Codec 2

- open source speech codec
- low bit rate (2400 bit/s and below)
- applications include digital speech for HF and VHF radio
- fills gap in open source speech codecs beneath 5000 bit/s

Why Open Source?

- Ham radio is an experimental service
- we need to be able to experiment, understand, and modify
- open source means no license fees, e.g. include in SDR systems for free

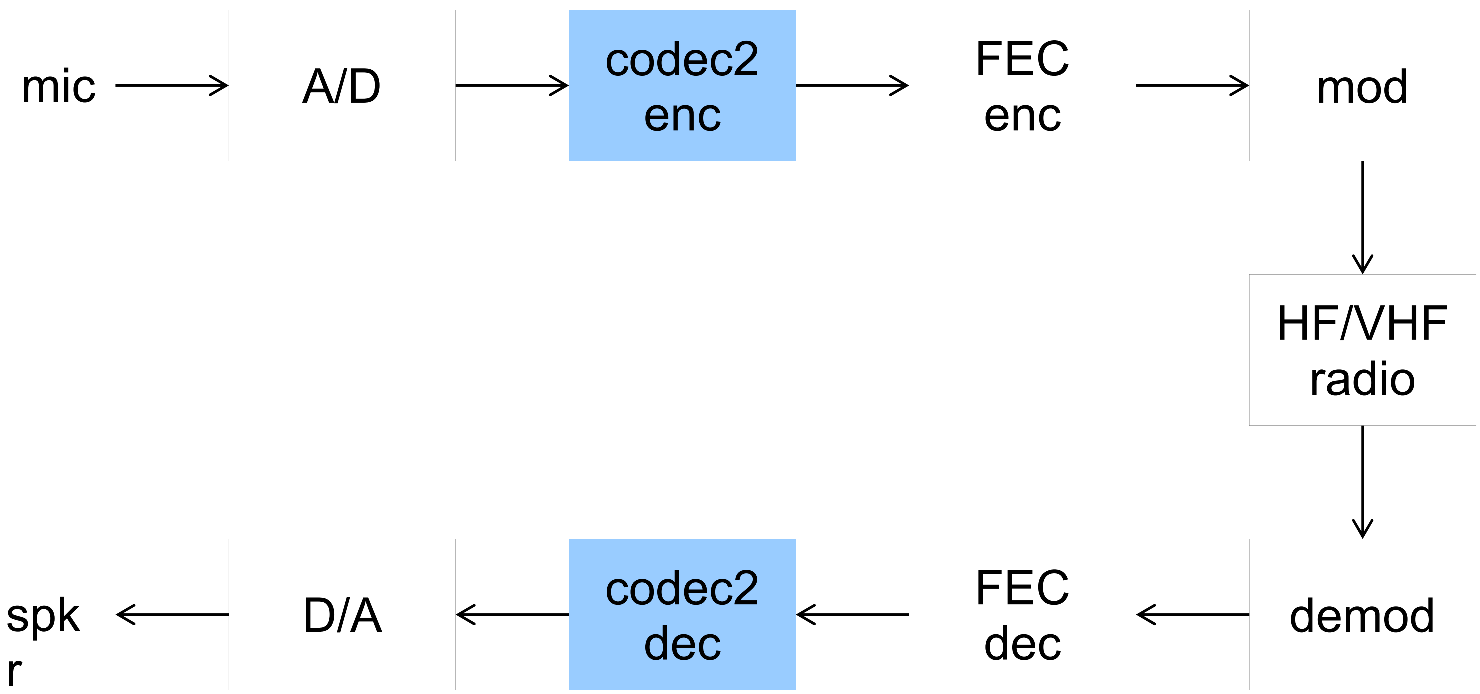
Proprietary Codecs

- come in hardware or licensed software form
- difficult to distribute
- they cannot be modified
- understanding how they work is discouraged
- modification may actually be illegal under the license

Codec 2 Author - David Rowe

- Adelaide, South Australia
- VK5DGR, first licensed over 30 years ago at age 13
- PhD in speech coding (1999)
- Built some of the first real time speech codecs in the late 1980's on early DSP chips
- Now work full time on open software/open hardware for developing world communications
- http://rowetel.com

Digital Voice Radio System



Patents and Codecs

- The authors of proprietary/patented codecs borrowed heavily from the public domain
- Perhaps 5% of the algorithms they use are original and patented
- 95% of the algorithms in these codecs are public domain algorithms
- To build an equivalent codec, we simply need alternatives for the 5% that is patented

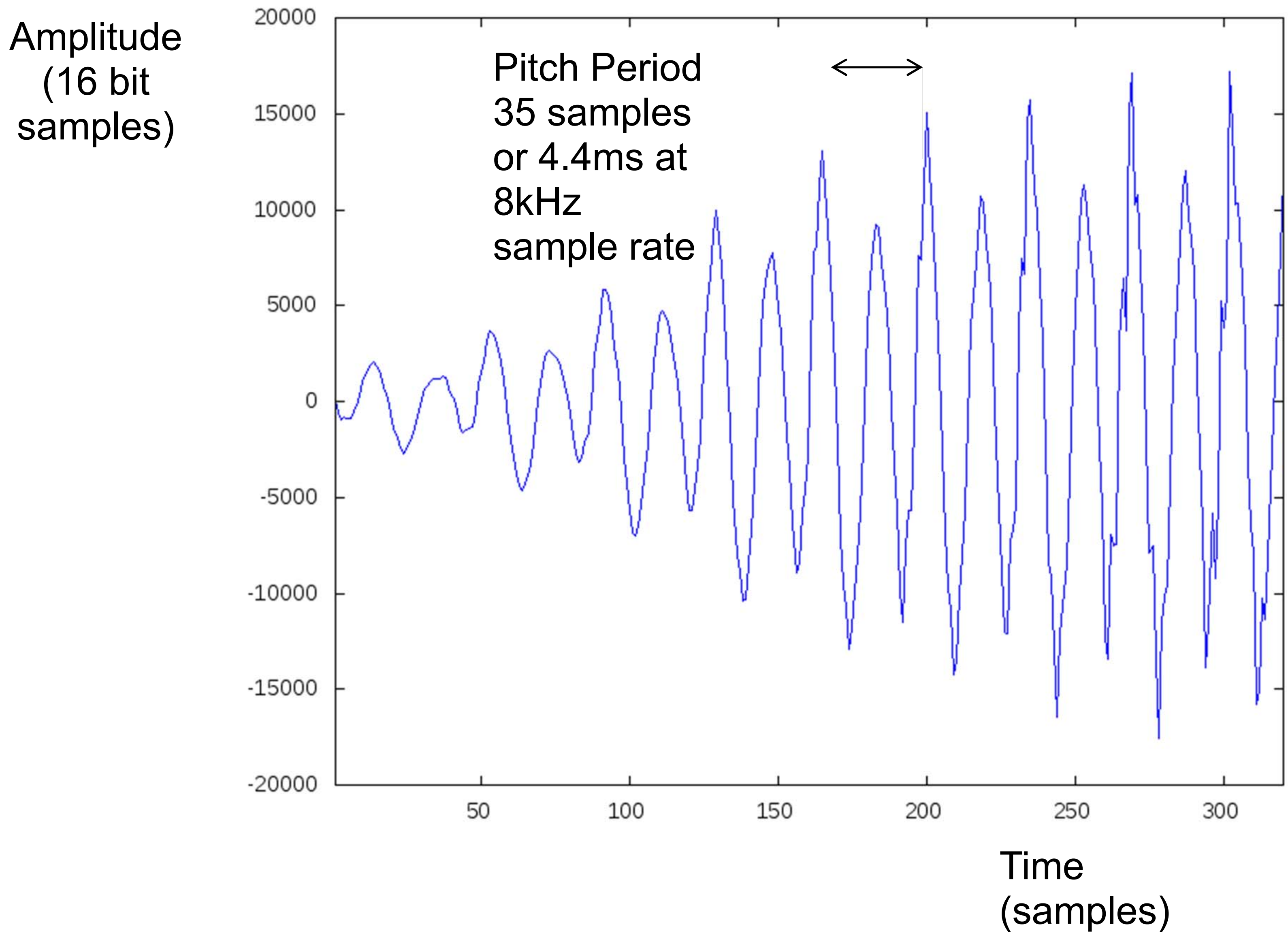
Speech Coding

- Take speech samples (e.g. 16 bit samples at 8 kHz sampling rate)
- Convert to 2400 bit/s
- What can we throw away?
- Retain intelligible speech
- Retain natural speech
- Use a model of speech, send model parameters

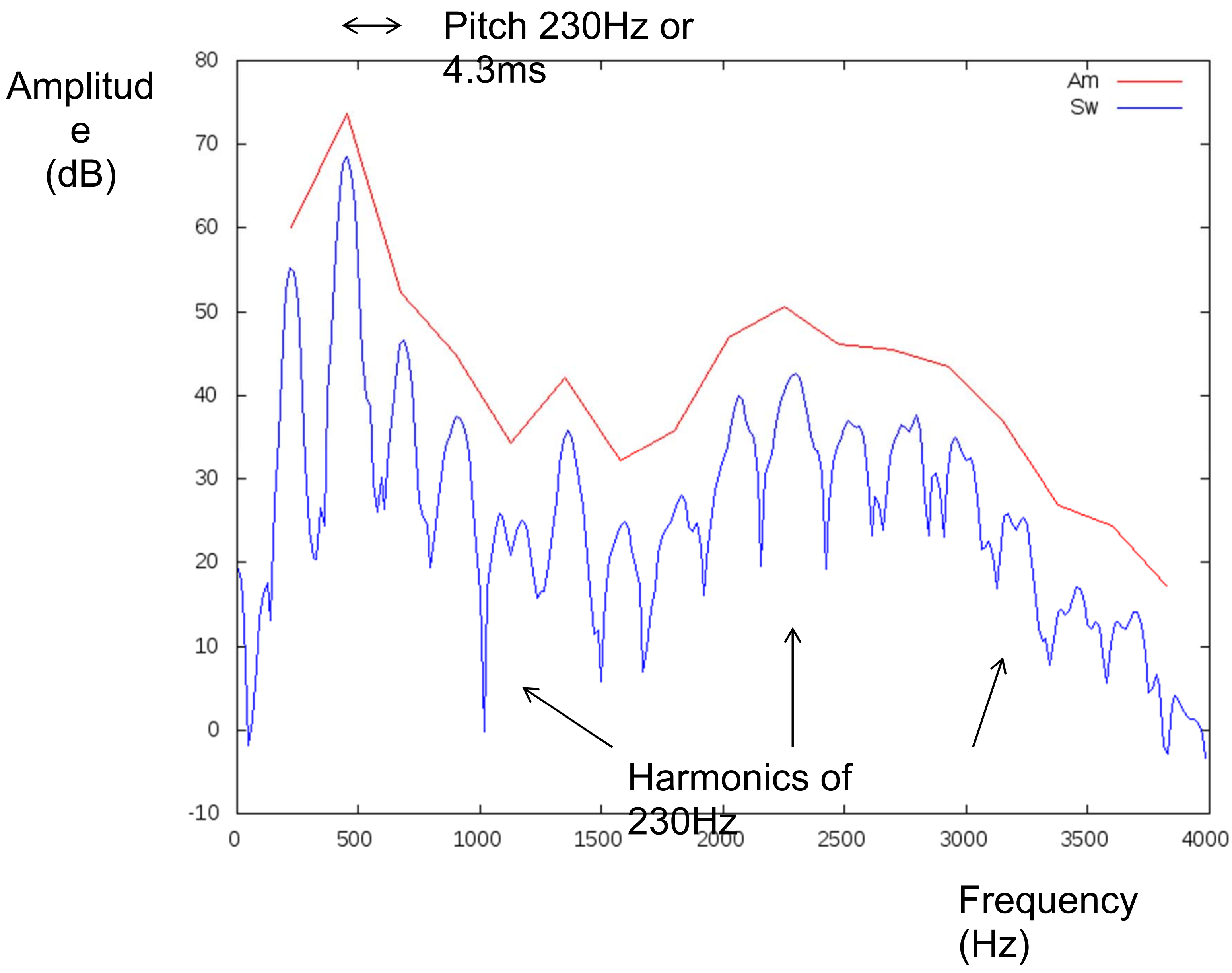
Model Parameter

- example of a model parameter is pitch
- for humans in the range 50 to 500 Hz
- can be quantised to 7 bits
- updated every 20 ms
- so $7/0.02 = 350$ bit/s to represent pitch

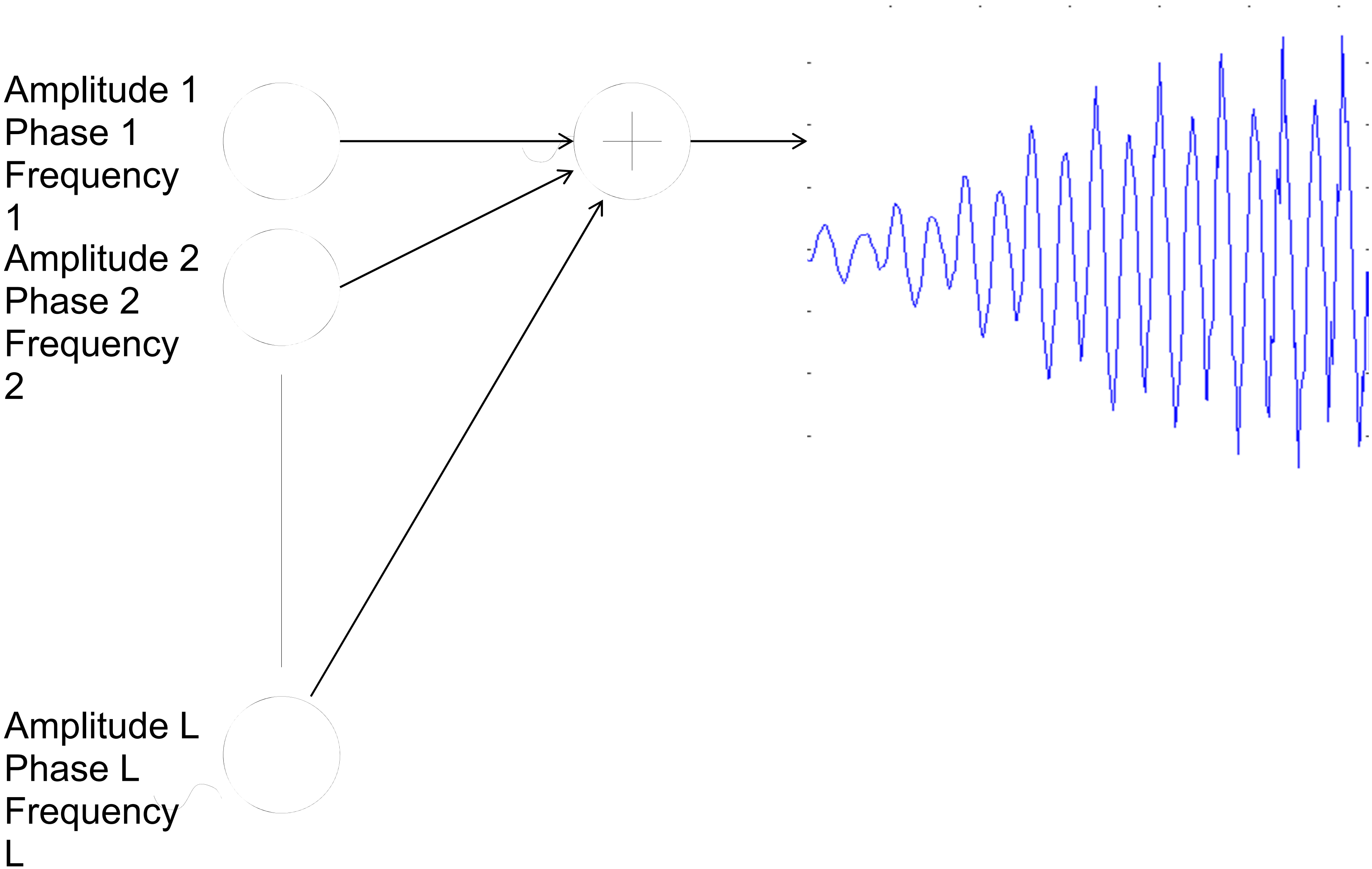
Sinusoidal Speech Coding



Sinusoidal Speech Coding



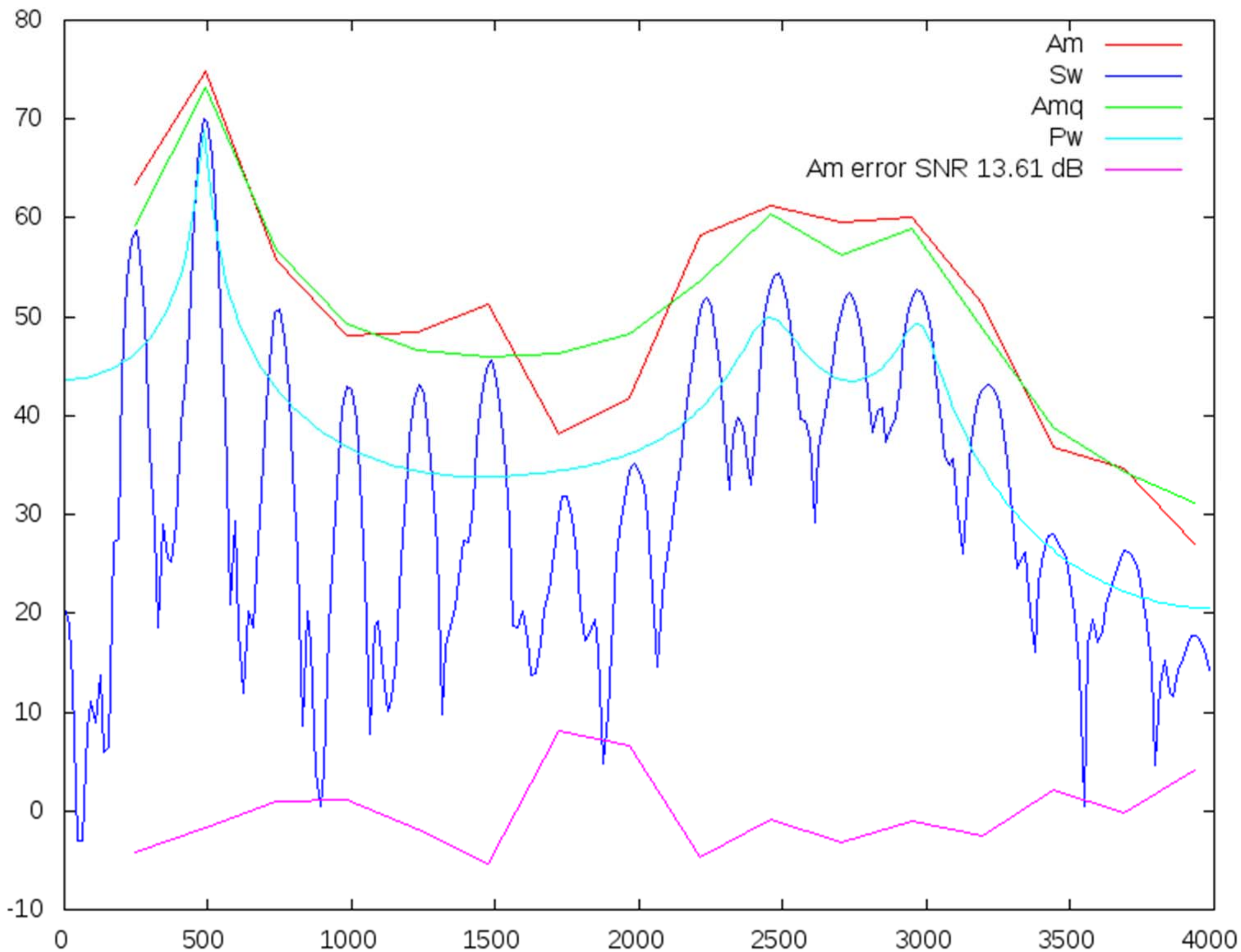
Sinusoidal Speech Model



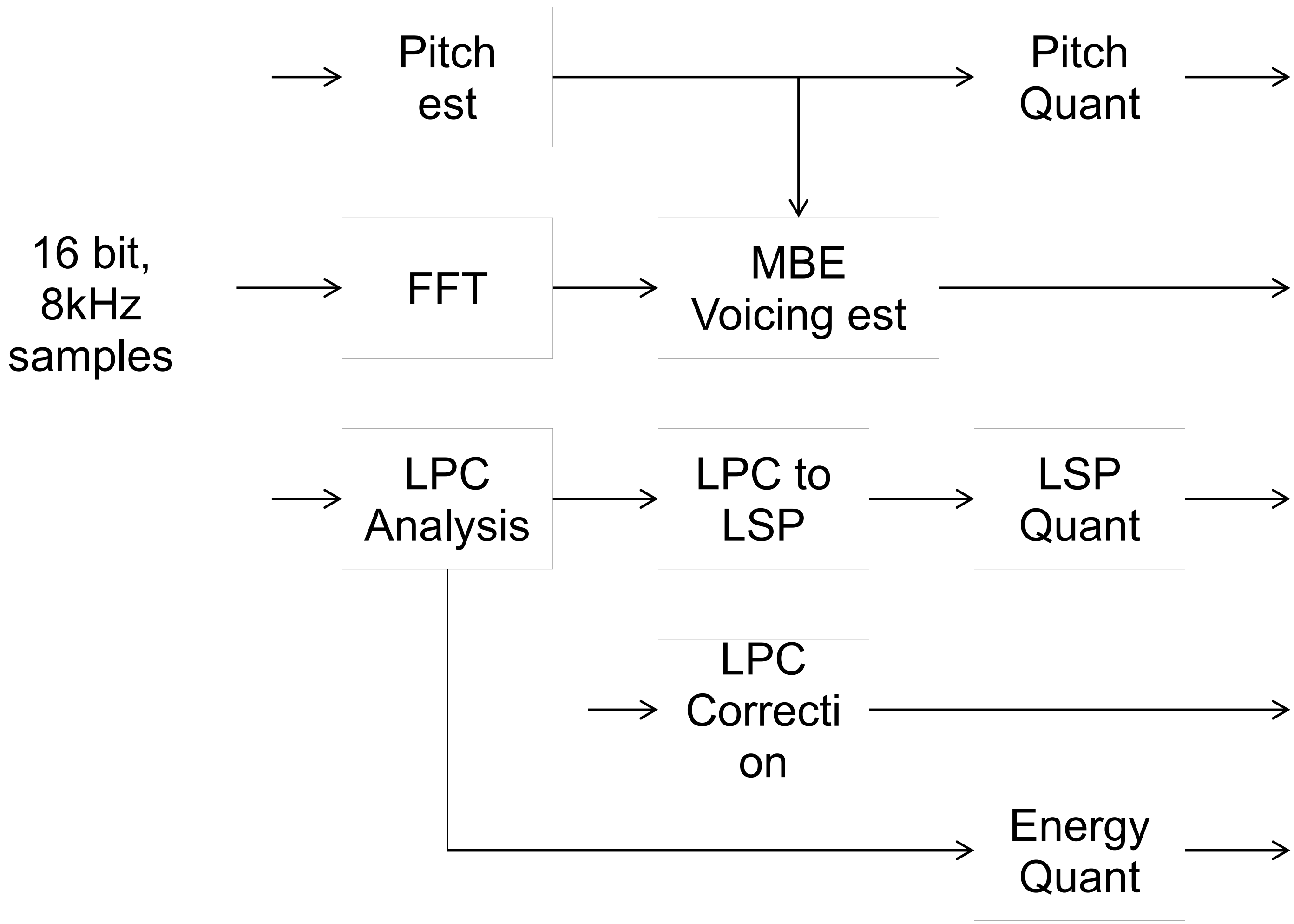
Amplitude Modelling

- Adjacent amplitudes have similar values
- This leads to coding efficiencies
- We use LPC to represent amplitudes
- fixed number of parameters
- LPC envelope approximates amplitudes
- Sampled at the decoder to recover amplitudes

Amplitude Modelling



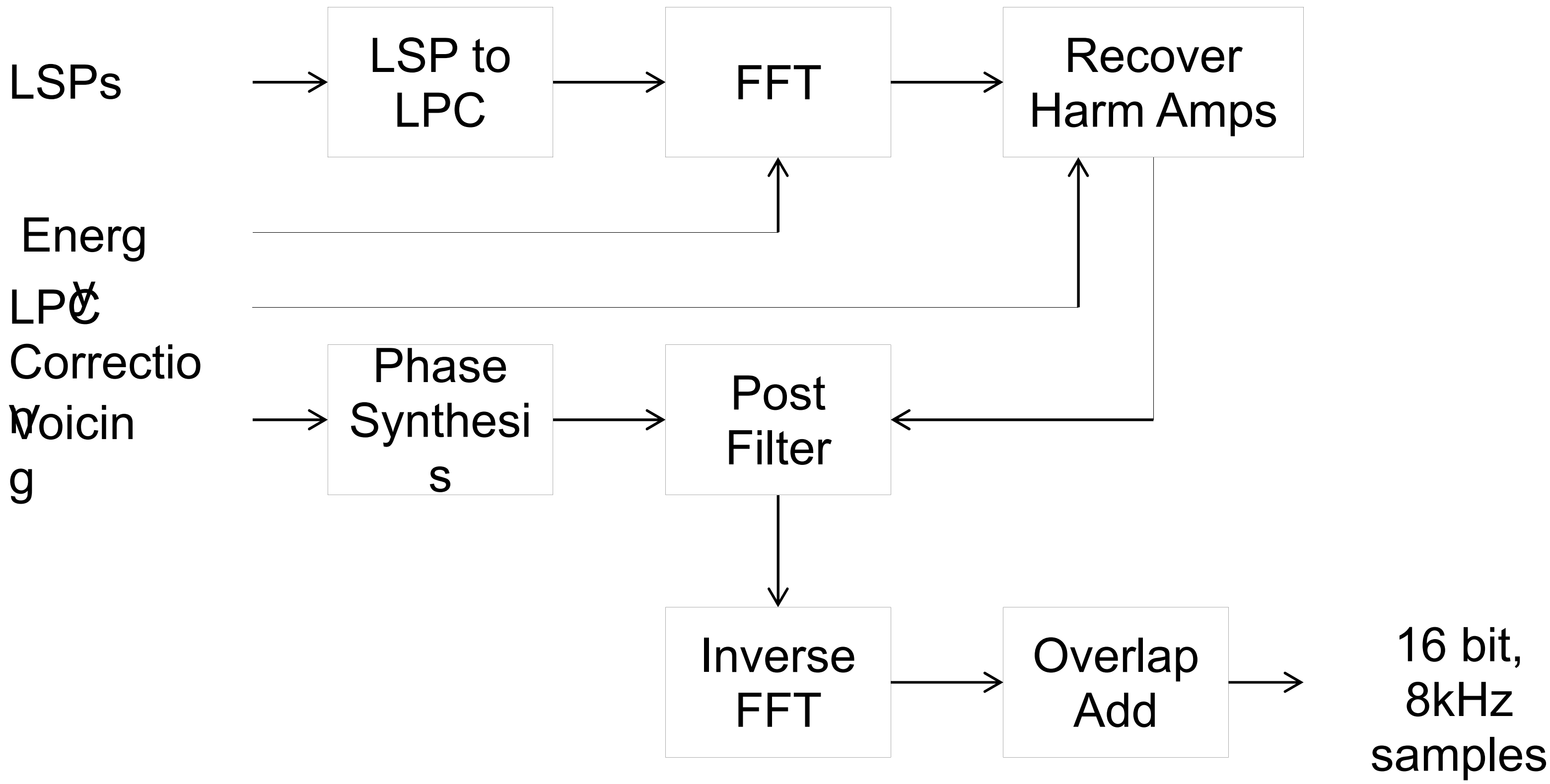
Encoder Block Diagram



Bit Allocation

- Alpha V0.1 codec, subject to rapid change
- 51 bits per 20ms frame, or 2550 bit/s

Decoder Block Diagram



Prior Art Summary

- Sinusoidal Coding, Mcaulay & Quatieri, 1984
- Linear Predictive Coding, Makhoul, 1975
- Line Spectrum Pairs, Itakura, 1975
- MBE Voicing, Griffin & Lim, 1988
- Overlap Add, Tribolet & Crochiere, 1979
- NLP Pitch Estimation, Rowe, 1999
- LPC Amplitude Recovery (algorithm used here), Rowe, 1991, 1999, 2009
- Post Filter, Rowe, 2009

Further Work

- Better phase model and voicing estimator
- Toll quality at 2000 bit/s
- Lower bit rate, 2400, 1200 bit/s
- Better background noise performance
- FEC and non-redundant error correction
- Integration with modem and test over radio channels
- Fixed point and DSP chip implementation

Brainstorms

- what can we do with Codec 2. HF rather than VHF?
- how can we get people using it?
- work with others to integrate with modem and FEC code
- create a digital voice application that can run on a laptop
- novel combinations of codec, FEC, modulation
- PSK31 low bit rate voice mode
- Better than SSB on HF?