LoRa Signal Decoding with and without Machine Learning

José Alamos, Thomas C. Schmidt

http://inet.haw-hamburg.de | t.schmidt@haw-hamburg.de

What is LoRa?

Wireless modulation technique

- Long range (up to 15 km)
- Low power consumption (mJ)
- Low data rate (bytes/s)

Popular in low power IoT deployments

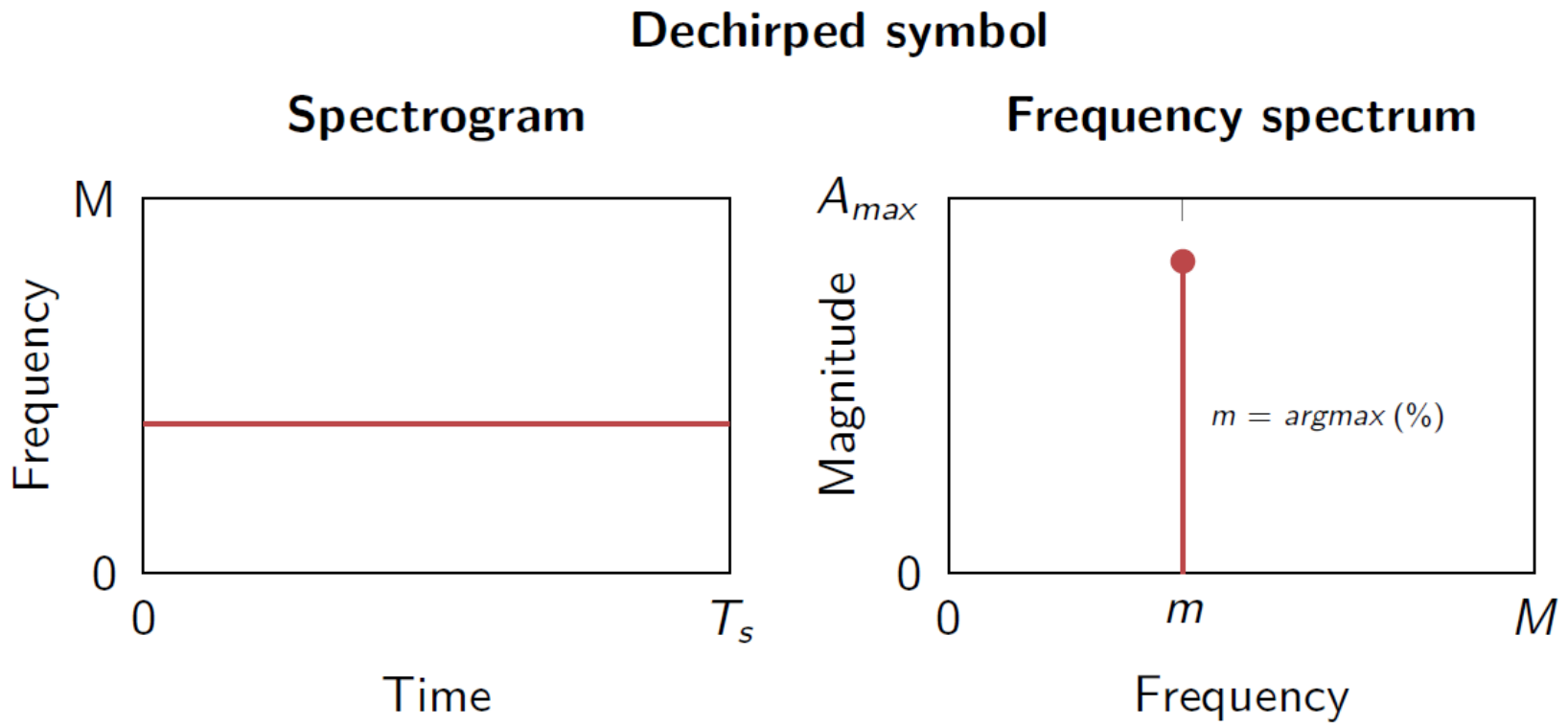
- Can run on batteries for years
- Collision-prone due to long time on air
- Improved LoRa on RIOT

Outline

- 🕐 Decoding LoRa
- 🕑 A deep Learning with ML
- 🕒 Augmenting Analytics with some Learning
- 🕓 Lessons Learned

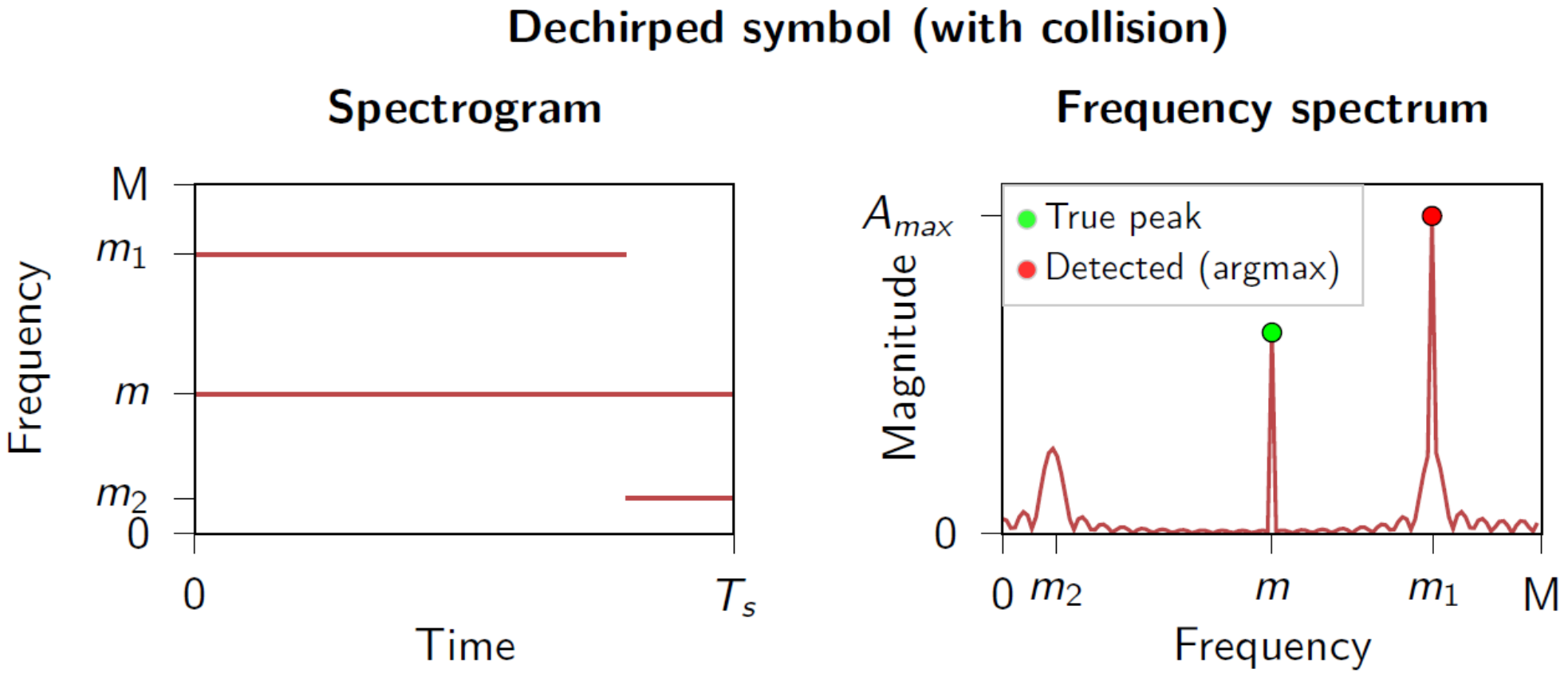
LoRa Decoding

Frequency bin with the highest magnitude.



When LoRa Frames Collide

LoRa Collision



Goal

Find the full sine wave in the dechirped symbol.

Decoding Colliding Frames

Baseline LoRa decoder may fail under collisions

Active area of research

- Colora – INFOCOM'20
- CIC – SIGCOMM'21
- NELoRa – SenSys'21
- TnB – CoNEXT'22

Machine learning techniques considered promising method to boost signal recovery

Data Samples

Software Defined Radio (SDR)

- Reuse existing deployment.
- Capture real-world symbol data.

Simulated data

- Generate symbol data with known parameters.
- Model symbol as complex chirp with white gaussian noise.

First Insight

- Finding the full sine wave is harder than expected
  - LoRa collisions yield a complex frequency spectrum
  - Hard to unravel using conventional signal processing techniques

- Potential of Machine Learning techniques for decoding LoRa frames

Convolutional Neural Networks - CNN

Good at finding patterns in data … let's train a CNN
to find the longest sine wave in the dechirped symbol

CNN Symbol Classifier Evaluation:

- Time domain as input
  - Does not converge
- Spectrogram (STFT) as input
  - Worse than baseline decoder
- Frequency spectrum (FFT) as input
  - Does not detect symbols with collisions
  - But performs slightly better than baseline decoder

Need more features?

Wavelet Transform

Wigner-Ville Distribution

Synchrosqueezing Transform

Hilbert-Huang Transform

Empirical Mode Decomposition

Fractional Fourier Transform

etc.

Aftermath of CNN Classifier

Some features yield slightly better accuracy
than the baseline decoder

- At the cost of high computational complexity
- The classifier works best for symbols without collision

Gains are not enough to justify the complexity

Adjust the Focus to Something Promising

Peak Classification

- Retrain CNN to peak probabilities
- Expensive … but worse than baseline decoder

Denoiser Autoencoder

- Train a neural network to remove noise from frequency domain
- Improves SNR … but distorts phase

Last Hope: Simple Math

We observe

- A true signal is either symmetric or anti-symmetric
  - Half-Period Discriminator (HDP) identifies symmetry properties

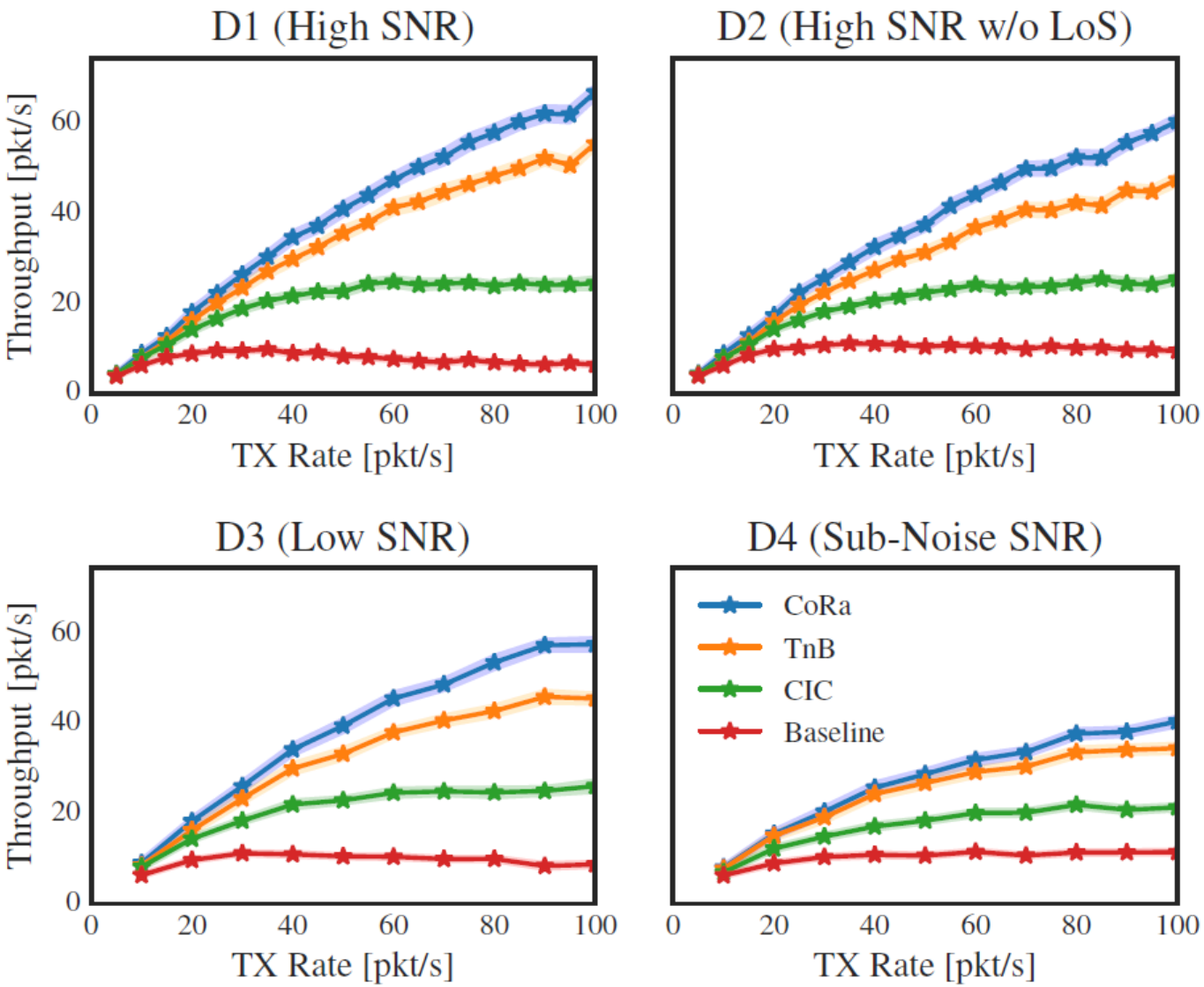
- Magnitude of a true peak is similar to preamble peaks
  - Peak Magnitude Deviation (PMD) evaluates the differences in peak heights

- Bayesian classifier serves as likelihood estimator from HDP and PMD
  - Posterior probability derived from simulations

Results

Decoding of recorded traffic (CIC dataset)

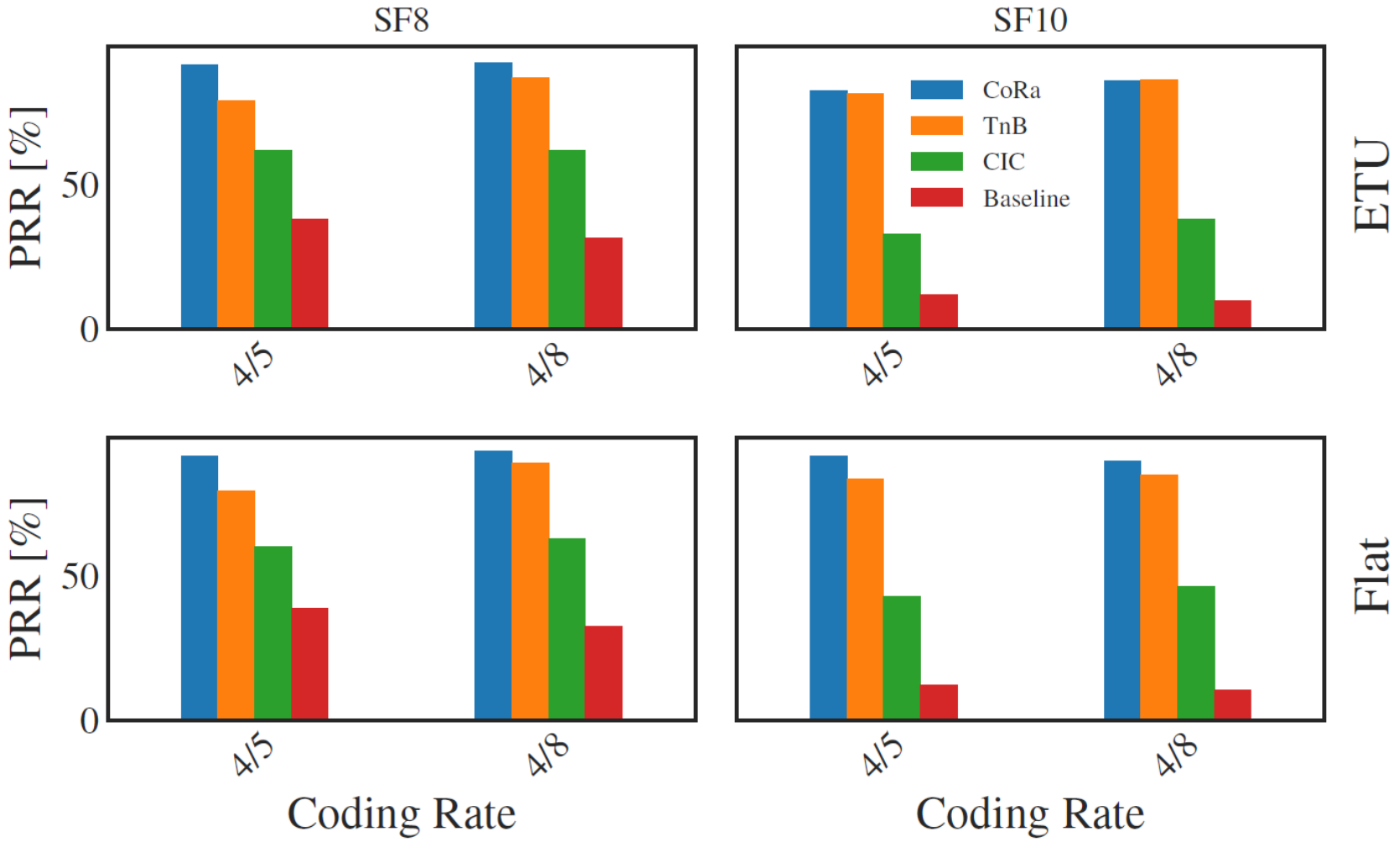
- Spreading factor 8
- Coding rate 4/5
- Bandwidth: 250 KHz



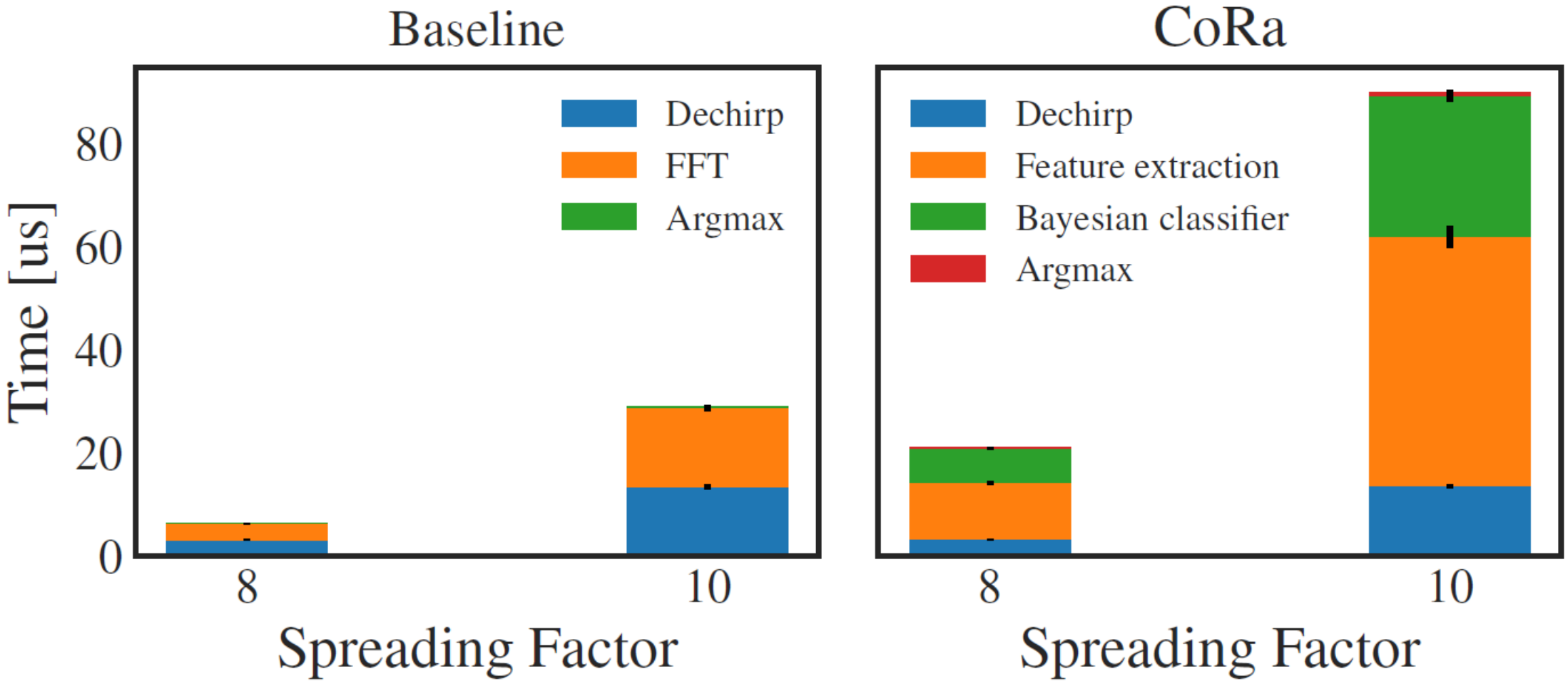
Results (2)

Simulated LTE channel models

- Extended Typical Urban (ETU)
- Flat fading
- Bandwidth: 125 KHz
- 15 pkt/s



Operational Complexity



Post Mortem Analysis

Overall, we spent six months in exploring, training, and twisting ML models

Was it worth it?

Yes!

- Learned much about Deep Learning
- Acquired much deeper insight into the problem from analyzing the flops
- Can reasonably claim:
  ML is Not a Silver Bullet

ML is Not a Silver Bullet

But for certain tasks, Deep Learning is likely the best approach.

- Large Language Models
- Denoising
- Sketching Santa Claus drinking Glühwein in Hamburg.

Conclusions

Intense struggle with data helped us to deeply learn about LoRa signals

Insides inspired analytic approach

Simple math approach did the job better and faster

But: With the Bayesian estimator, we can still claim to use ML!

Backup: Bayesian Posterior Probability