

Ten Times Faster: Error correction at the speeds required for 5G mobile telephony

FPGA implementation of the Fully-Parallel Turbo Decoder (FPTD)

An Li, Taihai Chen, Robert G. Maunder

Email: {al4e08, tc1m13, rm}@ecs.soton.ac.uk

Introduction

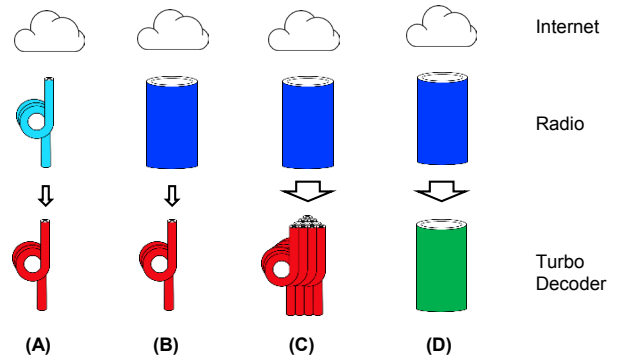
- Turbo decoders are used in 3G and 4G mobile telephony for correcting communication errors caused by poor signal quality.
- Compared to 4G, the target for 5G is to achieve ten times reductions in download times and in the lag experienced in video chat and online games.
- Lag will be reduced by decomposing chat and gaming streams into short messages, comprising tens or hundreds of bits.
- However, **conventional turbo decoders** are not fast enough to keep up with these 5G requirements, even when many of them work together.

- This is because the data-dependencies within the **conventional turbo decoder** require each of its steps to be performed one-at-a-time and in the correct order.
- This invention is the world's first general-purpose **fully-parallel turbo decoder**, which breaks the data dependencies and allows the steps to be performed at the same time.
- For the first time, our **fully-parallel turbo decoder** is fast enough to meet the requirements of 5G.

Analogy

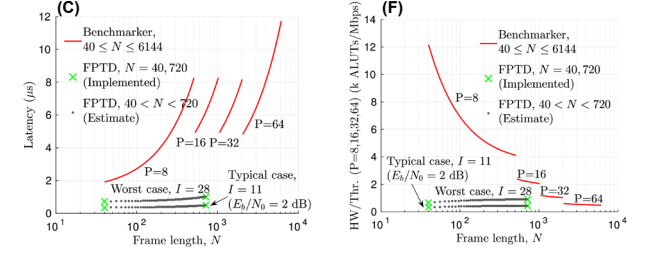
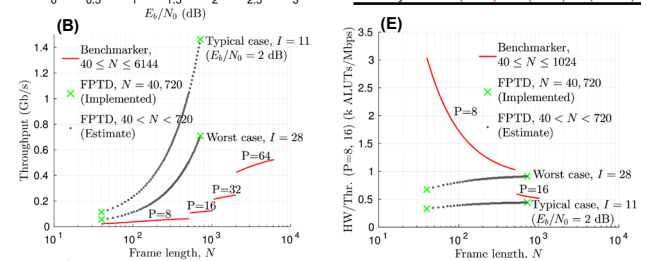
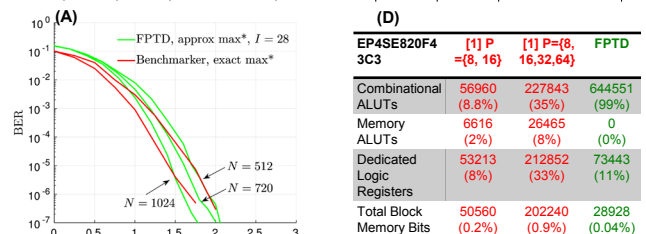
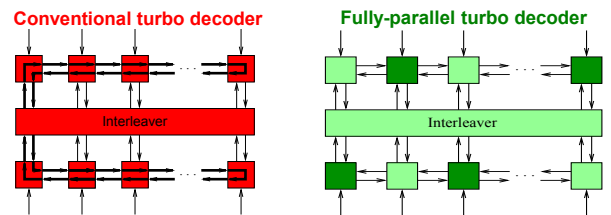
- The radio and turbo decoder parts of 4G and 5G systems each have associated throughputs and latencies, which can be represented by pipes:
 - Throughput - Wider pipes allow more information to flow per second, reducing download times;
 - Latency - Shorter pipes deliver the information to the other end with less delay, reducing lag in video chat and on-line games.

- 4G radio and conventional turbo decoder.** This has a peak throughput of **1 Gbit/s** and a latency of **10 ms**.
- 5G radio and conventional turbo decoder.** The latter bottlenecks peak throughput to **1 Gbit/s** and latency to **10 ms**.
- 5G radio and multiple conventional turbo decoders.** This removes the throughput bottleneck, giving **10 Gbit/s**, but does not remove the latency bottleneck of **10 ms**.
- 5G radio and fully-parallel turbo decoder.** Only this gives the desired peak throughput of **10 Gbit/s** and latency of **1 ms**.



Technical Perspective

- Conventional turbo decoder [1]**
 - Data dependencies require blocks to be operated one-at-a-time, in the order indicated by the bold arrows in the schematic to the right.
 - 100s of clock cycles are required to decode each message.
 - $P=\{8, 16\}$ version supports message lengths of up to $N=1024$ bits at run time.
 - $P=\{8, 16, 32, 64\}$ version supports message lengths of up to $N=6144$ bits at run time.
- Fully-parallel turbo decoder [2,3,4]**
 - Blocks having the same shading in the schematic to the right are operated at the same time, in the same clock cycle.
 - Only 10s of clock cycles are required to decode each message.
 - ASIC performance – better than 10-fold improvement in throughput and latency, with up to 20% improvement in energy efficiency and 30% improvement in area efficiency.
 - GPGPU performance – up to 3 times faster.
- FPGA Implementation (EP4SE820F4 3C3)**
 - I/Os: 562 input pins (5 control pins), 92 output pins.
 - Clock frequencies: Core @ 65 MHz ($N=720$) - 90 MHz ($N=40$), Memory @ 333 MHz
 - Supports different message lengths of up to $N=720$ bits at design time.
- FPGA Performance** in comparison with Benchmark [1]:
 - Error correction performance** – Equal to that of [1], when performing up to $I=28$ iterations, regardless of the message length N . However, only an average of $I=11$ iterations are required for a typical channel quality value, eg. $E_b/N_0 = 2$ dB.
 - Throughput** - 6.1 (worst case) – 12.6 (typical case) times better
 - Latency** - 6.1 (worst case) – 12.6 (typical case) times better
 - E. F. Hardware efficiency** - 1.7 times worse (worst case), 1.2 times better (typical case), compared with that of the $P=\{8, 16\}$ version of [1]. 2.4 times better (worst case) – 5 times better (typical case), compared with that of the $P=\{8, 16, 32, 64\}$ version of [1].
- Future work**
 - Run time support for different message lengths.
 - Pipelined decoder design, which potentially doubles throughput, latency and hardware efficiency.
 - Optimise



References

- L. Gonzalez-Perez, L. Yllescas-Calderon, R. Parra-Michel, "Parallel and configurable turbo decoder implementation for 3GPP-LTE", ReConFig, 2013
- R. G. Maunder, A. Li and I. Perez-Andrade, "Receiver, communications device and method of receiving", U.K. Patent Application 1 414 376.2, 2014
- R. G. Maunder, "A fully-parallel turbo decoding algorithm", IEEE Trans. Commun., 2015, <http://eprints.soton.ac.uk/368984>
- R. G. Maunder, Marketing video for fully-parallel turbo decoder, <https://youtu.be/hUx1uC9ZXsq>