Chapter 4 High Frequency Reconstruction Method by Linear Exploration

The method is a frequency-domain approach since we reconstruct the high frequency signals in the frequency domain. Let $X[k]$ be the spectrum signals at some time frame. The method reconstructs the high frequency signals with a linear extrapolation on the magnitude with logarithm scale. We adopt the logarithm scale due to the fitting with the magnitude absorption model [24]. On the other hand, the frequency scale will be in linear model due to the harmonic extension in linear scale. On the assumption, we try to find the envelope of the high frequency through the linear extrapolation of signals with frequencies lower than the reconstructed point, say k_c , and then replicate low frequency spectrum to high frequency fitting the envelope defined.

Figure *24* illustrates the concept.

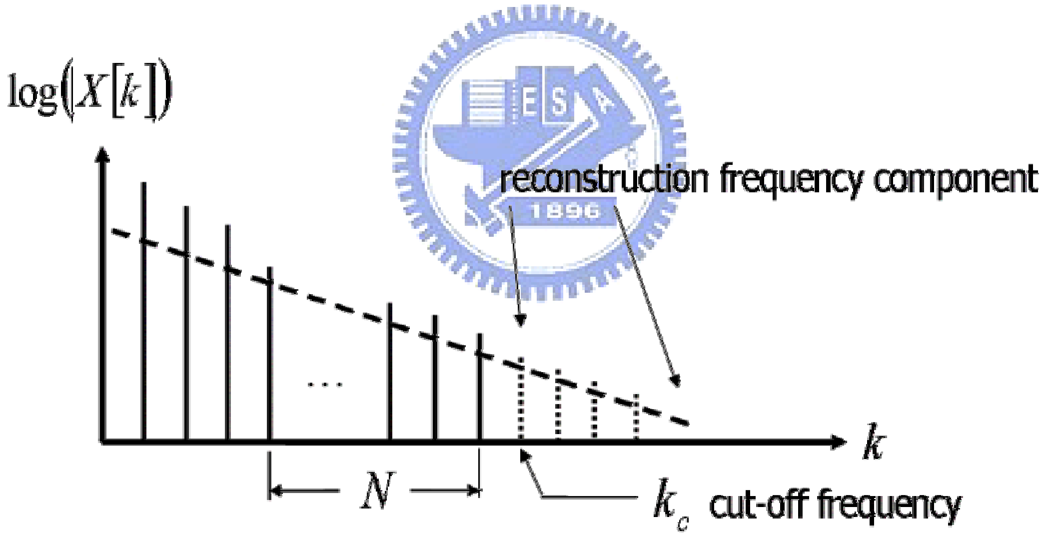

Figure 24: Linear extrapolation on the magnitude with logarithm scale.

4.1 Least Squares Method by Linear Model

The envelope is basically evaluated by the following theorem:

Theorem Given a set Γ consists of N frequency lines with logarithm magnitude; that is

$$\Gamma = \{ \ln(|X[k_c - N]|), \ln(|X[k_c - (N-1)]|), \ldots, \ln(|X[k_c - 1]|) \}. \tag{73}$$

Let L: $\ln|X[k]| = a_{opt} \cdot k + b_{opt}$ be the linear approximation with the least-square method on the N frequency lines. Then

$$a_{opt} = \frac{12}{(N-1)N(N+1)} \cdot \ln\left\{ \prod_{i=1}^{\frac{N-1}{2}} \left[\frac{|X[k_c - i]|}{|X[k_c - (N+1-i)]|} \right]^{\left(\frac{N+1}{2} - i\right)} \right\}, \tag{74}$$

and

$$b_{opt} = \frac{\ln\left(\prod_{i=1}^{N} |X[k_c - i]| \right)}{N} - \left(k_c - \frac{N+1}{2} \right) a_{opt}. \tag{75}$$

<Proof>

We need to find b_{opt} and a_{opt} such that the summation

$$\sum_{i=1}^{N} [b + (k_c - i) \cdot a - \ln|(X[k_c - i])|]^2 = \sum_{i=1}^{N} [b + (k_c - i) \cdot a - X'[k_c - i]]^2$$
$$= \left\| \begin{bmatrix} 1 & k_c - 1 \\ 1 & k_c - 2 \\ \vdots & \vdots \\ 1 & k_c - N \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} - \begin{bmatrix} X'[k_c - 1] \\ X'[k_c - 2] \\ \vdots \\ X'[k_c - N] \end{bmatrix} \right\|^2 \tag{76}$$

has the minimum value, where $X'[k_c - i] = \ln(|X[k_c - i]|)$.

The problem can be solved by solving a normal equation; that is,

$$\begin{bmatrix} 1 & 1 & ,\ldots, & 1 \\ k_c - 1 & k_c - 2 & ,\ldots, & k_c - N \end{bmatrix} \begin{bmatrix} 1 & k_c - 1 \\ 1 & k_c - 2 \\ \vdots & \vdots \\ 1 & k_c - N \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 1 & 1 & ,\ldots, & 1 \\ k_c - 1 & k_c - 2 & ,\ldots, & k_c - N \end{bmatrix} \begin{bmatrix} X'[k_c - 1] \\ X'[k_c - 2] \\ \vdots \\ X'[k_c - N] \end{bmatrix} \tag{77}$$

This is equivalent to solve the equation (78).

$$\begin{bmatrix} N & Nk_c - \frac{N(N+1)}{2} \\ Nk_c - \frac{N(N+1)}{2} & Nk_c^2 - N(N+1)k_c + \frac{N(N+1)(2N+1)}{6} \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{N} X'[k_c - i] \\ k_c \cdot \sum_{i=1}^{N} X'[k_c - i] - \sum_{i=1}^{N} i \cdot X'[k_c - i] \end{bmatrix} \tag{78}$$

By Gaussian-Jordan elimination method, (78) can be reduced to

$$\begin{bmatrix} 1 & k_c - \frac{N+1}{2} \\ 0 & \frac{(N-1)N(N+1)}{12} \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} \frac{\sum_{i=1}^{N} X'[k_c - i]}{N} \\ \sum_{i=1}^{N} \left(\frac{N+1}{2} - i \right) X'[k_c - i] \end{bmatrix} \tag{79}$$

The optimum solution b_{opt} and a_{opt} can be found by solving (79). ■

The complexity to calculate a_{opt} is $O(N^2)$ and to calculate b_{opt} is $O(N)$, where N is the number of frequency lines to predict the envelope. In next section, a fast computing method is proposed.

4.2 Fast Computing method for a_{opt} and b_{opt}

Assume integer N is odd and $N > 1$. We denote Y_i and W_i in (74) according to

$$Y_i = X[k_c - i]; \quad for\ i = 1, 2, ..., \frac{N-1}{2}. \tag{80}$$

and

$$W_i = X[k_c - (N+1-i)]; \quad for\ i = 1, 2, ..., \frac{N-1}{2}. \tag{81}$$

Substituting (80) and (81) to (74) yields

$$a_{opt} = \frac{12}{(N-1)N(N+1)} \cdot \left\{ \ln\left[\prod_{i=1}^{\frac{N-1}{2}} |Y_i|^{\left(\frac{N+1}{2} - i\right)} \right] - \ln\left[\prod_{i=1}^{\frac{N-1}{2}} |W_i|^{\left(\frac{N+1}{2} - i\right)} \right] \right\}. \tag{82}$$

That is

$$a_{opt} = \frac{12}{(N-1)N(N+1)} \cdot \left\{ \ln\left[\left|\prod_{i=1}^{\frac{N-1}{2}} Y_i^{\left(\frac{N+1}{2}-i\right)}\right|\right] - \ln\left[\left|\prod_{i=1}^{\frac{N-1}{2}} W_i^{\left(\frac{N+1}{2}-i\right)}\right|\right] \right\}. \tag{83}$$

Furthermore, we define the product of a series of Y_j as Z_i, that is

$$Z_i = \prod_{j=1}^{i} Y_j\,; \quad for\ i = 1, 2, ..., \tfrac{N-1}{2}. \tag{84}$$

Taking a recursive way to calculate Z_i leads to

$$Z_i = Z_{i-1} \cdot Y_i\,; \quad for\ i = 1, 2, ..., \tfrac{N-1}{2}. \tag{85}$$

Z_0 =1. Similarly, we define the product of a series of W_j as V_i,

$$V_i = \prod_{j=1}^{i} W_j\,; \quad for\ i = 1, 2, ..., \tfrac{N-1}{2}. \tag{86}$$

Taking a recursive way to calculate V_i leads to

$$V_i = V_{i-1} \cdot W_i\ ; \quad for\ i = 1, 2, ..., \tfrac{N-1}{2}. \tag{87}$$

V_0 =1. The recursive forms in (85) and (87) can be derived as

$$\prod_{i=1}^{\frac{N-1}{2}} Y_i^{\left(\frac{N-1}{2}-i\right)} = \prod_{i=1}^{\frac{N-1}{2}} Z_i \quad , \tag{88}$$

and

$$\prod_{i=1}^{\frac{N-1}{2}} W_i^{\left(\frac{N-1}{2}-i\right)} = \prod_{i=1}^{\frac{N-1}{2}} V_i \quad . \tag{89}$$

Substituting (87) and (88) to (83) yields

$$\mathrm{a}_{\mathrm{opt}} = \frac{12}{(\mathrm{N}-1)N(N+1)} \cdot \ln\left(\left|\frac{\prod_{\mathrm{i}=1}^{\frac{\mathrm{N}-1}{2}} \mathrm{Z_i}}{\prod_{\mathrm{i}=1}^{\frac{\mathrm{N}-1}{2}} \mathrm{V_i}}\right|\right). \tag{90}$$

From (90), the complexity of computing the values of Z_i needs $(N–3)/2$ multiplications. To compute the product of Z_i, it also needs $(N–3)/2$ multiplications. Hence, computing $\prod_{i=1}^{\frac{N-1}{2}} Z_i$ totally needs $N-3$ multiplications. Similarly, to compute the value of $\prod_{i=1}^{\frac{N-1}{2}} V_i$ needs $N-3$ multiplications. Using (90) to calculate a_{opt} needs *2N–6* multiplications. Thus, computing (90) leads to a linear complexity and needs only one logarithm, division and absolute operation, respectively. On the other hand, computing b_{opt} needs a constant complexity due to

$$Z_{\frac{N-1}{2}} \cdot V_{\frac{N-1}{2}} \cdot X\left(k_c - \frac{N+1}{2}\right) = \prod_{i=1}^{N} X\left(k_c - i\right) \tag{91}$$

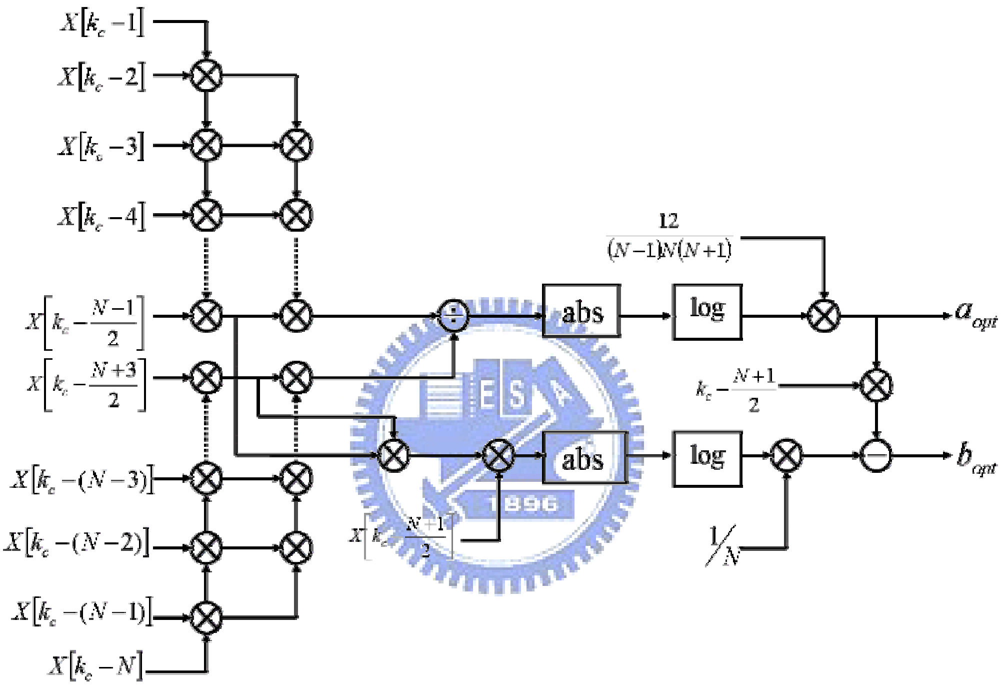

Figure 25: Signal flow diagram of the fast computing method.

4.3 Reconstruction Algorithm

This section presents the algorithm for high frequency reconstruction based on the linear model above. The algorithm consists of three components that include envelope extractor module, spectrum duplication module, and envelope adjustment module. The block diagram of the reconstruction algorithm is illustrated in Figure 26. At first, based on the low frequency, envelope extractor module will determine a suitable cut-off frequency point k_c and calculate a_{opt} and b_{opt} by fast computing method. In turn, the spectrum duplication module will generate high frequency by duplicate low frequency. Ultimately, the envelope adjustment module will adjust the high frequency to fit the defined envelope. These following subsections will exploit the three components in detail.

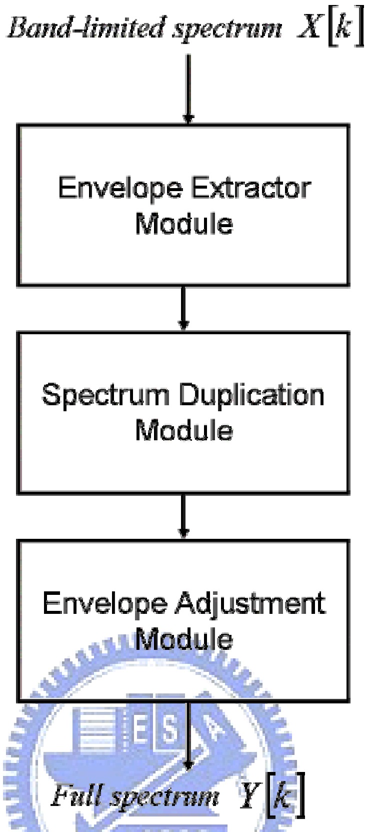

Figure 26: The block diagram of high frequency reconstruction method.

4.3.1 Envelope Extractor Module

In order to extract the two parameters a_{opt} and b_{opt} of the spectral envelope, the determining of cut-off frequency should be considered at first. In encoder, to truncate the high frequency component, the time-domain audio signal is usually filtered by a low pass filter with appropriate cut-off frequency before encoding, otherwise, as an equivalent method without filters, the high frequency signal over the cut-off frequency is reset as zero prior to bit allocation and then no bits are allocated to high frequency component. If the cut-off frequency is fixed, the reconstruction algorithm can correctly reconstruct high frequency. However, once the cut-off frequency is unknown or not fixed, an automatic determining mechanism should be incorporated into the envelope extractor module to offer a correct cut-off frequency dynamically.

We define the initial candidate cut-off frequency k'_c as the first frequency such that $X[k]$ is zero for all $k \geq k'_c$, and the selection unit relative to k'_c as a spectral band that consists of $X[k]$ for $k_c - U$ to $k_c - 1$. Whenever a candidate cut-off frequency is chosen, the two parameters a_{opt} and b_{opt} can be calculated by the fast computing method. Furthermore, the relative selection unit of the candidate cut frequency will be measured its quality. A shattered unit due to compression should be replaced by a pseudo unit after the high frequency reconstruction to enhance the perceptual quality and to ease the artifact of connection of spectrum duplication. A ratio, say selection ratio φ, of the summation of the frequency magnitudes on the reconstructed unit and the relative summation of estimated pseudo magnitudes is calculated.

$$Selection\ \ Ratio\ \varphi = \frac{\sum_{i=1}^{U} \left| X[k_c - i] \right|}{\sum_{i=1}^{U} X_p[k_c - i]} \qquad (92)$$

where

$$\sum_{i=1}^{U} X_P[k_c - i] = \sum_{i=1}^{U} \exp^{b_{opt} + a_{opt}(k_c - i)} \qquad (93)$$

Substituting (93) into (92) leads to (94),

$$\varphi = \begin{cases} \dfrac{\exp^{a_{opt}} - 1}{\exp^{b_{opt} + a_{opt}(k_c - U)} \cdot \left(\exp^{a_{opt} U} - 1 \right)} \cdot \sum_{i=1}^{U} \left| X[k_c - i] \right|, & if\ a_{opt} \neq 0. \\ \dfrac{\sum_{i=1}^{U} \left| X[k_c - i] \right|}{U \exp^{b_{opt}}}, & if\ a_{opt} = 0. \end{cases} \qquad (94)$$

If selection ratio φ is lower than a threshold like 0.25, the candidate cut-off frequency should be decreased a unit length U and check whether the new candidate cut frequency is available again. On the other hand, because the first candidate selection unit is usually imperfect due to low pass filtering, the selection threshold can be raised as a stricter criterion for the first time of judgment. Once the selection ratio passes through the threshold, the high frequency reconstruction algorithm will continue to the next module with the ultimately determined cut-off frequency. Figure 27 illustrates the conception of the determining mechanism.

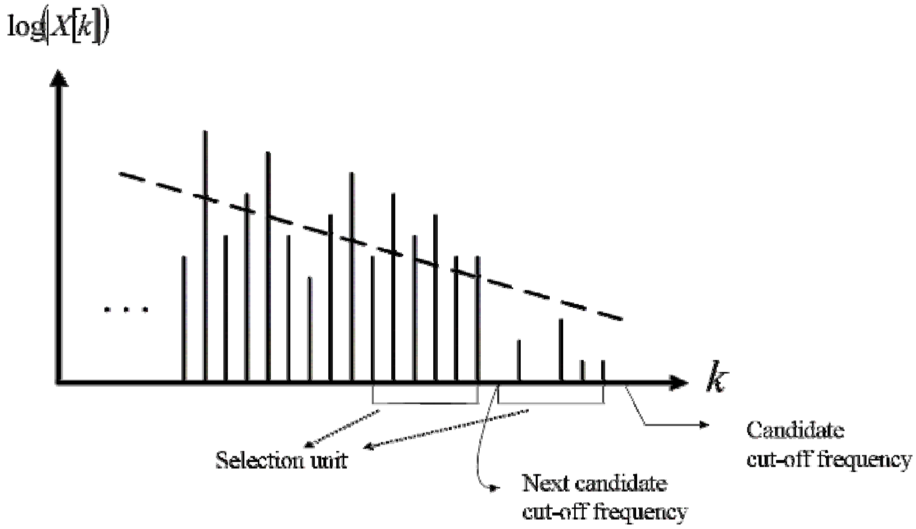

Figure 27: An automatic determining mechanism of cut-off frequency.

There are still two calibrations required for the module. The first calibration is on the dithering of the zero magnitude to avoid the undefined problem of the logarithm operation of zero. The zero magnitudes of frequency lines are replaced with a small random number μ. Let μ be $\varepsilon \times \gamma$. γ is a random number between -1 to 1. ε needs to adapt with the audio frames. A too large or small ε will affect the evaluation of envelope. This thesis calculates the minimum average magnitude of the three pieces of spectrum that consist of nine reconstruction units relative to initial candidate cut-off frequency k'_c to have ε. More precise,

$$\varepsilon = \min\left[\frac{\sum_{k=k'_c-3m}^{k'_c-1}|X[k]|}{3m}, \frac{\sum_{k=k'_c-6m}^{k'_c-3m-1}|X[k]|}{3m}, \frac{\sum_{k=k'_c-9m}^{k'_c-6m-1}|X[k]|}{3m}\right]. \quad (95)$$

The second calibration is on the envelope parameter a_{opt}. a_{opt} should be constrained to be non-positive. Hence, we set those positive a_{opt} to 0 to avoid the increasing envelope. And then b_{opt} is computed as the average of the logarithm magnitude of the N frequency lines. On the other hand, when cut frequency is too low (for example, lower then 13.5 kHz), there are fewer information available to predict the envelope. Under the condition, the predicted slope a_{opt} of the linear model usually approaches zero and is insufficient to responses the actual variation of the envelope. Such a_{opt} results in a too slow decay rate of the envelope, such that the energy of the reconstructed high frequency is excessive and lead to very annoy "metal friction" noise. Hence, we constrain a_{opt} to be *-0.3* under the condition to reduce the risk of reconstruction if a_{opt} is larger then *-0.3*. However, if b_{opt} is computed by (75) directly, in the future, it will result a fault envelope illustrated in Figure 28. Once original a_{opt} is modified as a'_{opt}, b_{opt} should be modified as b'_{opt} to adjust the envelope as Figure 29.

$$b'_{opt} = b_{opt} + (a_{opt} - a'_{opt}) \cdot k_c \qquad (96)$$

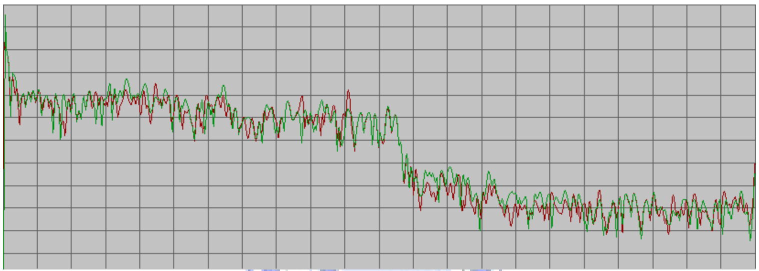

Figure 28: A fault envelope due to a modified model that only adjusts a_{opt}.

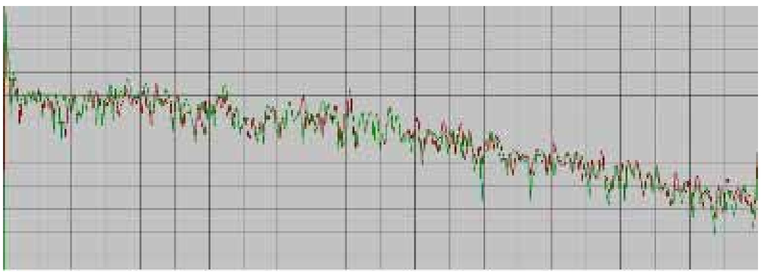

Figure 29: A suitable envelope reconstructed by a modified model that adjusts both a_{opt} and b_{opt}.

4.3.2 Spectrum Duplication Module

The objective of the module is to regenerate high frequency spectrum by duplicate low frequency. Assume k_e is reconstruction-ended frequency. The module will duplicate a long piece of the low frequency spectrum which bandwidth is bw to high frequency from $X[k_c - bw]$ to $X[k_c - 1]$, where $bw = \min\{N, k_e - k_c + 1\}$. More precise,

$$X[k] = X[k - bw],\ for\ k = k_c \sim k_e. \tag{97}$$

Figure 30 illustrate a result after the module, where a spectrum with excessive energy is reconstructed on high frequency and hence envelope adjustment is necessary.

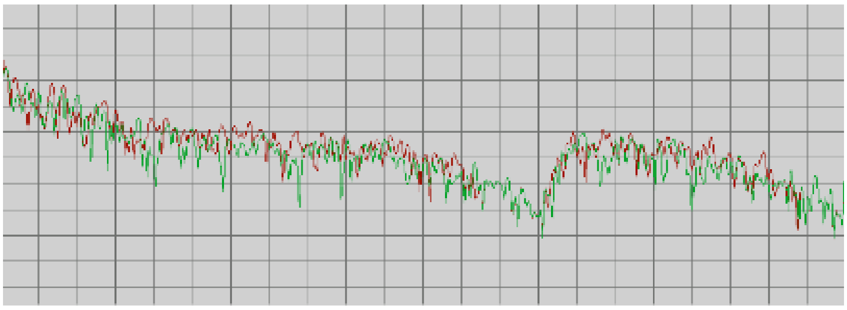

Figure 30: A spectrum with an unsuitable envelope on high frequency after spectrum duplication module.

4.3.3 Envelope Adjustment Module

The module will adjust the energy of the reconstructed high frequency after spectrum duplication module to fit the envelope defined. We separate the reconstructed high frequency M sections, refers to reconstruction units, which consist of m frequency lines individually. The adjustment is based on the units, not line level, to keep the harmonic of the spectrum in detail. For an envelope adjustment unit, the relative adjustment ratio α is defined as (100).

$$\alpha_u = \sqrt{\frac{E_u}{S_u}},\ for\ u = 0 \sim M - 1 \tag{98}$$

where S_u is the energy of the reconstructed high frequency on the uth unit and E_u is the relative pseudo energy. S_u and E_u are defined as (99) and (100), respectively.

$$S_u = \sum_{k=k_c+mu}^{k_c+m(u+1)-1} X[k]^2,\ for\ u = 0 \sim M - 1 \tag{99}$$

and

$$E_u = \sum_{k=k_c+mu}^{k_c+m(u+1)-1} \left(\exp^{b_{opt}+a_{opt}k}\right)^2,\ for\ u = 0 \sim M - 1 \tag{100}$$

Furthermore, (100) can be reduced to (101).

$$E_u = \begin{cases} \left[\exp^{2b_{opt}+2a_{opt}k_c} \cdot \frac{1-\exp^{2a_{opt}m}}{1-\exp^{2a_{opt}}} \right] \cdot \left(\exp^{2a_{opt}m}\right)^u, \ if \ a_{opt} \neq 0 \\ m\exp^{2b_{opt}}, if \ a_{opt} = 0. \end{cases} \quad (101)$$

(101) suggests an efficient recursive computing method for E_u with complexity $O(M)$, instead of $O(mM)$.

$$E_u = E_{u-1} \cdot \exp^{2a_{opt}} = E_{u-1} \cdot \rho, for \ u = 1 \sim M-1 \quad (102)$$

where the ratio $\exp^{2a_{opt}}$ is referred to unit decay ratio ρ, and E_0 is defined as (103).

$$E_0 = \begin{cases} \left[\exp^{2b_{opt}+2a_{opt}k_c} \cdot \frac{1-\rho}{1-\exp^{2a_{opt}}} \right], \ if \ a_{opt} \neq 0 \\ m\exp^{2b_{opt}}, if \ a_{opt} = 0. \end{cases} \quad (103)$$

On the other hand, as E_u is smaller then a threshold, we let E_u be E_{u-1} to keep the energy not decay to zero. The eventual adjusted high frequency $X'[k]$ is given as (104).

$$X'[k] = X[k] \cdot \alpha_u, \quad for \ k = k_c + um \sim k_c + (u+1)m - 1, and \ u = 0 \sim M-1 \quad (104)$$

After the module, the reconstruction algorithm is completed.

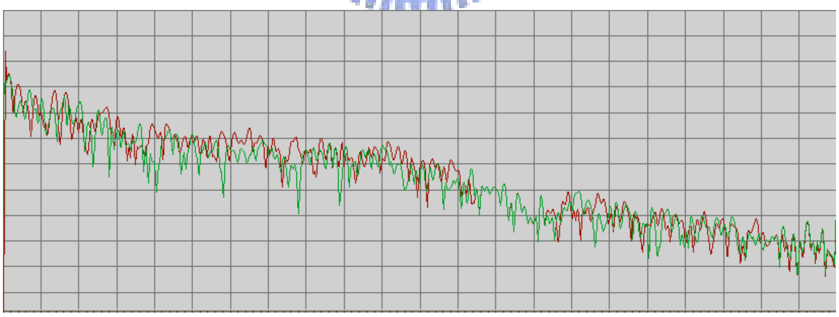

Figure 31: The unsuitable spectrum in Figure 30 is adjusted after envelope adjustment module.

4.4 Reconstruction Algorithm

The algorithm can be summarized as follows:

Input data: The basic sources to extend bandwidth are described below. (a) $X[k]$: Frequency-domain audio signal (b) k_e: Reconstruction-ended frequency

(c) N : An odd number of frequency lines for prediction
(d) U : Selection unit length
(e) m : Reconstructed unit length
There are total fourteen steps of the algorithm expressed as follow:
Step1: Initial cut frequency k'_c finding.
Step2: Zero sample replacing.
Step3: Calculate Z_i and V_i recursively

(a) Let $Z_0 = 1$ and $V_0 = 1$

(b) Let $Z_i = Z_{i-1} \cdot X[k_c - i]$ and $V_i = V_{i-1} \cdot X[k_c - (N+1-i)]$ for i =1 to N.

Step4: Calculate $P_Z = \prod_{i=1}^{\frac{N-1}{2}} Z_i$ and $P_V = \prod_{i=1}^{\frac{N-1}{2}} V_i$ respectively.

Step5: Calculate $a_{opt} = \frac{12}{(N-1)\cdot N \cdot (N+1)} \cdot \ln\left(\left|\frac{P_Z}{P_V}\right|\right)$.

Step6: If a_{opt} >0, let a_{opt} =0.

Step7: Calculate $b_{opt} = \frac{\ln\left(Z_{\frac{N-1}{2}} \cdot V_{\frac{N-1}{2}} \cdot X\left(k_c - \frac{N+1}{2}\right)\right)}{N} - \left(k_c - \frac{N+1}{2}\right) \cdot a_{opt}$.

Step8: Calculate Selection Ratio φ

$$\varphi = \begin{cases} \frac{\exp^{a_{opt}} - 1}{\exp^{b_{opt} + a_{opt}(k_c - U)}\left(\exp^{a_{opt}U} - 1\right)} \cdot \sum_{i=1}^{U} |X[k_c - i]| & if\ a_{opt} \neq 0 \\ \frac{\sum_{i=1}^{U} |X[k_c - i]|}{U \exp^{b_{opt}}} & if\ a_{opt} = 0 \end{cases}$$

If φ < threshold, k_c is subtracted U, and go to step 3.Otherwise, go to Step 9.

Step9: If k_c < low bound of cut frequency and $a_{opt} > a'_{opt}$, then

(1) Let b_{opt} be modified as $b'_{opt} : b'_{opt} = b_{opt} + (a_{opt} - a'_{opt})k_c$.

(2) Let a_{opt} be modified as a'_{opt} .

Step10: Duplicate low frequency to high frequency.

$$X[k] = X[k - bw],\ for\ k = k_c \sim \left(m \cdot \left\lfloor \frac{(k_e - k_c + 1)}{m} \right\rfloor \right).$$

$$bw = \min\{N, k_e - k_c + 1\}.$$

Step11: Calculate $S_u = \sum_{k=k_c+mu}^{k_c+m(u+1)-1} X[k]^2$, for u=0~M. ($M = \left\lfloor \frac{(k_e - k_c + 1)}{m} \right\rfloor$)

Step12: Calculate pseudo energy E_u.

1. Calculate unit decay ratio $\rho = \exp^{2a_{opt}}$.

2. Calculate $E_0 = \begin{cases} \left[\exp^{2b_{opt}+2a_{opt}k_c} \cdot \frac{1-\rho}{1-\exp^{2a_{opt}}} \right], if\ a_{opt} \neq 0 \\ m\exp^{2b_{opt}}, if\ a_{opt} = 0. \end{cases}$.

3. Calculate $E_u = E_{u-1} \cdot \rho$, for u=1~M.

Step13: Calculate adjustment ratio $\alpha_u = \sqrt{\frac{E_u}{S_u}}$, for u=1~M.

Step14: Adjust the reconstructed spectrum.

$X[k]' = X[k] \cdot \alpha_u$, for $k = k_c + um \sim k_c + (u+1)m - 1$ and $u = 0 \sim M - 1$.

The associated flow chart of the algorithm is illustrated by Figure 32.

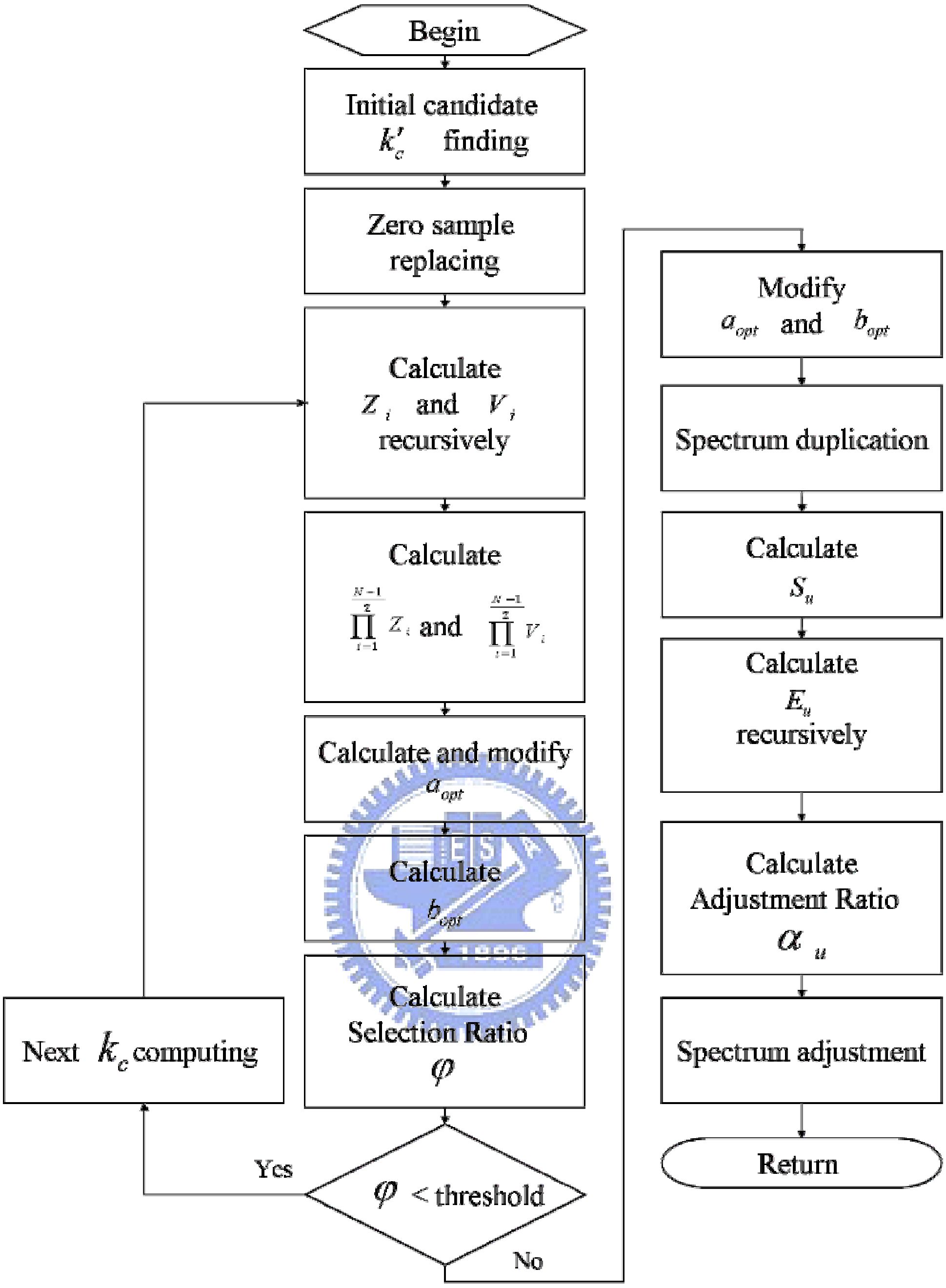

Figure 32: The flow chart of high frequency reconstruction.