Waves Neural Networks

Clarity Vx Pro

Real-Time Noise Reduction for Voice

Powered by Waves Neural Networks®

User Guide

Contents

Introduction

Thank you for choosing Waves! To get the most out of your new Waves plugin, please take a moment to read this user guide. Installing software and managing your licenses require a free Waves account. Sign up at www.waves.com. With a Waves account you can keep track of your products, renew your Waves Update Plan, participate in bonus programs, and keep up to date with other important information.

We suggest that you become familiar with the Waves Support pages, www.waves.com/support, where you will find technical articles about installation, troubleshooting, specifications, contact information, and more.

Neural Networks

A neural network is, in essence, a learning machine. It uses algorithms inspired by the brain to store information and make choices. Before a neural network can do anything useful, such as distinguishing ambience from voice and acting on that information, it must analyze millions of relevant samples. As it learns more and more from these diverse samples, the neural network creates a set of associations. When an input signal is sent through the network, it is analyzed, and a set of weights and biases are created. This is the information that is used to control processing. The more samples it learns from, the better its decision-making skills. The choice of material used to teach neural network greatly influences its decisions.

Waves Clarity Vx Pro

Clarity Vx Pro uses a deep neural network to adaptively isolate even the most complicated background sounds and separate them from the voice. It can give you voice without ambience, ambience without voice, or anything in between, and provide powerful background noise reduction with almost no impact on the voice. For basic use, a single knob controls the mix of the fully separated voice and the ambience.

This is often all you need. Advanced controls can be used to adjust the process amount and gain in four adjustable bands. You can select how source material is analyzed, adjust how unvoiced regions are treated, hear the delta (DIFF), and control the stereo image of the output.

Clarity Vx Pro has mono and stereo components. It can be used in real time as a channel insert or as an offline processor.

Quick Start

This quick start will teach you how to isolate ambience from voice sound with minimum impact on either. Follow these steps and you'll achieve a high-quality separation of voice from ambience. Refer to the rest of this user guide to learn more about Clarity Vx Pro controls and how to use them for best results.

Understanding the Mix

The key to superior results lies in controlling the mix between the input signal and the output of the neural network. There are three controls that govern this mix; it is important to understand how each control works and how they interact.

Main Control Knob (Ambience/Voice)

The Main Control knob sets the range of how much ambience will be separated from the voice. Let's say you are using Clarity Vx Pro to remove ambience and you set the Main Control Knob to 80%. No matter how the other controls are set, the voice processing mix will never exceed 80%. For maximum flexibility, we suggest that you initially set this control to 100%.

Band Process faders

Each band in the FFT display has a draggable fader bar. This lets you control how the signal is processed by the neural network, per band. It lets you *reduce* processing per band, not *add* it. At the topmost setting, you are hearing 100% of the mix of the Main Control Knob in that band. As you lower the fader, you hear increasingly more of that band's ambience (i.e., you are moving back toward the bypass/input signal).

Process Amount

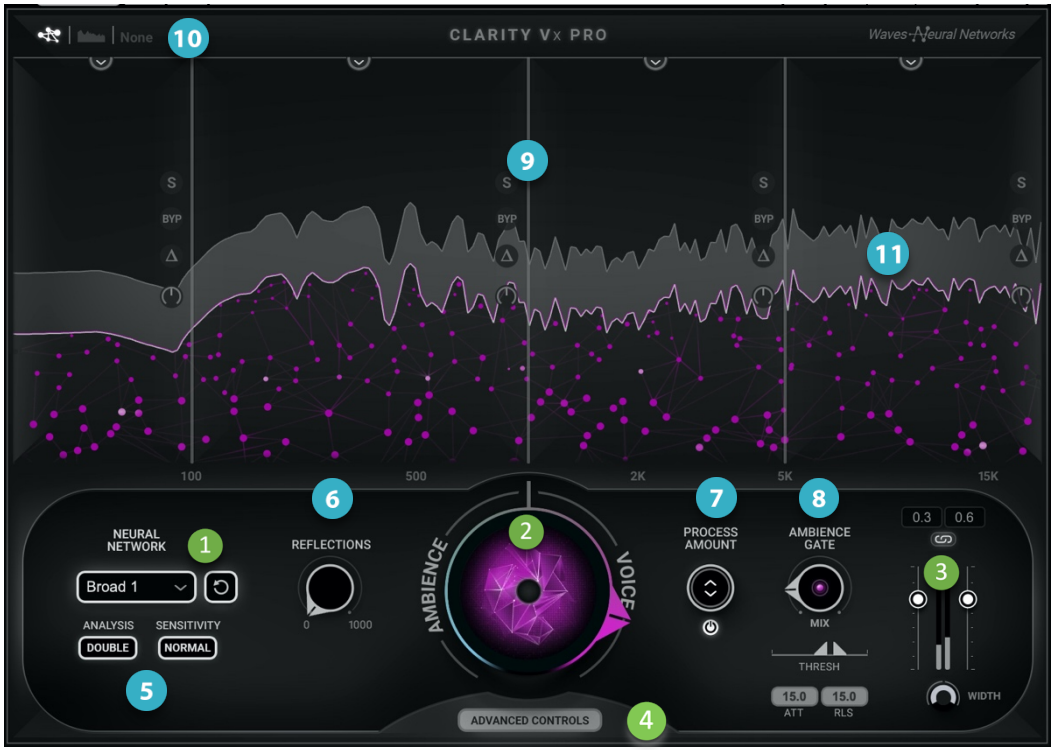
This is an intelligent VCA that links all the Band Process faders so that you can move them together. This lets you set up your band-by-band processing and then move them as a group.

Bottom Line

The individual Band Process faders, which are linked together by the Process Amount control, determine the amount of processing applied to each band. The Main Control knob is a mix between the input signal and the neural network output, as defined by the other controls.

Interface

(Basic and Advanced controls shown.)



Basic Controls

1. Neural Network Selector and Reset
2. Main Control: voice vs. ambience
3. Output Section
4. Advanced Controls On/Off

Advanced Controls

5. Input Signal settings: Analysis and Sensitivity
6. Reflections: restores natural reflections
7. Process Amount Master: controls all bands together
8. Ambience Gate (amount of ambience left in unvoiced parts)
   - Mix, Threshold, Attack, and Release
9. Graphic FFT Display
10. Graphic FFT View Options
11. Input, Output, and Reduction graph

Tooltips

Hover over a control for a brief description of its function. The setting of the control is shown in a nearby value box. Turn tooltips on or off in in the WaveSystem Toolbar drop-down menu, next to the Save button.

Step-by-Step

Basic Controls

1. Start with the **Main Control** knob. It sets the range of processing. Turn the knob to the right for more voice, turn it to the left for more ambience. For basic processing, find a setting that isolates the desired amount of voice or ambience. Some Advanced controls are limited by this setting, so if you plan to use the Advanced controls for band-by-band adjustments, we suggest that you begin with the maximum Main Control knob setting of 100.
2. There are currently three neural networks. Each has a different effect on the balance of the voice signal. Use the **Neural Network** menu to select one.
   The neural networks used by Clarity Vx Pro are each trained differently. Its training determines what a neural network is best suited for. This could be isolating voice from an especially harsh ambience or focusing on a foreground voice while eliminating other voices (or keeping them), or many other instances. Use the tooltip descriptions on the Neural Network menu to find the neural network that best suits your needs. Try more than one neural network since the results are content-dependent.
3. Consult the Meter to set the output levels using the output **Faders** and adjust the size of the output stereo image with the **Width** control.

Advanced Controls

Use the Advanced Controls to adjust the processing amount and mix across four bands, enhance the sidechain for improved analysis, and control the post-processing reflections. The Advanced Controls view/hide button is beneath the Main Control knob.

Advanced Controls—Graphic Section

4. Focus first on the four Band Process faders, the horizontal lines initially at the top of the graphic display. When a Band Process fader is at its highest setting, you are hearing the sound of the Main Control knob, at that frequency band. Drag a fader downward to offset the processing. The lower you drag the line, the less processing you'll hear in that band. Note that these band settings are dependent on the Main Control settings; a fader can reduce processing in a band, but not increase it. If you are not achieving enough ambience reduction with the Band Process faders, increase the value of the Main Control knob.

5. Drag an adjustable **Crossover** left or right to adjust the range of a band—find the best isolation without cutting into delicate parts of the voice. You can **Solo** or **Bypass** one or more bands, and you can monitor the difference between the input and processed signals with the **Delta** control.
6. Once you're satisfied with the wet/dry process amount of a band, use the **Pure Gain** control to raise or lower the level of the band with respect to the other bands.

Advanced Controls—Bottom Panel

7. Drag **Process Amount Master** to group the faders of all four bands and move them together. This lets you adjust each band and then move them as a group while maintaining the ratios between them.
8. **Analysis and Sensitivity**. Analysis controls how a stereo signal is analyzed before neural network processing: either two separate stereo tracks or a sum of the two. Sensitivity dynamically changes the input signal-to-noise ratio for improved analysis under certain circumstances. These settings do not change the audio signal, they only change the sidechain going into the neural network.
9. Use the **Reflections** control to change the release time of the processor. This restores the natural reflections of the input signal, which may have been reduced during processing. Higher settings will restore more reflections, but may also bring back some noise during the reflections tail if the input signal has loud ambience.
10. The **Ambience Gate** allows you keep some ambience in the voiced section but attenuate or completely remove it during unvoiced sections. You cannot use this control to *increase* the level of ambience in unvoiced sections.
    - **Mix** controls the amount of ambiance; the higher the setting, the more ambience is *removed* between voices.
    - The **Threshold** sliders set the level at which ambience reduction will begin and end.
    - **Attack** sets the residual ambience reduction speed.
    - **Release** sets the residual ambience recovery speed.

All Clarity Vx Pro processing is carried out on your computer. It is not sent to a cloud for processing, so all operations can be carried out without internet connection.

Controls

This section describes all Clarity Vx Pro controls, so you'll need to be in Advanced Controls view to follow along. The controls are discussed in two parts: the Bottom Panel and the Graphic Display.

Bottom Panel

Before a signal can be separated into voice and ambience, it first must be analyzed. The Analysis and Sensitivity controls are used to optimize the input signal so that it is best suited for neural network processing. These controls do not affect the actual signal, just the detection. Compare this to a sidechain, which can equalize, compress, and gate the signal to provide better information to the dynamics processor—without changing the input signal.

Analysis Mode (stereo component only)

In the stereo component, left and right channels can be analyzed independently to create a unique set of weights and biases for each channel, or they can be summed together for a single analysis. Weights and biases provide instructions for neural network processing; two analyses mean that the neural network runs two processes.

- **Single** sums the left and right input channels before analysis. If the ambience and voice are similar on both channels, you don't need to analyze or process the input signal twice. In Single mode, both channels are processed using the same analysis data. Single mode consumes considerably less CPU than Double.
- **Double** Use this mode when the left and right channels are significantly different (i.e., very different ambience or voice). Double mode analyses the left and right channels independently and processes them separately. This results in more precise processing for each channel, but it requires substantially more CPU. You may need to use the Width control to further adjust the stereo image.

Note that in some circumstances, Single will not yield the same degree of voice/ambience separation as Double. When in doubt, experiment with both since this is very content-dependent.

Sensitivity

Sensitivity changes the signal-to-noise ratio of the input signal to aid analysis. Normal sensitivity is usually the appropriate choice. But in cases where the signal-to-noise ratio is very low, try the High setting. This enhances the sidechain dynamics, which can improve the separation of certain noises and ambiences, especially high-end hiss and similar High-Frequency profile ambiences.

> **!** Remember, the Analysis and Sensitivity settings do not affect the input and output signals. They affect *only* the analysis. In a stereo signal, both left and right signals are always processed, and the sound image is not altered.

REFLECTIONS

Reflections controls the release times of the neural network process, after analysis. Low settings can yield a sound that is very accurate, but the voice's space and natural reflections may be reduced or cut into. A high Reflections setting will maintain the natural tail of the original signal, but it may bring back some of the original ambience. Settings above 150 ms can result in a sound that is pleasing, but there's risk of smearing, like a plate reverb. Appropriate settings typically lie between 30 ms and 200 ms

Range: 0 ms to 1000 ms

Neural Network Selector

There are currently three Clarity Vx Pro neural networks. They differ in the statistics, signal-to-noise ratios (which were introduced to it during training), and weights and biases (which they apply during processing). Choose a network based on the input source, as well as the voice focus that you are seeking.

Switching between neural networks is very quick (approximately 150 ms). In the real-time component, the network name will be grayed out while a network is loading. When running the plugin in offline processing, networks will be grayed out, even though they are ready to be used.

Broad 1	Isolate primary and secondary voices from ambience. This preserves the main voice, as well as some secondary voices, and isolates them from the ambience.
Broad 1 HF	Separate primary and secondary voices from ambience. This is similar to Broad 1, but with a greater high-frequency focus. This is not a shelf, but rather, it reflects different training biases.
Broad 2	Isolate the primary voice from secondary voices and ambience. This choice focuses on the main speaker/singer, separating it from the ambience and the secondary sounds (e.g., background voices, noises such as vacuum cleaners, traffic, or construction sites).

Reset

Clarity Vx Pro features deep neural networks that are capable of adaptive and continuous processing. As a neural network learns, it accumulates a “history” that influences its future decisions. When playing a region in a loop, the neural network adapts, improving its performance with each pass. This can result in progressively better results with each loop cycle. If you encounter a particularly problematic section of your program material, we suggest that you create an uninterrupted loop and record several passes (if your DAW permits). It is likely that after two or three passes you will achieve optimal results.

On rare occasions, playing a loop or moving to a new start point will cause degradation. If this occurs, click Reset to clear the neural network history and start anew. It is best to reset the neural network during an unvoiced region or silence. This can be automated. Clarity Vx Pro automatically resets each time playback stops. It does not affect your settings.

Main Control Knob: Voice vs. Ambience

This sets the mix between voice and ambiance, and therefore the overall amount of processing.

As you turn the knob clockwise, you will gradually hear less ambience, and the control will become more purple. At the highest setting, only voice is heard. Turn in the other direction to gradually hear less voice; the knob becomes increasingly blue and more intense. These colors are reflected in the graphic display to illustrate the voice/ambience balance.

The Main Control knob is after the Band Process faders in Clarity Vx Pro signal flow, so once you establish a good mix in the graphic display, you can use this knob to refine the overall mix. Hover over or touch the control knob to display its current value.

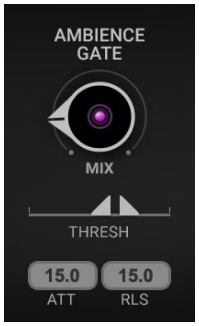
Range: -100 to +100

Process Amount

This links all the individual Band Process faders in the Graphic display. When the Process Amount control is moved up or down, the process amounts for all the bands will move together, maintaining the ratios between them. The controls converge as they approach 0 or 100 and will return to their previous positions as the Process Amount control moves them back to the center. You never lose your starting point.

Click the button below to disable all Band Process faders and revert to the Main Control’s settings. This control cannot be automated, but you can automate each of the four Band Process faders. Band-based mixing is discussed in the next section. Range: 0 to 100

Ambience Gate Section



This section controls how much ambience is left within the gaps between the voiced parts. This affects the signal *after the Main Control knob,* so ambience between voiced sections can be reduced, but not increased.

> The Main Control knob must be set to 0 or above *and* be less than 100 for the Ambience Gate section to be active.

The Ambience Gate section consists of several controls:

Ambience Gate Mix sets the amount of ambience attenuation within non-voiced areas, so as you increase the setting, more ambience is removed. Range: 0–100

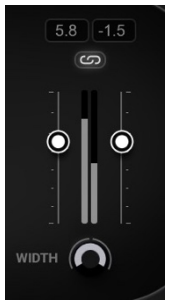
The **Threshold Open** and **Threshold Close** markers set the levels between which ambience attenuation will occur during unvoiced sections. Range: -90 dB to 0 dB

Attack sets the residual ambience reduction speed (in milliseconds). This is the speed at which the ambience will be reduced during unvoiced sections.

Release sets the residual ambience recovery speed (in milliseconds). This is the speed at which the ambience will return at the end of unvoiced sections.

Attack and Release Ranges: 0 ms to 500 ms

Output Section



Full-scale meters (stereo or mono) indicate the plugin output. Peak hold values are shown above the meters. A red clip LED indicates a level of 0 dBFS. Click on the meter to clear peak hold and clip indictors.

Faders can move separately or together in the stereo component. Click the Link button above the meter to link them. If there is an offset between the two faders, it will be maintained when linked faders are moved.

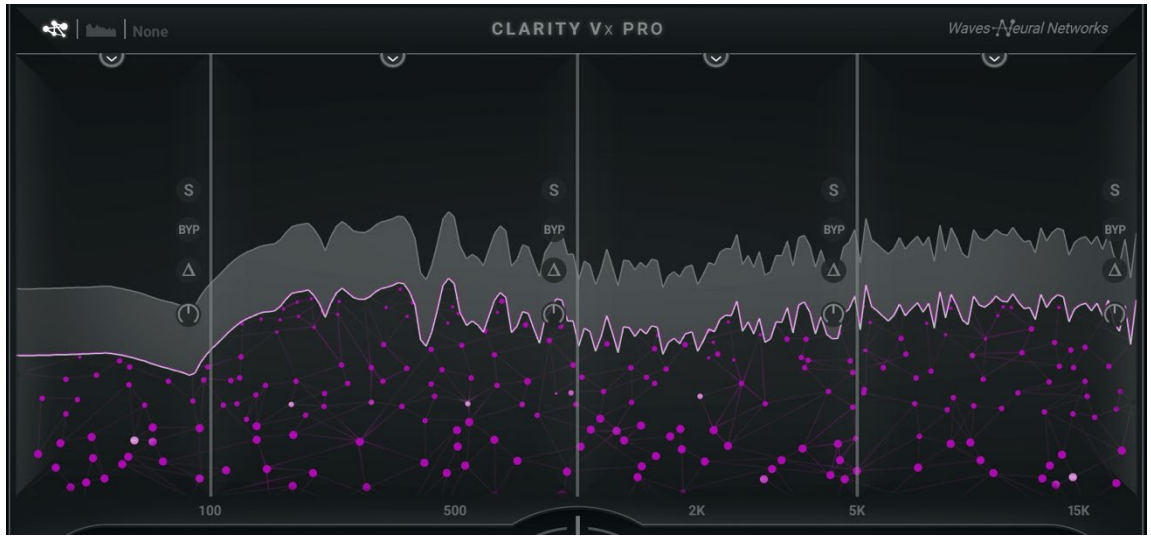
Range: -inf to +24 dB

Use the **Width** control to adjust the size of the stereo image.

Range: 100 (original stereo width) to 0 (mono)

The Show Advanced Controls button controls only the display—settings and controls are not affected. If you have changed any of the Advanced controls from their default positions, and then switched the Advanced button to Off, you will see an asterisk. This indicates that there has been a change in the “hidden” Advanced controls.

Graphic Display



The graphic FFT display shows the Clarity Vx Pro input and output levels, as well as the difference between them (reduction). The display is divided into four frequency bands whose ranges are set with draggable crossovers. Each band has a Band Process faders that is used to offset the processing of Main Control knob processing. As you drag a band’s process control downward, you reduce the neural network’s process strength for that band and bring back some of the ambience.

There are three ways to view the graphic display. Select a view with the buttons at the top of the window.

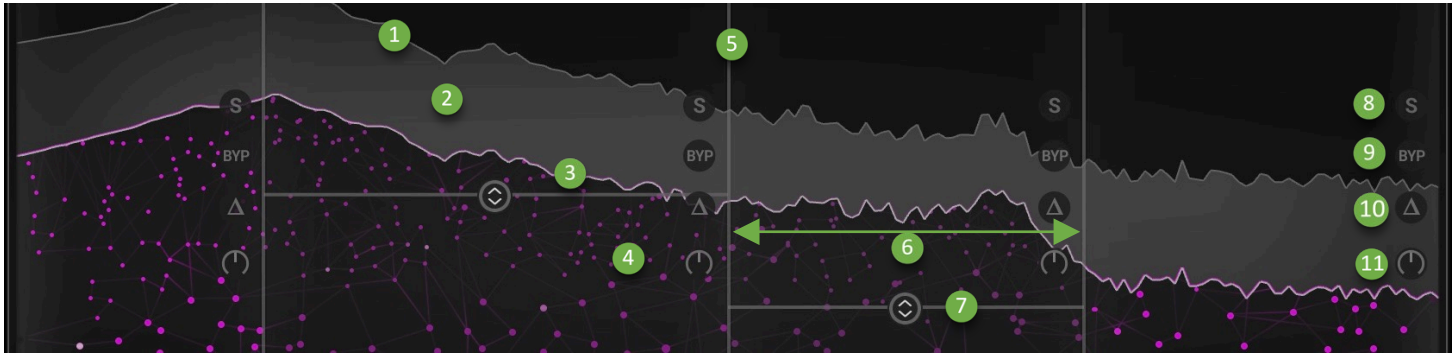
Neural Network FFT shows input (white outline on top), reduction (grey) and output signals (colorful). The intensity of the background indicates the mix level for each band, ranging from the most colorful (all neural network output) to grayed out. A grayed-out bottom area may indicate a zero setting on the Band Process fader or band bypass. This is the most CPU-heavy view.

Traditional FFT is the Neural Network FFT view, but without the sparkly bottom section.

FFT Off turns off all FFT graphing. All band mix controls remain active.

Controlling the process amounts per band

Using the Main Control Knob to set the *Voice vs. Ambience* mix will usually provide excellent noise reduction without impacting the voice. To enable greater control, use the four-band process controls in the Graphic display to adjust processing amounts in defined frequency bands.



1. Input signal
2. Difference between the input signal and the output signal after processing (i.e., reduction meter)
3. Output signal after processing.
4. Area that illustrates the overall output signal
5. Adjustable crossover
6. Frequency band
7. Band Process fader
8. Band solo
9. Band bypass (revert to input signal for this band)
10. Delta (hear only the reduction for this band)
11. Pure Gain control, per band

Band Process Controls

Crossovers

There are three crossovers. Drag a crossover to adjust the boundaries of two adjacent mix bands. It can be helpful to solo a mix band while adjusting the crossover. The crossover frequency is displayed when the crossover bar is clicked or moved.

Default crossover values are 100 Hz, 1 kHz and 5 kHz

Band Solo, Bypass, and Delta

Each band can be soloed or bypassed so that you can better understand how each band is contributing to the mix.

The **Solo** button mutes all other bands. Solo and Bypass buttons are latching, so several bands can be soloed and/or bypassed at the same time. When a band is soloed, the other bands will dim but their FFT displays and other controls will still be active.

Bypass bypasses the band processing. This lets you compare the current band mix setting with the input signal.

Delta lets you hear only the input signal minus output signal (reduction) of the respective band.

Pure Gain

Use the band **Pure Gain** control to change the gain of the band, post-neural-network processing. The process amounts do not change when gain is applied; only the level of the band with respect to the other bands is changed.

> Pure Gain between bands has a Q of infinity. This means that there is no slope and no bleed between bands. This enables surgical precision, without the need for smoothing or the fear of clicks or sweeps.

WaveSystem Toolbar

Use the bar at the top of the plugin to save and load presets, compare settings, undo and redo steps, and resize the plugin. To learn more, click the icon at the upper-right corner of the window and open the WaveSystem Guide.

Full Reset

To reset Clarity Vx Pro to factory default settings, open the WaveSystem presets menu and select “Full Reset.”

Light Mode

The Clarity Vx Pro display is, by default, dark. Click the “Light Mode” button on the left side of the WaveSystem Toolbar to display a light version of the display.

CPU

To avoid running out of CPU in your DAW, don’t open more instances of Clarity Vx Pro than you need. Please also note that Double analysis mode uses substantially more CPU than the Single mode. Therefore, if you are using several instances of Clarity in Single mode and you begin switching them to Double, you may encounter CPU spikes. All of this depends on your host computer, as well as the other plugins that you are running.

Unsupported Hosts

Clarity Vx Pro currently does not support Native ARM (M1) processors. It can, however, be run under Rosetta, which enables a Mac with Apple silicon to use apps built for a Mac with an Intel processor.