Overlap-Add Windows with Maximum Energy Concentration for Speech and Audio Processing

Tom Bäckström, Aalto University, Department of Signal Processing and Acoustics, Espoo, Finland

Background

- Windowing for *analysis* is well-understood.
- Reconstruction of signals needed for *processing* applications.
- *Overlap-add* is frequently used for reconstruction.
- Optimal windows for overlap-add have not been available.

Overlap-add (OLA)

Objective: Minimize spectral leakage.
Constraints:

- Perfect reconstruction.
- Uniform output noise.

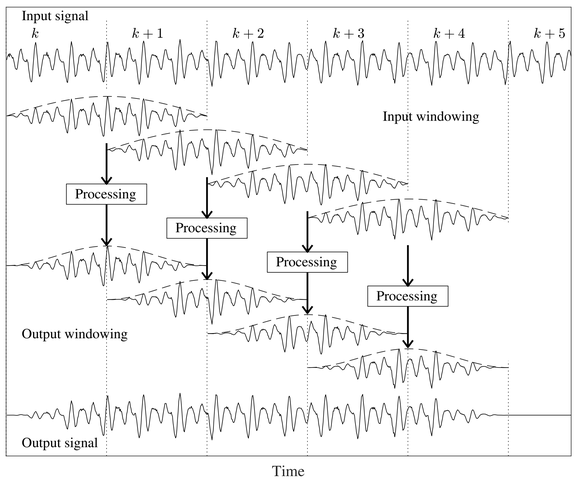

Figure: The overlap-add process.

Constraints for OLA

Let ω_n be a window with $n \in [0, L-1]$.
Constraints satisfied when

$$\omega_n^2 + \omega_{n+L/2}^2 = 1.$$

(The Princen-Bradley condition.)
For $w = [\omega_0, \ldots, \omega_{L-1}]^T$, equivalently

$$w^T \Lambda_n w = 1, \qquad \forall n \in [0, \tfrac{L}{2}) .$$

where

$$\Lambda_n = \mathrm{diag}(\underbrace{0, \ldots, 0}_{n-1}, 1, \underbrace{0, \ldots, 0}_{L/2-1}, 1, \underbrace{0, \ldots 0}_{L/2-n}).$$

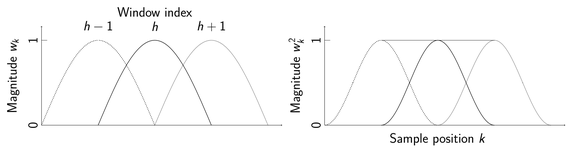

Figure: OLA with perfect reconstruction.

Energy Concentration

Energy concentration measure

$$\tau = \frac{\int_{-\delta}^{\delta} |W(f)|^2 df}{\int_{-\infty}^{\infty} |W(f)|^2 df} = \frac{w^T T w}{\|w\|^2}$$

where $T_{k,h} = \frac{L \sin\left(\frac{\pi}{L}\alpha(k-h)\right)}{(k-h)}$.

The classic Slepian / DPSS window is found by

$$\max w^T T w \quad \text{such that } \|w\| = 1.$$

Gives minimal spectral leakage.

OLA-DPSS window

Solution to the problem

$$\max w^T T w \quad \text{such that } w^T \Lambda_n w = 1$$

with $n \in [0, \frac{L}{2})$.

$\Rightarrow$ Quadratically constrained quadratic problem (QCQP).

$\Rightarrow$ Solvable with numerical methods.

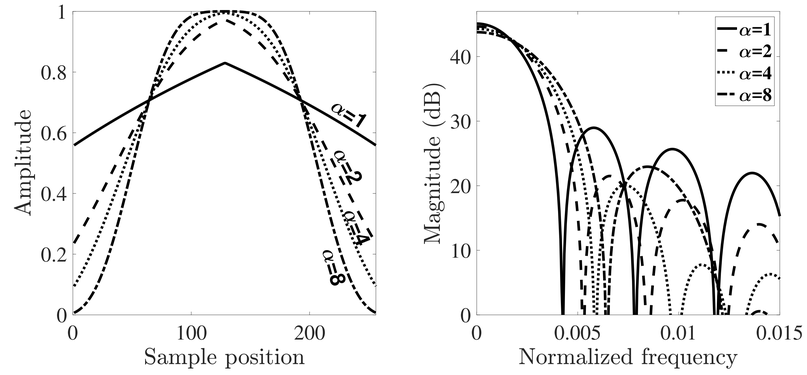

Figure: OLA-DPSS window shapes for $L = 256$.

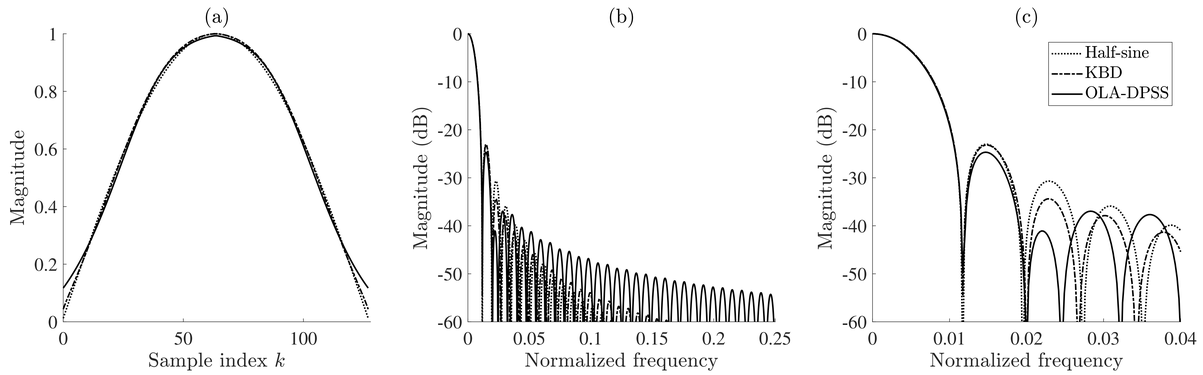

Figure: Windows for $L = 128$ where KBD has $\alpha = 4.25$ and OLA-DPSS has $\alpha = 2.75$.

Conclusion

- Optimal windows for overlap-add.
- Maximizes energy concentration = Minimal spectral leakage.
- Perfect reconstruction.
- Uniform output noise.
- Also low-overlap windows available.
- For any processing applications.