7.3.3 Overlap-Add (OLA) Method

Just as the FBS method was motivated from the filtering view of the STFT, the OLA method is motivated from the Fourier transform view of the STFT [1],[13],[20]. The simplest method obtainable from the Fourier transform view is, in fact, not the OLA method. It is instead a method known as the Inverse Discrete Fourier Transform (IDFT) method. In this method, for each fixed time, we take the inverse DFT of the corresponding frequency function and divide the result by the analysis window. This method is generally not favored in practical applications because a slight perturbation in the STFT can result in a synthesized signal very different from the original.[3] For example, consider the case where the STFT is multiplied by a linear phase factor of the form $e^{j\omega n_o}$ with n_o unknown. Then the IDFT for each fixed time results in a shifted version of the corresponding short-time section. Since the shift n_o is unknown, dividing by the analysis window without taking the shift into account, introduces a distortion in the resulting synthesized signal. In contrast, the OLA method, which we describe in this section, results in a shifted version of the original signal without distortion.

In the OLA method, we take the inverse DFT for each fixed time in the discrete STFT. However, instead of dividing out the analysis window from each of the resulting short-time sections, we perform an overlap and add operation between the short-time sections. This method works provided the analysis window is designed such that the overlap and add operation effectively eliminates the analysis window from the synthesized sequence. The intuition here is that the redundancy within overlapping segments and the averaging of the redundant samples remove the effect of windowing.

The OLA method is motivated by the following relation between a sequence and its discrete-time STFT:

$$x[n] = \frac{1}{2\pi W(0)} \int_{-\pi}^{\pi} \sum_{p=-\infty}^{\infty} X(p, \omega) e^{j\omega n} d\omega, \tag{7.19}$$

where

$$W(0) = \sum_{n=-\infty}^{\infty} w[n].$$

This synthesis equation can be thought of as the synthesis equation in Equation (7.8), averaged over many short-time segments and normalized by $W(0)$. For the derivation of this synthesis equation, the reader is referred to [13]. The OLA method carries out a discretized version of the operations suggested on the right of Equation (7.19). That is, given a discrete STFT $X(n, k)$, the OLA method synthesizes a sequence $y[n]$ given by

$$y[n] = \frac{1}{W(0)} \sum_{p=-\infty}^{\infty} \left[\frac{1}{N} \sum_{k=0}^{N-1} X(p, k) e^{j\frac{2\pi}{N}kn} \right].$$

The term inside the rectangular brackets is an inverse DFT which for each p gives us

$$f_p[n] = x[n]w[p - n],$$

[3] One might consider averaging out this effect by summing many inverted DFTs, with the analysis window divided out, at each time n. This is in fact the strategy of the OLA method, but without the need to divide out each IDFT by the window shape.