Black-Box Modelling of Nonlinear Audio Circuits

Craig Rollo

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo, Finland 14.02.2018

Thesis supervisor:

Prof. Vesa Välimäki

Thesis advisor:

Fabián Esqueda, M.Sc.

Aalto University
School of Electrical
Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE
MASTER'S THESIS

Author: Craig Rollo

Title: Black-Box Modelling of Nonlinear Audio Circuits

Date: 14.02.2018 Language: English Number of pages: 6+43

Department of Signal Processing and Acoustics

Professorship: Audio Signal Processing (S-89)

Supervisor: Prof. Vesa Välimäki

Advisor: Fabián Esqueda, M.Sc.

This Thesis details various black-box modelling techniques for nonlinear audio circuits. The aim of this work was to compare the different approaches and determine the current most accurate model. The models analysed were: Chebyshev polynomials, Hammerstein, polynomials of best fit and a parametric model. These implementations were compared based on their spectral accuracy and time-domain accuracy.

From the analysis of these models, a new method is proposed. This proposed model restructures the determined most accurate black-box model by placing a nonlinearity between a high-pass and a low-pass filter. The filters emulate the linear response of the nonlinear circuit. These filters also have a reduced filter order in comparison to the linear response. The proposed method results in a model that performs close to the current best but does result in a massive reduction in computational time, up to 95% less time required and 99% fewer operations.

Keywords: Acoustic signal processing, Differential equations, Filters, Nonlinear distortion, Nonlinear circuits

Preface

I would like to thank my supervisor Prof. Vesa Välimäki and my advisor Mr. Fabián Esqueda for their guidance and patience(!) throughout this Thesis work. Our regular meetings were helpful, informative and fun.

I am also grateful for the support and encouragement from my parents and my sister during this Thesis and my time at University.

Otaniemi, Espoo, Finland 14.02.2018

Craig A. Rollo

Contents

Symbols and Abbreviations

Symbols

A	Gain
C	Capacitance in F
f_0	Fundamental frequency of sine wave in Hz
f_1	Start frequency of the sine-sweep in Hz
f_2	Stop frequency of the sine-sweep in Hz
f_3	Minor third frequency of sine wave in Hz
F_s	Sample rate in Hz
G	Control for distortion level in Tube Screamer circuit
I_s	Reverse saturation current of diode in A
L	Rate of frequency increase in the synchronised sine method
R	Resistance in Ω
T_R	Exponential rate of frequency increase
T_s	Sample period in seconds
V_i	Input voltage in V
V_o	Output voltage in V
V_t	Thermal voltage of diode in V
η	Diode ideality
ω_1	Start frequency of the sine-sweep in rad/s
ω_2	Stop frequency of the sine-sweep in rad/s

Abbreviations

ESR	Error-to-signal ratio
FE	Forward Euler
FIR	Finite impulse response
IIR	Infinite impulse response
LM	Levenberg-Marquardt
LTI	Linear Time-Invariant
LTV	Linear Time-Varying
NLB	Nonlinear block
TR	Trapezoidal Rule
TRR	Trust-Region-Reflective
WDF	Wave Digital Filter

1 Introduction

The desire to model nonlinear audio circuits arises from musicians' and audio engineers' preference of old analogue devices. These older devices, such as vacuum tube guitar amplifiers, have various issues including their price, size, weight, and difficulty finding replacement parts. This resulted in the development of virtual analogue modelling. The aim of this field of study is to accurately simulate the characteristics of an analogue device through the use of signal processing.

There are two main approaches to virtual analogue modelling of nonlinear audio systems: white-box and black-box models. White-box models use knowledge of the circuit and aim to model the components exactly. These include state space methods [1, 2, 3, 4, 5], wave digital filters [2, 4, 6, 7, 8] and physical models [9, 10]. These have been used to model amplifiers [11, 12, 13, 14], guitar distortion circuits [9, 15] and nonlinear aspects of musical instruments [1, 16]. Black-box models do not require any knowledge of the circuitry and are based on measurements. These include Volterra series [17, 18, 19], Hammerstein/Wiener models [20, 21, 22] and static nonlinearities [9, 23]. This approach has been used to model loudspeakers [20], guitar distortion pedals [22, 23, 24], and guitar amplifiers [25]. Measuring of the reference device is usually carried out using the swept-sine, or sine-sweep, technique [20, 24]. This is used to extract the impulse response from the system [26, 27, 28]. Another modelling approach is a combination of these two methods called grey-box models. These models are based off of measurements but use knowledge of the circuit in an attempt to improve the accuracy of the model [25, 29, 30].

A novel method is introduced based on the results of previous models. This proposed method changes the structure of a previous model. It consists of a high-pass filter, a nonlinearity and a low-pass filter. The filters simulate the linear response of a nonlinear audio system. The model is intended to reduce the computation time of the previous model.

The Thesis is structured as follows. Section 2 will introduce various black-box approaches for modelling nonlinear systems. Section 3 presents two distortion circuits and their physical model implementation. The circuits modelled were the diode clipper and the Tube Screamer. This section also details and compares the numerical methods used in the models. A comparison of the black-box models in relation to the physical models is given in Section 4. The proposed method is presented in Section 5 with comparisons to a previous black-box model and the physical models. Finally, Section 6 concludes the Thesis.

2 Black-box Models

The modelling of nonlinear musical systems has been an active research topic since the 1970's [31, 32, 33]. This section will introduce various black-box approaches to model such systems, as well as nonlinear audio circuits. The first two models are the Volterra series and the swept-sine technique. These serve as the basis for the Chebyshev polynomials and Hammerstein models discussed after. The final two methods are a static nonlinearity, using a polynomial of best fit, and a parametric model with a nonlinear mapping function. Comparisons between the models will be made in Section 4.

2.1 Volterra Series

The Volterra series was developed from the functionals defined by Vito Volterra in 1887. They are used to analyse nonlinear systems to create an explicit input–output relationship of ordinary nonlinear differential equations. The continuous time Volterra series is described by [18, 19]:

$$\begin{aligned} y(t) = h_0 + \int_{-\infty}^{\infty} h_1(\tau_1)x(t-\tau_1)d\tau_1 + \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)d\tau_1 d\tau_2 + \ldots + \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} h_p(\tau_1,\tau_2,\ldots,\tau_p)x(t-\tau_1)x(t-\tau_2)\ldots x(t-\tau_p)d\tau_1 d\tau_2\ldots d\tau_p + \ldots \end{aligned} \tag{1}$$

where $y(t)$ is the output, $x(t)$ is the input, τ is the time lag, and h_p is the Volterra kernel. This series is similar to the Taylor series; however, the Volterra series is able to model the memory of a nonlinear system [17]. The discrete time truncated Volterra series is given in [17, 18]:

$$\begin{aligned} y[n] = \sum_{m_1=0}^{N-1} h_1[m_1]x[n-m_1] + \\ \sum_{m_2=0}^{N-1}\sum_{m_1=0}^{N-1} h_2[m_1,m_2]x[n-m_1]x[n-m_2] + \ldots + \\ \sum_{m_p=0}^{N-1}\cdots\sum_{m_1=0}^{N-1} h_p[m_1,\ldots,m_p]x[n-m_1]\ldots x[n-m_p]. \end{aligned} \tag{2}$$

The h_0 kernel can be assumed to equal 0 as it does not affect the adaptive filter structure [17]. This model becomes increasingly complex for highly nonlinear systems as they require higher values of N and p and the number of coefficients are proportional to N^p [2, 17]. Due to this, there have been developments in making

the Volterra series more efficient which include using least-mean square, recursive least-squares and measuring with sine sweeps [18, 34].

The use of Volterra series in modelling audio-related systems include guitar amplifiers [34], nonlinear propagation in a pipe for brass instruments [35], strings [36], and the Moog ladder filter [19, 37].

2.2 Swept-Sine

The use of a sine-sweep for acoustical measurements was first proposed in [38] for a method known as time delay spectrometry. This used a linear swept-sine and the method was utilised for loudspeaker and room frequency responses [38, 39, 40, 41]. This technique was then modified to use a logarithmic sine-sweep to obtain room impulse responses in [26, 27], which was later developed for extracting impulse responses from nonlinear systems in [28]. The logarithmic sine-sweep is defined as:

$$x[n] = A \sin\left(\omega_1 T_R \left[\exp\left\{\frac{n}{T_R}\right\} - 1\right]\right), \tag{3}$$

where ω_1 is the start frequency of the sweep in rad/s, A is the gain of the sweep,

$$T_R = \frac{N}{\ln\left(\frac{\omega_2}{\omega_1}\right)}, \tag{4}$$

is the exponential rate of frequency increase of the sine-sweep, N is the desired length of the sweep in samples, and ω_2 is the stop frequency of the sweep in rad/s.

Novak et al. [20] resdesigned the generation of the sine-sweep so that the resulting higher harmonics are synchronised. This method is called the Synchronised Swept Sine [21, 42]. The equation for this method is:

$$x[n] = A \sin\left(\omega_1 L \left[\exp\left\{\frac{n}{L}\right\}\right]\right), \tag{5}$$

where

$$L = \left[\text{round}\left(\frac{f_1}{\ln \frac{f_2}{f_1}}\hat{T}\right)\right] \Big/ f_1 \tag{6}$$

is the rate of frequency increase, f_1 and f_2 are the start and stop frequencies of the synchronised sweep in Hz, and $\hat{T}$ is the desired time duration of the sweep in seconds. The synchronised time duration is given as:

$$T = L \ln\left(\frac{f_2}{f_1}\right). \tag{7}$$

This swept-sine is fed into the nonlinear system under test to produce a distorted output signal $y[n]$. This output is used in the convolution with the inverse filter of the input signal. The inverse filter is the time reversed and amplitude modulated version of the sweep. It is obtained analytically as [42, 43]:

$$X_{\text{inv}}(f) = 2\sqrt{\frac{f}{L}}\exp\left\{-j2\pi f L\left[1-\ln\left(\frac{f}{f_1}\right)\right]+j\frac{\pi}{4}\right\}. \tag{8}$$

The advantage of using the analytical form is that it does not require scaling the impulse responses [43]. The convolution is then performed as:

$$h[n] = \sum_{m=1}^{M} h_m[n+\Delta n_m] = \mathcal{F}^{-1}\left[H(f)\right] = \mathcal{F}^{-1}\left[Y(f)X_{\text{inv}}(f)\right], \tag{9}$$

where $H(f)$ and $h[n]$ are the impulse response of the nonlinear system in the frequency- and time-domain, respectively, $h_m[n]$ is the impulse response of the m^{th} harmonic up to M total harmonics, and Δn_m is the discrete time index of the m^{th} harmonic impulse produced via:

$$\Delta n_m = \frac{N}{2} - L\ln(m). \tag{10}$$

The results of this convolution for an arbitrary system are shown in Figure 1. The first impulse response, at Δn_1, is the linear response of the system. These impulse responses, or harmonics, can be windowed out and used as filters in various black-box modelling methods [21, 22, 24, 43].

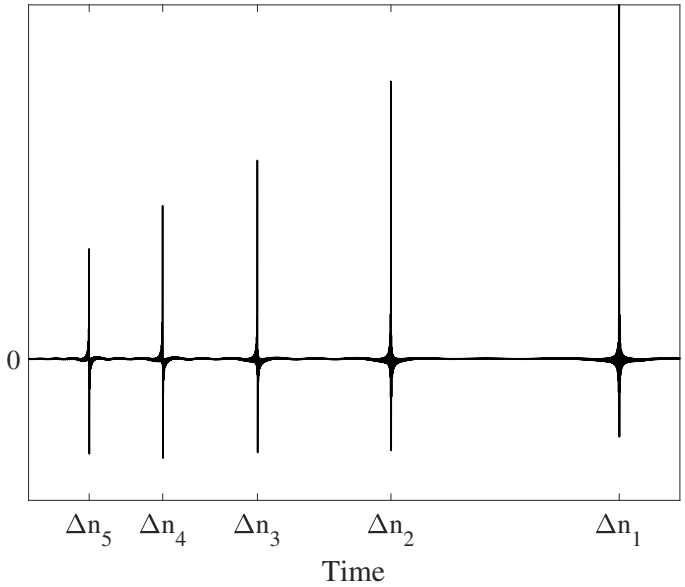


Figure 1: Nonlinear impulse response after convolution with the inverse filter.

2.3 Chebyshev Polynomials

The use of Chebyshev polynomials of the first kind to model nonlinear aspects of music was first documented by Schaefer [31]. This was based on the fact that Chebyshev polynomials produce the exact harmonics of a sinusoidal input. This was quickly improved by [32] to take into account the amplitude of the input. This used the coefficients defined in [31] and normalising in relation to powers of the input gain. This was further refined in [44] and [45].

These previous approaches focused more on waveshaping rather than modelling distortion. The use of these polynomials to model distortion was proposed by [21] using the impulse responses of the harmonics from the nonlinearity system to be modelled. The responses are extracted using the sine-sweep method and then they are convolved with the output from the polynomials. These polynomials are defined by:

$$T_0(x[n]) = 1,\ T_1(x[n]) = x[n], \tag{11}$$

and by the recurrence relation:

$$T_k(x[n]) = 2x[n]T_{k-1}(x[n]) - T_{k-2}(x[n]),\ k = 2, 3, ... \tag{12}$$

This model follows a parallel block structure and is shown in Figure 2. The input is transformed up to the M^{th} Chebyshev polynomial, which represents the highest harmonic to be modelled. The $h_m[n]$ filters are the harmonics from the reference system up to the M^{th} harmonic. The use of this model is limited, however, as the Chebyshev polynomials work for sinusoidal inputs only.

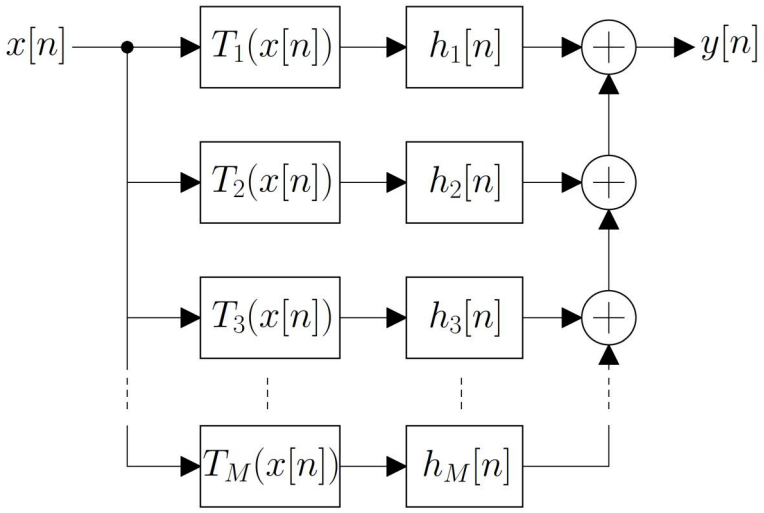


Figure 2: Block diagram for the Chebyshev polynomial model.

2.4 Hammerstein/Wiener

The Wiener model was introduced as an alternate representation of the Volterra model [18]. The Wiener model consists of a linear filter followed by a nonlinearity and generally has a parallel block structure. The Hammerstein model is similar in design but the nonlinearity precedes the filter. A general block diagram of a Hammerstein model is given in Figure 3.

This model is the same as that of the model in Figure 2; however, the Chebyshev polynomials have been replaced by the power series. The model consists of M parallel branches where each branch has a nonlinearity where the input is transformed to the power of the m^{th} branch. This results in a static monomial nonlinearity [43]. For every even power, this results in an output which contains all the even harmonics of the input up to the power of the branch. Similarly, odd powers will create all the odd harmonics to the current power. This means that the harmonic impulse responses from the swept-sine method cannot be used as is, as they do not influence the same harmonics generated via the nonlinearity. Therefore, the linear filters following the nonlinearity are the Hammerstein kernels which were created to deal with impulse response issue.

The Hammerstein kernels (in this implementation) were generated using the Hammerstein Kernels Identifications by Sine Sweep (HKISS) method detailed in [43]. This method uses the sine-sweep to obtain the harmonics of the system. After they have been obtained, they are transformed by a matrix that is developed from trigonometric power formulas of sine.

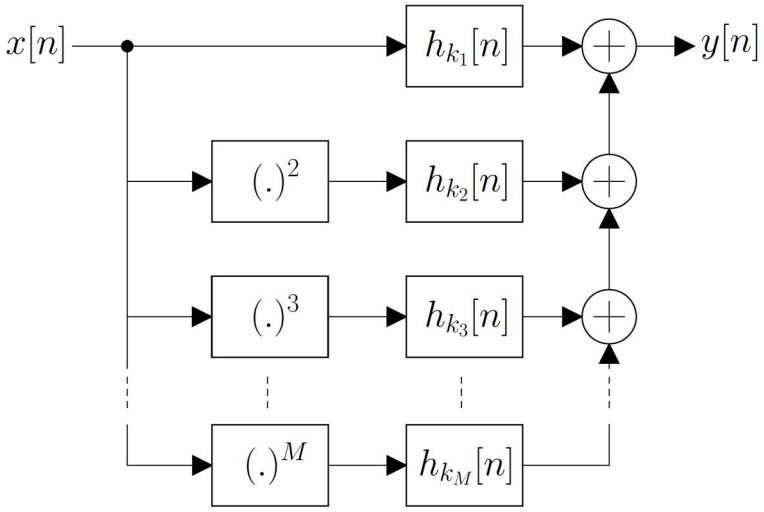


Figure 3: Block diagram for the Hammerstein model.

2.5 Static Nonlinearity

This method aims to model the nonlinearity by approximating the relationship between the input and output of the nonlinear system. This approximation results in a memoryless nonlinearity. This method was proposed in [9] but the idea has existed in other areas of research, for example, in neural response models [46]. This approximation can be acquired in numerous ways, via measurements or by deriving a nonlinear ordinary differential equation (ODE) of the circuit by replacing the capacitors with open circuits. The result can be stored and used as a lookup table or by fitting a polynomial to the input–output relationship.

Figure 4 compares the diode clipper and a polynomial of best fit, while Figure 5 compares the Tube Screamer to another polynomial of best fit. The grey lines in each figure are produced from white-box analysis which is detailed in Section 3. The order for both of these polynomials was 11, as this order is low enough to not produce an ill-conditioned polynomial but results in sufficient harmonics to perceive an accurate sound.

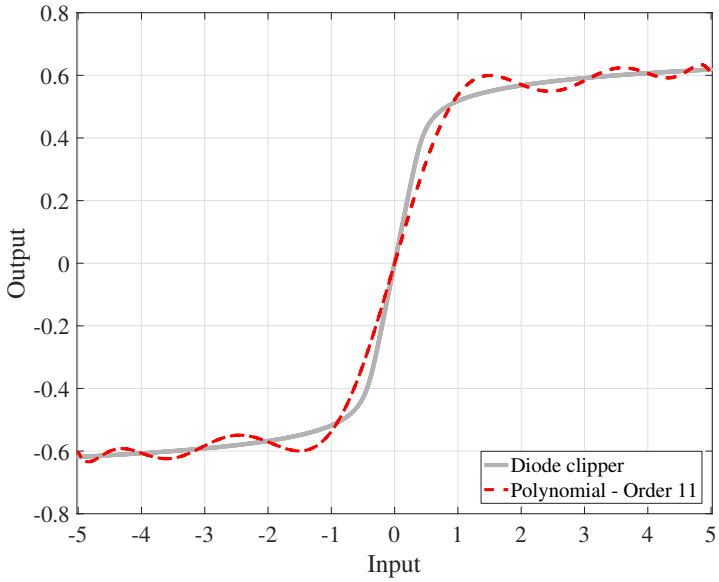


Figure 4: Input–output relationship comparison of the diode clipper and polynomial of best fit.

The method also consists of a filter which precedes the nonlinearity. This filter estimates the linear response of the circuit. For the diode clipper, the filter implemented was a first-order low-pass with a cut-off frequency of a sixth of the sampling rate, F_s. The magnitude response of the low-pass filter in comparison to the linear response is depicted in the top plot of Figure 6. The linear response of the Tube Screamer was estimated by cascading a DC blocker and a low-pass FIR filter with cut-off frequencies of $\frac{F_s}{70}$ and $\frac{F_s}{3.5}$, respectively. The result is a fifth-order band-pass

filter. The response of this filter in comparison to the Tube Screamer linear response is displayed in bottom plot of Figure 6.

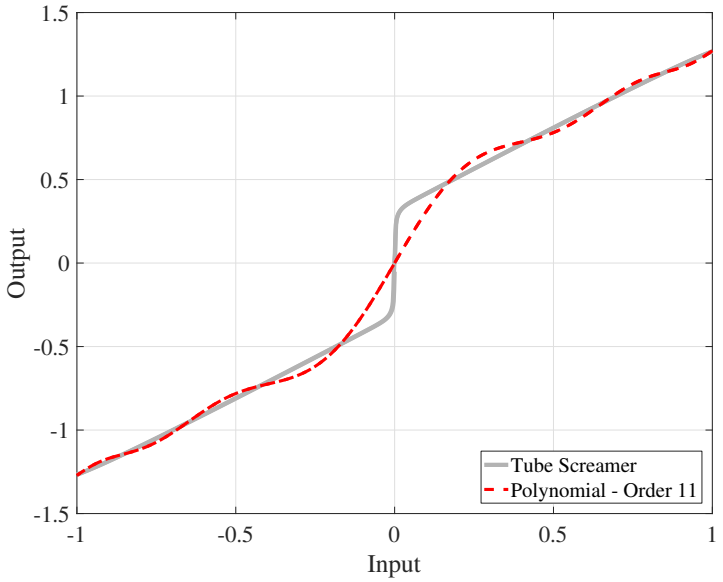


Figure 5: Input–output relationship comparison of the Tube Screamer and polynomial of best fit.

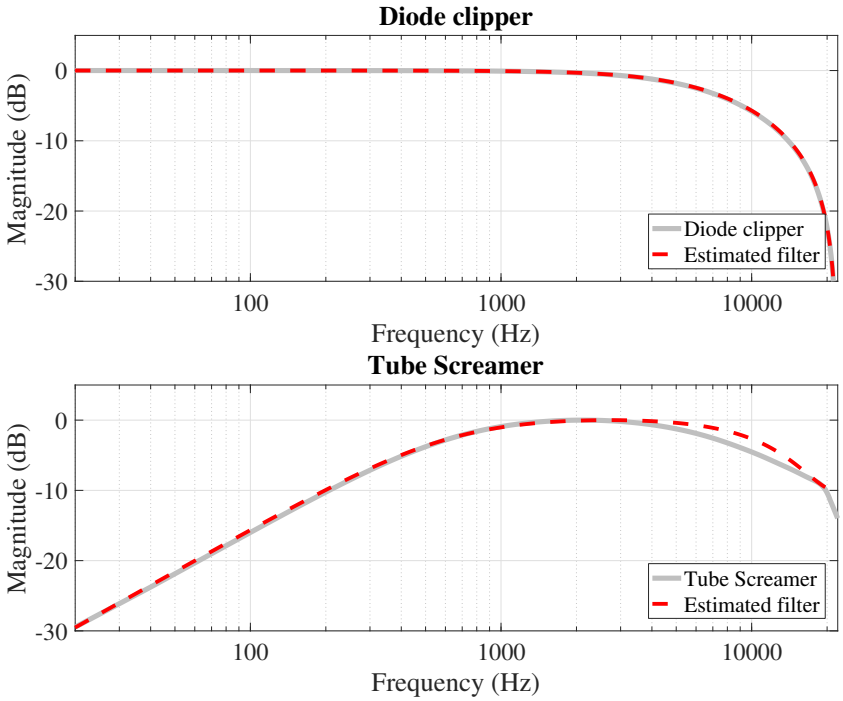


Figure 6: Comparison of linear response of the circuits to estimated filters. Top: diode clipper. Bottom: Tube Screamer.

2.6 Eichas Parametric Model

This model was first proposed in [23] as a parametric Wiener-Hammerstein model. It consists of a linear time-invariant (LTI) block, a nonlinear block (NLB), and followed by another LTI block. The LTI block is a band-limited equaliser whose coefficients can be parameterised. The model was developed further in [24] which changed the LTI block to an FIR filter where the coefficients of the filter are from the measured linear response of the nonlinear system to be modelled, obtained via the swept-sine method. This development meant that the second LTI block was no longer required. Figure 7 displays the outline of the block diagram for this model as a simplified Wiener model.

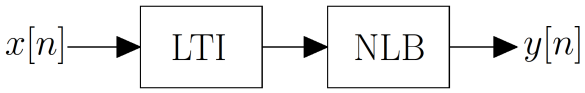


Figure 7: Block outline of the full Eichas' parametric model.

The nonlinear block consists of a side-chain envelope detector, a dry/mix stage and a mapping function. The envelope detector models the time-variant behaviour that nonlinear systems with memory exhibit. It also compensates for the bias voltage that is introduced by operational amplifiers in some nonlinear circuits. The envelope detector computes the absolute value of the input signal, low-pass filters this and then subtracts the result from the main signal path. The low-pass filter is a biquad filter with a cut-off frequency of 5 Hz. The various gains for each stage are stored in a parameter vector. The block diagram for this model is shown in Figure 8.

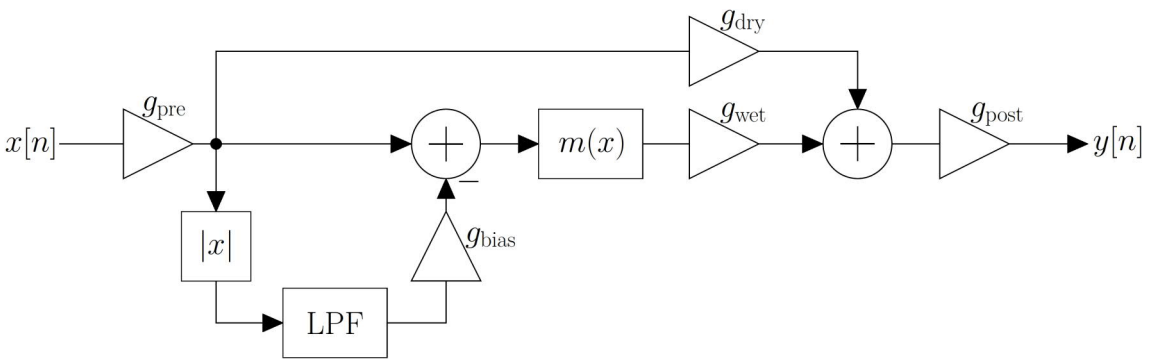


Figure 8: Block diagram for the nonlinear block of Eichas' parametric model. Adapted from [23].

The path that includes the $|x|$ and the LPF is the envelope detector and $m(x)$ is the mapping function. This mapping function is composed of three hyperbolic tangent functions which control the sharpness of the saturation in the input–output relationship. They are determined by four parameters: k_p and k_n, which are the positive and negative knee points, respectively; and g_p and g_n, which affect the smoothness of the knee point transitions. The mapping function is defined in Eq. 13.

$$m(x) = \begin{cases} \tanh(k_p) - \left[\dfrac{\tanh(k_p)^2 - 1}{g_p} \tanh(g_p(x - k_p))\right] & \text{if } x > k_p \\ \tanh(x) & \text{if } -k_n \leq x \leq k_p \\ -\tanh(k_n) - \left[\dfrac{\tanh(k_n)^2 - 1}{g_n} \tanh(g_n(x + k_n))\right] & \text{if } x < -k_n \end{cases} \tag{13}$$

The final parameter vector is thus

$$\mathbf{p} = (g_{\text{pre}},\ g_{\text{bias}},\ k_p,\ k_n,\ g_p,\ g_n,\ g_{\text{wet}},\ g_{\text{post}})\,. \tag{14}$$

The parameters were limited to $[-1, 1]$, where g_{pre}, g_p, g_n, and g_{post} are scaled and converted to dB, in a given range. In this implementation the range was [−80 dB, 80 dB].

The parameters were optimised using the Levenberg-Marquardt (LM) algorithm or the Trust-Region-Reflective (TRR) algorithm. The LM method, introduced in [47] and improved upon in [48], is a gradient-based optimisation procedure which minimises the residual between the output of the nonlinear system under test and the nonlinear block. The TRR technique uses a model function to minimise the function to be solved [49, 50]. This model function simulates the function in a subset at a given point. This subset is the trust-region. TRR was used when the LM algorithm produced parameters that exceeded the limits as the LM does not allow for set limits.

The input–output relationship of the model after optimisation for the diode clipper is shown in Figure 9 and for the Tube Screamer in Figure 10. The input for optimisation was a single-tone sine wave at 1000 Hz. The amplitude of this sine wave logarithmically decreased from 10 to 10^{-5}, to cover a wide range of potential inputs.

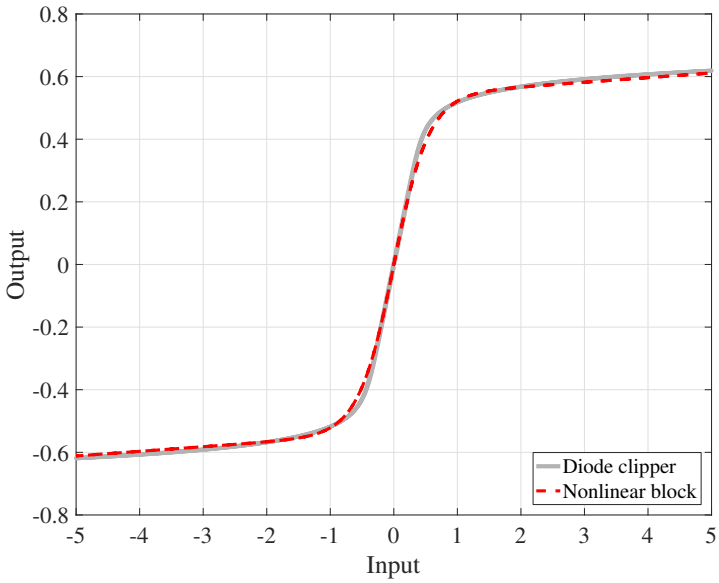


Figure 9: Input–output relationship comparison between the diode clipper and the nonlinear block.

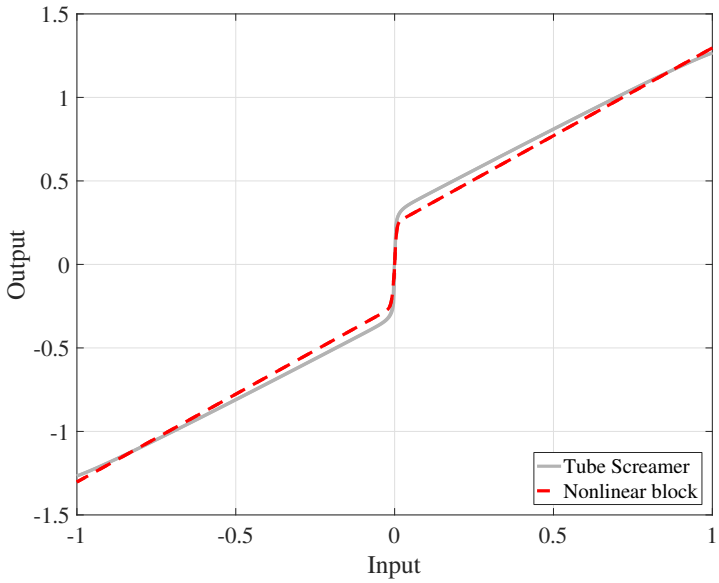


Figure 10: Input–output relationship comparison between the Tube Screamer and the nonlinear block.

3 Physical Models

This section will discuss the two circuits that will be used as references for which the different black-box models detailed in the previous section are evaluated against. The circuits were modelled as physical models, using nonlinear ordinary differential equations, which is a white-box modelling technique. The numerical methods used to model these circuits will be discussed, as well as their accuracy and stability. Only the distortion of these devices were modelled, while the filtering stages of the devices were ignored.

3.1 Circuits

3.1.1 Diode Clipper

The basis of most guitar distortion pedals is the diode clipper (or diode limiter) [10]. The diode clipper consists of anti-parallel diode pair coupled with an RC low-pass filter. The simplified circuit schematic of the diode clipper is shown in Figure 11. The input and output voltages are defined by V_i and V_o in the schematic, respectively.

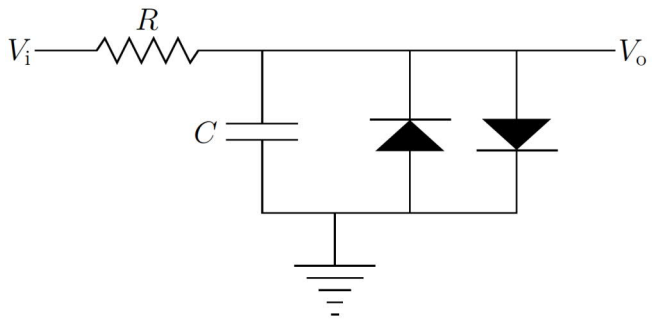


Figure 11: Simplified circuit diagram of the diode clipper.

The diode pair limits the maximum voltage to a set value, which is determined by the diodes used, across the capacitor in either direction. As this circuit has a capacitor, the nonlinearity has memory and therefore the input–output relationship has to be the approximated by replacing this capacitor with an open circuit. This relationship is depicted in Figure 12 with a comparison to the hyperbolic tangent, a commonly used saturating nonlinear function to model distortion [9].

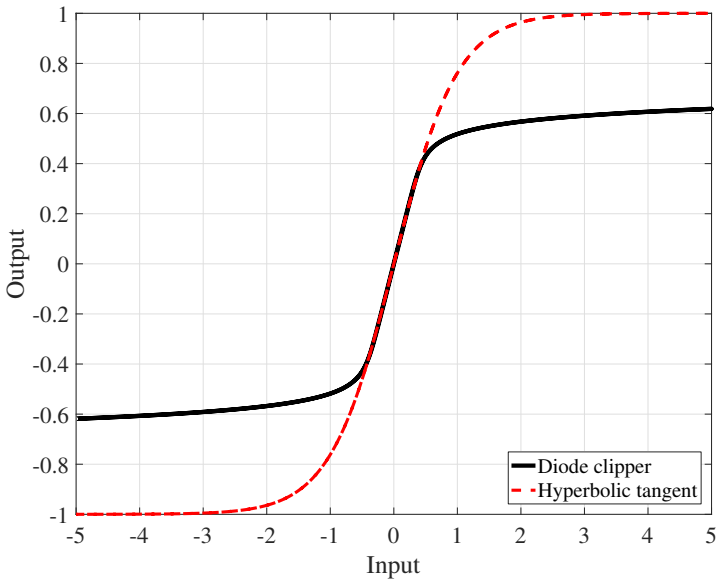


Figure 12: Approximation of the input–output relationship of the diode clipper.

The range of the input represents the peak range of this physical model. The diode clipper is commonly preceded by an non-inverting operational amplifier (not shown in the schematic in Figure 11). This functions as a linear gain, which in this instance is a gain of 5.

This circuit can be simulated via the use of an ODE [9]. This ODE can be derived by using Kirchoff's laws and the Shockley diode equation:

$$\frac{dV_\mathrm{o}}{dt} = \frac{V_\mathrm{i} - V_\mathrm{o}}{RC} - 2\frac{I_\mathrm{s}}{C}\sinh\left(\frac{V_\mathrm{o}}{\eta V_\mathrm{t}}\right), \tag{15}$$

where $R = 2.2\,\mathrm{k\Omega}$ and $C = 10\,\mathrm{nF}$ are the component values of the resistor and capacitor, and $I_\mathrm{s} = 2.52\,\mathrm{nA}$, $V_\mathrm{t} = 26\,\mathrm{mV}$ and $\eta = 1.752$ are the saturating current, the thermal voltage and ideality factor related to the diodes, respectively.

3.1.2 Tube Screamer

The Tube Screamer is a popular guitar pedal developed by Ibanez to simulate the sound of vintage tube amplifiers [51]. It has also been modelled numerous times due to its popularity, using physical models [10], wave digital filters [15, 52] or with a black-box mapping function [24]. The circuit diagram of the clipping section of the Tube Screamer is illustrated in Figure 13.

In this circuit, the diodes limit the voltage across the operational amplifier to keep it within ideal op-amp conditions. This results in equal voltage at the positive and negative terminals of the op-amp. For this implementation, the V_bias, which is

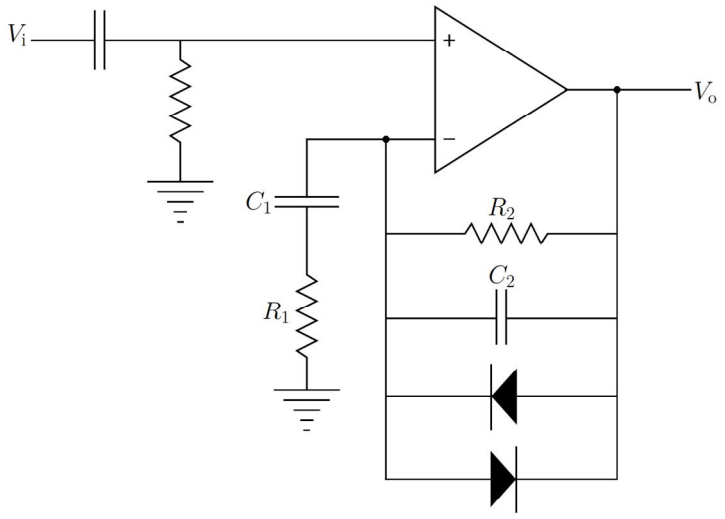


Figure 13: Circuit diagram of the clipping section of the Tube Screamer.

usually incorporated as half of the supply voltage to the op-amp, was ignored as the op-amp was assumed to have a bipolar power supply.

This circuit can also be simulated using an ODE derived by Kirchoff's laws and the Shockley diode equation to produce the following equations:

$$\frac{dV}{dt} = \frac{I}{C_2} - \frac{V}{R_2 C_2} - \frac{2I_\mathrm{s}}{C_2} \sinh\left(\frac{V}{\eta V_\mathrm{t}}\right), \tag{16}$$

where $R_2 = 51\,\mathrm{k\Omega} + G500\,\mathrm{k\Omega}$ with G controlling the distortion level and $V = V_\mathrm{i} - V_\mathrm{o}$. The current, I, is calculated in [10] in the s-domain as:

$$I = V_\mathrm{i} \frac{s}{R_1\left(s + \dfrac{1}{R_1 C_1}\right)}, \tag{17}$$

which can be converted to the z-domain via the bilinear transform:

$$s = \frac{2}{T_\mathrm{s}} \frac{1 - z^{-1}}{1 + z^{-1}}, \tag{18}$$

where T_s is the sampling period, defined as the inverse of the sampling rate, to produce:

$$I = V_\mathrm{i} \frac{b_0 + b_1 z^{-1}}{a_0 + a_1 z^{-1}}, \tag{19}$$

with

$$b_0 = 2C_1,\ b_1 = -2C_1 \tag{20}$$

$$a_0 = T_\mathrm{s} + 2R_1 C_1,\ a_1 = T_\mathrm{s} - 2R_1 C_1. \tag{21}$$

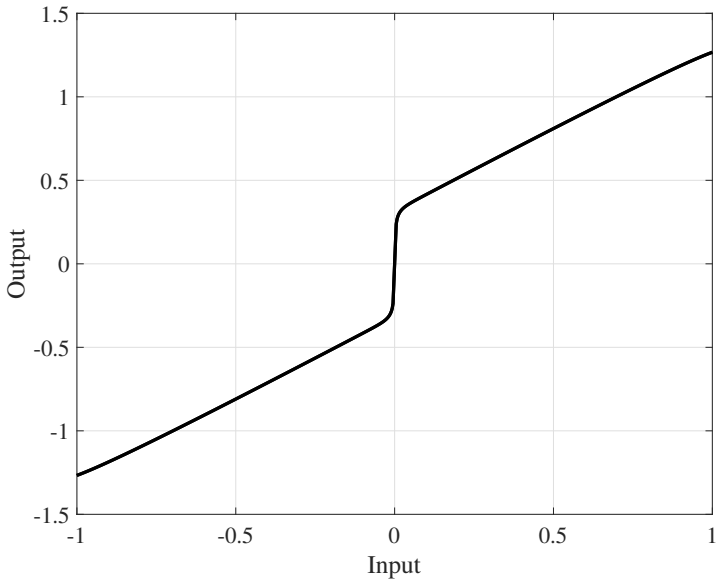


Figure 14: Approximation of the input–output relationship of the Tube Screamer.

Figure 14 shows an approximation of the input–output relationship. In order to achieve this relationship the system was fed the first quarter of a sine wave, containing only positive amplitudes and then storing the result. This was repeated for the third quarter of a sine wave, containing only negative amplitudes. The final result was obtained by ordering these results to recreate a full distorted sinusoidal output. This was done due to the differential and the resulting change in direction of the waveform. The input range is not boosted as the Tube Screamer is not preceded with a linear gain stage.

3.2 Numerical Methods

Numerical methods use numerical integration to solve equations of the form

$$\frac{dV_\mathrm{o}}{dt} = \dot{V}_\mathrm{o} = f(t, V_\mathrm{o}, V_\mathrm{i}), \tag{22}$$

where V_o is the system state and $f(t, V_\mathrm{o}, V_\mathrm{i})$ is a nonlinear function that calculates the time derivative of the system state depending on the input and the current system state.

There are two different methodologies to solving these methods: explicit and implicit. Explicit methods calculate the output solely on the state at previous time steps. Implicit methods depend on the current state to calculate the output. This requires an additional iterative solver to compute the ODE if it is nonlinear. The Newton-Raphson method is the most common solver for this method [9].

3.2.1 Forward Euler

Forward Euler is a first-order explicit method and is defined as:

$$V_\mathrm{o}[n] = V_\mathrm{o}[n-1] + T_\mathrm{s}\dot{V}_\mathrm{o}[n-1], \tag{23}$$

where $v[n]$ is the system state at discrete time index n, $\dot{V}_\mathrm{o}[n-1]$ is the derivative of the system state at the previous time index, and T_s is the sampling period, as before.

The equation for the diode clipper (Eq. 15) can be substituted into the above equation to produce the explicit recursion:

$$V_\mathrm{o}[n] = V_\mathrm{o}[n-1] + \frac{T_\mathrm{s}}{RC}\left(V_\mathrm{i}[n-1] - V_\mathrm{o}[n-1]\right) - 2\frac{T_\mathrm{s}I_\mathrm{s}}{C}\sinh\left(\frac{V_\mathrm{o}[n-1]}{\eta V_\mathrm{t}}\right). \tag{24}$$

3.2.2 Trapezoidal Rule

The trapezoidal rule is a second-order implicit method as it uses the average of the system state derivatives at the current time index n and the previous $n-1$. The trapezoidal rule is given as:

$$V_\mathrm{o}[n] = V_\mathrm{o}[n-1] + \frac{T_\mathrm{s}}{2}(\dot{V}_\mathrm{o}[n] + \dot{V}_\mathrm{o}[n-1]). \tag{25}$$

Substituting Eq. 15 into the above results in:

$$\begin{aligned} V_\mathrm{o}[n] = V_\mathrm{o}[n-1] &+ \frac{T_\mathrm{s}}{2RC}(V_\mathrm{i}[n] + V_\mathrm{i}[n-1] - V_\mathrm{o}[n] - V_\mathrm{o}[n-1]) \\ &- \frac{I_\mathrm{s}T_\mathrm{s}}{C}\left(\sinh\left(\frac{V_\mathrm{o}[n]}{\eta V_\mathrm{t}}\right) + \sinh\left(\frac{V_\mathrm{o}[n-1]}{\eta V_\mathrm{t}}\right)\right). \end{aligned} \tag{26}$$

3.2.3 Comparison of FE and TR

The two numerical methods were implemented and compared based on the need for oversampling, the time-domain error and their respective magnitude spectra. The minimum oversampling requirement to produce an output and the time-domain result for the methods is shown in Figure 15. The input used for this comparison was a two-tone sine wave, at frequencies 147 Hz and 220 Hz at a base sample rate of $F_\mathrm{s} = 44100$ Hz. The time-domain error was computed by using a reference output created with 32 times oversampling, with the same input signal. The time-domain error is depicted in Figure 16.

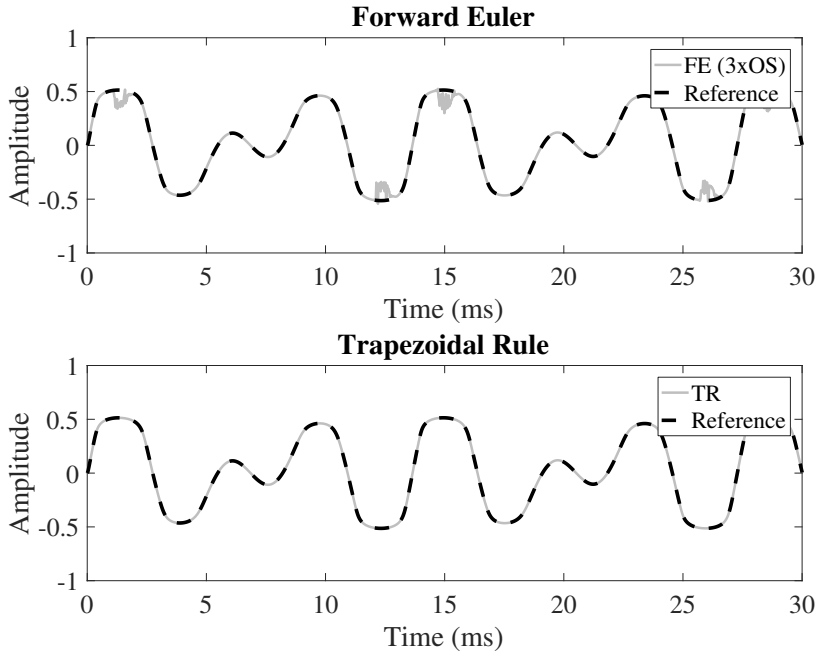


Figure 15: Time-domain comparison for the diode clipper with a two-tone 147 Hz and 220 Hz input. Top: Forward Euler. Bottom: Trapezoidal Rule.

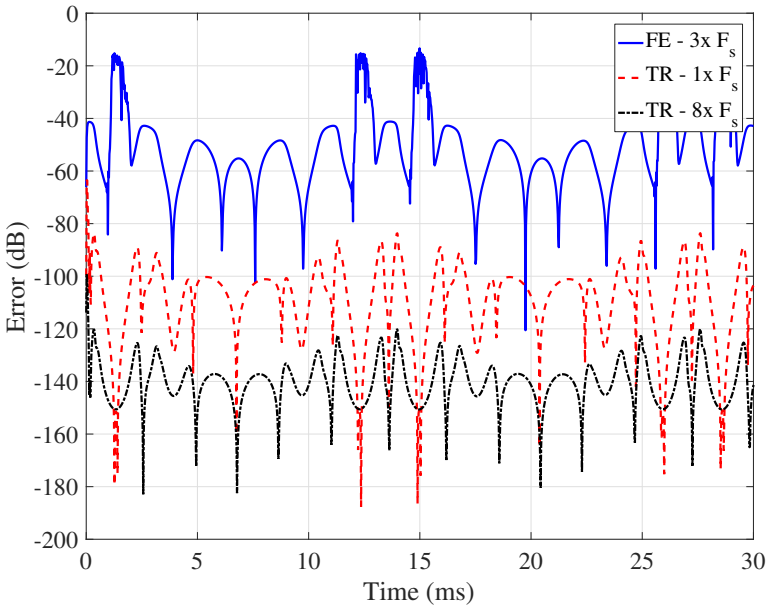


Figure 16: Time-domain error for the diode clipper with a two-tone 147 Hz and 220 Hz input.

Due to stability issues, the minimum oversampling factor required to create Figure 15 and Figure 16 using the FE method was 3 times oversampling, whereas none was required when using TR. In Figure 15 there are clear artefacts in the FE waveform, while the other methods are indistinguishable from one another. However, in Figure 16 the TR with 8 times oversampling does produce a more accurate result in comparison to the TR with no oversampling.

The stability of the methods was computed for a high frequency as well. The input for this test was a single-tone sine wave with a frequency of 10067 Hz and an input gain of 5. The comparisons in the time-domain are given in Figure 17. The magnitude spectra results were compared and normalised to a 64 times oversampled reference. The methods were tested using the same base sample rate of 44100 Hz and the results are given in Figure 18.

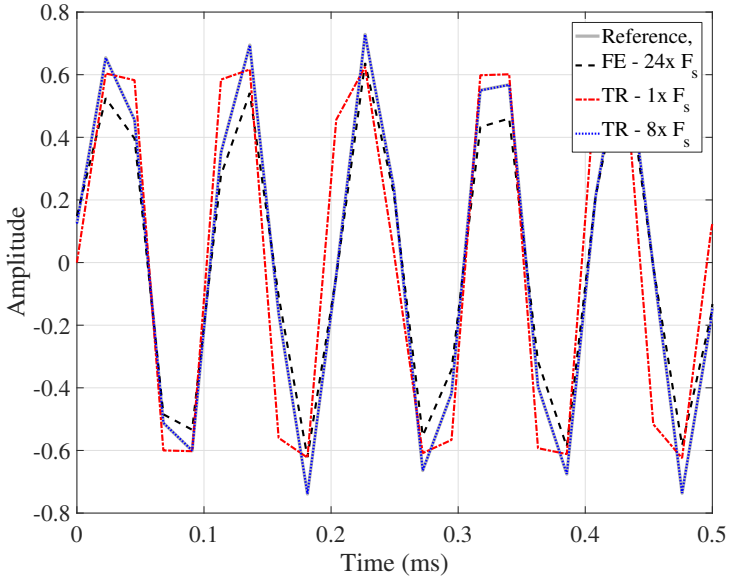
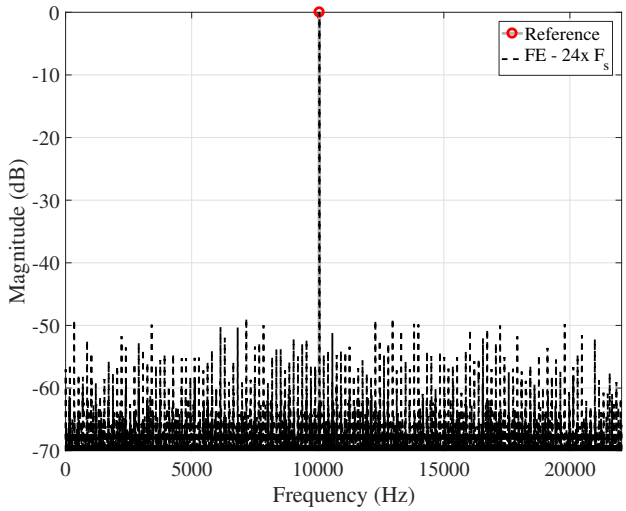


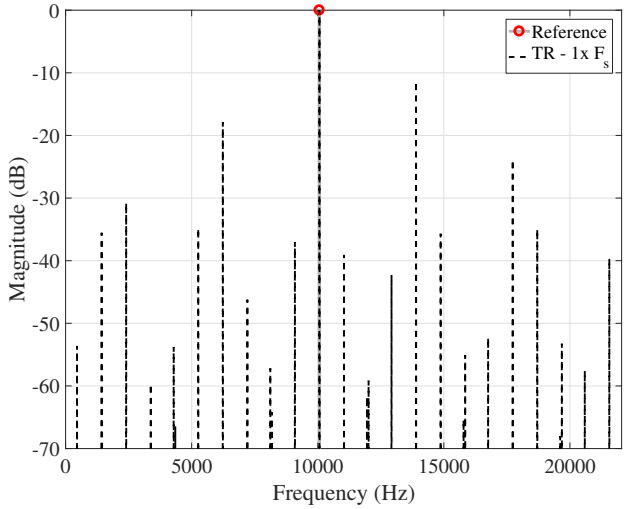
Figure 17: Time-domain waveform comparison for the diode clipper with a 10067 Hz input.

The TR method with no oversampling does not perform as well at higher frequencies at higher gain, evidenced from Figure 17. The FE method requires a minimum oversampling factor of 24 times to produce a stable result. Comparing the spectra for the methods in Figure 18 shows that the FE and the TR with no oversampling suffer from aliasing. The only frequency that should appear in the spectrum modelled is marked with a red circle in the figures. From this, the TR with 8 times oversampling was used in all further tests for the diode clipper.

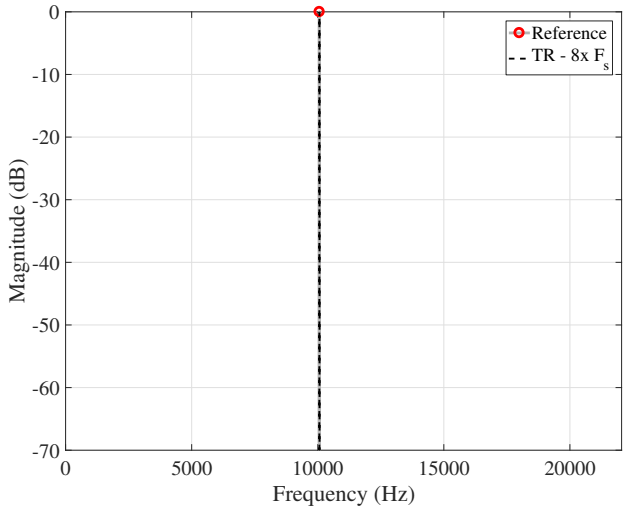
These tests were also performed for the Tube Screamer. It was decided that the FE method could not produce an accurate result for the circuit. The same oversampling factor for the TR method in the diode clipper physical model was used in all further comparisons.



(a) Forward Euler with 24 times oversampling



(b) Trapezoidal Rule with 1 times oversampling



(c) Trapezoidal Rule with 8 times oversampling

Figure 18: Magnitude spectra for the diode clipper with a 10067 Hz input.

4 Comparison of Black-box Models

The black-box models detailed in Section 2 are compared against the physical models described in Section 3. They are evaluated based on the accuracy of their harmonic spectra on single and multi-tone inputs with different gains. They are also tested using two objective evaluation methods to determine their time-domain accuracy.

4.1 Spectra Comparisons

Each method was compared using the following test cases. A single-tone sine wave with a specific input gain for each model. The Chebyshev and Hammerstein models had a peak gain for the input of 1, as the models are expected to work with this gain. For the static nonlinearity the gain of the input was the same as that used for the input to the system to which the polynomial was fitted to. This was the peak amplitude for each system, 5 for the diode clipper and 1 for the Tube Screamer. The Eichas' parametric model used an input gain of 1, resulting in a peak amplitude of 5 for the diode clipper model. The fundamental frequency, f_0, of the sine wave used in the tests was 220 Hz, which is the note A3.

The second test case was a single-tone sine wave but with a different gain. In all cases the input gain was reduced to 0.5 for the diode clipper and 0.1 for the Tube Screamer. The fundamental frequency of the sine wave was the same as that of the first test case.

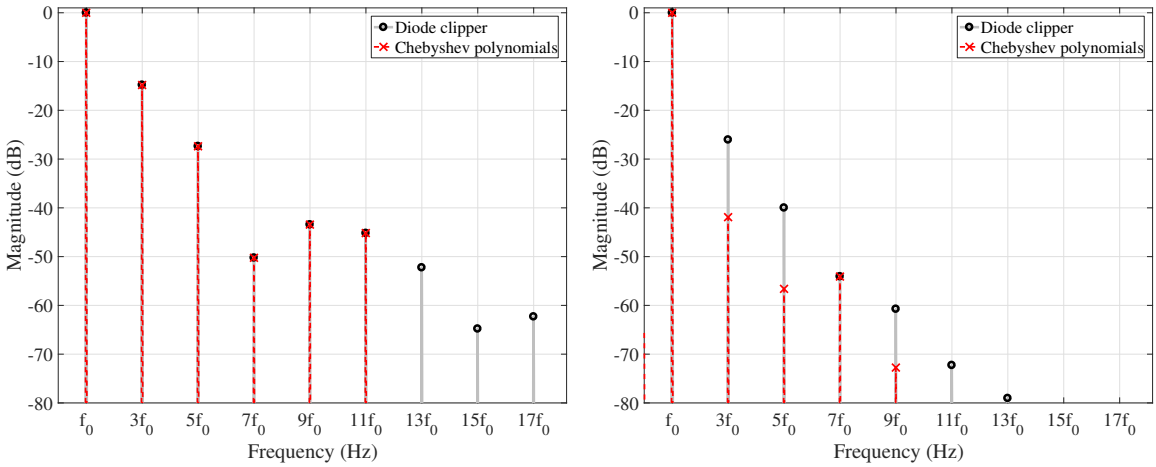
The third test case was a multi-tone sine wave with the same peak gains used in the first test case. The frequencies of the sine wave, f_0 and f_3, were 220 Hz and 261 Hz, which are notes A3 and C4, creating a minor third. The Eichas' parametric model was also tested with a multi-tone sine wave at the same different gain used in the second test.

For all cases the fundamental frequency of the model and the reference were normalised to 0 dB. The total number of harmonics, i.e. impulse responses and kernels, measured was 11 for the Chebyshev and Hammerstein models.

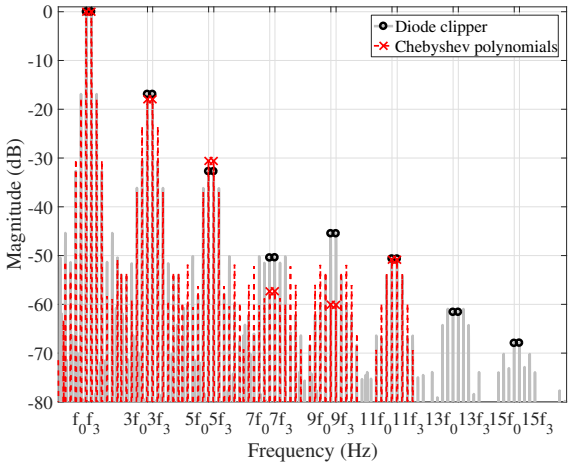
The different test cases were all created with a sample rate of $F_\mathrm{s} = 44100$ Hz. The output of the reference systems were downsampled to this sample rate before their spectra was calculated.

4.1.1 Chebyshev Polynomials

The results for the diode clipper and the Tube Screamer are displayed in Figure 19 and Figure 20, respectively.



(a) Single-tone sine wave (same gain) **(b)** Single-tone sine wave (different gain)

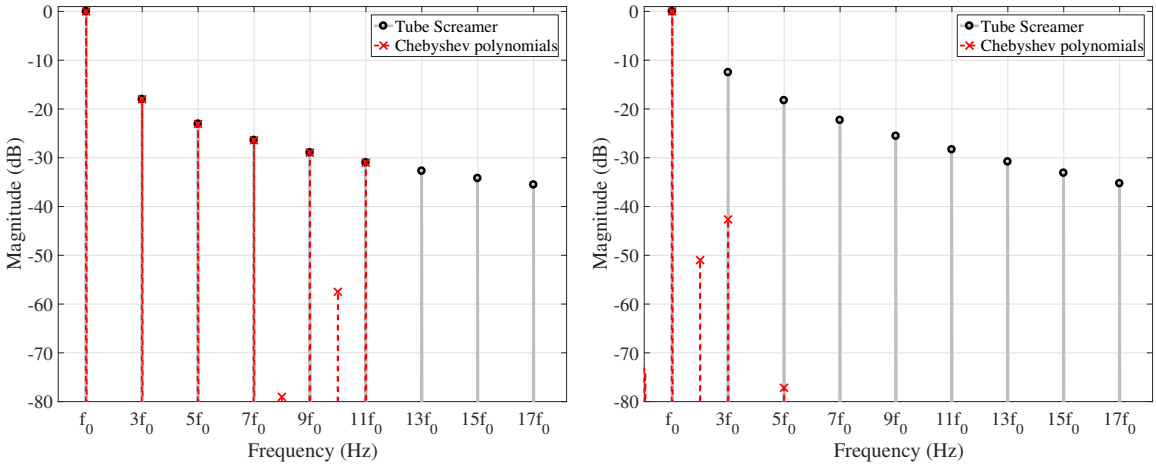


(c) Multi-tone sine wave (same gain)

Figure 19: Spectra comparisons between the diode clipper and the Chebyshev polynomial model.

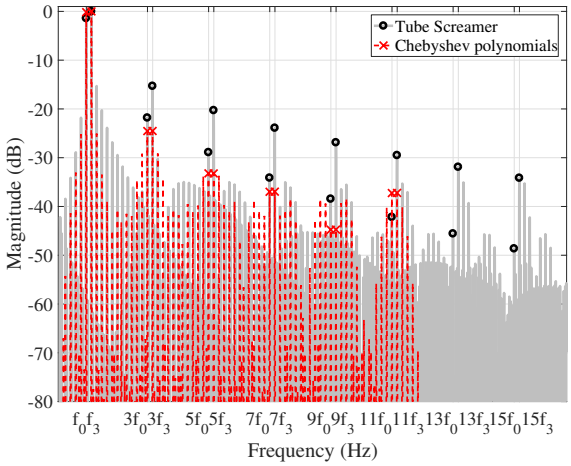
The Chebyshev polynomials exactly match when a single-tone sine wave with the same gain is used for the diode clipper as depicted in Figure 19a. They also model a multi-tone sine quite accurately, although in Figure 19c the 7^{th} and 9^{th} harmonics are several decibels lower than the reference. The model fails once a different gain is used as observed in Figure 19b.

The multi-tone plot only highlights the multiples of the fundamental, the rest of the energy plotted is intermodulation components. This is generated via the distortion and is not aliasing. This is kept consistent for all multi-tone plots presented throughout.



(a) Single-tone sine wave (same gain)

(b) Single-tone sine wave (different gain)

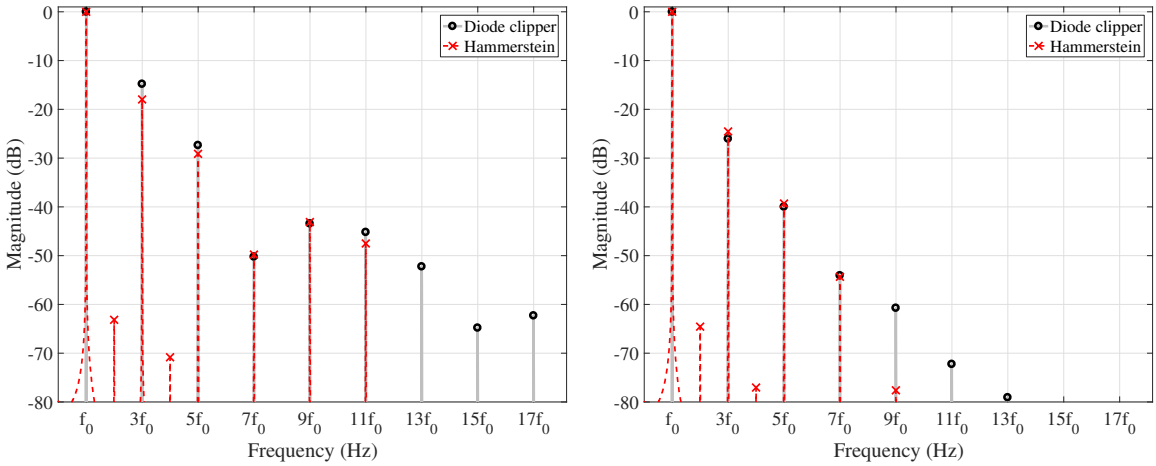


(c) Multi-tone sine wave (same gain)

Figure 20: Spectra comparisons between the Tube Screamer and the Chebyshev polynomial model.

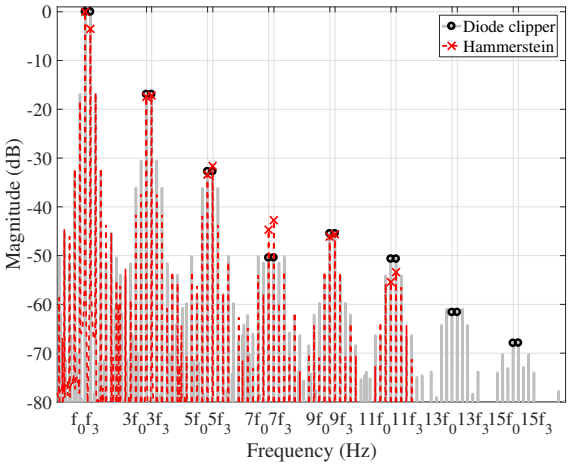
When modelling the Tube Screamer, the Chebyshev polynomials are not as accurate as when modelling the diode clipper. This could be due to the more complex nonlinearity leading to a loss of information when windowing out the impulse responses. The single-tone in Figure 20a is accurate with the only noticeable issue occurring between the 9^{th} and 11^{th} harmonic. Again, the model fails once a different input gain is used in Figure 20b. The model also does not perform well with multi-tone inputs, the magnitude of the harmonics are equal for the two frequencies, which is not the case in Tube Screamer and are also several dB too low. This is most prevalent at the 9^{th} harmonic in Figure 20c.

4.1.2 Hammerstein



(a) Single-tone sine wave (same gain)

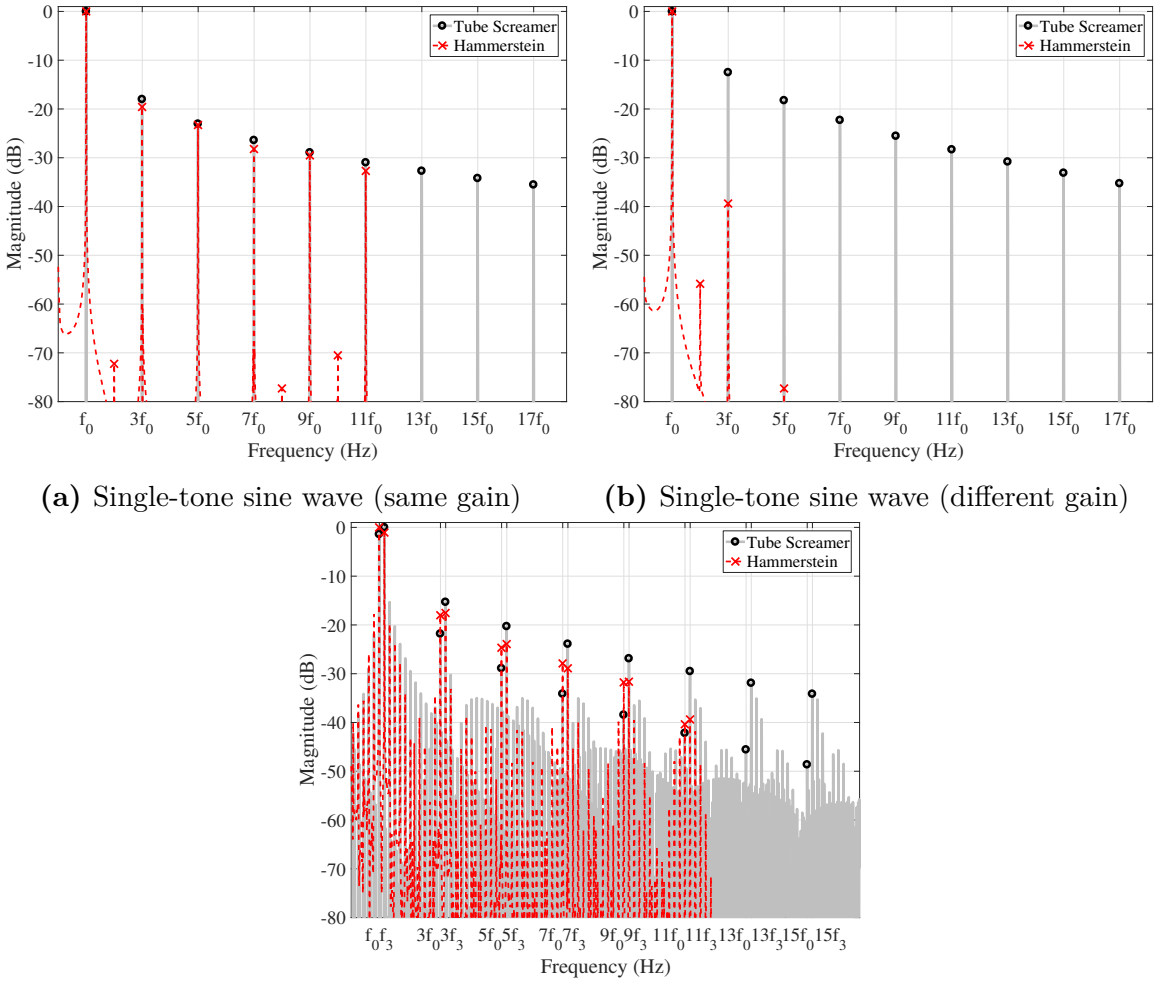
(b) Single-tone sine wave (different gain)



(c) Multi-tone sine wave (same gain)

Figure 21: Spectra comparisons between the diode clipper and the Hammerstein model.

The single-tone same gain input result for the Hammerstein model in Figure 21a is nearly identical to that of the Chebyshev model, with only a maximum of 2 dB difference in a couple harmonics. Surprisingly, the input with a different gain in Figure 21b produces an accurate result but decays faster after the 7^{th} harmonic than the reference. The multi-tone input has a similar accuracy with the Hammerstein in comparison to the Chebyshev model result. The 9^{th} harmonic in Figure 21c is now accurate but the 7^{th} harmonic still suffers.



(a) Single-tone sine wave (same gain)

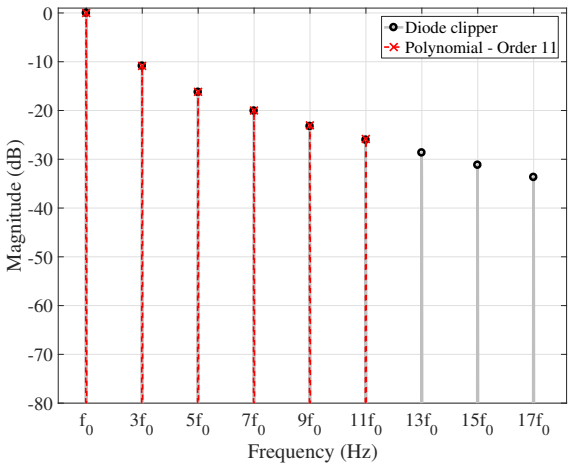
(b) Single-tone sine wave (different gain)

(c) Multi-tone sine wave (same gain)

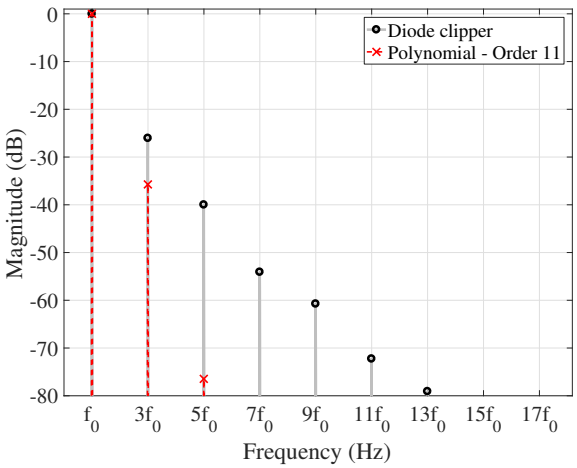
Figure 22: Spectra comparisons between the Tube Screamer and the Hammerstein model.

The Hammerstein model for the Tube Screamer performs as well as the Chebyshev model for the single-tone same gain input in Figure 22a. The multi-tone sine wave in Figure 22c performs slightly better than Figure 20c; however, it does not model the differing magnitudes of the harmonics of the fundamental frequencies. The different gain input fails similarly to the Chebyshev polynomial model, which is expected, in contrast to the Hammerstein diode clipper result. Schmitz and Emcrechts [43] suggested a potential fix for this by using an amplitude factor function such that the kernels can be weighted depending on the input gain.

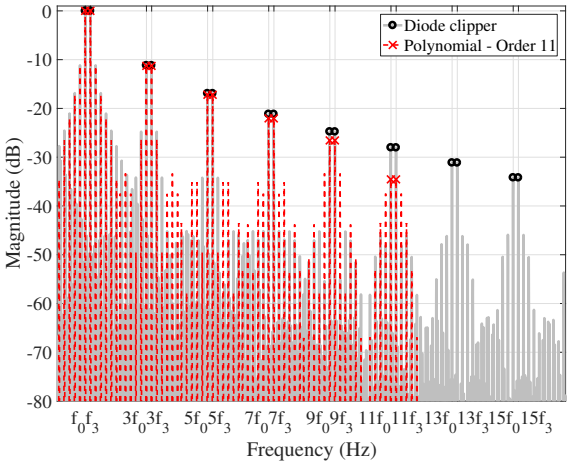
4.1.3 Static Nonlinearity



(a) Single-tone sine wave (same gain)



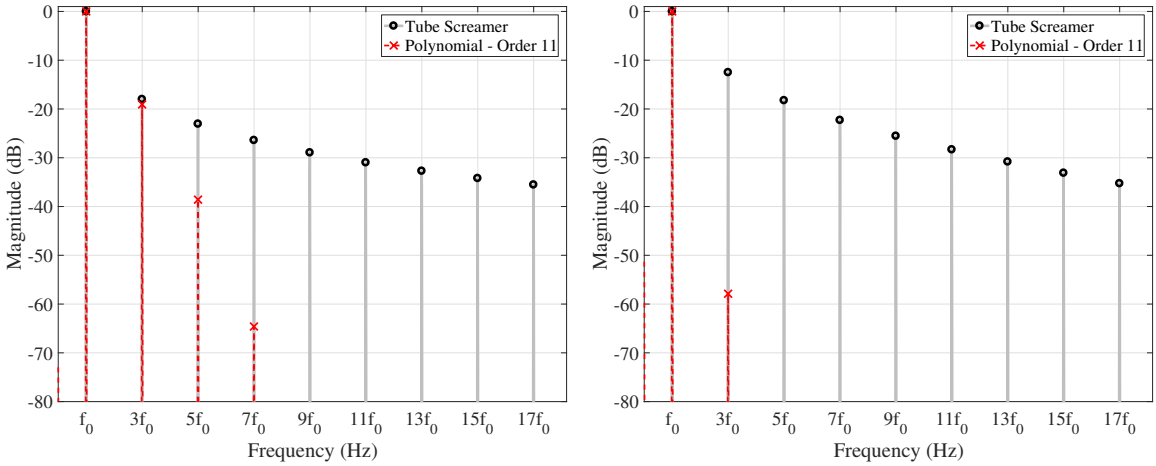
(b) Single-tone sine wave (different gain)



(c) Multi-tone sine wave (same gain)

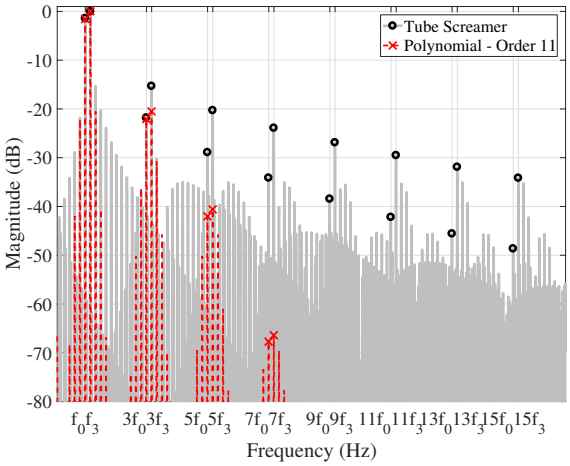
Figure 23: Spectra comparisons between the diode clipper and a polynomial of best fit.

The polynomial of best fit performs just as accurately as the Chebyshev and Hammerstein models, for the single-tone same gain input in Figure 23a. It outperforms those models for the multi-tone input in Figure 23c. This model, however, does not accurately simulate the magnitude response when using a different gain input, shown in Figure 23b.



(a) Single-tone sine wave (same gain)

(b) Single-tone sine wave (different gain)



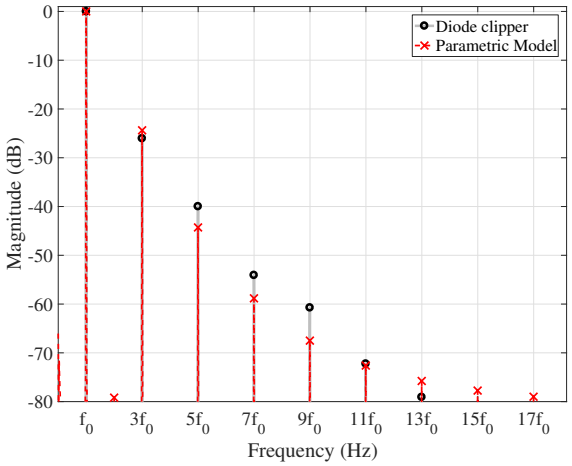
(c) Multi-tone sine wave (same gain)

Figure 24: Spectra comparisons between the Tube Screamer and a polynomial of best fit.

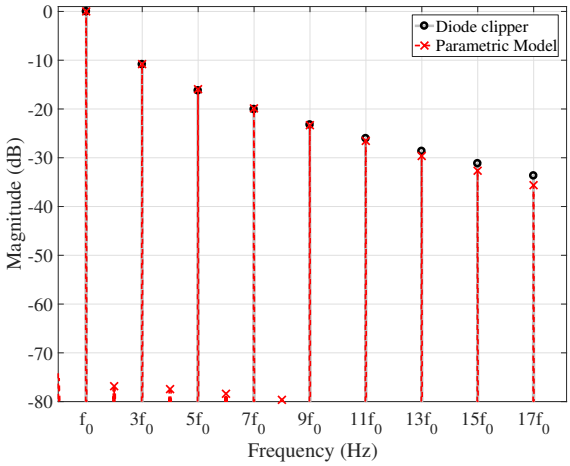
Despite the accuracy of this model for the diode clipper, it fails to model the high nonlinearity of the Tube Screamer. Increasing the polynomial order did not do much to improve the accuracy of the model. The polynomial was fitted to the approximation described in Section 3.1.2 which may explain why this polynomial is not very accurate, in comparison to the diode clipper polynomial. Another approach could be to use piecewise polynomials for the central region of the input–output and then fit ramps to end of the knee points.

4.1.4 Eichas Parametric Model

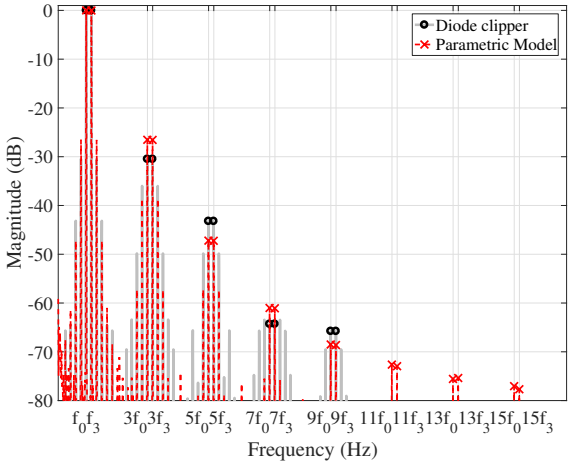
In the case of the parametric model, the two different gains were treated as low gain and high gain.



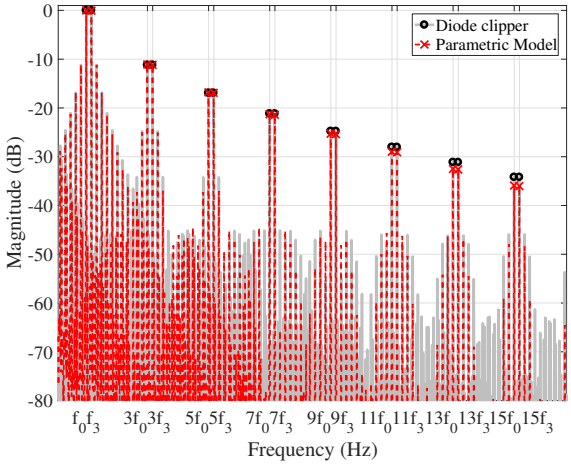
(a) Single-tone sine wave (low gain)



(b) Single-tone sine wave (high gain)



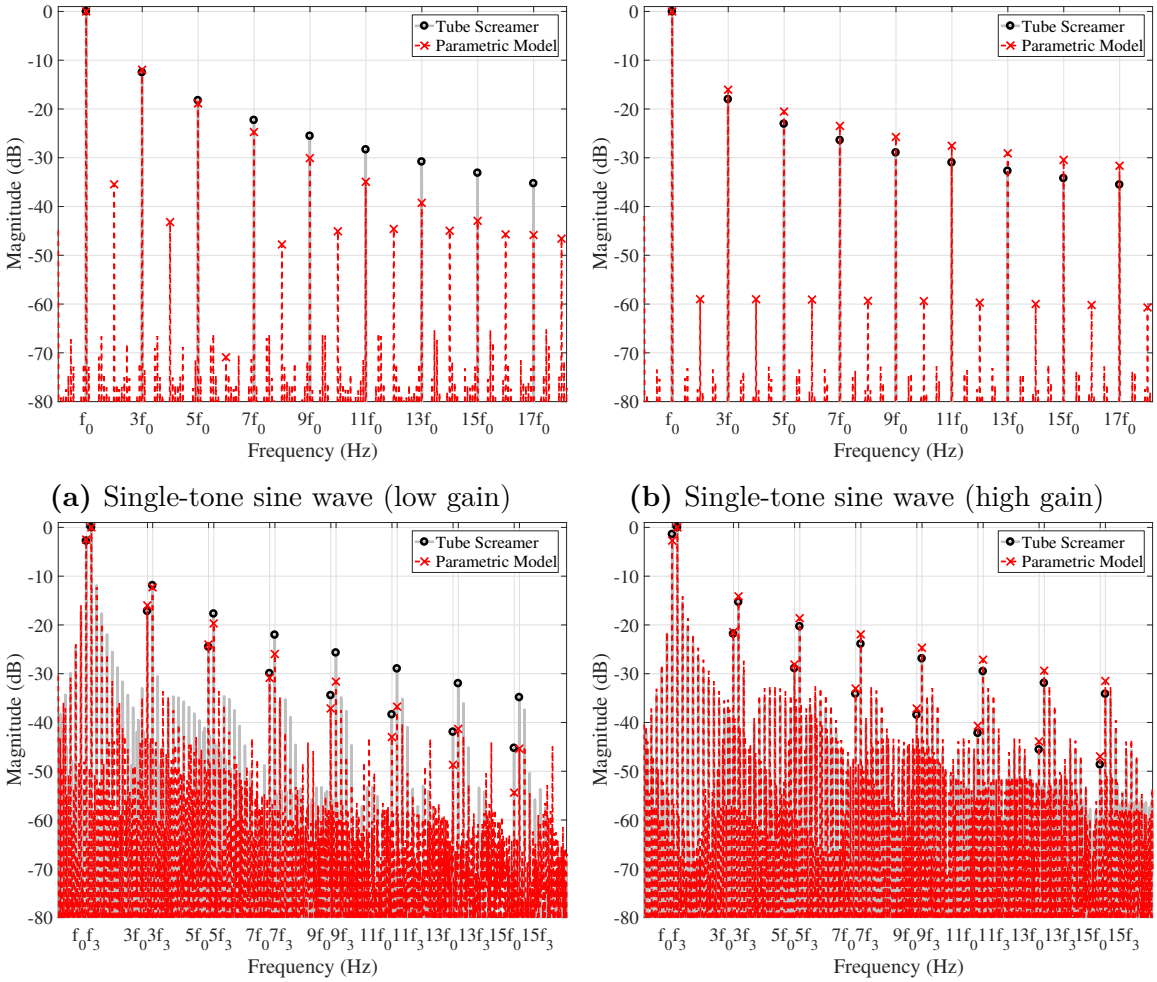
(c) Multi-tone sine wave (low gain)



(d) Multi-tone sine wave (high gain)

Figure 25: Spectra comparisons between the diode clipper and Eichas' parametric model.

Comparing the single-tone sine wave at the different gains in Figure 25a and Figure 25b indicates that the parametric model for the diode clipper is more accurate when the distortion is greater. This result is consistent when using a multi-tone sine wave, as shown in Figure 25c and Figure 25d. However, the intermodulation components around the 9^{th}-15^{th} harmonics of the frequencies are several dB too low.



(a) Single-tone sine wave (low gain)

(b) Single-tone sine wave (high gain)

(c) Multi-tone sine wave (low gain)

(d) Multi-tone sine wave (high gain)

Figure 26: Spectra comparisons between the Tube Screamer and Eichas' parametric model.

The results presented in Figure 26 for the parametric model of the Tube Screamer suggests that this model is highly accurate at high gains. The high gain inputs in Figure 26b and Figure 26d follow the same shape of the reference and the largest difference in magnitude occurs at the 17$^{\text{th}}$ harmonic in Figure 26b by about 2-3 dB. The low gain inputs in Figure 26a and Figure 26c are noticeably less accurate. This is due to the input signal used to optimise the nonlinear block, as the optimisation was biased from the higher amplitudes. The prominent even harmonics in the single-tone inputs could be due to different values for k_p and k_n which would result in a non-symmetric input–output relationship. The parametric model models the differing magnitudes of the harmonics of the two frequencies in both multi-tone plots, which none of the previous models achieved accurately.

4.2 Evaluation Scores

The models were also judged by using two different objective measures, the error-to-signal ratio (ESR) and the correlation coefficient, from [24]. The ESR is the percentage of error energy of a recorded guitar track played through the model and the reference system. It is defined as the ratio of error energy to the energy of the reference,

$$\text{ESR} = \frac{E_{\text{res}}}{E_{\text{ref}}} = \frac{\sum_{n=1}^{N} |y_{\text{ref}}[n] - y_{\text{model}}[n]|^2}{\sum_{n=1}^{N} |y_{\text{ref}}[n]|^2}, \tag{27}$$

where $y_{\text{ref}}[n]$ and $y_{\text{model}}[n]$ are the outputs to the reference system and the model, respectively and N is the length of the guitar track in samples. The closer the ESR is to 0 the more accurate the model.

The correlation coefficient is a measure of the linear dependence of two variables,

$$\rho_{y_{\text{ref}}, y_{\text{model}}} = \frac{\text{cov}(y_{\text{ref}}, y_{\text{model}})}{\sigma_{y_{\text{ref}}} \sigma_{y_{\text{model}}}}, \tag{28}$$

where cov is the covariance between the reference and model output, $\sigma_{y_{\text{ref}}}$ and $\sigma_{y_{\text{model}}}$ are the standard deviations of the reference and model output, respectively. When $\rho = 1$ the variables follow the same movement. When $\rho = -1$ the variables are exactly inverse of one another, i.e. when there is an increase in one, there is a decrease in the other. Finally, for $\rho = 0$ there is no relation between the variables.

The low gain for the diode clipper was 0.5 and the Tube screamer was 0.1. The high gain for the diode clipper was 5 and the Tube Screamer was 1, however, the high gain used for the Chebyshev and Hammerstein models was 1 for both systems. This is because these models performed best with this gain and are designed for this input gain.

Table 1: Objective results for modelling the diode clipper with a low gain input.

(a) Rock riff

Model	**ESR**	ρ
Chebyshev	3.2939×10^8	0.9989
Hammerstein	3.8347×10^{10}	0.8876
Polynomial	0.0950	**0.9999**
Parametric	**0.0110**	0.9949

(b) Jazz riff

Model	**ESR**	ρ
Chebyshev	3.2335×10^8	0.9981
Hammerstein	5.9874×10^{10}	0.9422
Polynomial	0.0943	**1.0000**
Parametric	**0.0145**	0.9927

The ESR results for the diode clipper in Table 1a and Table 1b suggests that the parametric model is the most accurate result, followed by the polynomial of best fit.

The ρ for the polynomial of best fit is an almost exact match. The high ESR values for the Chebyshev and Hammerstein models are due to summation of the harmonics without normalisation. The correlation coefficients for the models are unaffected by normalisation and from this the Chebyshev model performs quite well.

Table 2: Objective results for modelling the Tube Screamer with a low gain input.

(a) Rock riff

Model	**ESR**	ρ
Chebyshev	8.1963×10^{5}	0.5959
Hammerstein	2.1439×10^{27}	0.5171
Polynomial	0.8258	0.8348
Parametric	**0.1037**	**0.9954**

(b) Jazz riff

Model	**ESR**	ρ
Chebyshev	3.4050×10^{5}	0.6845
Hammerstein	8.6936×10^{26}	0.5498
Polynomial	0.8714	0.8463
Parametric	**0.1153**	**0.9947**

For the Tube Screamer results in Table 2 the parametric model performs the best with lowest ESR and highest correlation coefficient. The polynomial does not perform as well for the Tube Screamer as it did for the diode clipper. The Chebyshev model, while having a smaller ESR than the diode clipper results, the correlation coefficient is lower. The results for the Hammerstein model indicate it is highly inaccurate.

Table 3: Objective results for modelling the diode clipper with a high gain input.

(a) Rock riff

Model	**ESR**	ρ
Chebyshev	1.3948×10^{8}	0.9691
Hammerstein	5.0124×10^{8}	0.9734
Polynomial	0.0450	0.9877
Parametric	**0.0178**	**0.9914**

(b) Jazz riff

Model	**ESR**	ρ
Chebyshev	1.1241×10^{8}	0.9843
Hammerstein	5.5604×10^{8}	0.9733
Polynomial	0.0675	0.9889
Parametric	**0.0143**	**0.9942**

Table 4: Objective results for modelling the Tube Screamer with a high gain input.

(a) Rock riff

Model	**ESR**	ρ
Chebyshev	2.2739×10^{8}	0.6506
Hammerstein	1.0560×10^{9}	0.6717
Polynomial	0.4652	0.8095
Parametric	**0.1291**	**0.9730**

(b) Jazz riff

Model	**ESR**	ρ
Chebyshev	8.5605×10^{7}	0.6344
Hammerstein	5.9814×10^{8}	0.6684
Polynomial	0.5659	0.8003
Parametric	**0.1079**	**0.9906**

The ESR results in Table 3, again, show that the parametric model is the most accurate model for the diode clipper, followed by the polynomial of best fit. This is reinforced by the correlation coefficient results. The Chebyshev and Hammerstein models perform similarly to the low gain inputs.

The models do not perform as well for the Tube Screamer at the higher gain, as indicated in Table 4. For both riffs, the parametric model performs the best. Comparing the results from the rock to jazz riff for the parametric model suggests that this model performs better for slower inputs. Again, the Chebyshev and Hammerstein models are highly inaccurate.

5 Proposed Method

Following from the results in Section 4, Eichas' parametric model was deemed the most accurate black-box model due to the close match in spectral content and the flexibility regarding input gain. The purpose of the proposed method is to maintain this accuracy but reduce the computational time of the parametric model. This model was specifically designed to simulate the Tube Screamer.

5.1 Novel Model

The proposed method changes the structure of Eichas' parametric model shown in Figure 7 of Section 2.6. This new method consists of a high-pass filter, followed by the nonlinearity block, then a low-pass filter. The block diagram for this novel model is illustrated in Figure 27.

Figure 27: Block diagram for the proposed model.

The high-pass filter used was the DC Blocker and the low-pass was an FIR filter, indicated by $H_{\mathrm{DC}}(z)$ and $H_{\mathrm{LP}}(z)$, respectively. Figure 28 gives a comparison for the magnitude response of the filters to the linear response of the Tube Screamer.

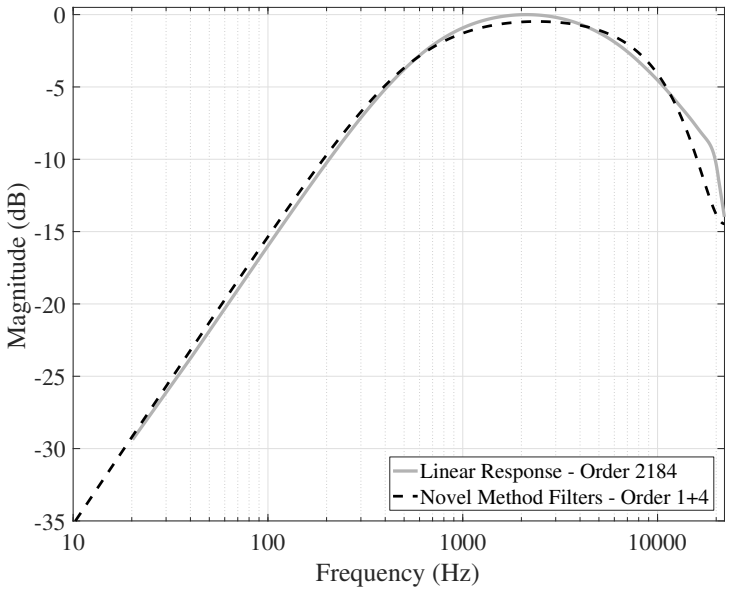


Figure 28: Magnitude response of the novel method filters against the linear response of the Tube Screamer.

5.1.1 DC Blocker

The DC blocker was used to model the steep high-pass effect of the linear response of the Tube Screamer. The transfer function for the DC blocker is defined as [53]:

$$H_{\mathrm{DC}}(z) = \frac{1+p}{2}\frac{1-z^{-1}}{1-pz^{-1}}, \tag{29}$$

where p is the real pole of the filter ($0 < p < 1$). The pole is determined by the cut-off frequency, f_c, and the sampling rate:

$$p = \tan\left(\frac{\pi}{4} - \frac{\pi f_c}{F_{\mathrm{s}}}\right). \tag{30}$$

The first term in Eq. 29 restricts the filter such that it does not produce a positive gain on the signal. The cut-off frequency was determined as the low frequency $-3\,\mathrm{dB}$ point of the Tube Screamer linear response.

5.1.2 Low-Pass FIR Filter

The low-pass FIR filter was designed as a fourth-order filter with a cut-off frequency determined by the $-3\,\mathrm{dB}$ point to the high frequencies from the Tube Screamer linear response. The transfer function for this filter is:

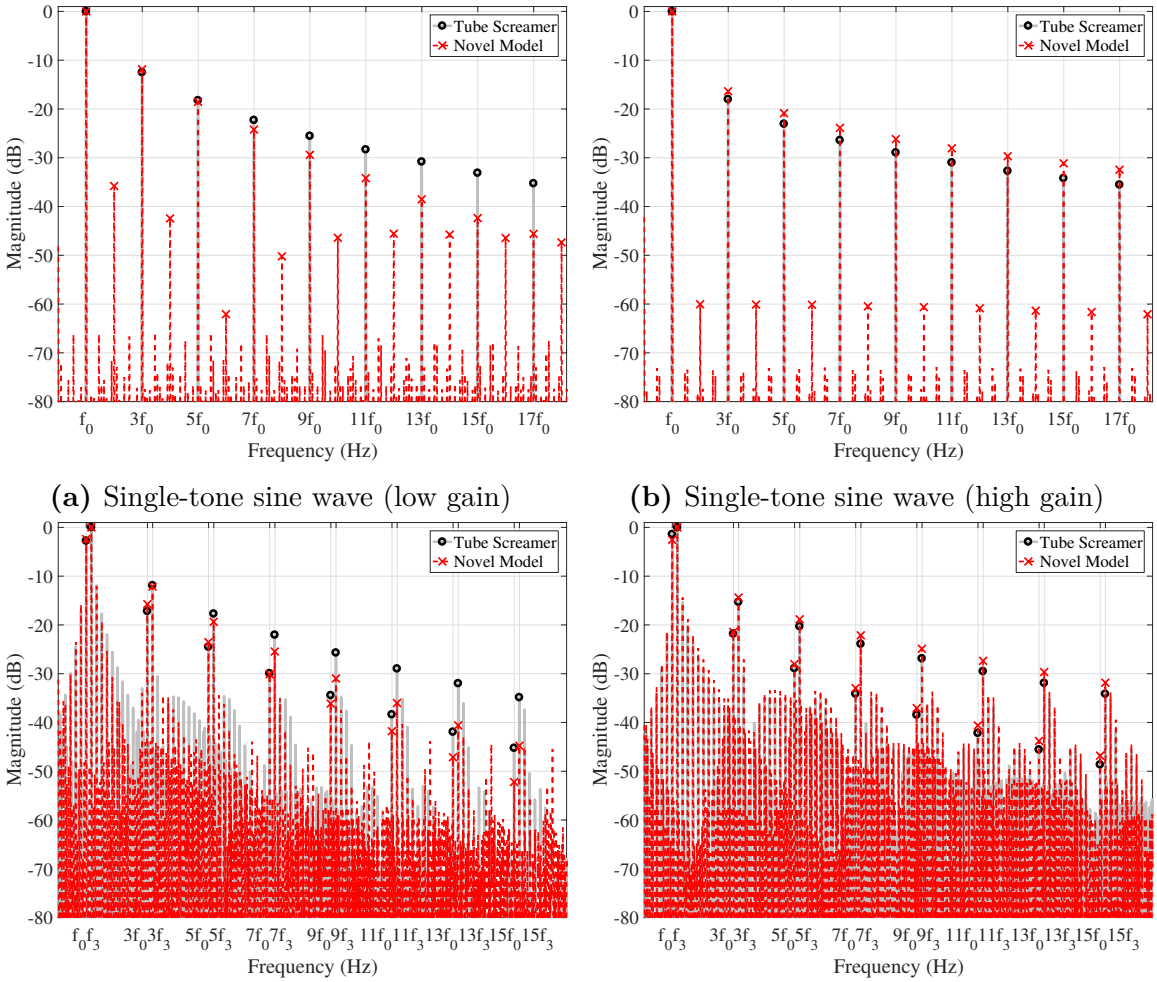
$$H_{\mathrm{LP}} = b_0 + b_1 z^{-1} + b_2 z^{-2} + b_1 z^{-3} + b_0 z^{-4}. \tag{31}$$

5.2 Results

The results for the model were produced using the novel model described above with the nonlinear block optimised for results of the Eichas' parametric model in Section 4. The spectra of the model was compared against the Tube Screamer. The model was evaluated using the objective measures with comparisons to Eichas' parametric model. The final test was to determine the computation time of each model.

5.2.1 Spectra Comparisons

This method was tested using the same test cases as in Section 4.1 for Eichas' parametric model. The results from the test cases are displayed in Figure 29.



(a) Single-tone sine wave (low gain)

(b) Single-tone sine wave (high gain)

(c) Multi-tone sine wave (low gain)

(d) Multi-tone sine wave (high gain)

Figure 29: Spectra comparisons between the Tube Screamer and the proposed model.

The results above show that model is accurate and produces extremely similar plots to the Eichas parametric model, in Figure 26 of Section 4.1.4. The novel model suffers from the same issues as the Eichas' model due to using the same optimised nonlinear block.

5.2.2 Evaluation Scores

The novel model was also tested for the ESR and correlation coefficient. In the following tables the Eichas' paramteric model is referred to as Eichas'. The inputs to test the models were the same guitar recordings as used in Section 4.2, using the same low gain and high gain cases.

Table 5: Objective results comparisons for Eichas' and the proposed method for modelling the Tube Screamer with a low gain input.

(a) Rock riff

Model	**ESR**	ρ
Eichas'	**0.1037**	**0.9954**
Proposed	0.1432	0.9666

(b) Jazz riff

Model	**ESR**	ρ
Eichas'	**0.1153**	**0.9947**
Proposed	0.1552	0.9653

Table 6: Objective results comparisons for Eichas' and the proposed method for modelling the Tube Screamer with a high gain input.

(a) Rock riff

Model	**ESR**	ρ
Eichas'	**0.1291**	**0.9730**
Proposed	0.1731	0.9424

(b) Jazz riff

Model	**ESR**	ρ
Eichas'	**0.1079**	**0.9906**
Proposed	0.1691	0.9477

The results for the proposed model with a low gain input in Table 5 show that it is only slightly less accurate than Eichas' model. The difference in ESR is +4% and for the correlation coefficient is −0.029. This is also held true for the high gain results in Table 6; however, the range of difference has increased. The ESR now differs in the range of +4%–6% and the correlation coefficient increased to −0.030–0.043.

5.2.3 Computational Costs

The novel method and Eichas' parametric model were also compared on the number of multiplications, additions, number of states, and the average runtime of the filters for each model. The tests were ran on a Windows 8.1 64-bit operating system with an Intel Core i7 @ 2.80 GHz laptop CPU and a NVIDIA GeForce GTX 860M. The results of these tests are presented in Table 7.

Table 7: Number of operations per output sample.

	Eichas'	**Proposed**
Multiplications	2185	7
Additions	2184	6
States	2184	5

The number of multiplications, additions and states are drastically reduced using the proposed method in comparison to Eichas' model. The proposed model requires 312 times less multiplications, 364 times less additions, and 436 times less states. This is equivalent to a 99% reduction in multiplications, additions and states. It

should be noted that the length of the impulse response used for the Eichas' model is not fixed. It is possible to reduce this number; however, in doing so would also reduce the frequency resolution of the filter.

The average runtime was calculated by running the filtering process of each model 1000 times and taking the average of the time for each run. Four different input signals were used to test the runtime. The first was a one-second long impulse and the second was single-tone sine wave of the same length. The final two inputs were recordings of different guitar riffs, one at a length 4 seconds and another at a length of 20 seconds. The sampling rate for all inputs was $F_s = 44100\,\text{Hz}$. Table 8 shows the average runtimes for these results.

Table 8: Runtime of the models in milliseconds.

	Eichas'	**Proposed**
Impulse	17.6 ms	4.18 ms
Sine	9.45 ms	0.416 ms
Short riff	46.7 ms	2.19 ms
Long riff	249 ms	9.99 ms

The runtime for the novel model shows a significant reduction in each input tested in comparison to Eichas' model. The impulse test is interesting as it required longer than a single-tone sine wave, despite being the same length. The reduction in time for the impulse was 76%, while all the musical inputs saw a reduction in 95%.

6 Conclusion

This Thesis presented various black-box modelling techniques. The definition and implementation of the models were detailed. The models were compared based on their spectral accuracy, their error in time and their correlation with respect to the reference. From these results it was determined that the Eichas' parametric model performed the best.

A new method was proposed that was based on the previous parametric model. This model restructured the filtering and nonlinearity stages. It also used filters to simulate the linear response of the nonlinear audio system, in an effort to reduce the order of the filtering stage. The model performed with a similar accuracy to the Eichas model, but was significantly faster. The runtime required was reduced by 95% and the number of operations was reduced by 99%.

For future developments on black-box modelling, one approach would be to place the nonlinearity in a feedback loop. This would more closely follow the coupling of the filtering and nonlinear behaviour of a real audio circuit. This was attempted by using the nonlinear block from Eichas' parametric model and creating a filter that could also be parametrised. This resulted in an unstable filter while optimising and requires further research.

Another possible advancement could be to replace the LTI filter of Eichas' parametric model, with a linear time-varying (LTV) filter. An LTV filter which effectively changes the cut-off frequency of the filter, depending on the input, would match the warping of the frequencies of highly nonlinear circuits. This was another method that was attempted but the effective cut-off frequencies of the filters changed too dramatically. This resulted in a flat frequency response at high gains, despite the harmonic spectra results showing some band-pass effect should occur. This was due to following the power spectrum of the Tube Screamer, which may not be an accurate method for analysing the warping effect of the distortion pedal.

Evaluating the models via user perception and interaction would be beneficial. Performing listening tests to determine what humans perceive as the most accurate model and to identify the most important elements to simulate accurately in the models. Another possible evaluation would be to have guitarists play with the models. This could help to discover the model that has the closest "musical feel" to the physical pedals.

References

[1] G. Borin, G De Poli, and D Rocchesso. Elimination of delay-free loops in discrete-time models of nonlinear acoustic systems. *IEEE Trans. Speech and Audio Processing*, 8(5):597–605, Sept. 2000.

[2] D. T. Yeh. *Digital Implementation of Musical Distortion Circuits by Analysis and Simulation*. PhD thesis, Stanford University, Stanford, California, June 2009.

[3] D. T. Yeh, J. S. Abel, and J. O. Smith III. Automated physical modeling of nonlinear audio circuits for real-time audio effects-part I: Theoretical development. *IEEE Trans. Audio, Speech and Language Processing*, 18(4):728–737, May 2010.

[4] R. C. D. Paiva. *Circuit Modeling Studies Related to Guitars and Audio Processing*. PhD thesis, Aalto University, Espoo, Finland, Sept. 2013.

[5] M. Holters and U. Zölzer. A generalized method for the derivation of non-linear state-space models from circuit schematics. In *Proc. 2015 23rd European Signal Processing Conference*, EUSIPCO, pages 1073–1077, Nice, France, Aug. 31-Sept. 4, 2015.

[6] A. Fettweis. Digital filters related to classical structures. *AEU: Archive für Elektronik und Übertragungstechniky*, 25(6):79–89, Feb. 1971.

[7] A. Fettweis. Wave digital filters: Theory and practice. *Proceedings of the IEEE*, 74(2):270–327, Feb. 1986.

[8] D. Fränken and K. Ochs. Synthesis and design of passive Runge-Kutta methods. *AEU: International Journal of Electronics and Communications*, 55(6):417–425, Jan. 2001.

[9] D. T. Yeh, J. S. Abel, and J. O. Smith III. Simulation of the diode limiter in guitar distortion circuits by numerical solution of ordinary differential equations. In *Proc. Intl. Conf. Digital Audio Effects*, DAFx, pages 197–204, Bordeaux, France, Sept. 10-15, 2007.

[10] D. T. Yeh, J. S. Abel, and J. O. Smith III. Simplified physically-informed models of distortion and overdrive guitar effects pedals. In *Proc. Intl. Conf. Digital Audio Effects*, DAFx, pages 10–15, Bordeaux, France, Sept. 10-15, 2007.

[11] M. Karjalainen and J. Pakarinen. Wave digital simulation of a vacuum-tube amplifier. In *Proc. Intl. Conf. Acoustics, Speech, and Signal Process*, ICASSP, pages 153–156, Toulouse, France, May 14-19, 2006.

[12] J. Pakarinen, M. Tikander, and M. Karjalainen. Wave digital modeling of the output chain of a vacuum-tube amplifier. In *Proc. Intl. Conf. Digital Audio Effects*, DAFx, pages 55–59, Como, Italy, Sept. 1-4, 2009.

[13] D. T. Yeh. Automated physical modeling of nonlinear audio circuits for real-time audio effects–part II: BJT and vacuum tube examples. *IEEE Trans. Audio, Speech and Language Processing*, 20(4):1207–1216, May 2012.

[14] S. D'Angelo and V. Välimäki. Wave-digital polarity and current inverters and their application to virtual analog audio processing. In *Intl. Conf. Acoustics, Speech, and Signal Process*, ICASSP, pages 469–472, Kyoto, Japan, Mar. 2012.

[15] R. C. D. Paiva, S. D'Angelo, J. Pakarinen, and V. Välimäki. Emulation of operational amplifiers and diodes in audio distortion circuits. *IEEE Trans. Circuits and Systems*, 59(10):688–692, Oct. 2012.

[16] G. Borin, G De Poli, and D Rocchesso. Block-based physical modeling for digital sound synthesis. *IEEE Signal Processing Magazine*, 24(2):42–54, Mar. 2007.

[17] V. J. Mathews. Adaptive polynomial filters. *IEEE Signal Processing Magazine*, 8(3):10–26, July 1991.

[18] T. Ogunfunmi. *Adaptive Nonlinear System Identification: The Volterra and Wiener Model Approaches*. Springer, New York, NY, USA, 2007. ISBN 978-0-387-26328-1.

[19] T. Hélie. Volterra series and state transformation for real-time simulations of audio circuits including saturations: Application to the Moog ladder filter. *IEEE Trans. Audio, Speech and Language Processing*, 18(4):747–759, May 2010.

[20] A. Novak, L. Simon, P. Lotton, and F. Kadlec. A new method for identification of nonlinear systems using MISO model with swept-sine technique: Application to loudspeaker analysis. In *Proc. 124th Convention of the Audio Engineering Society*, AES, Amsterdam, The Netherlands, May 17-20, 2008.

[21] A. Novak, L. Simon, P. Lotton, and J. Gilbert. Chebyshev model and synchronized swept sine method in nonlinear audio effect modeling. In *Proc. Intl. Conf. Digital Audio Effects*, DAFx, pages 423–426, Graz, Austria, Sept. 6-10, 2010.

[22] R. C. D. Paiva, J. Pakarinen, and V. Välimäki. Reduced-complexity modeling of high-order nonlinear audio systems using swept-sine and principal component analysis. In *Proc. AES 45th Conf. Applications of Time-Frequency Processing in Audio*, AES, Espoo, Finland, Mar. 2012.

[23] F. Eichas, S. Möller, and U. Zölzer. Block-oriented modeling of distortion audio effects using minimization. In *Proc. Intl. Conf. Digital Audio Effects*, DAFx, pages 243–248, Trondheim, Norway, Nov. 30-Dec. 3, 2015.

[24] F. Eichas and U. Zölzer. Black-box modeling of distortion circuits with block-oriented models. In *Proc. Intl. Conf. Digital Audio Effects*, DAFx, pages 39–45, Brno, Czech Republic, Sept. 5-9, 2016.

[25] F. Eichas, S. Möller, and U. Zölzer. Block-orientated gray box modelling of guitar amplifiers. In *Proc. Intl. Conf. Digital Audio Effects*, DAFx, pages 184–191, Edinburgh, UK, Sept. 5-9, 2017.

[26] A. J. Berkhout, D. de Vries, and M. M. Boone. A new method to acquire impulse responses in concert halls. *Journal of the Acoustical Society of America*, 68(1):179–183, July 1980.

[27] A. Farina. Simultaneous measurement of impulse response and distortion with a swept-sine technique. In *Proc. 108th Convention of the Audio Engineering Society*, AES, Paris, France, Feb. 19-22, 2000.

[28] A. Farina. Non-linear convolution: A new approach for the auralization of distorting systems. In *Proc. 110th Convention of the Audio Engineering Society*, AES, Amsterdam, Netherlands, May 12-15, 2001.

[29] K.C. Tan and Y. Li. Grey-box model identification via evolutionary computing. *Control Engineering Practice*, 10(7):673–684, July 2002.

[30] J. Pakarinen. A review of digital techniques for modeling vacuum-tube guitar amplifiers. *Computer Music Journal*, 32(2):85–100, Summer 2009.

[31] R. A. Schaefer. Electronic musical tone production by nonlinear waveshaping. *Journal of the Audio Engineering Society*, 18(4):413–417, Aug. 1970.

[32] C. Y. Suen. Derivation of harmonic equations in nonlinear circuits. *Journal of the Audio Engineering Society*, 18(6):675–676, Dec. 1970.

[33] T. Letowski. Difference limen for nonlinear distortion in sine signals and musical sounds. *Acta Acustica*, 34(2):106–110, Dec. 1975.

[34] J. S. Abel and D. P. Berners. A technique for nonlinear system measurement. In *Proc. 121st Convention of the Audio Engineering Society*, AES, San Francisco, CA, USA, Oct. 5-8 2006.

[35] T. Hélie and V. Smet. Simulation of weakly nonlinear propagation in a straight pipe: Application to a real-time brassy audio effect. In *Proc. 2008 16th*

Mediterranean Conference on Control and Automation, IEEE, pages 1580–1585, Ajaccio, France, June 25-27, 2008.

[36] T. Hélie and D. Roze. Sound synthesis of a nonlinear string using Volterra series. *Journal of Sound and Vibration*, 314(1-2):275–306, July 2008.

[37] T. Hélie. On the use of Volterra series for real-time simulations of weakly nonlinear analog audio devices: Application to the Moog ladder filter. In *Proc. Intl. Conf. Digital Audio Effects*, DAFx, pages 7–12, Montreal, Canada, Sept. 18-20, 2006.

[38] R. C. Heyser. Acoustical measurements by time delay spectrometry. *Journal of the Audio Engineering Society*, 15(4):370–382, Oct. 1967.

[39] R. C. Heyser. Loudspeaker phase characteristics and time delay distortion: Part 1. *Journal of the Audio Engineering Society*, 17(1):30–41, Jan. 1969.

[40] J. Vanderkooy. Another approach to time-delay spectrometry. *Journal of the Audio Engineering Society*, 7/8(34):523–538, July-Aug. 1986.

[41] M. A. Poletti. Linearly swept frequency measurements, time-delay spectrometry, and the Wigner distribution. *Journal of the Audio Engineering Society*, 36(6): 457–468, June 1988.

[42] A. Novak, P. Lotton, and L. Simon. Synchronized swept-sine: Theory, application and implementation. *Journal of the Audio Engineering Society*, 63(10): 786–798, Oct. 2015.

[43] T. Schmitz and J. Emcrechts. Hammerstein kernels identification by means of a sine sweep technique applied to nonlinear audio devices emulation. *Journal of the Audio Engineering Society*, 65(9):696–710, Sept. 2017.

[44] M. Le Brun. Digital waveshaping synthesis. *Journal of the Audio Engineering Society*, 27(4):250–266, Apr. 1979.

[45] D. Arfib. Digital synthesis of complex spectra by means of multiplication of nonlinear distorted sine waves. *Journal of the Audio Engineering Society*, 27 (10):757–768, Dec. 1979.

[46] T. O. Sharpee, K. D. Miller, and M. P. Stryker. On the importance of the static nonlinearity in estimating spatiotemporal neural filters with natural stimuli. *Journal of Neurophysiology*, 99(5):2496–2509, May 2008.

[47] K. Levenberg. A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, 2(2):164–168, 1944.

[48] D. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 11(2):431–441, 1963.

[49] D. C. Sorensen. Newton's method with a model trust region modification. *SIAM Journal on Numerical Analysis*, 19(2):409–426, Apr. 1982.

[50] A. R. Conn, N. I. M. Gould, and P. L. Toint. *Trust Region Methods*. MOS-SIAM Series on Optimization, 2000. ISBN 978-0-89871-460-9.

[51] D. Hunter. *Guitar Effects Pedals - the Practical Handbook*. Backbeat Books, 2004. ISBN 978-0-87930-806-3.

[52] K. J. Werner, N. Vaibhav, A. Bernardini, J. O. Smith III, and A Sarti. An improved and generalized diode clipper model for wave digital filters. In *Proc. 139th Audio Engineering Society Convention*, AES, New York, USA, Oct. 29-Nov. 1, 2015.

[53] S. K. Mitra and J. F. Kaiser. *Handbook for Digital Signal Processing*. John Wiley & Sons, New York, USA, 1st edition, 1993.