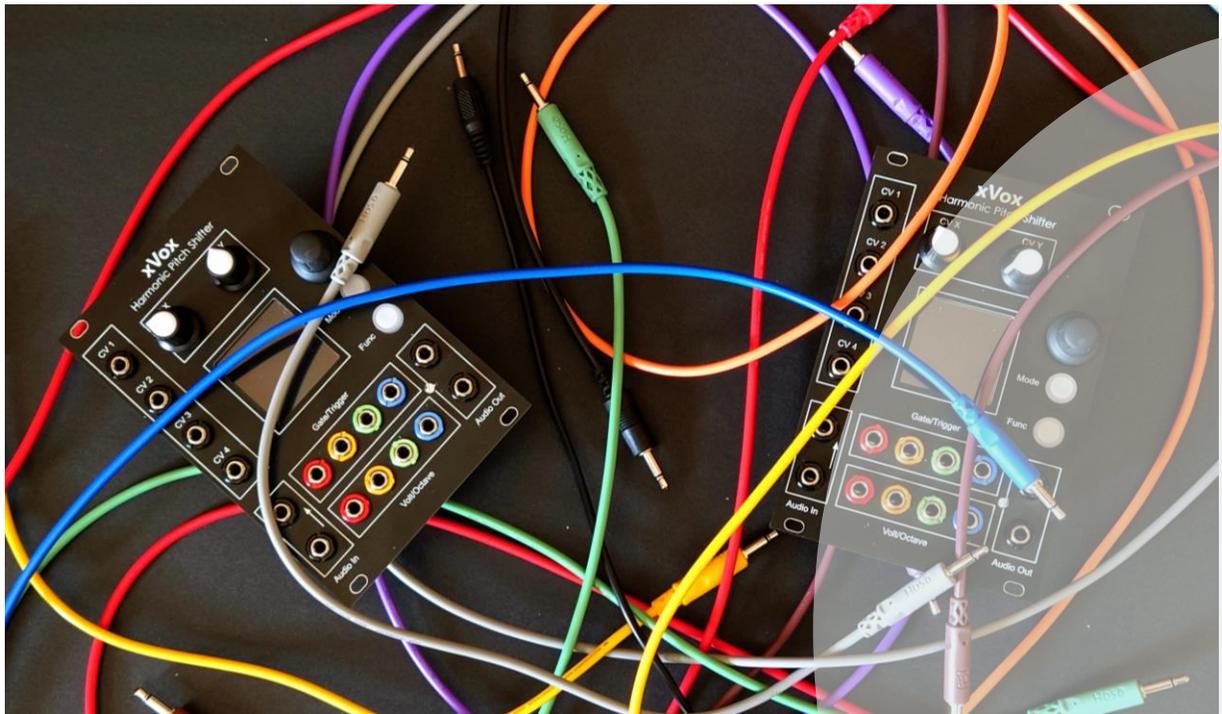




gregsbrain
EURORACK MODULAR



xVox

Four Voice
Harmonic
Pitch Shifter
User Manual

Note: this User Manual is for xVox firmware version 1.7 and later

Contents

Features	2
Introduction	3
Front Panel Layout	4
User Interface.....	5
Page Controls	6
Audio Modes	7
Pitch Shift Mode.....	8
7th Chord Mode	9
Chord Progression Mode	10
The Tuning Function	11
Function Pages	13
Gate Configuration	14
Envelope Generator	16
Audio Levels	17
Filter Controls	18
Audio Options	19
CV Configuration.....	21
System Settings	22
Flow diagram	25
Firmware Update.....	26
Release Change History	27
Implementation Notes	28
VCV Rack Addendum	30

Features

Front panel

- Two audio inputs, primary normalled to secondary.
- Stereo audio output.
- Four VPO and four gate/trigger inputs.
- Four assignable CV sockets and two assignable knobs.
- Pushbutton rotary encoder and two LED push buttons.
- 1.5-inch 128x128 RGB OLED display.

Four voice channels

- Pitch control over each channel depends on pitch-shift mode.
- Individually triggered or triggered in unison.

Three pitch shift modes

- Open tuning or quantization to scales, semitones, or octaves.
- 7th chord pitch shift for root, 3rd, 5th, and 7th.
- Chord progressions based on circle-of-fifths.

Envelope generator

- Attack/Decay, ADSR, or OFF
- Solo, unison, or strum trigger options or on note change.
- Selectable attack and decay curves
- Attack and decay times from several milliseconds to minutes

Other functions

- Vocoder cross-modulation effect
- Sample and hold option when not using envelope generator.
- Tuner with automatic tuning feature.
- Filter with controls for HP, LP, and slope.
- Audio input and output level controls with soft-clipping and AGC.

Color-coded visualizations

- Piano keyboard shows notes as they are triggered.
- Spectrum shows real-time harmonics of each voice.

Dimensions

- 17HP wide, 32mm deep.

Power consumption

- **+12V** 144mA (133mA with display dimmed)
- **-12V** 8.4mA
- **+5V** 0mA

Introduction

A pitch shifter changes the musical pitch of an incoming audio signal. In other words, it can move a musical note up or down without changing the tone. Importantly, the ratios of the harmonics that make up the audio signal are largely preserved so the pitch shifted audio sounds the same, just a higher or lower note. This is different from a frequency shifter which moves all harmonics up or down by some fixed amount often dramatically changing the timbre of a note, which is either a good thing or a bad thing depending on what you are trying to do.

The xVox Eurorack module is a four-voice pitch shifter. It can take in one note and pitch shift it to generate a four-note chord. It does this in real-time so changes in the audio input have an immediate effect on the sound coming out.

There are four volt-per-octave (VPO) inputs to control the pitch shift amount for each voice. The default input voltage range is -5 Volts to +5 Volts giving a maximum pitch shift range of plus or minus five octaves. For compatibility with sequencers and other CV sources that only provide a 0 to 5 Volt output range there is an option to set the input range to 0 to 5 Volts. With this option the input voltage for no pitch change is 3 Volts, so the maximum pitch shift range via the VPO inputs is three octaves down and two octaves up. Extreme pitch shifts in either direction will change the sound as lower harmonics get shifted up out of audible range or higher harmonics shifted down start to overlap, this is explained in the Implementation Notes section.

xVox has three different pitch shift modes that output notes that are musically related. This is why we call xVox a “harmonic” pitch shifter. The three modes are Pitch Shift, 7th Chord, and Progression.

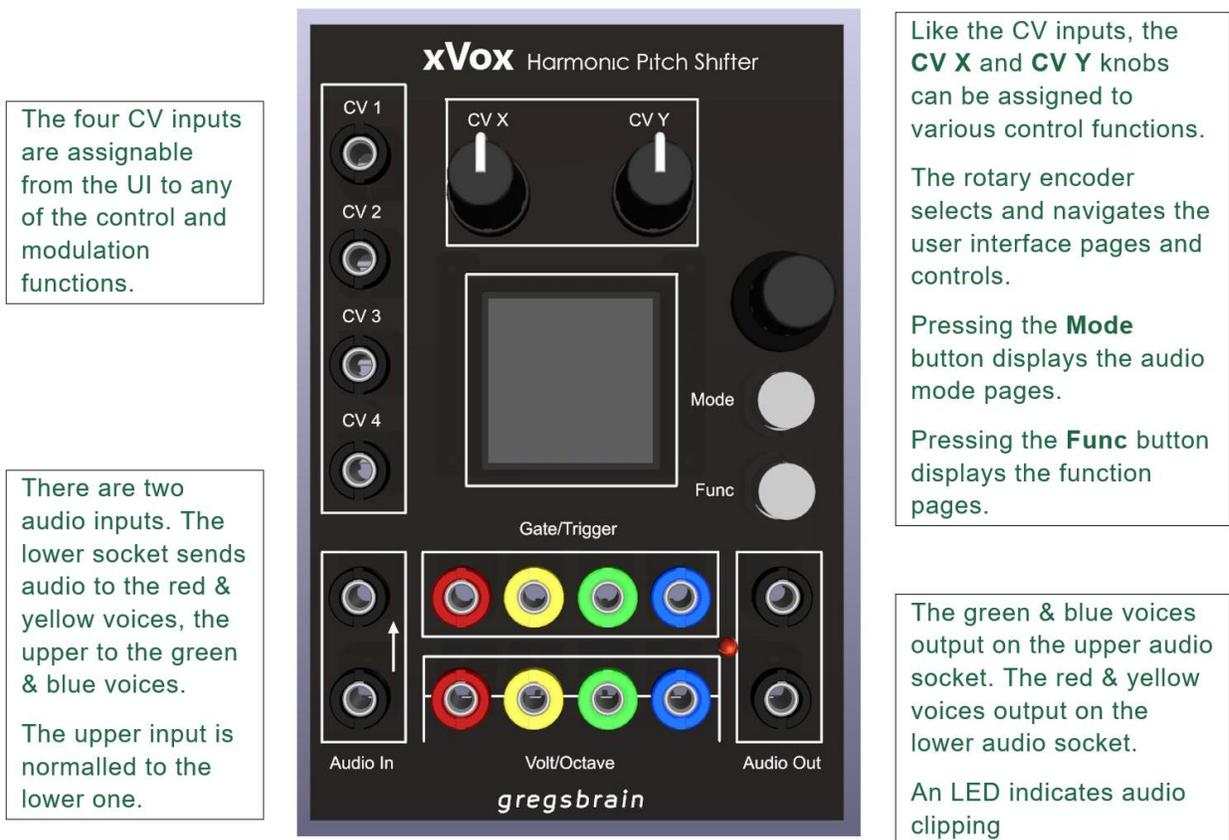
In Pitch Shift mode the VPO inputs are quantized so the output notes conform to a chosen musical scale such as major or minor. The complete set is listed in the Pitch Shift Mode section.

In 7th Chord mode the audio input is pitch shifted so the output notes are the root, 3rd, 5th, and 7th notes of a chosen chord structure. Chord inversions and drop chords are supported as covered in detail in the 7th Chord Mode section.

Progression mode, as the name suggests, generates notes that conform to a chord progression. This mode is based on the musical theory of the Circle of 5ths and covered in detail in the Progression Mode section of this user guide.

Front Panel Layout

xVox is 17HP wide, has two audio inputs, two audio outputs, four CV inputs, two control knobs, four VPO inputs and four gate/trigger inputs. In the center is a 1.5-inch 128x128 pixel RGB OLED display and a rotary encoder and two illuminated push buttons for interacting with the module.



For each voice channel there is a Gate/Trigger and a Volt/Octave input. The input socket jacks are color coded to match information shown in the OLED display. The Volt/Octave inputs are configurable from the user interface to be either bipolar -5V to +5V with a 0V center point, or unipolar 0V to 5V with a center point at 3V. The Trigger/Gate inputs can trigger an internal envelope generator or a sample and hold.

User Interface

The user interface is organized as pages. Pressing the **mode** button opens the page for the current audio mode, pressing the **func** button opens the most recently selected function page. Pressing the **mode** button a second time opens the Audio Mode menu page, pressing the func button opens the Functions menu page. The corresponding button flashes while a menu page is being displayed. Use the rotary encoder to choose an audio mode or functions page from the list and click the encoder button to open it.



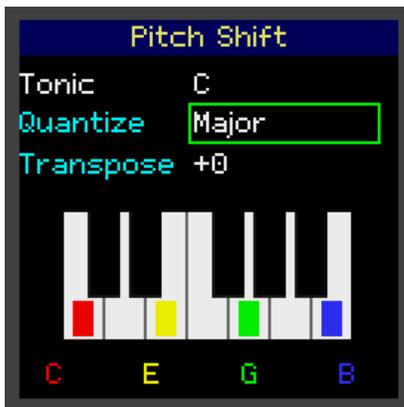
Selecting **System Settings** on the Functions menu page brings up another menu page that lists infrequently used functions such as changing the display settings, recalibrating the VPO inputs or viewing the firmware version.



Most changes to settings made via the user interface, including the most recent audio mode, are stored to flash memory, and restored when the module is next powered on. Selecting **Factory Reset** from the System Settings menu page sets all configurable settings to their original defaults.

Page Controls

The images below show examples of the Pitch Shift page with two different configurations. In pitch shift mode there are two parameters controlling this mode: **Quantize** and **Transpose**. In the left example both parameters are controlled by the rotary encoder, **Quantize** and **Transpose** are displayed in light-blue, and the green box shows that focus is on the **Quantize** parameter. Rotate the encoder to navigate between the two parameters. Click the encoder button and the box turns yellow indicating that the parameter has been selected. A selected parameter is changed by rotating the encoder, click a second time to confirm the selection and box returns to green. Options that only have two values such as check boxes need only be clicked to change the value; they do not transition to selected.



In the right-hand example both the **Quantize** and **Transpose** parameters have been configured for CV control. Assigned parameter names appear in yellow and the numbers at the right-hand side of the display show which CV input is controlling that parameter. CV assignments are set from the **CV Configuration** page and the selected assignments are saved to flash memory.

The text at the top showing the tonic note, a **C** in this case, is informational. The note displayed is based on the pitch of the note coming in offset up or down by the number of semitones set by the **Transpose** parameter. See the Tuning Function section for details on pitch detection. The piano keyboard at the bottom of the page is a real-time display of the pitch shifted notes for each of the four voices, in this case a C Major 7th chord. An option described in the Display Options section can be set to show a real-time harmonic spectrum of the notes instead of the piano keyboard.

Audio Modes

There are four audio modes, Pitch Shift, 7th Chord, Progression, and Tuner. The first three apply different algorithms to change the pitch of the output notes.



Pitch Shift mode provides independent control of the output pitch of the four voices with optional quantization to a selected scale.

7th Chord mode quantizes the output pitch of the four voices to a root note, 3rd, 5th, and 7th interval of a selected chord.

Progression mode quantizes the output pitch of the four voices to the notes of a chord selected from the chords of one of classical musical modes.

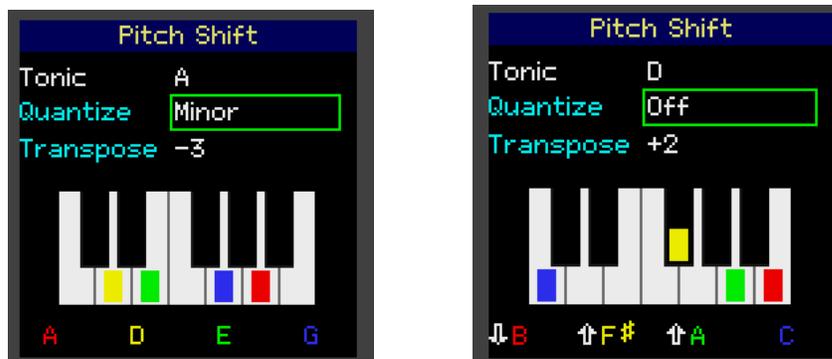
Tuning is not a pitch shift mode but is accessed from the same page because the tuning function and the audio modes are mutually exclusive. The module is either playing an audio mode or running the tuner. The tuner affects all the audio modes.

There is an [Audio Options](#) page for enabling some optional features that apply to all audio modes; see the *Function Pages* section below for more information.

The following sections describe audio modes and the tuning function in detail

Pitch Shift Mode

The Pitch Shift page selects the pitch shift audio mode. In this mode the pitches of each of the four voices are independently controlled by the four volt-per-octave inputs. This page has two parameters: **Quantize** and **Transpose**.



Quantize

Quantization constrains the pitches of the output notes to fit within a selected scale. The quantization options and scales currently supported are shown below.

Quantization Options

- Major*
- Minor*
- Harmonic Minor*
- Melodic Minor*
- Phrygian*
- Lydian*
- Mixolydian*
- Blues*
- Minor Blues*
- Major Blues*
- Chromatic*
- Octaves*
- Microtonal*
- Off**

Off disables quantization completely and pitches are set based solely on the VPO value modified by the transposition value. If the piano keyboard is being displayed, pitches that fall between notes show an up-arrow or down-arrow indicating the pitch is higher or lower than the note shown as shown in the right-hand example above.

Octave quantizes the pitch to the nearest octave up or down. This creates big fat sounds and when the inputs are modulated produces some lovely effect.

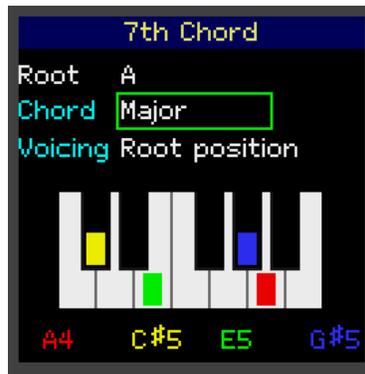
Microtonal changes the input scaling to Volt/Semitone which is useful for chorus and other types of effect where a small pitch spread is required.

Transpose

The transpose parameter adjusts the output pitch of all the voices up or down by a specified number of semitones. The transposition range is -12 to +12 semitones.

7th Chord Mode

The 7th Chord page selects the 7th chord audio mode. This mode generates chords based on a chosen chord structure. In 7th chord mode the red voice channel sets the root note of the chord, the other voice channels are pitch shifted to the 3rd, 5th, and 7th intervals of the chord. If voltage is applied to the yellow, green, or blue voice channels, the pitch of those voices are quantized to the nearest 3rd, 5th, or 7th interval preserving the chord structure. This page has two parameters: **Chord** and **Voicing**.



Chord

The chord parameter specifies the chord structure for the chord. The chords are listed below.

<i>7th Chords</i>	
<i>Major</i>	To get triads rather than 7 th chords mute the blue voice. Voices can be muted on the <u>Audio Levels</u> page. If the internal envelope generator is enabled with solo triggers only notes that are triggered will play. This is explained in the Gate Configuration section.
<i>Minor</i>	
<i>Diminished</i>	
<i>Half Diminished</i>	
<i>Augmented</i>	
<i>Dominant</i>	
<i>Augmented Major</i>	
<i>Minor Major</i>	

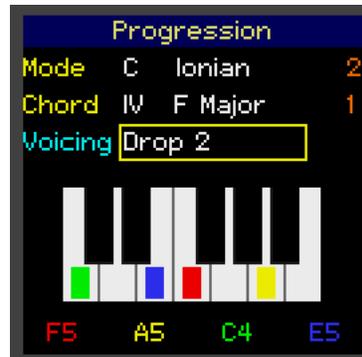
Voicing

Voicing changes the sound of a chord by raising or lowering notes within the chord by an octave. The voicing parameter options are:

1st Inversion, 2nd Inversion, 3rd Inversion, Drop 1, Drop 2, Drop 3, and Root position.

Chord Progression Mode

The Progression page selects the chord progression audio mode. A chord progression is a series of harmonically related chords played in sequence. Like 7th chord mode in progression mode the red voice channel sets the root note of the chord, the other voice channels complete the chord. If a voltage is applied to the yellow, green, or blue voice channels, the pitch of those voices are quantized to stay within the chord structure. This page has three control parameters: **Mode**, **Chord**, and **Voicing**.



Mode

These are the seven musical modes in modern usage:

Progression Modes

- Lydian
- Ionian
- Mixolydian
- Dorian
- Aeolian
- Phrygian
- Locrian
- Minor Major

Ionian mode corresponds to the major scale and Dorian the natural minor scale.

The Progression page also displays the root note of the chord modified up or down in pitch by the voltage on the red voice channel. In the graphic above, control of the **Mode** parameter has been assigned to CV 1 and the **Chord** parameter assigned to CV 2, **Voicing** is under control of the rotary encoder.

Chord

xVox follows the musical tradition of using Roman numerals to label the chords in a scale. The **Chord** parameter specifies which of the seven chords in the mode will play. The display shows the chord name for informational purposes.

Voicing

Voicing has the same options as for 7th Chord mode:

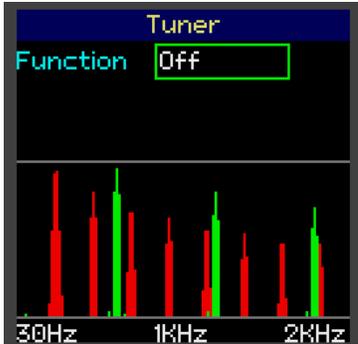
1st Inversion, 2nd Inversion, 3rd Inversion, Drop 1, Drop 2, Drop 3, and Root position.

The Tuning Function

To display accurate information about the notes xVox will output it needs to know what the input notes are. The Tuner page is one of the audio mode options. Unless configured for automatic tuning the tuner passes the input audio unchanged to the audio output.

The tuner has four methods of operation, **Off**, **Automatic**, **Manual**, and **Live**.

The **Off** function is self-explanatory, the tuner does nothing. However, it does



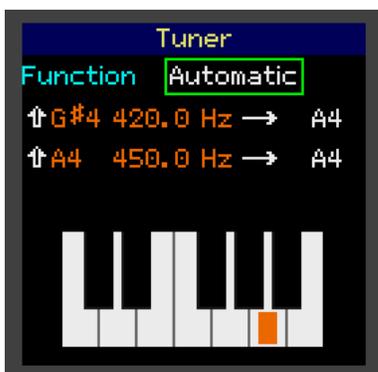
report to the audio modes that the input note was a C, so the piano keyboard and other page sections can display the relative pitch of the shifted notes.

The page displays the spectrum of the two incoming audio signals. See the Display Options section in this guide for more information about the spectral display.

The audio inputs are passed directly through to the audio outputs. The **Off** function can be useful for checking that the module is receiving the expected audio input.

When the **Automatic** tuning function is selected the incoming frequency and note closest to that frequency are displayed for each audio input.

The piano keyboard at the bottom of the page shows the target note which is also displayed to the right of the frequency display. When possible, the input audio is pitch-shifted to the selected target note.



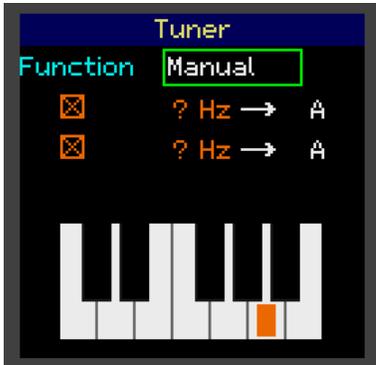
Automatic tuning is only possible when the input audio signal is harmonically stable and lasts long enough for the tuner to obtain a sample to analyze. Drones work well, but the tuner cannot tune to noise or short pings and may be unable to lock onto low frequencies and audio signals that have rapidly changing harmonics.

When automatic tuning is selected it is best to tune the input oscillator to be close to the required target note, the octave does not matter, xVox will make the minimal pitch adjustment up or down. Note that in the example the input frequencies are already close to A4 at 440Hz so require only a minor pitch shift adjustment.

Octave estimation is sometimes a problem with tuners because higher or lower harmonics of the fundamental note may appear dominant to the tuner but not the ear. xVox is typically quite good at getting the octave correct but other tuners might give different results.

Automatic tuning does assume that the input frequency will remain stable. If the oscillator is tuned to a different pitch after performing automatic tuning the Tuner page must be opened again to adjust to the new input frequency.

The **Manual** function is for tuning audio signals that for some reason cannot be

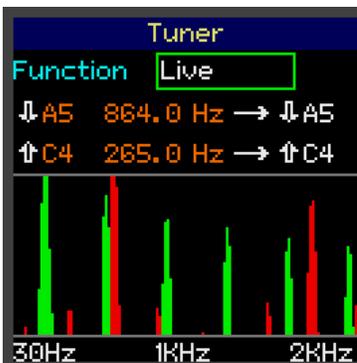


tuned automatically. The user interface is the same as for automatic tuning but there is no pitch shifting to the target frequency. In other words, the user is simply telling the tuner what the input note is.

When manually tuning, the octave is treated as unknown and octave information is not displayed. Note and frequency information are still shown. In the example screen on the left the input was a noise

signal so there is no usable frequency information for the tuner to lock onto so the notes and frequency are unknown.

The **Live** function tracks the input frequencies and displays the notes in real-



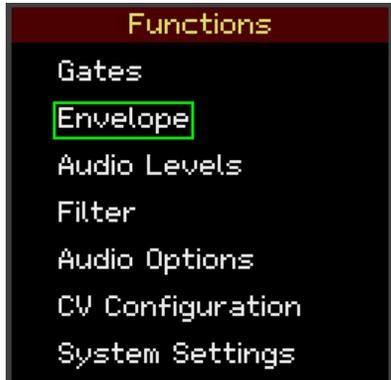
time but the tuner does perform any pitch shifting. This function is useful for cases where the input oscillator or oscillators are in tune and are being sequenced or otherwise externally controlled. In the example on the left, note that the arrows indicate if the frequency is too high or too low.

When the **Live** function is selected, rather than showing a piano keyboard the spectrum of the input signal is displayed. See the Display Options section in this guide for more information about the spectral display.

A feature unrelated to the tuning function is that, unless automatic tuning is selected, in tuning mode the audio is passed through unchanged. This means the tuner offers a quick way to listen to the dry input signal.

Function Pages

The function pages control various configurable settings and functions of the module. The chosen settings apply to all audio modes. There are six top level function pages listed on the [Functions](#) menu page and a link to the System



Settings menu page for accessing rarely used functions described in the System Settings section.

The [Gates](#) page selects how the internal envelope generator is enabled and triggered.

The [Envelope](#) page is for setting the attack, decay, and sustain parameters of the internal envelope.

The [Audio Levels](#) page sets the audio input and output gain for each of the audio inputs and outputs.

It also has controls for managing clipping, automatic gain control, and for muting selected output voices.

The [Filter](#) page gives control of high and low pass cutoff frequencies of the pitch-shifted audio.

The [CV Configuration](#) page is for assigning CV inputs to the various control functions of the audio modes, and many of the functions on the function pages listed above.

The [Display](#) page controls aspects of the user interface display and has controls for setting the brightness of the OLED display.

Click the encoder on System Settings to open the [System Settings](#) menu page.

Gate Configuration



The Gates page configures how and when the internal envelope generator or sample-and-hold (S&H) are triggered. The **Envelope** parameter has three values: **OFF** which disables the envelope generator, **Attack/Decay** to configure the envelope generator as an attack/decay envelope, and **ADSR** to configure the internal envelope generator as an attack/decay/sustain/release envelope.

The envelope generator and sample-and-hold function are mutually exclusive. However, in both cases the Volt/Octave values for a voice channel are only read when the corresponding gate is triggered. If the envelope generator and sample-and-hold are both set to **OFF** pitch changes are continuously read.

When the Envelope parameter is set to **Attack/Decay** or **ADSR** the envelope generator is enabled. The **Trigger** parameter controls how the envelope is triggered. There are four trigger options: **Unison**, **Solo**, **Pitch Change**, and **Strum**.

Unison Trigger

In this setting when a trigger or gate signal is sent to the Gate/Trigger socket jack for the red voice channel all voice envelopes are triggered simultaneously. A signal on the yellow Gate/Trigger socket jack triggers all voice envelopes except the red voice channel. A signal sent to the green Gate/Trigger socket jack triggers the envelopes for the green and blue voice channels. A signal sent to the blue Gate/Trigger input jack only triggers the envelope for the blue voice channel.

Solo Trigger

With this setting the envelope for each voice is independently triggered when a signal is sent to the Gate/Trigger socket jack for that voice channel.

Pitch Change Trigger

This setting triggers an envelope for a voice channel when a pitch change is detected on the VPO input jack for that channel. Envelopes for all voices are triggered simultaneously if there is a chord or scale change. Depending on the current audio mode this is any of the **Transpose**, **Chord**, **Scale**, or **Voicing** parameters.

Strum Trigger

The **Strum** setting triggers envelopes for each of the four voices with a delay between each trigger point set by the **Speed** parameter. A signal on any of the four Gate/Trigger inputs will start the strum. The four voices are triggered in round-robin fashion starting with the voice channel that received the trigger input signal. The strum rate ranges from about 30 milliseconds to 0.5 seconds. CV control can be assigned to the **Speed** parameter from the [CV Configuration](#) page.

The internal envelope generator will be most often used when the audio input is a continuous drone. If the input signal is already shaped, for example a ping, or plucked sound, the sample-and-hold function may be more appropriate. The **S&H** parameter has three values: **OFF**, **Solo**, and **Unison**. When both the envelope generator and sample-and-hold are **OFF** the VPO levels are sampled at about every 5.3 milliseconds.

Unison Sample and Hold

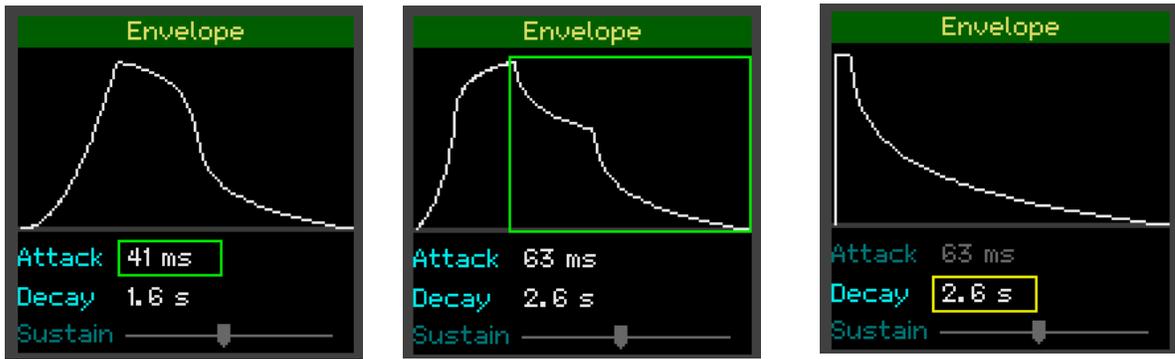
With this setting the VPO levels for all voices are sampled when there is a trigger on the red Gate/Trigger socket jack. A trigger on the yellow Gate/Trigger socket jack samples all VPO levels except the red one. A trigger on the green Gate/Trigger socket jack samples only the green and blue VPO levels, and a trigger on the blue Gate/Trigger socket jack only samples the blue VPO level. This is the same cascading behavior as for the envelope case.

Solo Sample and Hold

With this setting the VPO levels for each voice are independently sampled when a signal is sent to the Gate/Trigger socket jack for that voice channel.

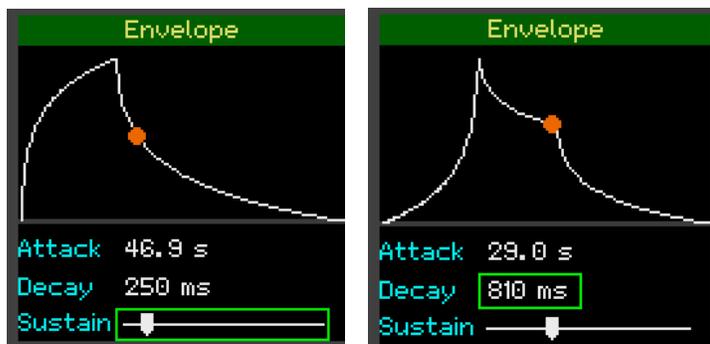
Envelope Generator

The **Envelope** page is where the shapes of the attack curve, the decay curve, the attack rise time, and the decay fall times for the internal envelope generator are set. There are seven attack curves and eight decay curves. The curves are selected using the rotary encoder.



The attack time ranges from 6 milliseconds to 2 minutes, the decay time ranges from 11 milliseconds to 2 minutes, there is also a zero-rise-time attack curve shown in the right-hand graphic.

When the **Envelope** parameter on the Gates page is set to **ADSR** the **Sustain** slider is activated and can be selected using the encoder. The **Sustain** parameter



does not set a specific sustain level, rather it sets the point on the decay curve where the output level will be held while the gate input is high. The sustain point is indicated by a round orange marker on the decay/release part of the curve.

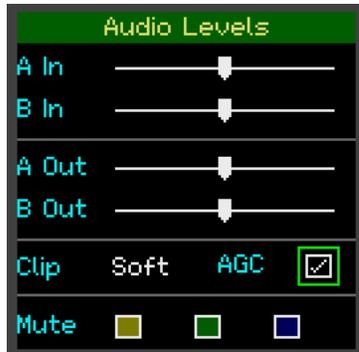
The section before the orange dot is the decay component of the ADSR curve and the section after is the release portion. In this case the **Decay** parameter sets the minimum length of the decay/release phase. The total length depends on how long the gate signal is held high.

CV inputs or the X and Y knobs can be assigned to control the attack and decay times, and the sustain point from the CV Configuration page.

The most recently configured envelope settings are restored from flash memory when the module is powered on.

Audio Levels

The Audio Levels page controls settings related to the audio input and output levels and controls how audio clipping is handled. The input level settings are mainly for boosting low level inputs or for balancing the levels of two different



audio sources operating on the analog signal before the analog to digital conversion (ADC). The control marked **A In** corresponds to the upper audio input which feeds the green and blue voices. The control marked **B In** corresponds to the lower audio input and feeds the red and yellow voices. The **A In** and **B In** levels can be configured for CV control of the input levels from the CV Configuration page.

A Out and **B Out** attenuate the digital audio signal levels prior to the digital to analog conversion (DAC). One of the less predictable effects of pitch shifting is where the pitch shifted harmonics end up. If harmonics stack up in the same frequency band above the maximum level the DAC can handle the result is clipping. The LED to the left of the audio output socket jacks indicates when the audio output is clipping. Clipping options are **Hard** or **Soft**. Hard clipping truncates the digital audio signal at the maximum value and can sound harsh. Soft clipping applies a smoothing function, in this case an inverse square root sigmoid function, that makes for a more rounded sound. However, in either case clipping is a source of distortion. To prevent clipping the output sliders can be adjusted manually or automatically using the AGC function.

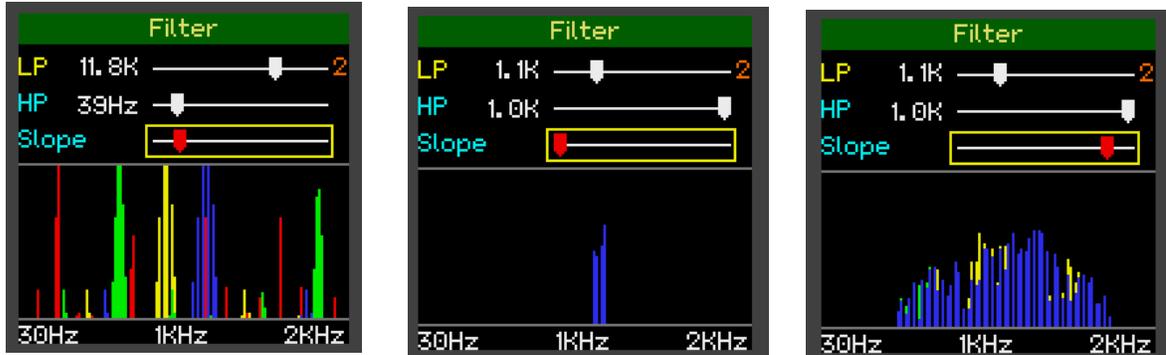
Setting the **AGC** checkbox engages an automatic gain control function that dynamically reduces the audio output levels when the module detects there is clipping. Output levels are slowly increased when there is no longer clipping.

The *Mute* checkboxes can be set to mute the yellow, green, and blue voice channels. The red voice channel is always active. Muting a channel also turns off the trigger input for that channel. For example, if the blue voice channel is muted chords will play as triads rather than 7th chords.

*Clipping is sometimes a desirable audio effect. In this case **AGC** is turned off, the **Clip** option set to **Hard**, and the input levels and output levels adjusted up to get the desired effect.*

Filter Controls

The Filter page has controls for setting low-pass (LP) and high-pass (HP) cutoff frequencies for filtering the audio input signal. The filter is applied before the pitch of the four voices are shifted.



The LP cutoff ranges from 100Hz to 24KHz, the maximum frequency for a 48KHz sample rate, obviously a lot higher than most humans can hear. The HP cutoff ranges from 0Hz to 1KHz. The user interface does not allow the cutoff frequencies to overlap. The Filter page also shows the spectrum of the pitch-shifted output after filtering has been applied. The range is deliberately limited to only show frequencies up to 2KHz or about note **B6** on a piano keyboard.

The **Slope** parameter controls how aggressively the filter rolls off as seen in the two right-hand graphics above. The input signal in both is the same white noise signal and the filter is set to only pass frequencies in the 1KHz to 1.1KHz range. It is clear how changing the **Slope** parameter affects which frequencies are passed in each case.

CV control can be assigned to all three parameters: **LP**, **HP**, and **Slope** from the CV Configuration page.

The core algorithm that does the pitch shifting in xVox uses the Fast Fourier Transform (FFT) function. In simple terms, the FFT operates on the digitized audio input signal to decompose the audio signal into a set of constituent sine waves or harmonics. At any instance in time xVox is recording the frequency and the magnitude of 1024 sine waves. Think of this as being like a graphic equalizer with 1024 sliders corresponding to frequencies across the entire audible range. When the audio signal is in this form filtering is very simple. A low-pass filter discards all sine waves above the low-pass cutoff frequency. A high-pass filter discards all sine waves below the high-pass cutoff frequency.

Audio Options

Note: The Audio Options page was new starting from firmware version 1.3.



The Audio Options page configures settings that are common across all audio modes. There are currently two options: Vocoder and Down Mix.

Future firmware releases will likely add more settings to this list.

Vocoder Option

A *vocoder*, also called a *speech vocoder*, was invented way back in 1938 at Bell Labs to synthesize human speech. This is a rather more sophisticated version of the same concept.

The vocoder option combines the audio signal from the upper and lower audio input jack sockets into a single audio signal sent to all four audio channels. In vocoder parlance the upper audio input is the carrier, the lower audio input is the modulator. The carrier input defines the harmonic content of the combined audio signal, so it works best with audio signals rich in harmonics such as sawtooth or square waves. Sending a sawtooth or square wave to the carrier input with a microphone patched into the modulator input recreates the classic “robot voice” vocoder effect. Patching two different oscillators to the two inputs generates some very interesting cross-modulation effects, particularly when the modulator signal is enveloped.

The bandwidth parameter controls the number of frequency bands for the modulator. This ranges from around 23Hz, corresponding to 1024 bands, to about 800Hz, corresponding to 28 bands. This parameter can be assigned to CV control.

All the pitch shift modes continue to function as before generating chords, and chord progression from the cross-modulated audio input.

For the vocoder to work there must be an audio signal patched into both audio input sockets. If there is no input patched into the upper socket (or both inputs are patched to the same source) and the bandwidth slider is all the way to the

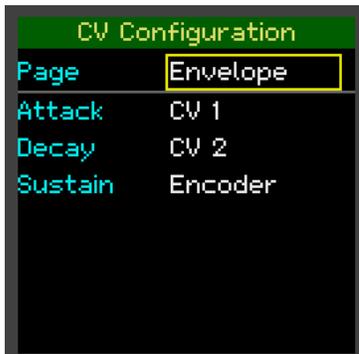
left, there is no cross-modulation and no vocoder effect. If the bandwidth slider is not all the way to the left, the input will self-modulate but the effect can be quite subtle.

Down Mix Option

The down-mix option is for convenience when separate audio output for the red/yellow and green/blue channels is not required. This option additively mixes the two pitch-shifted audio signals and sends the mixed output to both audio output jack sockets.

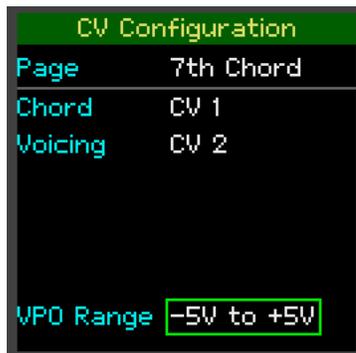
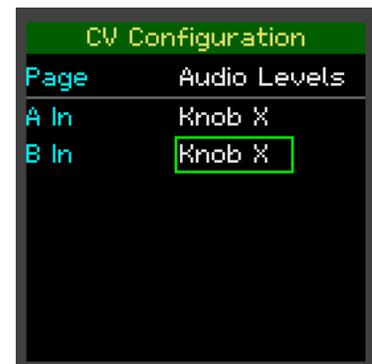
CV Configuration

There are currently fifteen parameters on the various audio and function pages that can be controlled by CV input. CV assignments are set from the CV Configuration page. The audio mode pages CV assignments only apply to a specific audio mode and are only active when that audio mode is selected. The function page CV assignments apply to all audio modes.



Using the rotary encoder, the **Page** control cycles through the list of pages that have CV controllable parameters. As each page name is displayed the parameters for that page and the current assignments are shown. Parameters that are not under CV control show as controlled by the **Encoder**. To assign a specific CV input of one of the two control knobs, select the parameter and use the encoder to assign the required CV input or control knob.

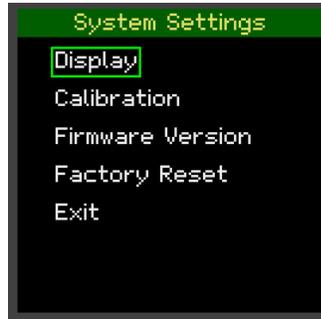
Note that the same control knob or CV input can be assigned to multiple parameters as shown on the right where audio input parameters **A In** and **B In** are both under control of **Knob X**.



The audio mode pages have one extra parameter, the **VPO Range** parameter is a common setting for all audio modes. Many sequencers and other controllers only output voltages in the range 0 to 5 Volts. The **VPO Range** setting can be set to either **-5V to +5V** or **0V to +5V**. When the range is set to 0V to 5V the VPO inputs are all *normalised* to 3V and the pitch shift range is reduced from -5/+5 octaves to -3/+2 octaves.

Because oscillators have frequency control knobs that can apply a negative voltage offset to the VPO input it is often not obvious that a sequencer is only emitting voltages in the 0 Volt to 5 Volt range. If the pitch shifted audio output from xVox seems to be all high notes, try selecting the 0V to + 5V setting.

System Settings

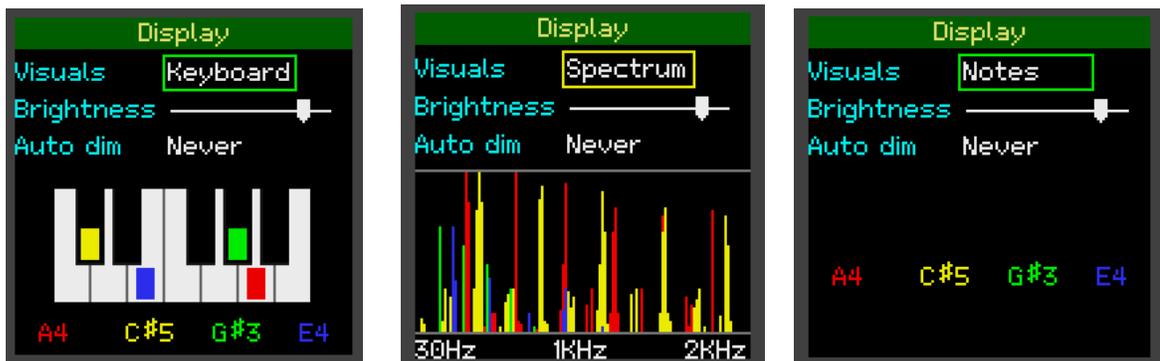


Select System Settings from the Functions index page to display the menu page for less frequently used system-level functions. These are **Display**, **Calibration**, **Firmware Version**, and **Factory Reset**. Selecting the **Exit** option or clicking the **Func** button returns to the **Functions** page.

Display

Note: The Display page was moved from the Functions menu page to the System Settings menu page in version 1.3 of the firmware.

The Display page configures how information is displayed on the other pages as well as setting the display brightness and dimming options. The **Visuals** parameter controls how the pitch shift information is displayed on the audio mode pages. There are four options: **Keyboard**, **Spectrum**, **Notes**, and **None**.



The **Keyboard** option displays a piano keyboard with the active notes shown in the colors corresponding to the VPO sockets that control the pitch for that voice channel. If the octaves are known the octave number is displayed along with the note being played. This assumes that the input notes have been set using the Tuner page. Voices that are muted are shown dimmed. If the internal envelope generator being used a note is only highlighted on the keyboard when the note is

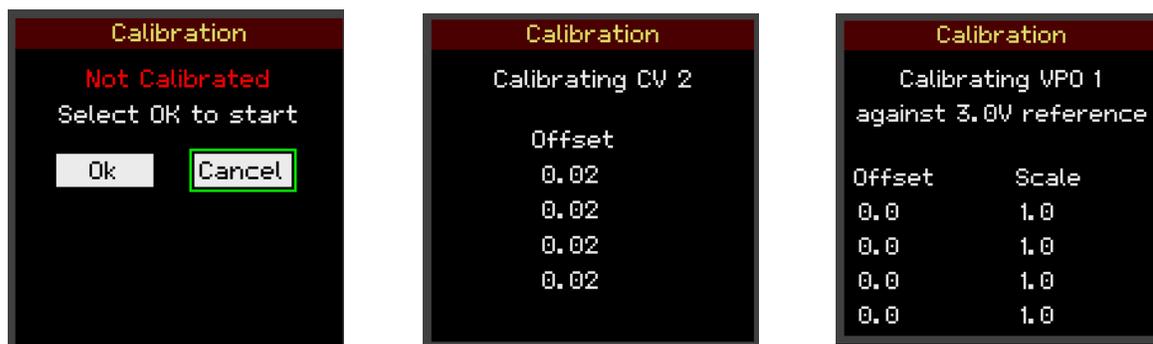
being triggered. The **Notes** option just displays the note names without the piano keyboard.

The **Spectrum** option shows the harmonics present in each of the pitch shifted notes, also colored for the corresponding voice channel. The spectrum shows the actual audio signal so will only show notes that are currently triggered.

The **None** option does not display any note or pitch information.

The **Brightness** slider controls the brightness of the OLED display and the push buttons. The **Auto dim** setting will dim the OLED display after a set period during which there has been no user interaction with the encoder knob or the push buttons. Rotating the encoder knob or pressing a push button restores the display to the currently set brightness.

Calibration



The calibration process ensures that the four VPO inputs are volt-per-octave accurate and have the same relative pitch, and the CV inputs are centered at 0 Volts. If the module is not calibrated this is shown in the calibration page and also on the startup splash screen. Calibration is required after a factory reset and usually after installing a major new firmware release, but recalibration can be done at any time. If you have built the xVox module from a kit you should perform calibration before using the module. It is preferable to wait a while for the module to warm up to normal operating temperatures before performing calibration.

With no patch cables in any of the socket jacks, open the Calibration page and use the rotary encoder to select and click on the **Ok** button. The four CV inputs are calibrated first, followed by the VPO inputs. xVox has its own internal reference voltage source, so calibration is a completely automated process. Because of small differences in the values of the components, mainly resistor values, you can expect to see small offsets from 0.0 and scale values close to but

unlikely to be exactly 1.0. The new calibration values are stored to flash memory. An error message is displayed if the module detects that patch cables are inserted when attempting to perform calibration.

Pressing the Func button before calibration completes will cancel the calibration before the new calibration values are stored.

Firmware Version

The Firmware Version page just displays the current firmware version number and build date. See the **Firmware Update** section below for instructions on how to update the firmware.



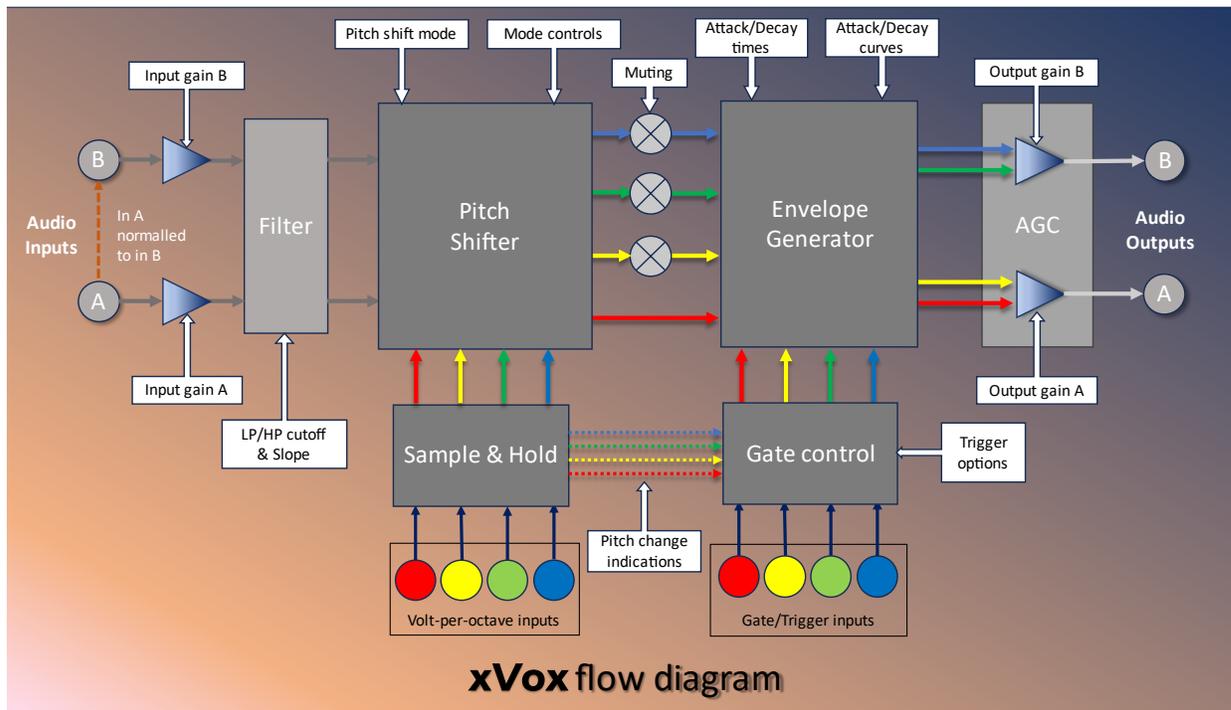
Factory Reset



The Factory Reset page is used to restore all configurations settings to their original values. This includes resetting all the CV assignments, setting the CV and VPO calibration offsets to 0.0 and scale factors to 1.0. This is probably something most users will rarely need to do. To do a factory reset, open the Factory Reset page and use the rotary encoder to select and click on the **Ok** button.

Flow diagram

The flow diagram below shows the main functional blocks described in the sections above.



The audio gain settings are applied to the analog inputs A and B before analog to digital conversion. All the functional blocks up to the final audio output are performed in software.

The filter, which has high-pass and low-pass cutoffs controls, is applied before the pitch shifter. The behavior of the pitch shifter depends on the selected pitch shift mode, the volt-per-octave inputs, and the settings of the mode control parameters.

The envelope generator is bypassed when envelopes are disabled, otherwise envelopes are triggered by the gate control block. The trigger options control whether envelopes are triggered by the inputs or by pitch changes. All four voice channels have the same envelope shape specified by the attack time, decay time, attack curve, and decay curve parameters.

The red and yellow channels are digitally mixed and converted to analog output A. The green and blue channels are digitally mixed and converted to analog output B. If the automatic gain control is enabled the audio signals are attenuated before the DAC.

Firmware Update

To update the xVox firmware, log on and download the firmware file from the download page on www.gregsbrain.co The name of the file will be in the form **xVox-M-R.bin** where the **M** is the major version number and **R** is the minor revision number. Copy the file to the top-level folder of a USB stick. For simplicity of operation the bootloader just looks for a file ending in **.bin** so make sure there are no other files in the top-level folder with that suffix.

With the module powered off, insert the USB stick into the USB A socket on the rear of the module and power the module on. The new firmware should automatically load and display the xVox splash screen. You can check on the [Firmware Version](#) page to verify that the new firmware version is listed. Power down the module and remove the USB stick. If the firmware was not updated as expected check the file **daisy_boot_log** written to the USB stick which will indicate if there was an error installing the firmware. If the log indicates that the file was not a valid binary file, check you don't have an encryption option set for external USB drives. The bootloader cannot load an encrypted file.

Note that the USB-A socket should only be used for firmware updates and should only be used with a USB thumb drive type device. It should never be used as a USB power source for other devices.

Release Change History

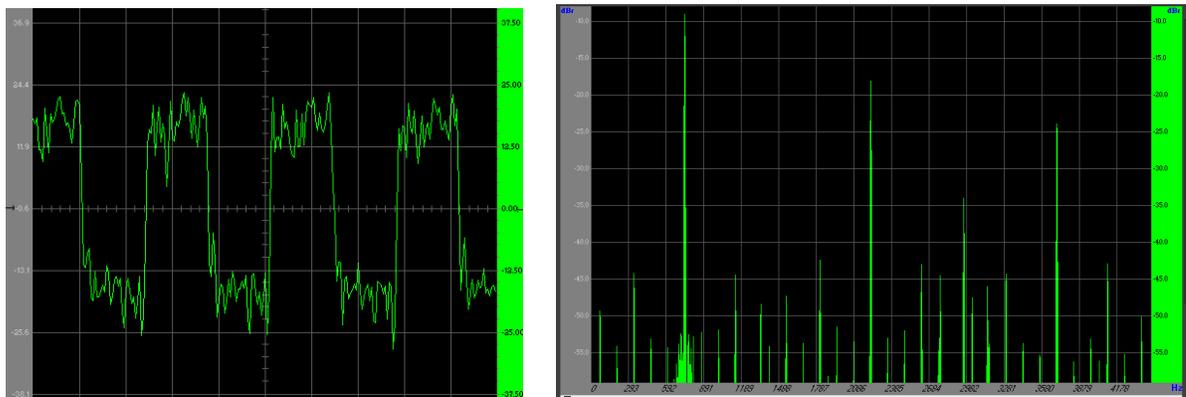
Change History

Version	Changes
1.0	Initial firmware release
1.1	Added a micro-tonal (semitone per octave) option to pitch shift mode.
1.2	Minor bug fixes
1.3	Added the <i>audio options</i> page with two new features: <ul style="list-style-type: none">• Vocoder effect• Down mix
1.4	Fixes bug where encoder was still actively controlling a parameter was under CV control. Adds a bandwidth parameter to the Vocoder effect.
1.5	Improve how the quantizer behaves when the VPO input is between notes.
1.6	Restructured the internal code in support of the VCV Rack plugin. No new features.
1.7	Bug fixes for the VCV Rack plugin to fix two crashes. Add acceleration to the slider widget. Improves response to short triggers on the hardware module.

Implementation Notes

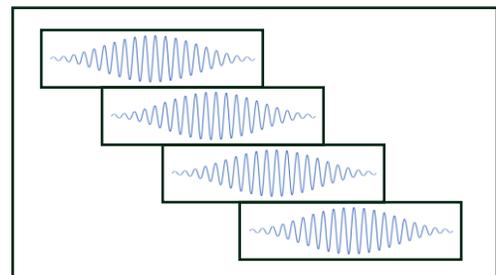
These notes are not intended to be an in-depth explanation of FFT, STFT, or pitch shifting in general, but some basic concepts help to set expectations of what xVox can and cannot do.

Pitch shifters operate in either the time domain or the frequency domain. The time domain is about changes in the audio signal over a span of time, basically what you see on an oscilloscope screen. The frequency domain is about the magnitude of the frequencies, or harmonics, that contribute to the audio signal. The graphics below show the same signal in the time domain on the left and frequency domain on the right.



The audio inputs and outputs of a digital module like xVox come in and leave as data in the time domain; a stream of values that represent the audio voltage level at short instant in time. In xVox that short instance in time translates to 48,000 values, or samples, per second.

Frequency domain pitch shifters require more computational power than time domain pitch shifters but generally perform better over a broader range of audio sources and pitch shift range. The xVox pitch shifter is based on the *Phase Vocoder* (Not to be confused with the vocoder option described on the [Audio Options](#) page above) implemented by the *Short-Term Fourier Transform* (STFT). The Fourier Transform (FT) being the mathematical transform function that converts a time domain signal into its constituent frequency components. The STFT operates on a series of overlapping blocks of audio data. The overlap is required for good audio quality. The minimum overlap for acceptable quality is 4, xVox uses an overlap factor of 8. The audio blocks are



2048 samples in length, the 8x overlap factor means that Fourier transforms are being done for every 256 new audio samples. xVox has two audio inputs and four audio outputs. At the xVox sample rate of 48KHz this means there are six Fourier transforms every 5.33 milliseconds, two forward transforms and four inverse transforms.

In the frequency domain the audio data is expressed as a series of sine waves of increasing frequency from DC up to a maximum frequency called the Nyquist frequency that is half the audio sample rate, so in this case 24KHz. A Fourier transform of 2048 samples generates 1024 sine waves, each with an amplitude and phase that depends on the audio signal that was input. These 1024 values, commonly referred to as *bins*, are each 23.4Hz wide. Many bins will have an amplitude of zero or very near zero (an empty bin) meaning that input audio signal did not include that frequency or harmonic.

Pitch shifting is a fairly simple process. To increase the pitch bin values are moved towards the last (Nyquist) bin. To decrease the pitch bin values are moved towards the first (DC) bin. However, the bins are not simply slid up or down, the bin movements scale according to the pitch change. To increase the pitch by an octave the bins positions are doubled, so bin 1 goes to bin 2, bin 2 goes to bin 4, bin 4 to bin 8 etc.

The Phase Vocoder (PV) is probably the most widely used process for pitch-shifting audio. The PV is good with a range of different audio inputs but has some well documented limitations. Some audio signals can be pitch shifted with good fidelity, others not so well. Audio signals with a fast attack like a kick drum tend to become smeared, softening the attack. Other audio signals develop a “phasiness”. Successful commercial products based on the PV use proprietary algorithms to improve the audio quality, especially for voice. The xVox PV is a standard implementation so there will be some artifacts. Within the context of Eurorack synthesizers these are not necessarily a bad thing.

VCV Rack Addendum

The VCV Rack implementation uses the same source code as the physical module so the xVox VCV Rack Plugin should behave the same as the hardware version but there are some minor differences.

In the plugin implementation there is no need to calibrate the volt-per-octave inputs, so the Calibration entry on the Systems menu page is greyed out.

The screen dimming function is present in the UI Config page but has no effect in the plugin implementation.

Using the rotary encoder knob to navigate the user interface was a little awkward so the plugin captures the mouse scroll button when the mouse is hovering over the encoder graphic. The mouse scroll button is also captured, as controls the encoder, when hovering over a flashing Mode or Func button. Normally the mouse scroll button scrolls the rack up and down, so capturing the scroll button in the plugin is non-standard behavior. There is an option in the plugin context menu to turn off scroll button capture if this non-standard behavior is annoying.