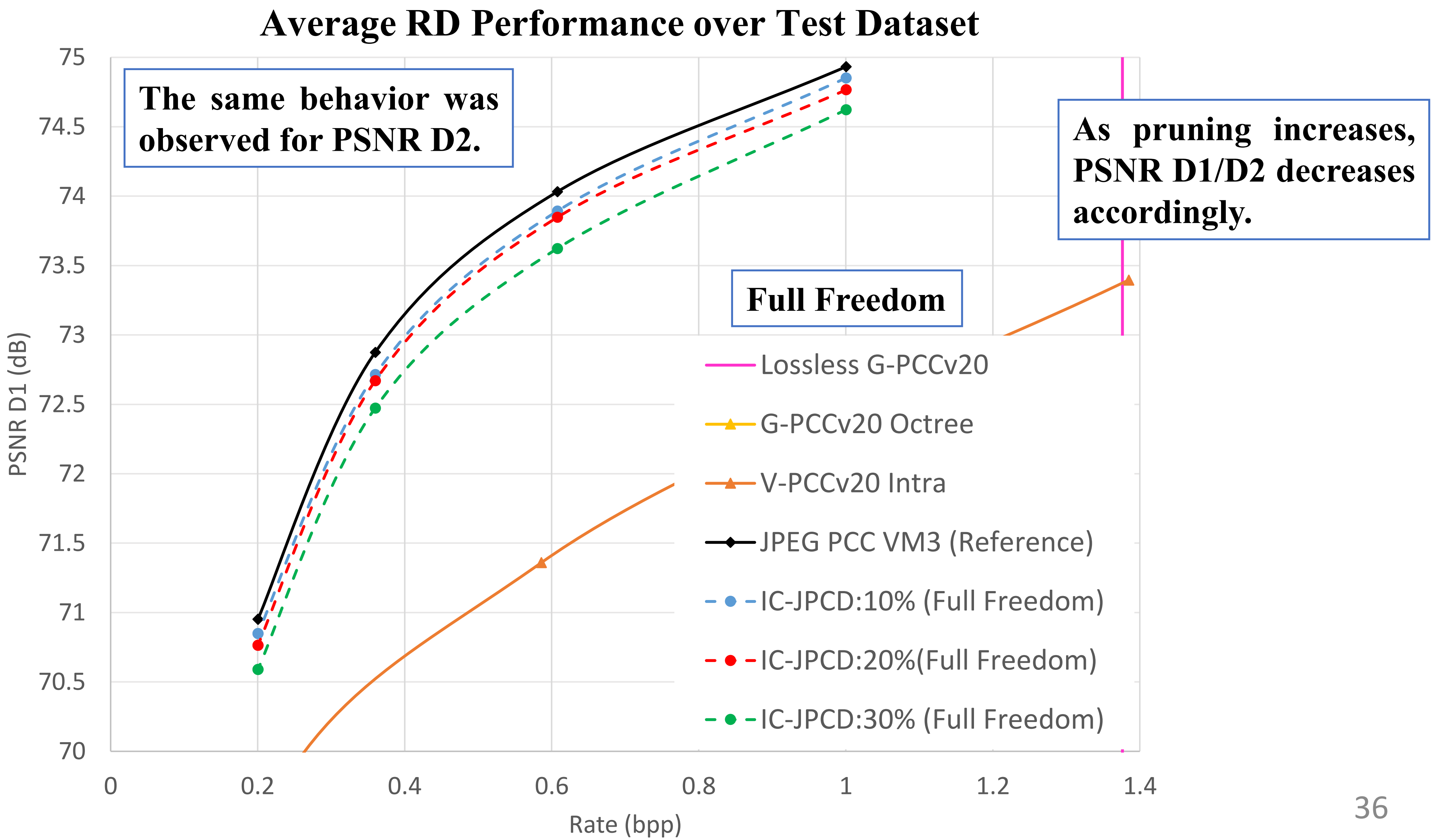




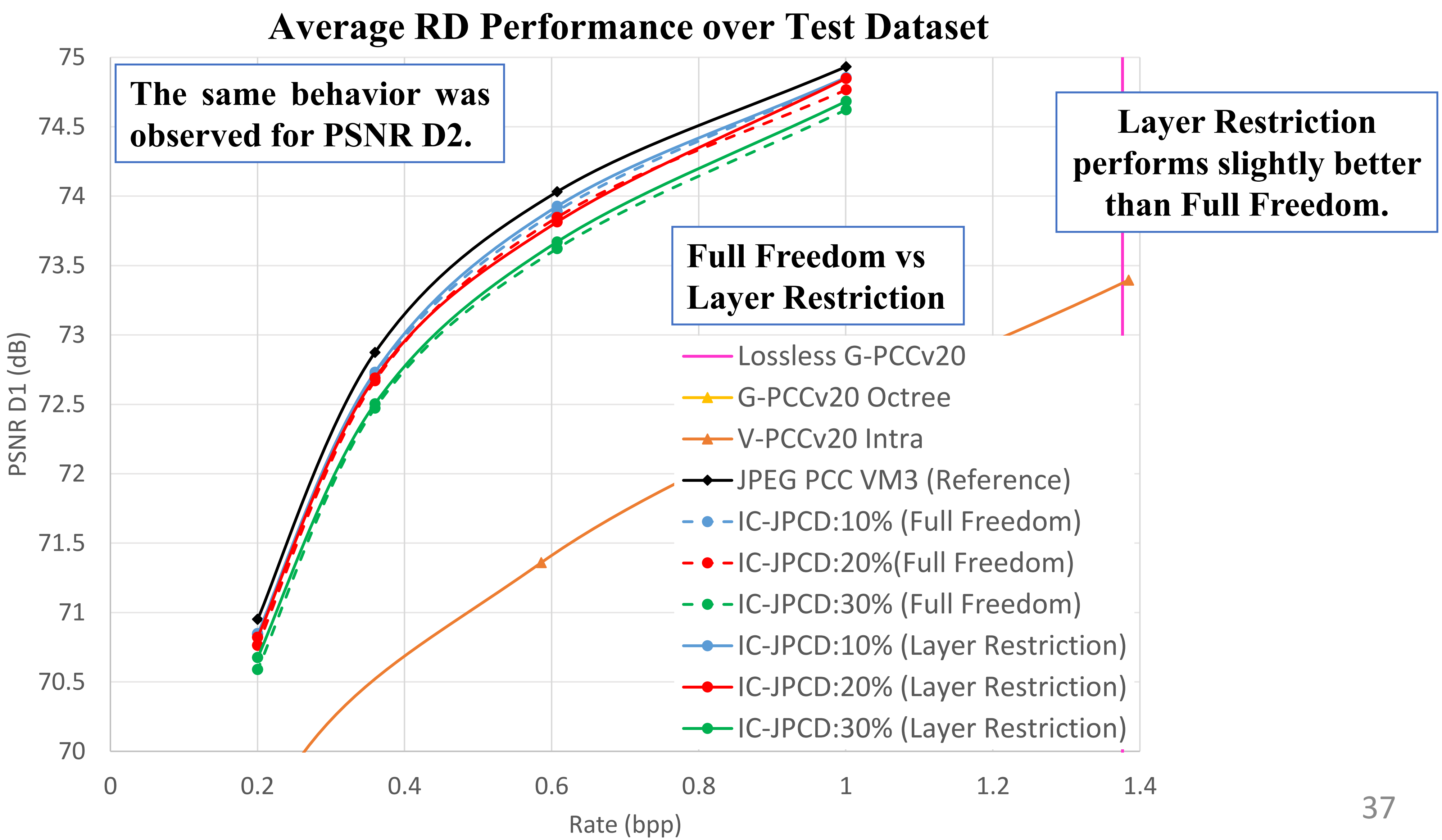
RD Performance: JPEG PCC vs MPEG PCC



RD Performance: Independent (IC-JPCD) Solution



RD Performance: Independent (IC-JPCD) Solution



BD-Rate, BD-PSNR and Complexity Comparison

JPEG PCC vs Independent Reduced Complexity Non-Compliant Decoders (avg. over all test dataset)

Solution	Pruning Strategy	PSNR D1		PSNR D2		Nº of Model Parameters and Memory Size			Decoding Time (GPU)
		BD-Rate	BD-PSNR	BD-Rate	BD-PSNR	Synthesis Transform	Full Decoder	All Models	
IC-JPCD: 10%	Full Freedom	4.8%	-0.14	5.3%	-0.18	-23.51%	-9.51%	+240.22%	-10.6%
IC-JPCD: 20%		6.8%	-0.19	7.8%	-0.25	-50.15%	-20.32%		-12.93%
IC-JPCD: 30%		14.3%	-0.39	14.1%	-0.45	-74.18%	-30.05%		-15.16%
IC-JPCD: 10%	Layer Restriction	4.1%	-0.12	5.8%	-0.18	-24.80%	-10.06%	+240.85%	-7.75%
IC-JPCD: 20%		6.9%	-0.18	7.0%	-0.22	-49.67%	-20.13%		-9.91%
IC-JPCD: 30%		13.1%	-0.34	12.5%	-0.39	-71.47%	-28.96%		-11.88%

instituto de telecomunicações

BD-Rate, BD-PSNR and Complexity Comparison

JPEG PCC vs Independent Reduced Complexity Non-Compliant Decoders (avg. over all test dataset)

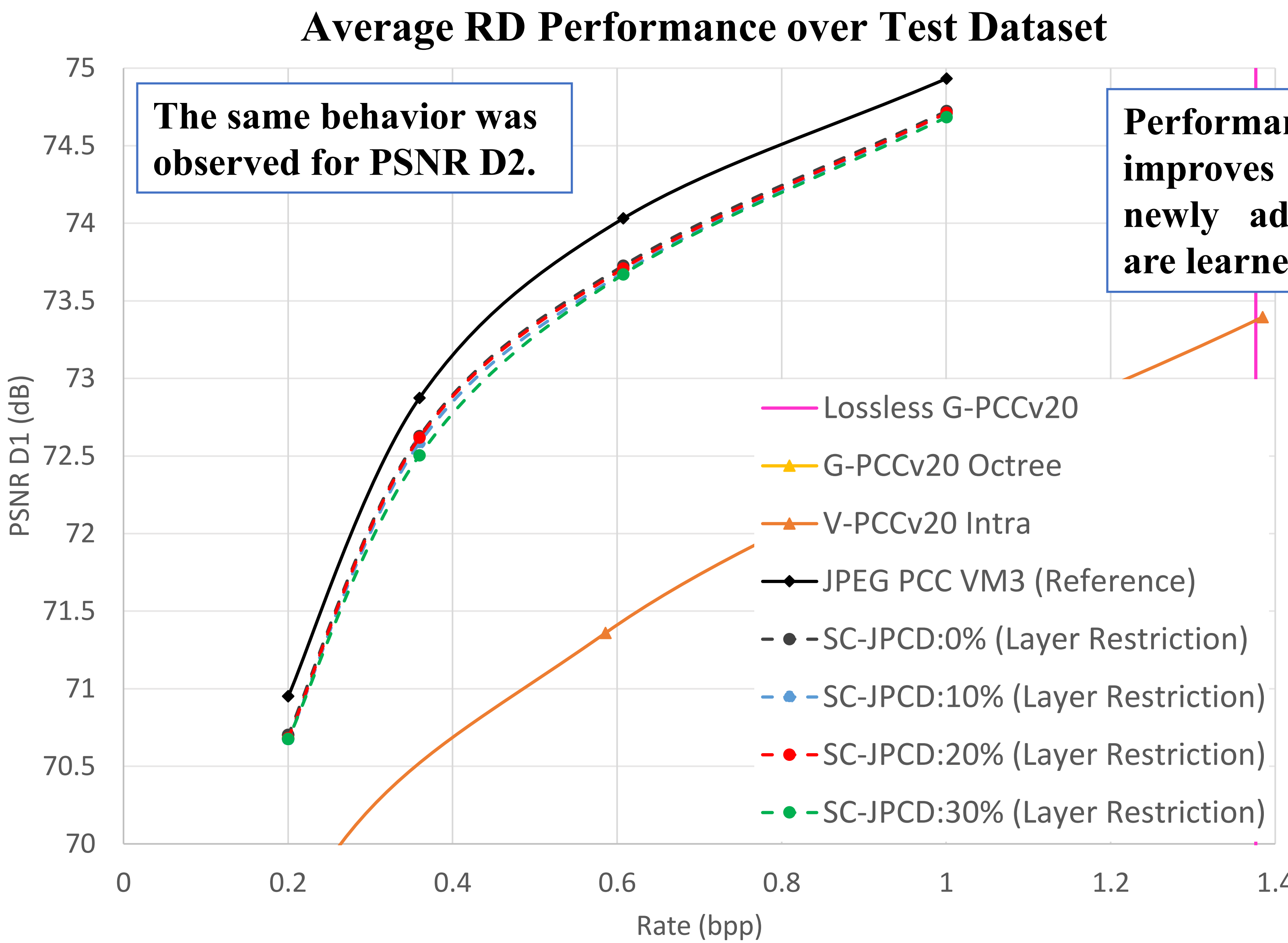
Smaller penalty in RD performance.

Lower complexity reduction.

Lower decoding time reduction.

Solution	Pruning Strategy	PSNR D1		PSNR D2		Nº of Model Parameters and Memory Size			Decoding Time (GPU)
		BD-Rate	BD-PSNR	BD-Rate	BD-PSNR	Synthesis Transform	Full Decoder	All Models	
IC-JPCD: 10%	Full Freedom	4.8%	-0.14	5.3%	-0.18	-23.51%	-9.51%	+240.22%	-10.6%
IC-JPCD: 20%		6.8%	-0.19	7.8%	-0.25	-50.15%	-20.32%		-12.93%
IC-JPCD: 30%		14.3%	-0.39	14.1%	-0.45	-74.18%	-30.05%		-15.16%
IC-JPCD: 10%	Layer Restriction	4.1%	-0.12	5.8%	-0.18	-24.80%	-10.06%	+240.85%	-7.75%
IC-JPCD: 20%		6.9%	-0.18	7.0%	-0.22	-49.67%	-20.13%		-9.91%
IC-JPCD: 30%		13.1%	-0.34	12.5%	-0.39	-71.47%	-28.96%		-11.88%

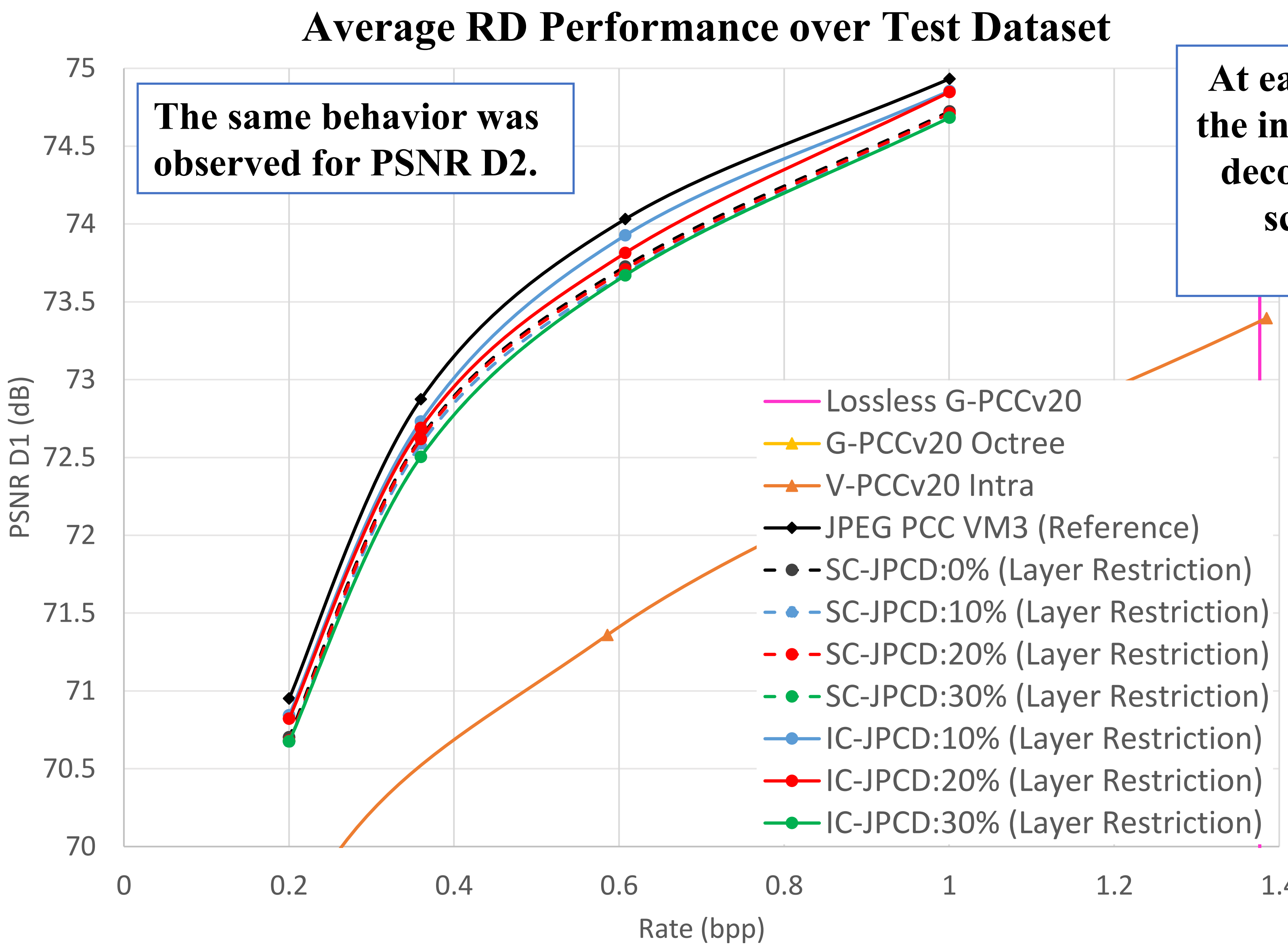
RD Performance: Scalable (SC-JPCD) Solution



The same behavior was observed for PSNR D2.

Performance slightly improves as the newly added filters are learned.

RD Performance: IC-JPCD vs SC-JPCD Solution



At each target complexity, the independent complexity decoder outperforms the scalable complexity decoder.

instituto de telecomunicações

BD-Rate, BD-PSNR and Complexity Comparison

JPEG PCC vs Scalable Reduced Complexity Non-Compliant Decoders (avg. over all test dataset)

Smaller penalty in RD performance.

Lower complexity reduction.

Lower decoding time reduction.

Solution	Pruning Strategy	PSNR D1		PSNR D2		Nº of Model rameters and Memory Siz			Decoding Time (GPU)
		BD-Rate	BD-PSNR	BD-Rate	BD-PSNR	Synthesis Transform	Full Decoder	All Models	
SC-JPCD: 0%	Layer Restriction	9.5%	-0.26	9.4%	-0.30	0%	0%	0%	-6.67%
SC-JPCD: 10%	Layer Restriction	10.9%	-0.30	11%	-0.36	-24.80%	-10.06%	0%	-7.54%
SC-JPCD: 20%	Layer Restriction	9.9%	-0.28	10.0%	-0.33	-49.67%	-20.13%	0%	-9.35%
SC-JPCD: 30%	Layer Restriction	13.1%	-0.34	12.5%	-0.39	-71.47%	-28.96%	0%	-11.88%

instituto de telecomunicações

BD-Rate, BD-PSNR and Complexity Comparison

The price of having a scalable complexity decoder.

Solution	Pruning Strategy	PSNR D1		PSNR D2		Nº of Model Parameters and Memory Size			Decoding Time (GPU)
		BD-Rate	BD-PSNR	BD-Rate	BD-PSNR	Synthesis Transform	Full Decoder	All Models	
SC-JPCD: 0%	Layer Restriction	**9.5%**	**-0.26**	**9.4%**	**-0.30**	**0%**	**0%**	0%	-6.67%
SC-JPCD: 10%		10.9%	-0.30	11%	-0.36	-24.80%	-10.06%		-7.54%
SC-JPCD: 20%		9.9%	-0.28	10.0%	-0.33	-49.67%	-20.13%		-9.35%
SC-JPCD: 30%		13.1%	-0.34	12.5%	-0.39	-71.47%	-28.96%		-11.88%

instituto de telecomunicações

BD-Rate, BD-PSNR and Complexity Comparison

JPEG PCC vs Lower Complexity Non-Compliant Decoders (avg. over all test dataset)

Solution	Pruning Strategy	PSNR D1		PSNR D2		Nº of Model Parameters and Memory Size			Decoding Time (GPU)
		BD-Rate	BD-PSNR	BD-Rate	BD-PSNR	Synthesis Transform	Full Decoder	All Models	
IC-JPCD: 10%	Full Freedom	4.8%	-0.14	5.3%	-0.18	-23.51%	-9.51%		-10.6%
IC-JPCD: 20%		6.8%	-0.19	7.8%	-0.25	-50.15%	-20.32%	+240.22%	-12.93%
IC-JPCD: 30%		14.3%	-0.39	14.1%	-0.45	-74.18%	-30.05%		-15.16%
IC-JPCD: 10%	Layer Restriction	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]5%		-7.75%
IC-JPCD: 20%		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]3%	+240.85%	-9.91%
IC-JPCD: 30%		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]5%		-11.88%
SC-JPCD: 0%	Layer Restriction	9.3%	-0.26	9.4%	-0.30	0%	0%		-6.67%
SC-JPCD: 10%		10.9%	-0.30	11%	-0.36	-24.80%	-10.06%		-7.54%
SC-JPCD: 20%		9.9%	-0.28	10.0%	-0.33	-49.67%	-20.13%	0%	-9.35%
SC-JPCD: 30%		13.1%	-0.34	12.5%	-0.39	-71.47%	-28.96%		-11.88%

Scalable complexity decoding offers a single full model with no memory size increase regarding the non-adaptive JPEG PCC model.

5 Conclusions

Conclusions

- ✶ The IC-JPCD and SC-JPCD decoders offer reduced independent and scalable complexity decoding solutions at the cost of a minimal RD performance penalty, while consuming the same JPEG PCC compliant bitstreams.
- ✶ Better RD performance could be achieved by re-training the full end-to-end model (including the encoder) for the target decoding complexity, but this would violate the compliance constraint for the bitstream.
- ✶ Scalable complexity decoders offer more versatility compared to the independent complexity decoders for the same memory size, as they can adapt based on available computational resources.

Thank you for your attention!

ACKNOWLEDGMENT
This work has been financially supported by the Fundação para a Ciência e a Tecnologia (FCT, Portugal) through the research project PTDC/EEI-COM/1125/2021, entitled "Deep Learning-based Point Cloud Representation".