Compression of sampled voltage and current values with Multiple-Models Coding scheme

C. Presvôts T. Prevost[1] M. Kieffer[2] P. Panciatici[1]

[1]RTE

[2]Laboratory of Signals and Systems (L2S)

NASPI Webinar 02/22/23

Outline

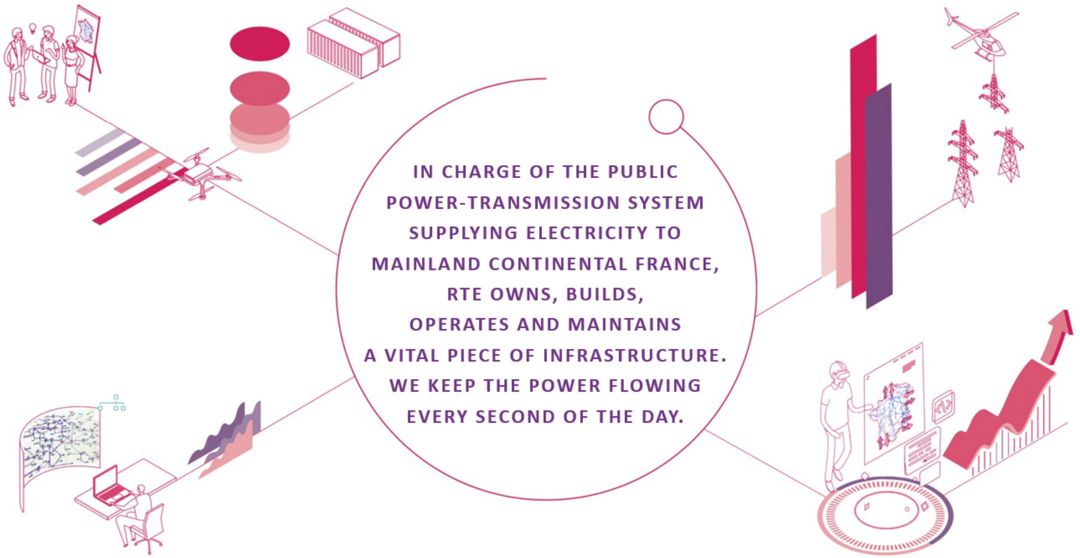
1. Context
   - RTE's presentation
2. Related work
   - Compression of PMU data
   - Sampled voltage and current values
3. Proposed approach
   - Parametric coding
   - Residual coding
   - Optimal rate-distortion
4. Results

Outline

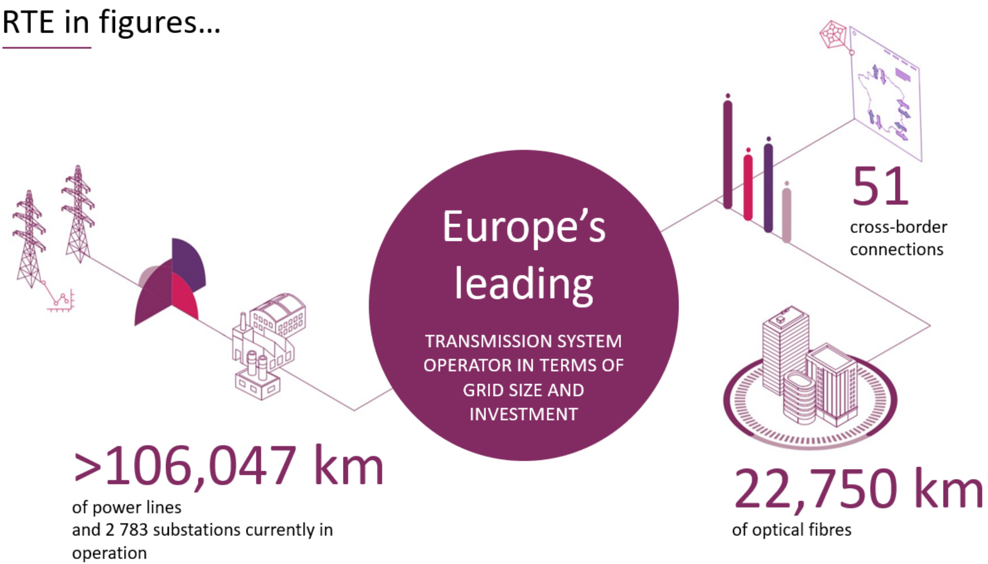
1. Context
    - RTE's presentation
2. Related work
    - Compression of PMU data
    - Sampled voltage and current values
3. Proposed approach
    - Parametric coding
    - Residual coding
    - Optimal rate-distortion
4. Results

Le réseau de transport d'électricité

Introduction to RTE

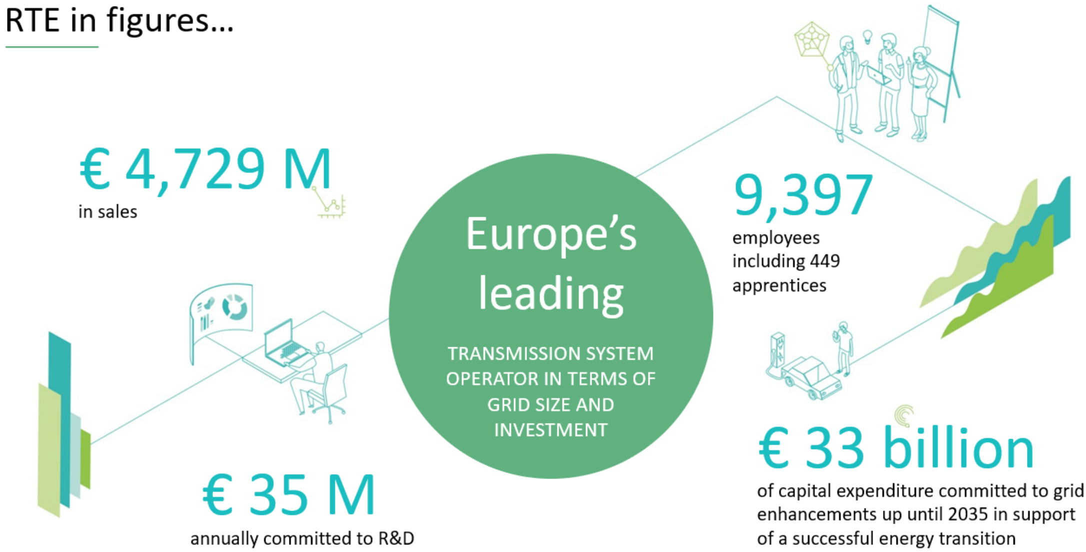


Introduction to RTE



Le réseau de transport d'électricité

Introduction to RTE



Times are changing for RTE

Massive integration of **renewable generation** resources into the electrical system

x3 in **10** years

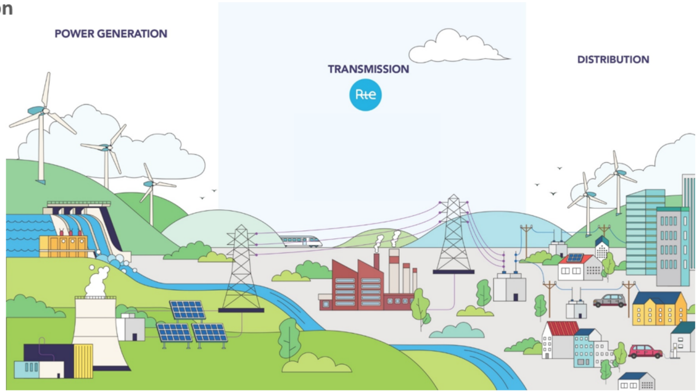
x5 in **15** years

(installed power)

Mainly (90%) on **distribution networks**

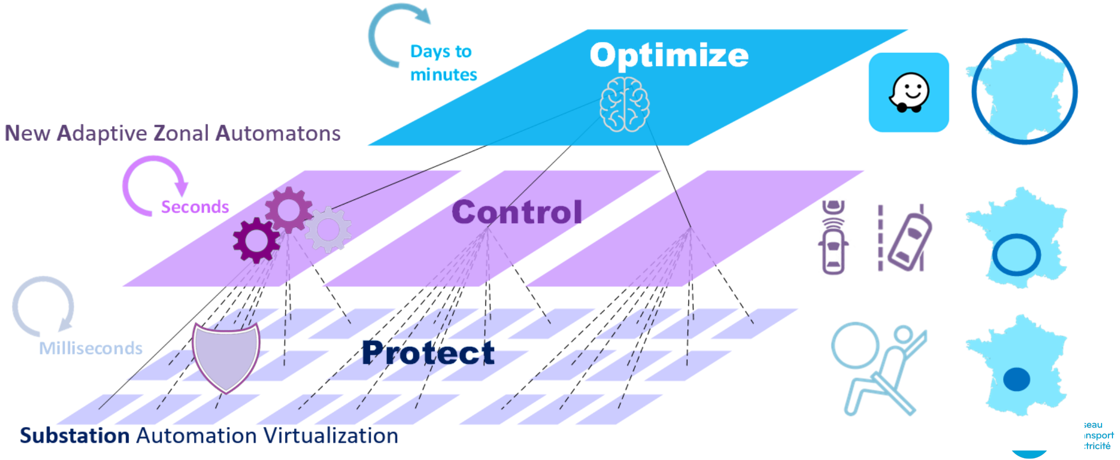
Numerous and generally **not very controllable**

High **uncertainties** on local production forecasts



Rte de transport d'électricité

A 3-Layer architecture to support the operators



Substation automation virtualization

Typical 225/63kV substation

X 11

1 bay
5 servers
4 switches

LF ENERGY SEAPATH

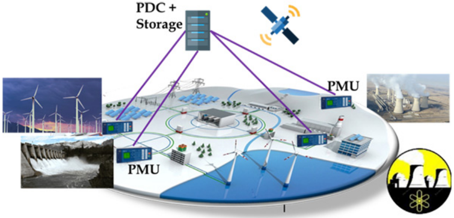
Cheaper and easier to deploy and upgrade

Rte Le réseau de transport d'électricité

Outline

1. Context
   - RTE's presentation
2. Related work
   - Compression of PMU data
   - Sampled voltage and current values
3. Proposed approach
   - Parametric coding
   - Residual coding
   - Optimal rate-distortion
4. Results

First approach to compress the PMU's data.



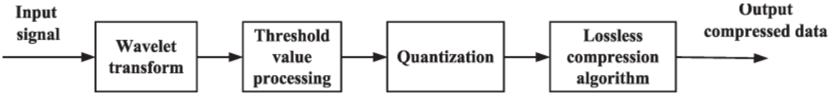
- 100 PMUs collecting 20 measurements at 30 Hz sampling rate generate over 50 GB of data per day[1].
- Lossless compression for PMU data szip, bzip, LZMA[2].
- J. Kraus *et al.* get a CR=2-3.

[1] R. Klump, P. Agarwal, J. E. Tate, and H. Khurana, "Lossless compression of synchronized phasor measurements", in *IEEE PES General Meeting*, pp. 1--7, 2010.

[2] J. Kraus, P. Štěpán, and L. Kukačka, "Optimal data compression techniques for Smart Grid and power quality trend data", in *IEEE 15th International Conference on Harmonics and Quality of Power* (2012), pp. 707--712, 2012.

Le réseau de transport d'électricité

Transform coding



- J. Khan *et al.*[3] get the compromise CR=10 / MSE=10^{-6}.

[3] J. Khan, S. M. A. Bhuiyan, G. Murphy, and M. Arline, "Embedded-Zerotree-Wavelet-Based Data Denoising and Compression for Smart Grid", IEEE *Transactions on Industry Applications 51*, pp. 4190--4200, 2015.

Exploitation of neighbouring PMUs

- S. Kirti *et al.*[4] found that the PMUs data are correlated.

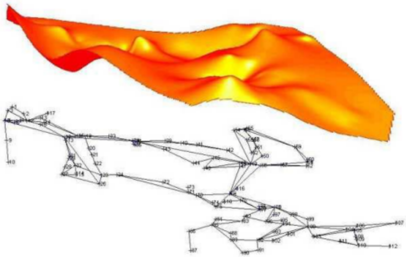


Fig. 9. The voltage angles in the IEEE 118 Bus Test Case in the steady state.

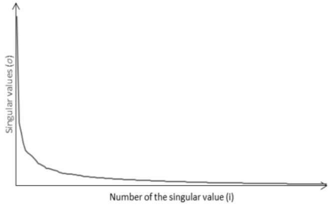
[4] S. Kirti, Z. Wang, A. Scaglione, and R. Thomas, "On the Communication Architecture for Wide-Area Real-Time Monitoring in Power Networks", in *40th Annual Hawaii International Conference on System Sciences*, pp. 119--119, 2007

Principal Component Analysis

- SVD: Transform multidimensional data into a linear combination of orthogonal components

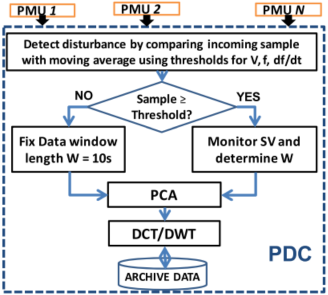
$$X_{K\times N} = \begin{pmatrix} P_1(1) & \cdots & P_N(k) \\ \vdots & \ddots & \vdots \\ P_1(K) & \cdots & P_N(K) \end{pmatrix}_{K\times N} = U_{K\times K} \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_N \\ 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 \end{pmatrix}_{K\times N} V^{*}_{N\times N}.$$

- P : voltage, frequency, angle or complex phasors (N PMUs, K samples).
- σ : singular values.



- With this method we get the compromise CR=6-10 / MSE=10^{-6}.

Improvement by exploiting the temporal correlation



- P. H. Gadde *et al.*[5] and M.Stacchini de Souza *et al.*[6] get the compromise CR=8-20 / MSE=10^{-6}.

[5] P. H. Gadde, M. Biswal, S. Brahma, and H. Cao, "Efficient Compression of PMU Data in WAMS", IEEE *Transactions on Smart Grid 7*, pp. 2406--2413, 2016.

[6] J. C. Stacchini de Souza, T. M. L. Assis, and B. C. Pal, "Data Compression in Smart Distribution Systems via Singular Value Decomposition", in IEEE *Transactions on Smart Grid 8*, pp. 275--284,2017.

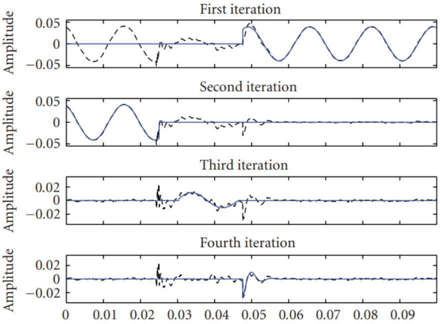
Outline

1. Context
   - RTE's presentation
2. Related work
   - Compression of PMU data
   - Sampled voltage and current values
3. Proposed approach
   - Parametric coding
   - Residual coding
   - Optimal rate-distortion
4. Results

Parametric coding

- Approximation by damped sinusoid

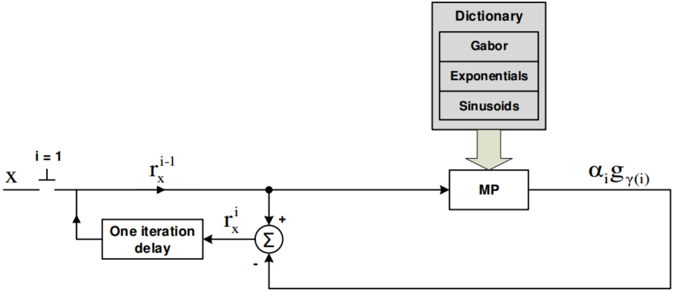
$$x = \sum_{m=1}^{M} \alpha_m g_\gamma(m), \text{ with } g_\gamma(n) = e^{-\rho(n-n^s)} \cos(\xi n + \phi) \times \left[u(n-n^s) - u(n-n^e)\right]$$



- An atom is represented by a six-tuple $(\alpha, \xi, \rho, \phi, n^s, n^e)$
- L. Lovisolo *et al.*[7] get the compromise CR=30 / MSE=10^{-3}.

[7] L. Lovisolo, M. P. Tcheou, E. A. B. da Silva, M. A. M. Rodrigues, and P. S. R. Diniz, "Modeling of Electric Disturbance Signals Using Damped Sinusoids via Atomic Decompositions and Its Applications", EURASIP *Journal on Advances in Signal Processing*, pp. 029507,

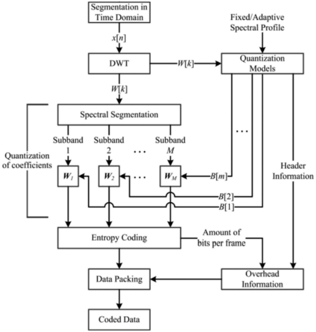
Parametric coding

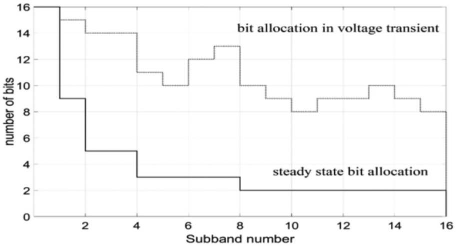


- To simplifiy the cost of calculation an Artificial Neural Networks is used to help to select the atom [8].

[8] A. Gabriel de Oliveira, M. P. Tcheou, and L. Lovisolo, "Artificial Neural Networks For Dictionary Selection in Adaptive Greedy Decomposition Algorithms With Reduced Complexity", in *International Joint Conference on Neural Networks*, pp. 1--8, 2018

Transform coding

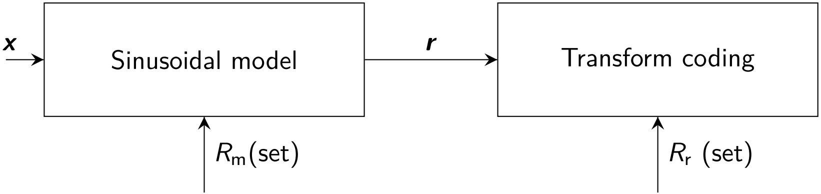




- Models : decreasing exponential, linear or square-root.
- F. A. de O. Nascimento *et al*[9]. get the compromise CR=20 \ MSE=10^{-6}.
- Advantage: no need to transmit the bit allocation vector.
- Disadvantage: we don't have an optimal bit allocation.

[9] F. A. de O. Nascimento, R. G. Saraiva, and J. Cormane, "Improved Transient Data Compression Algorithm Based on Wavelet Spectral Quantization Models", IEEE *Transactions on Power Delivery 35*, pp. 2222--2232, 2020.

Parametric coding and non parametric coding



- The parameters of the models are coded on a fixed number of bits.
- An optimal rate-distortion compromise is made to code each subband of the transformed residual.
- M. V. Ribeiro *et al.*[10] get the compromise CR=15 \ MSE=10^{-6}.

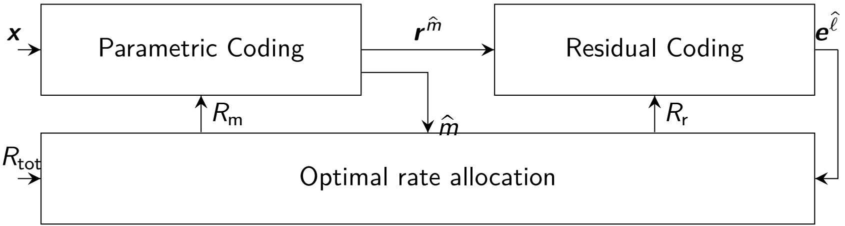
[10] M. V. Ribeiro, S. H. Park, J. M. T. Romano, and S. K. Mitra, "A Novel MDL-based Compression Method for Power Quality Applications", IEEE *Transactions on Power Delivery 22*, pp. 27--36, 2007.

Our goal

- Carrying out a compression of voltage and current signal driven by the applications
  - Exploiting the knowledge we have on electric waves via parametric models.
  - Being capable of refining the model if needed by coding refinement layers.

Le réseau de transport d'électricité

New features compared to the literature



- Add some models in the first stage such as predictive models[11]
- Add trainable transforms in the second stage: Variational Autoencoder (VAE)[12]
- Propose a rate-distortion compromise taking into account the two stages compression.

[11] D. O'Shaughnessy, "Linear predictive coding". IEEE *potentials*, 7(1), 29-32, 1988.

[12] J. Ballé, D. Minnen, S. Singh, S. J. Hwang, and N. Johnston, "Variational image compression with a scale hyperprior", *arXiv*, 2018.

Outline

1 Context
- RTE's presentation

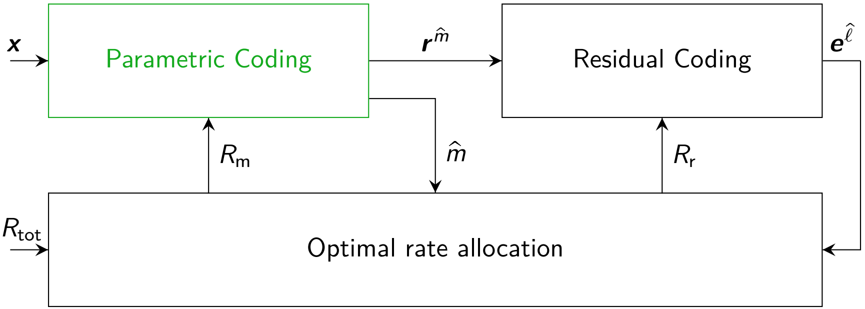
2 Related work
- Compression of PMU data
- Sampled voltage and current values

3 Proposed approach
- Parametric coding
- Residual coding
- Optimal rate-distortion

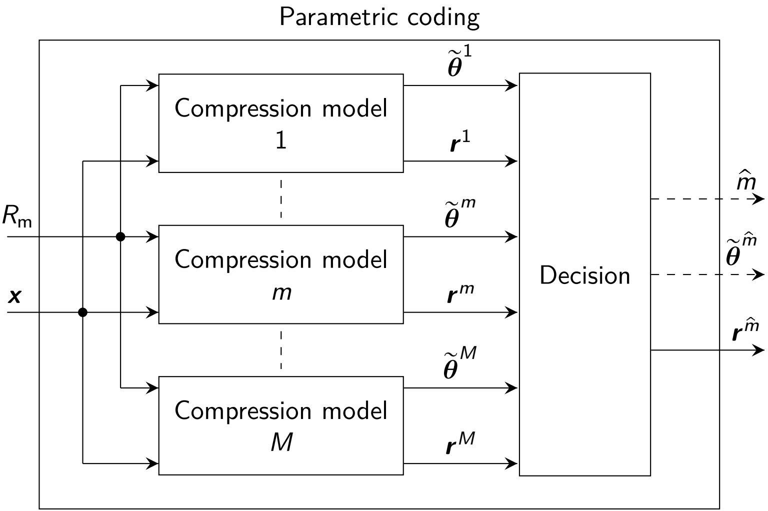
4 Results

Le réseau de transport d'électricité

Parametric coding



Details on parametric coding



Choice of the best model

$$\left(\hat{m}, \hat{\boldsymbol{\theta}}^{\hat{m}}\right) = \arg\min_{m,\boldsymbol{\theta}^m} \left\| \boldsymbol{x} - \boldsymbol{x}^m\left(\tilde{\boldsymbol{\theta}}^m\right) \right\|^2 \quad (1)$$

$$\text{s.t. } \tilde{\boldsymbol{\theta}}^m = Q\left(\boldsymbol{\theta}^m, \boldsymbol{\Delta}^m\right) \quad (2)$$

$$-\log_2\left(p_{\tilde{\boldsymbol{\theta}}^m}\left(\tilde{\boldsymbol{\theta}}^m\right)\right) \leqslant NR_{\mathrm{m}}, \quad (3)$$

- m index of model
- $\boldsymbol{\theta}$ vector of parameters
- $\boldsymbol{x}$ original signal of size N
- $\boldsymbol{x}^m$ reconstructed signal by the m-th model
- R_{m} rate to encode the model
- $\boldsymbol{\Delta}^m$ vector of quantization step
- Q quantizer
- $p_{\tilde{\boldsymbol{\theta}}^m}\left(\tilde{\boldsymbol{\theta}}^m\right)$ distribution of $\tilde{\boldsymbol{\theta}}^m$

Models

- m-th sinusoidal model

$$\mathcal{M}^m\left(\boldsymbol{\theta} = (a, f, \phi), n\right) = a\cos\left(2\pi f n T_\mathrm{s} + \phi\right)$$

  - T_s sampling period

- m-th polynomial model

$$\mathcal{M}^m\left(\boldsymbol{\theta} = (\theta_0, \ldots, \theta_K), n\right) = \sum_{k=0}^{K} \theta_k \left(n T_s\right)^k$$

  - K order of polynomial model

- Reconstructed signal by the m-th model

$$x_n^m = \mathcal{M}^m(\boldsymbol{\theta}, n), \quad n = 1, \ldots, N$$

Le réseau de transport d'électricité

Models

- m-th sinusoidal model

$$\mathcal{M}^m\left(\boldsymbol{\theta} = (a, f, \phi), n\right) = a\cos\left(2\pi f n T_\mathrm{s} + \phi\right)$$

  - T_s sampling period
- m-th polynomial model

$$\mathcal{M}^m\left(\boldsymbol{\theta} = (\theta_0, \ldots, \theta_K), n\right) = \sum_{k=0}^{K} \theta_k \left(n T_\mathrm{s}\right)^k$$

  - K order of polynomial model
- Reconstructed signal by the m-th model

$$x_n^m = \mathcal{M}^m(\theta, n), \quad n = 1, \ldots, N$$

Models

- m-th sinusoidal model

$$\mathcal{M}^m\left(\boldsymbol{\theta} = (a, f, \phi), n\right) = a\cos\left(2\pi f n T_\mathrm{s} + \phi\right)$$

  - T_s sampling period

- m-th polynomial model

$$\mathcal{M}^m\left(\boldsymbol{\theta} = (\theta_0, \ldots, \theta_K), n\right) = \sum_{k=0}^{K} \theta_k \left(n T_s\right)^k$$

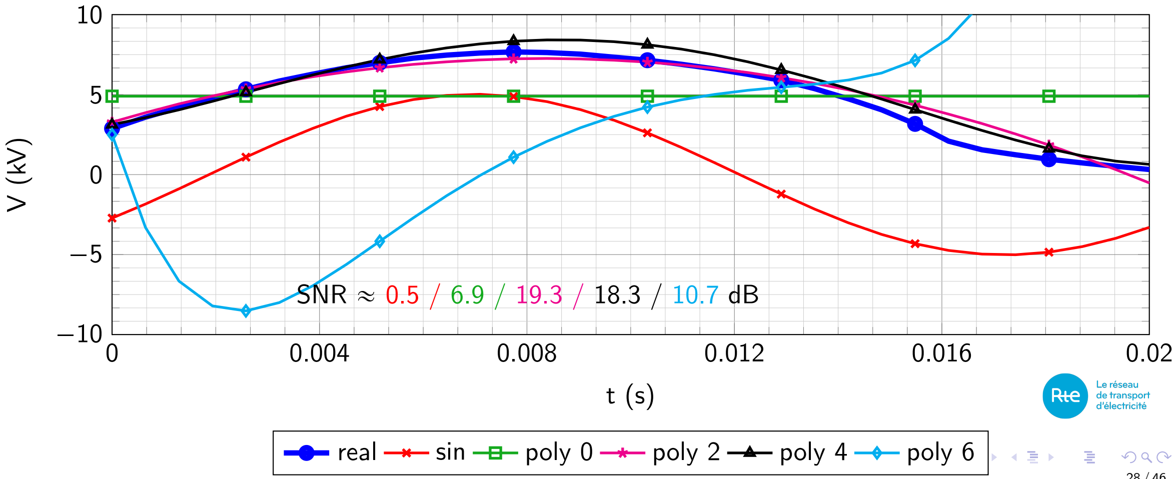
  - K order of polynomial model

- Reconstructed signal by the m-th model

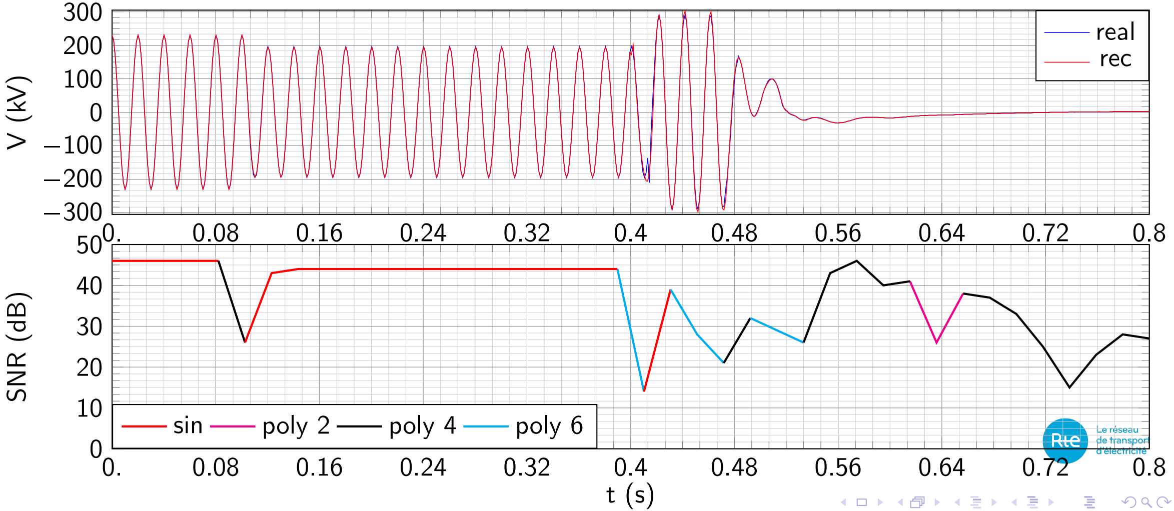
$$x_n^m = \mathcal{M}^m\left(\boldsymbol{\theta}, n\right), \quad n = 1, \ldots, N$$

Le réseau de transport d'électricité

Example of reconstructed models at the output of the first stage with $R_m = 0.2$ bps, N=128 samples



Example of a test signal with $R_\mathrm{m} = 0.2$ bps, N=128 samples per window



Predictive Models

- Parametric predictive models

$$\mathcal{M}^{m(i)}\left(\boldsymbol{\delta\theta}^{(i)}, n\right) = \mathcal{M}^{m(i-1)}\left(\widetilde{\boldsymbol{\theta}}^{(i-1)} + \boldsymbol{\delta\theta}^{(i)}, n\right)$$

  - i current window
  - $i-1$ previous window
  - $\widetilde{\boldsymbol{\theta}}^{(i-1)}$ previous quantized coefficients
- Samples predictive models

$$\mathcal{M}^{m}\left(\boldsymbol{\theta} = \left(\alpha_1, \ldots, \alpha_K, \eta\right), n\right) = \sum_{i=1}^{K} \alpha_i \overline{x}_{n-i-\eta}$$

  - $\eta \in \mathbb{N}$ shift
  - $\overline{x}_{n-i-\eta}$ previous encoded coefficients

Predictive Models

- Parametric predictive models

$$\mathcal{M}^{m(i)}\left(\boldsymbol{\delta\theta}^{(i)}, n\right) = \mathcal{M}^{m(i-1)}\left(\widetilde{\boldsymbol{\theta}}^{(i-1)} + \boldsymbol{\delta\theta}^{(i)}, n\right)$$

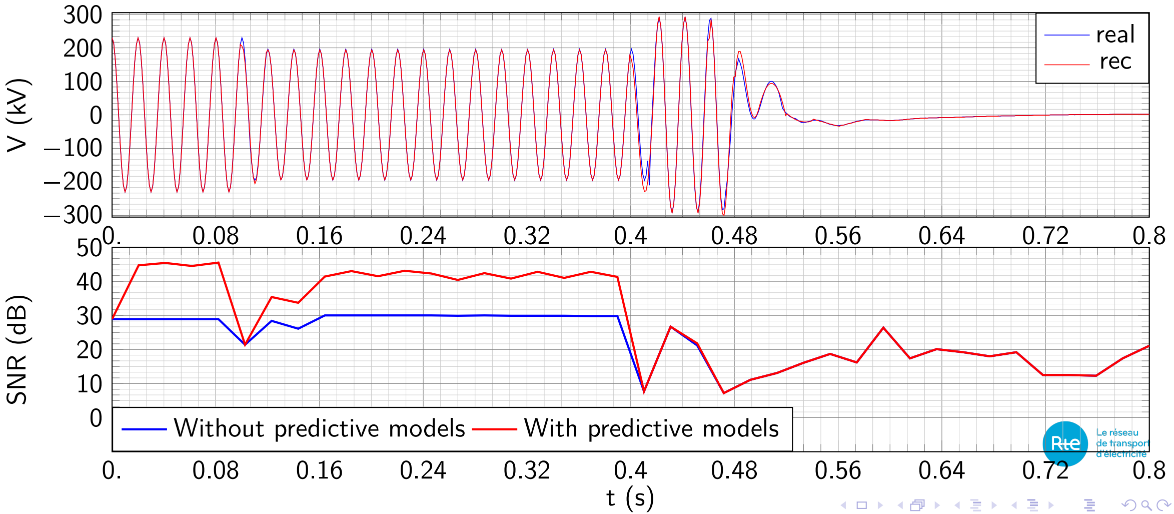
  - i current window
  - $i-1$ previous window
  - $\widetilde{\boldsymbol{\theta}}^{(i-1)}$ previous quantized coefficients

- Samples predictive models

$$\mathcal{M}^{m}\left(\boldsymbol{\theta} = (\alpha_1, \ldots, \alpha_K, \eta), n\right) = \sum_{i=1}^{K} \alpha_i \overline{x}_{n-i-\eta}$$

  - $\eta \in \mathbb{N}$ shift
  - $\overline{x}_{n-i-\eta}$ previous encoded coefficients

Reconstruction with the predictive models, $R_m = 0.1$ bps



Le réseau de transport d'électricité

Outline

1 Context
- RTE's presentation

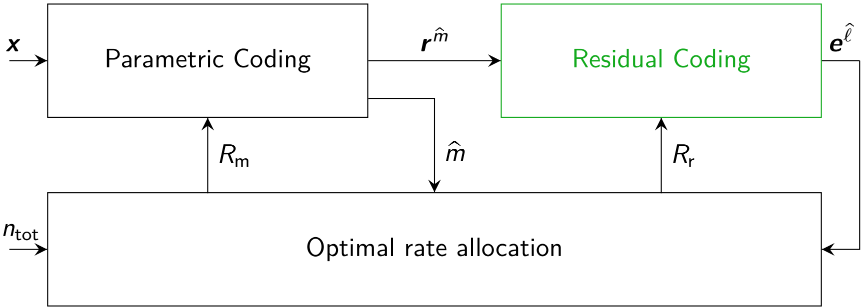
2 Related work
- Compression of PMU data
- Sampled voltage and current values

3 Proposed approach
- Parametric coding
- Residual coding
- Optimal rate-distortion

4 Results

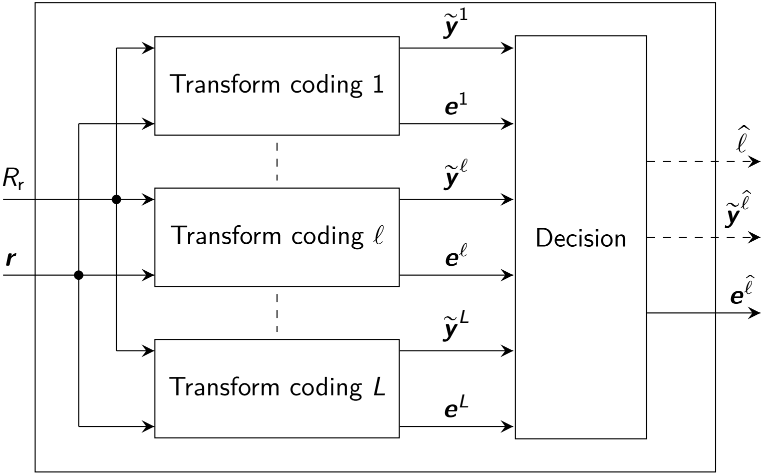
Le réseau de transport d'électricité

Residual coding



Detail of Transform coding

Residual compression



Determination of the best transform

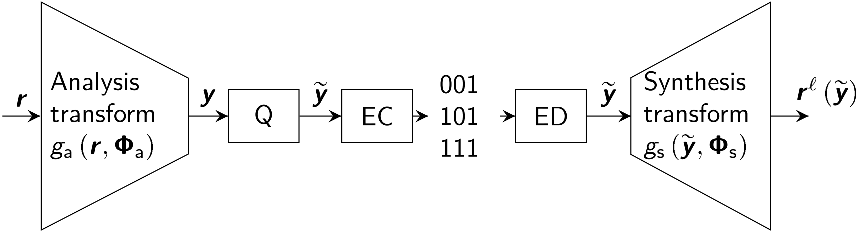
$$\widehat{\ell} = \arg\min_{\ell} \left\| \boldsymbol{r} - \boldsymbol{r}^{\ell}\left(\widetilde{\boldsymbol{y}}^{\ell}\right) \right\|^2 \tag{4}$$

$$\text{s.t. } \widetilde{\boldsymbol{y}}^{\ell} = Q\left(\boldsymbol{y}^{\ell}, \boldsymbol{\Delta}^{\ell}\right) \tag{5}$$

$$-\log_2\left(p_{\widetilde{\boldsymbol{y}}^{\ell}}\left(\widetilde{\boldsymbol{y}}^{\ell}\right)\right) \leqslant NR_{\mathrm{r}}, \tag{6}$$

- ℓ index of transform coding
- $\boldsymbol{r}$ original residual including N samples
- $\boldsymbol{y}^{\ell}$ transform coefficients
- $\boldsymbol{r}^{\ell}$ reconstructed residual
- Q quantizer
- $\boldsymbol{\Delta}^{\ell}$ vectors of quantization step for each coefficient
- R_{r} rate to encode the coefficients of $\widetilde{\boldsymbol{y}}^{\ell}$
- $p_{\widetilde{\boldsymbol{y}}^{\ell}}$ distribution of $\widetilde{\boldsymbol{y}}^{\ell}$

Trainable transforms



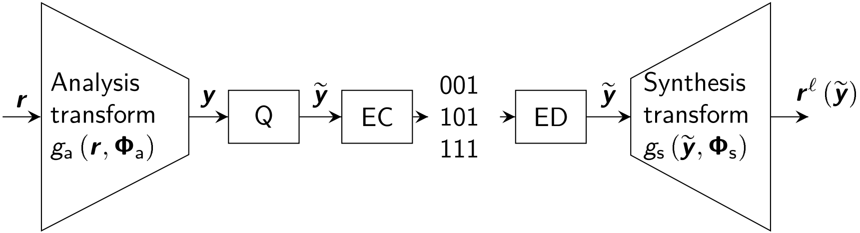
$$\hat{\boldsymbol{\Phi}}_a, \hat{\boldsymbol{\Phi}}_s = \arg\min_{\boldsymbol{\Phi}_a, \boldsymbol{\Phi}_s} \mathbb{E}_{p_r(\boldsymbol{r})}\left[\left\|\boldsymbol{r} - \boldsymbol{r}^\ell\left(\tilde{\boldsymbol{y}}^\ell\right)\right\|^2 + \lambda\left(-\log_2\left(p_{\tilde{\boldsymbol{y}}^\ell}\left(\tilde{\boldsymbol{y}}^\ell\right)\right)\right)\right]$$

$$\text{s.t. } \boldsymbol{y}^\ell = g_a(\boldsymbol{r}, \boldsymbol{\Phi}_a)$$

$$\tilde{\boldsymbol{y}}^\ell = Q\left(\boldsymbol{y}^\ell, \boldsymbol{\Delta}^\ell\right)$$

$$\boldsymbol{r}^\ell\left(\tilde{\boldsymbol{y}}^\ell\right) = g_s\left(\tilde{\boldsymbol{y}}^\ell, \boldsymbol{\Phi}_s\right)$$

Trainable transforms



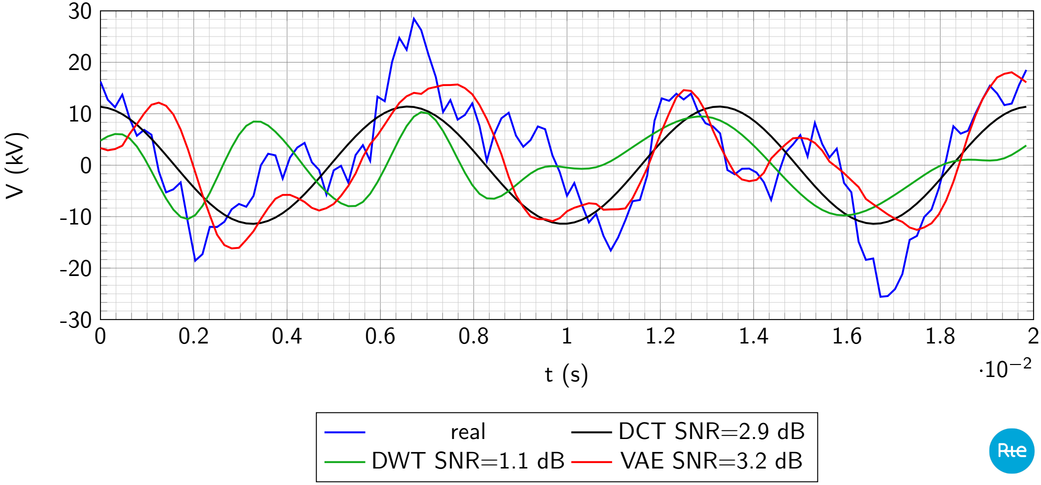
$$\hat{\boldsymbol{\Phi}}_{\mathrm{a}}, \hat{\boldsymbol{\Phi}}_{\mathrm{s}} = \arg\min_{\boldsymbol{\Phi}_{\mathrm{a}}, \boldsymbol{\Phi}_{\mathrm{s}}} \mathbb{E}_{p_{\boldsymbol{r}}(\boldsymbol{r})} \left[\left\| \boldsymbol{r} - \boldsymbol{r}^{\ell}\left(\tilde{\boldsymbol{y}}^{\ell}\right) \right\|^2 + \lambda \left(-\log_2 \left(p_{\tilde{\boldsymbol{y}}^{\ell}} \left(\tilde{\boldsymbol{y}}^{\ell} \right) \right) \right) \right]$$

$$\text{s.t.} \quad \boldsymbol{y}^{\ell} = g_{\mathrm{a}}\left(\boldsymbol{r}, \boldsymbol{\Phi}_{\mathrm{a}}\right)$$

$$\tilde{\boldsymbol{y}}^{\ell} = Q\left(\boldsymbol{y}^{\ell}, \boldsymbol{\Delta}^{\ell}\right)$$

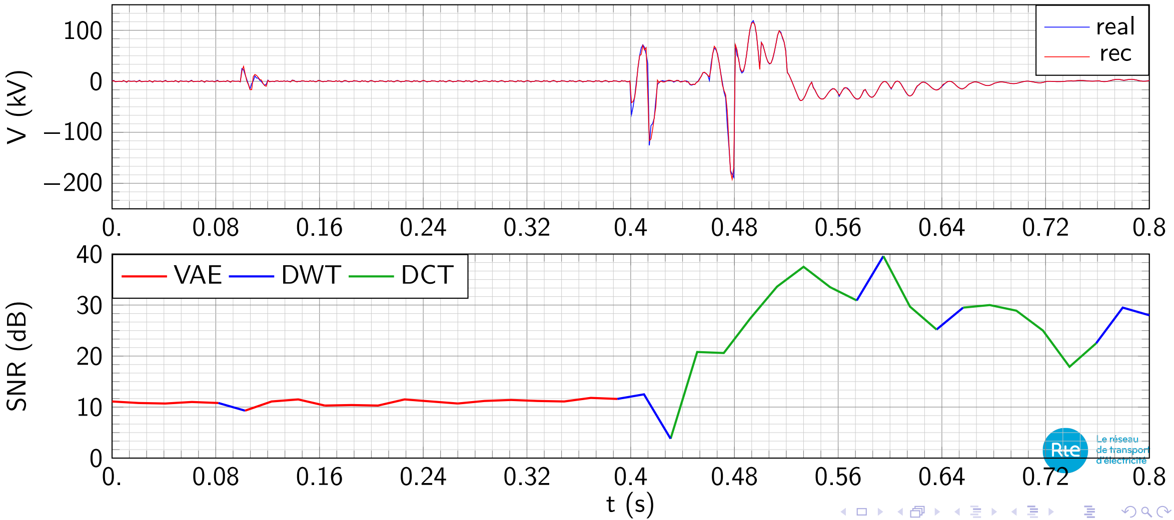
$$\boldsymbol{r}^{\ell}\left(\tilde{\boldsymbol{y}}^{\ell}\right) = g_{\mathrm{s}}\left(\tilde{\boldsymbol{y}}^{\ell}, \boldsymbol{\Phi}_{\mathrm{s}}\right)$$

Example of reconstructed residual with $R_r = 0.2$ bps



Le réseau de transport d'électricité

Example of a test residual signal with $R_r = 0.5$ bps

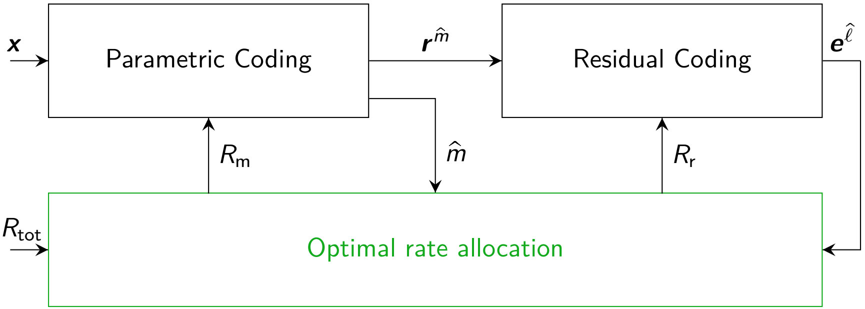


Outline

1. Context
   - RTE's presentation
2. Related work
   - Compression of PMU data
   - Sampled voltage and current values
3. Proposed approach
   - Parametric coding
   - Residual coding
   - Optimal rate-distortion
4. Results

Le réseau de transport d'électricité

Optimal rate-distortion



Le réseau de transport d'électricité

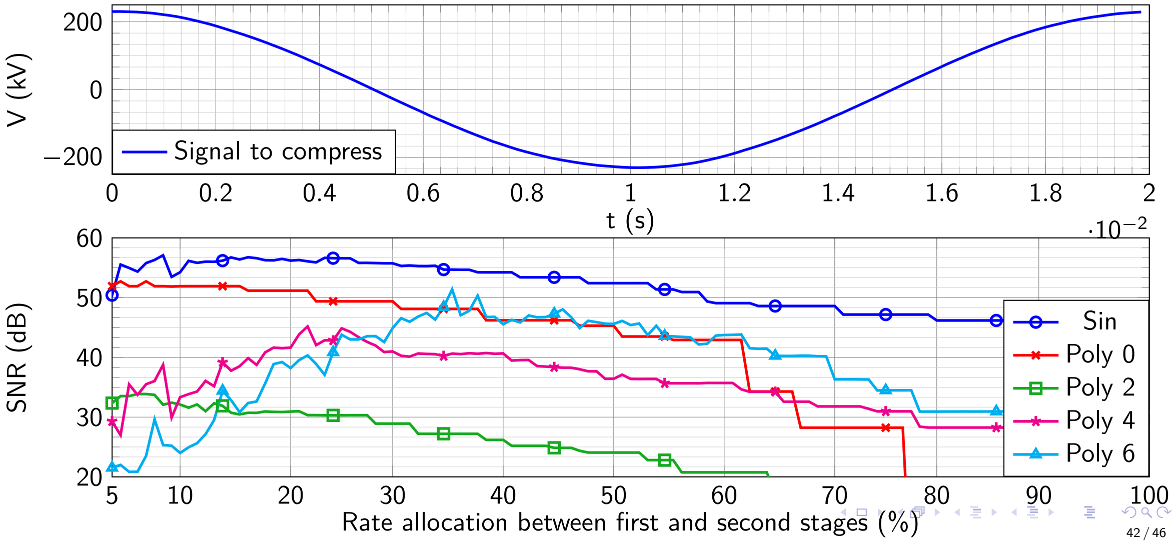
Optimal rate-distortion between the two stages

$$\left(\widehat{m}, \widehat{\ell}, \widehat{R}_{\mathrm{m}}, \widehat{R}_{\mathrm{r}}\right) = \arg \min_{m,\ell,R_{\mathrm{m}},R_{\mathrm{r}}} \left\| \boldsymbol{x} - \boldsymbol{x}^{m}\left(\widetilde{\boldsymbol{\theta}}^{m}, R_{\mathrm{m}}\right) - \boldsymbol{r}^{\ell}\left(\widetilde{\boldsymbol{y}}^{\ell}, R_{\mathrm{r}}\right) \right\|^{2}$$
$$\text{s.t.} \quad R_{\mathrm{h}} + R_{\mathrm{m}} + R_{\mathrm{r}} \leqslant R_{\mathrm{tot}}. \tag{7}$$

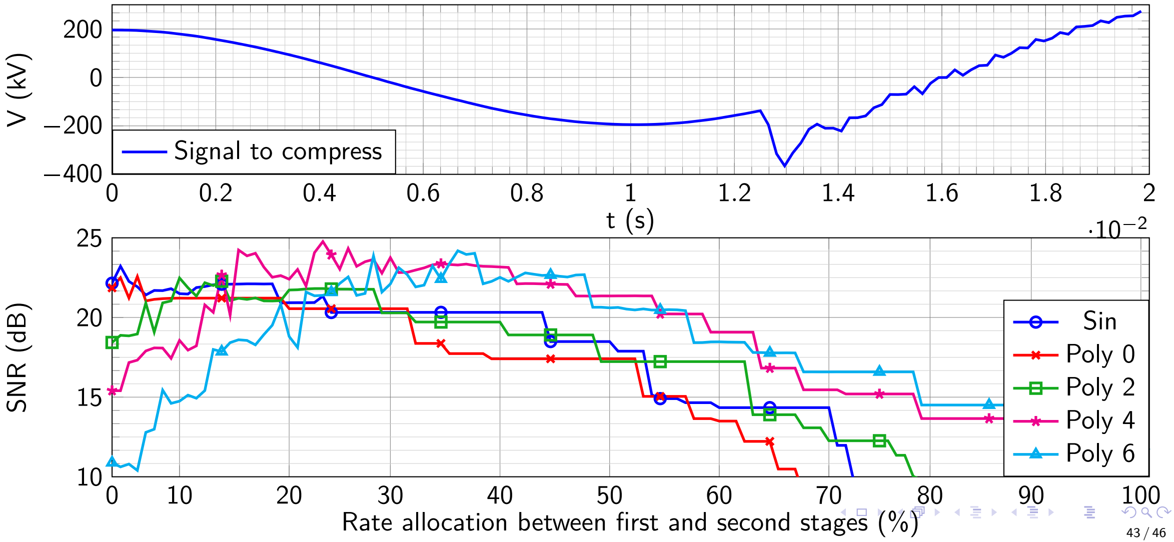
- m index of the model
- ℓ index of the transform coding
- $\boldsymbol{x}$ original signal including Nsamples
- $\boldsymbol{x}^{m}$ reconstructed model on the rate R_{m}bps
- $\boldsymbol{r}^{\ell}$ reconstructed model on the rate R_{r}bps
- $\widetilde{\boldsymbol{\theta}}^{m}$ quantified parameters for the m-th model
- $\widetilde{\boldsymbol{y}}^{m}$ quantified transform coefficients for the ℓ-th transform
- R_{tot} total rate to encode the signal $\boldsymbol{x}$
- R_{h} rate to encode the hyperparameters of our coder

Le réseau de transport d'électricité

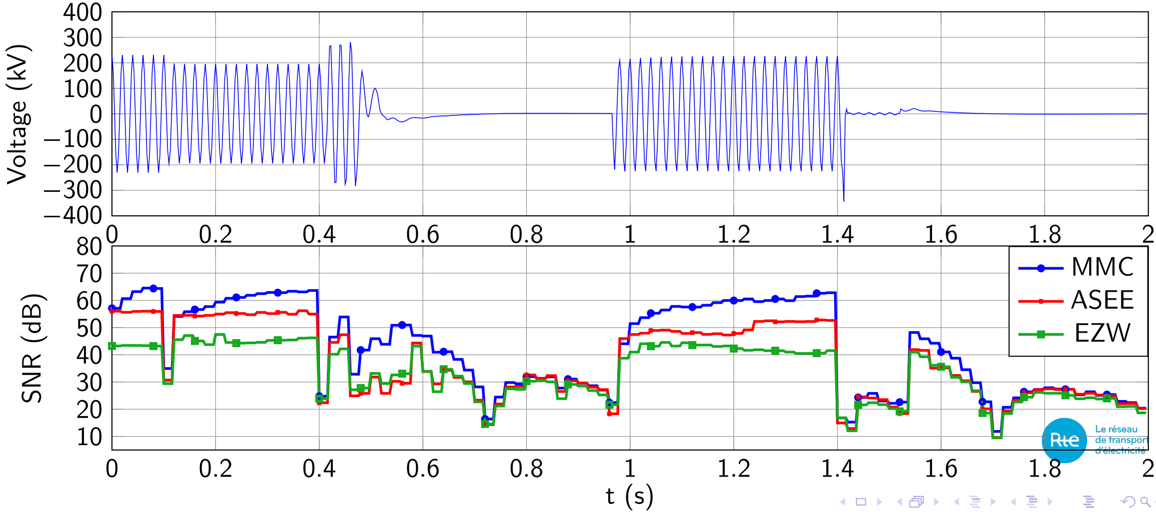
Example of optimal rate-distortion with $R_{tot} = 1$ bps



Second example of optimal rate-distortion with $R_{tot} = 1$ bps



SNR for each window with $R_{tot} = 1$ bps



Comparison on the EPRI database

- We selected 166 transient signal windows from Electric Power Research Institute (EPRI) database
- f=60Hz, f_s=15384.6 Hz
- N=256 samples, i.e. 97.89% of a period of the signal at 60 Hz.
- Comparative approach: Adaptive Spectral Estimation Envelope (ASEE) [13], Embedded Zerotree Wavelet (EZW) [14]

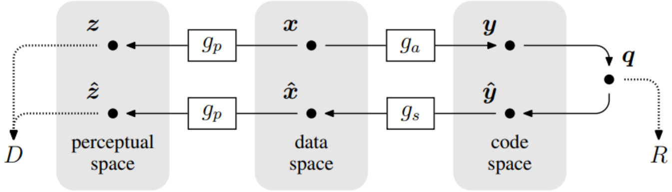
Bitrate (bps)	MMC SNR (dB)	ASEE SNR (dB)	EZW SNR (dB)
0.5	33.6	27.3	22.3
1	41.1	35.0	32.2
2	50.5	42.0	38.8

[13] F. A. de O. Nascimento, R. G. Saraiva, and J. Cormane, "Improved Transient Data Compression Algorithm Based on Wavelet Spectral Quantization Models", IEEE *Transactions on Power Delivery 35*, pp. 2222--2232, 2020.

[14] J. Khan,S. Bhuiyan,G. Murphy, and M. Arline, "Embedded zerotree wavelet based data compression for smart grid", In IEEE *industry applications society annual meeting*, pp. 1-8, 2013.

Conclusion

- Goal-oriented: the approach is adapted to the end-users.



- Applications
  - Fault location: z corresponds to the time of the fault
  - Fitting a model: z?
  - Inter area oscillation monitoring: z?
  - Network control: z?