BSS AUDIO application

How to use Acoustic Echo Cancellation [AEC]

Soundweb™ London

What is acoustic echo and why do I need to cancel it? Acoustic echo occurs in a conferencing system when the far-side speech played in the loudspeakers is picked up by microphones in the room and is transmitted back to the far side. This transmitted signal is a delayed version of the original, which causes the echo.

The received far-side signal does not transfer directly from the speaker to the microphone, but is subject to the artifacts of the room. This may include differing signal paths causing reverb, frequency filtering and attenuation. These effects are the transfer function of the room. The transfer function of the room is also dynamic, as objects in the room move or the microphone moves position.

To correctly subtract the required signal, the AEC therefore needs to simulate the dynamic room transfer function. It can then apply that transfer function to the received signal and correctly subtract the modified original signal.
Each Soundweb London AEC card consists of 4 AEC input channels.

Each channel offers the following features:

- Independent 20Hz - 8kHz algorithm
- Individual AEC references
- Automatic Gain Control (AGC)
- Noise Cancellation (NC)
- Adaptive (Speech Passing) Non-Linear Processing (NLP)
- Extremely fast convergence rates of 49dB/s

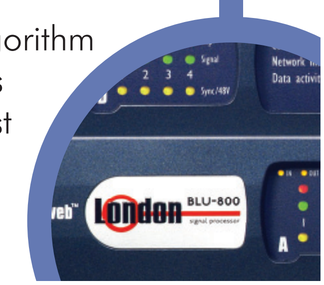
NOTE: AEC Input cards can only be used in BLU-800, BLU-320, BLU-160 or BLU-120 devices configured for 48kHz operation.

AEC DEFINITIONS

Before we continue let's review the definitions for some common terms used when discussing AEC.

Convergence Rate

Measures the speed of the linear processing component of the AEC algorithm and does not include the non-linear processing or suppression (NLP) as dictated by industry standards. This means this is a measure of how fast the algorithm can recognize and remove echo from the signal path.



HiQnet

How to use Acoustic Echo Cancellation

Double Talk
Both far and near side speech are present.

Echo Return Loss (ERL)
This is a measure of the coupling between the AEC reference signal and the AEC input signal.

Echo Return Loss Enhancement (ERLE)
This shows the loss through the linear AEC algorithm (not including the non-linear processing.)

Far Side (Reference)
This is the remote side of the conference which will be heard from the near-side speakers.

Gain Structure
Proper gain structure will provide an adequate signal to noise ratio and reasonable headroom for an input signal.

Near Side (Local)
This is the local side of the conference where the echo canceller is located.

Non-Linear Processing (NLP)
The non-linear processing increases the power of the echo cancellation for difficult acoustic environments.

Noise Cancelation (NC)
Noise cancellation removes ambient noise from the AEC signal (e.g. computer fan noise).

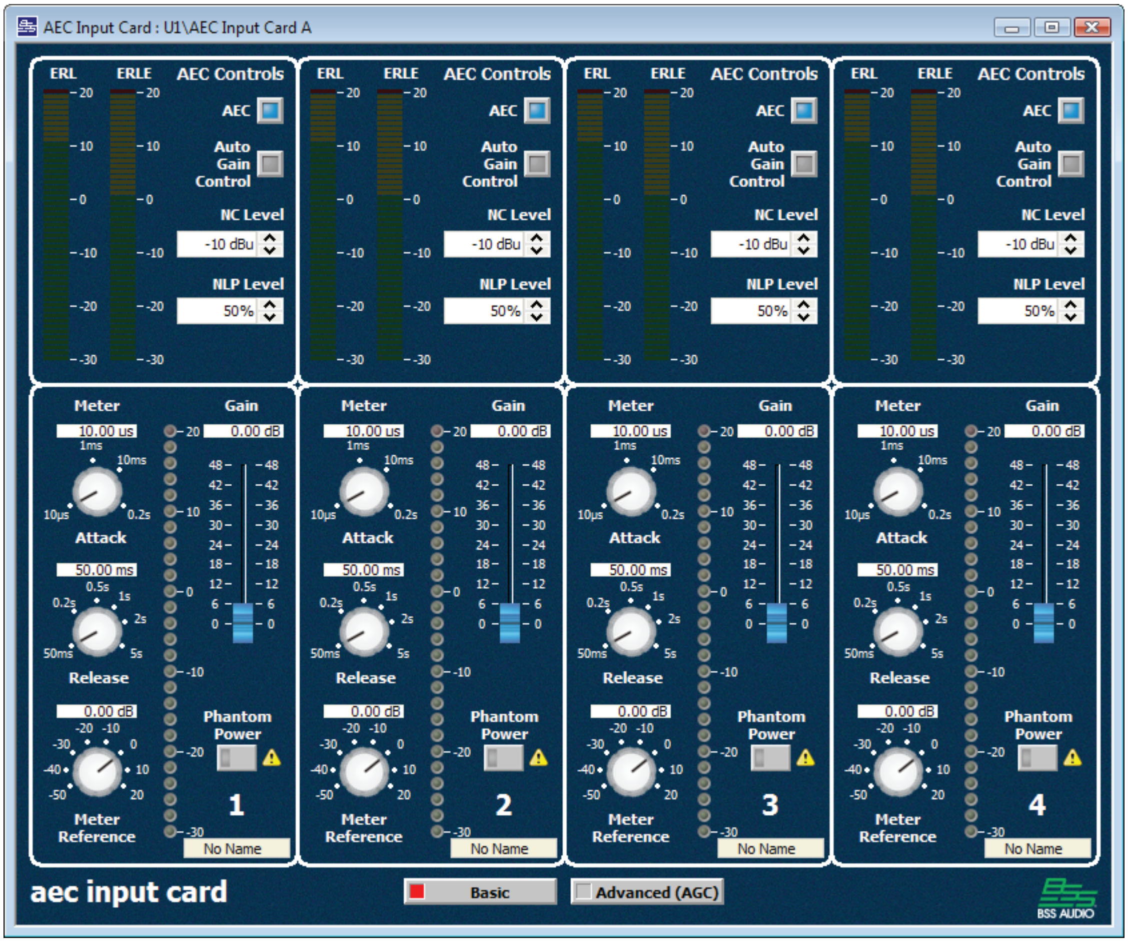
Voice Activity Detection (VAD)
Detects whether the audio is speech or silence/background noise.

AEC Card Control Panel
The AEC default control panel is ordered in two groups of controls for every input channel. The first group of controls are identical to the standard Soundweb London input cards and function in the same manner. These controls are the audio input meter (configurable as Pre or Post-AEC), input meter controls - Attack, Release, Reference, and Phantom Power - for each input channel. The second group of controls are the AEC controls.

How to use Acoustic Echo Cancellation

AEC Control Panel [Basic]

The basic AEC control panel allows enabling and disabling of AEC and AGC, and allows setting levels for noise cancellation and non-linear processing.



AEC

This button enables or disables AEC processing for each channel. When this button is enabled the AEC algorithm will remove the acoustic echo from the audio channel with linear processing and with a specified amount of non-linear processing. (See NLP Level below.)

ERL Meter

The Echo Return Loss (ERL) meter is a measure of the room's natural attenuation of the far-side audio as it leaves the speaker(s) and re-enters the microphone(s). This parameter is controlled by proper gain structure setup (ensuring a good signal to noise ratio and reasonable headroom for the AEC input signal). A proper gain structure is critical for

distortion free sound and optimal performance for AEC. This is the single most important parameter when setting up the AEC system.

The AEC algorithm will recognize and remove echo to its best extent when this meter is displaying in the 'green' range. The 'green' range is indicated on the control panel below 10dB. The algorithm will continue to converge over 10dB, but the convergence rate will decrease in that range.

This meter will not update during double-talk. It is updated based on far-side speech only.

ERLE Meter

The Echo Return Loss Enhancement (ERLE) Meter measures how much acoustic echo is being removed from the signal path. This measurement consists of the natural room attenuation as indicated by the ERL meter and the amount of echo removed by the AEC algorithm. A lower signal indicates more echo being removed. The lower the meter, the better.

NOTE: As dictated by industry standards, NLP contributions are not included in this reading. NLP contributions are made in addition to this meter's reading.

NLP Level

The Non-Linear Processing (NLP) setting determines the amount of non-linear suppression that will be applied in conjunction with the AEC algorithm for each channel. NLP will remove the residual echo not removed by the linear part of the AEC algorithm.

This parameter represents a trade-off between achieving good double-talk performance, with no suppression of the local speech signal, and very robust echo suppression, with no echo audible on the far side. At its most aggressive setting (NLP at 100%), the non-linear processing will remove any of the residual far-side echo picked up by the microphone. However, this is done with an increased risk that some of the near-side speech will be degraded as well, especially during double-talk. At its least aggressive setting (NLP at 0%), the non-linear processing is effectively disabled, which may let some echo through, but will allow for a more natural double-talk performance. The best setting for this parameter may depend on several factors, including the acoustic properties of the room and user preference. The default value of 50% may give a good balance between these two competing goals.

NC Level

The Noise-Cancellation (NC) setting will determine the amount of noise cancellation that will be applied to each channel. The noise cancellation algorithm is a very advanced algorithm that will remove steady-state noise without compromising the quality of speech

How to use Acoustic Echo Cancellation

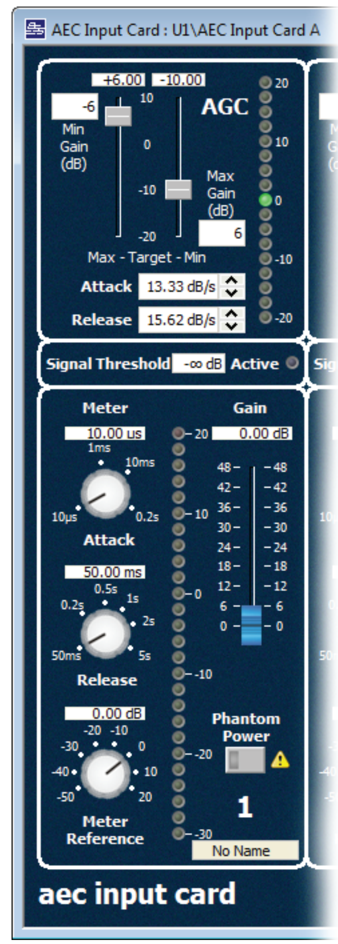
passing through the channel. This algorithm is great for removing projector noise, HVAC, and other unwanted background noise that can compromise speech intelligibility.

AEC Control Panel [Advanced]

The advanced panel gives access to controls for the Automatic Gain Control and Signal Threshold features.

AGC

The Automatic Gain Control is designed for voice applications. It is designed to compensate for varying distances between the speaker and their microphone as well as speech level variances at the near end. This provides the far end with a signal that will automatically be increased or decreased to maintain a consistent audio level. Setting the maximum gain too high can cause inconsistent gain structures and bring up the noise floor. The AGC will adjust the gain during near side speech only. This means that during pauses in near side speech, the noise floor will maintain a constant level, and will not grow to hit a target gain output. Only near side speech signals are used to control the gain.



To use the AGC, first define target levels for the transmitted speech signal. The default target levels for AGC are a maximum of 6dBu and a minimum of -10dBu, which define a target window with 16dB of dynamic range. If the speech level is within the target window already, then the AGC-applied gain will go to 0 dB.

If the speech signal is below the target window (i.e., below the minimum target level), then the AGC will increase the gain (to a limit) so that the signal level meets the minimum target level. The AGC will limit the gain it can add to a signal by a maximum gain setting. Once the AGC has adjusted its gain high enough to meet the maximum gain setting, it will stop adding gain, even if the minimum target level is not reached.

How to use Acoustic Echo Cancellation

This is useful to stop very weak speech signals, such as whispering, from driving the gain too high. Setting the maximum gain too high can cause inconsistent gain structures and bring up the noise floor.

Similarly, if the level of the speech signal is higher than the maximum target level, then the AGC will reduce the gain, by as much as the minimum gain setting, in an attempt to bring the speech level down to the maximum target level.

A generous range for the maximum gain and the minimum gain have been provided. Care should be taken, particularly with the maximum gain setting, to avoid extreme levels. Situations where the maximum gain setting should be set over 10dB will be rare. The maximum gain setting has the potential to break a gain structure, so set it carefully, especially if the setting is to be used over 10dB.

The attack and release rates for the AGC describe how fast it will adjust its gain. Because the AGC only adjusts gain during near-side speech signals (and not during unvoiced consonants like t, s, p, and f), the attack and release rates should be set higher than other typical AGC implementations.

The AGC meter shows the current amount of gain being applied to the signal.

Signal Threshold

In a conferencing system, some microphones may have a mute or push-to-talk feature built in. If a mic goes into or comes out of mute, then the characteristics of the conferencing system change instantly, and echo may leak through as the AEC re-converges. A signal threshold is defined to allow mics with mute or push-to-talk features to work seamlessly with AEC. Using the threshold, a level can be defined that is below the normal, ambient noise floor of the room. If the mic level goes below this level, then the AEC algorithm will treat the microphone as muted, and minimize any echo that would have occurred otherwise. The "Active" LED indicates that the microphone level is over the threshold, and the mic is not treated as being muted. When the LED is off, the mic level is below the threshold, and the mic will be treated as being muted.

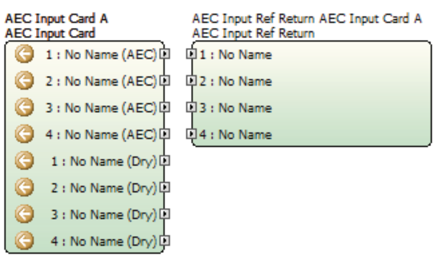
To set the threshold:

- Put the microphone in its muted mode.
- Set the threshold to a level where the LED turns itself on and off randomly.
- Raise the threshold from this level by 3 to 6 dB. The LED should now be off with no flickering.
- Take the microphone out of its muted mode and the LED should illuminate.

This process may need to be repeated if the microphone's preamp gain is adjusted. To disable the mute feature based on signal threshold, simply set the threshold to its minimum value.

London Architect Configuration Symbol

For each recognized AEC card in a Soundweb London unit the following configuration symbols will appear in the Default Configuration view:



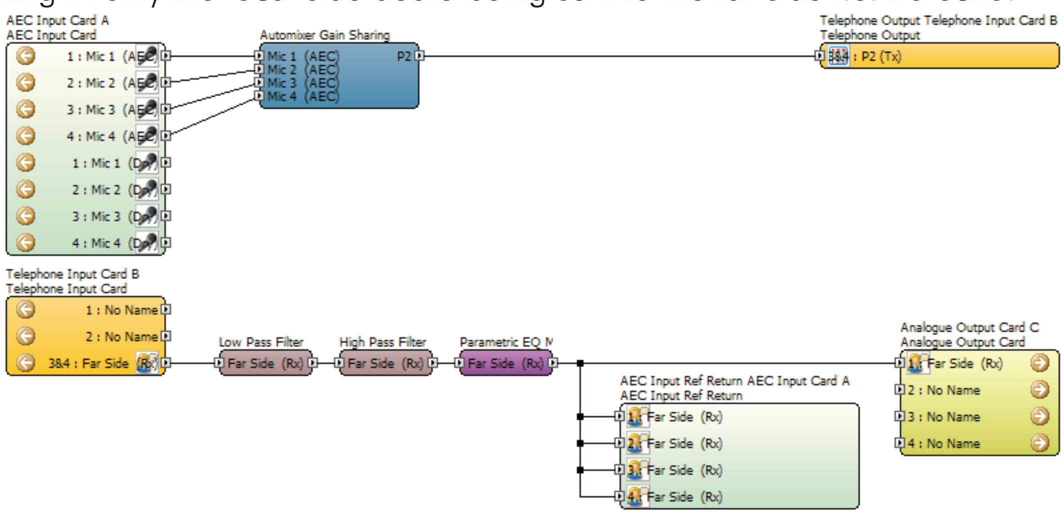
The left hand 'AEC Input Card' block functions like a standard Soundweb London input card. This block contains the 4 channels of processed AEC audio as well as the 4 channels of 'dry' (unprocessed) input audio being fed into the Soundweb London AEC card. The right hand block 'AEC Input Ref (REFERENCE) Return' is used to provide the REFERENCE signal for each AEC algorithm. The Reference signal is the signal that will be removed by the AEC algorithm from the signal path. The Reference signal should be taken from as close to the output as possible. This will provide the AEC algorithm with the most accurate representation of the signal (to be cancelled) and will provide the best AEC performance.

How to use Acoustic Echo Cancellation

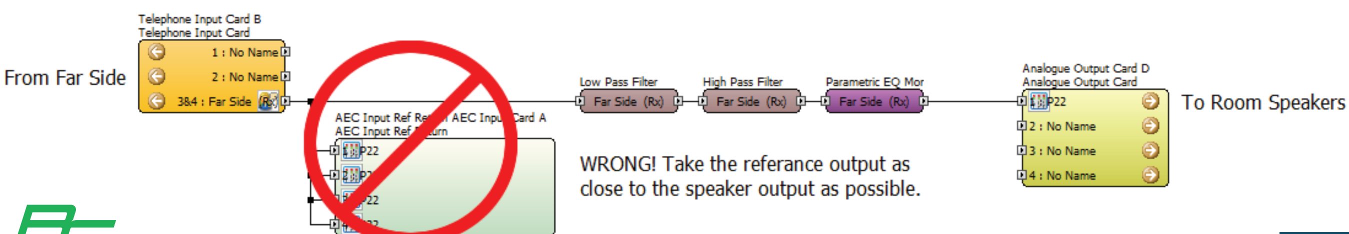
Example: Basic Conferencing without Local Sound Reinforcement

Local Sound Reinforcement refers to a design where the local microphones feed both the far-side and the local room speakers. This typically applies in large rooms where other participants located in the same room can't hear the person speaking.

This example shows 4 microphones feeding the local audio to the far-side via a 'Telephone Hybrid'. The far side audio is received via the 'Telephone Hybrid' and processed by the London's Low Pass, High Pass, and Parametric EQ's before being sent to the local room's speaker(s) for the local participants to hear. Once the far-side audio leaves the local-side's speakers the signal will bounce around the room, re-enter the local microphone (mixing with the local side's speech), and the far-side signal will be sent back to the far-side resulting in 'echo'. To prevent this echo, the far-side signal is sent to the Reference inputs of the AEC card where the AEC algorithm will compare this signal with the input signals of the AEC card (the microphones) and remove the Reference signal (far side signal) from the input signal path resulting in only the local side audio being sent to the far-side i.e. no echo.



The image below is the same design with the Reference signal wired incorrectly. Because the Reference is taken before the room processing blocks, the AEC algorithm will not understand that certain frequencies were cut/boosted intentionally and will not be able to model the room to its full ability.

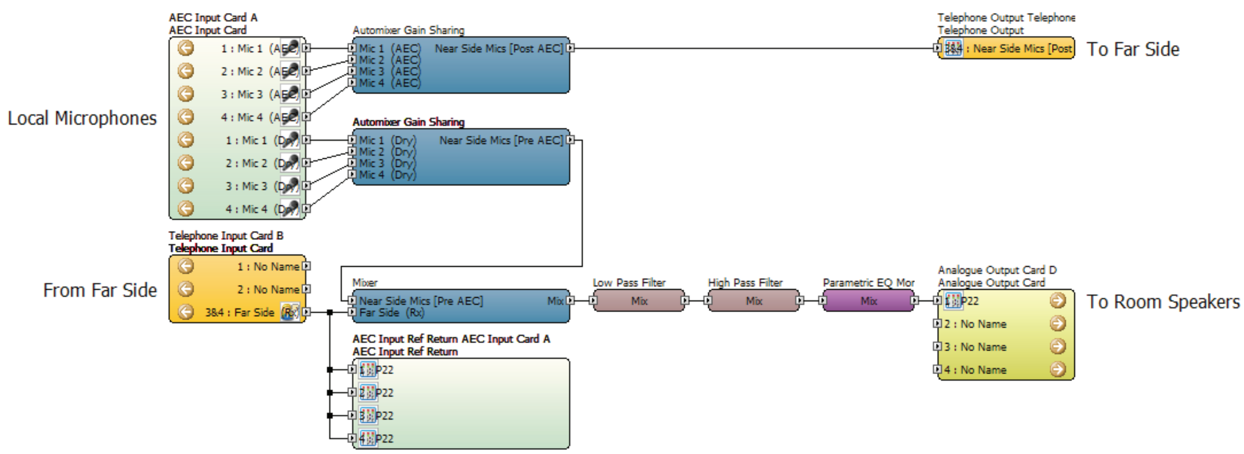


How to use Acoustic Echo Cancellation

Example: Basic Conferencing with Local Sound Reinforcement

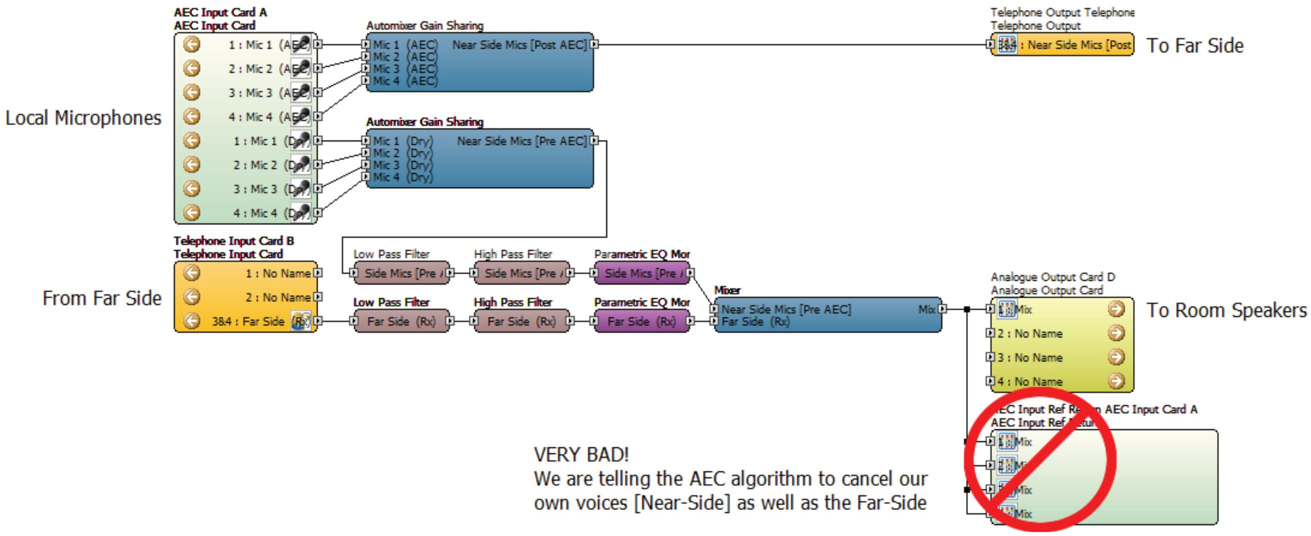
This example shows 4 microphones feeding the audio to the far side via a 'Telephone Hybrid' as well as feeding the local speakers for local sound reinforcement. Signal mixing is performed using the Gated Automixer processing object. The best method for this type of design incorporates a mix-minus setup to maintain proper gain structure, and to prevent the speaker directly above the person talking from transmitting a - room-colored - copy that will re-enter the open microphone and be transmitted to the far side along with the original voice signal.

The design below shows both the far side and near side signals feeding the local room speakers. This design works, but as explained above, since the Reference signal is not being fed after the room processing blocks (as close to the speaker output as possible) the AEC algorithm won't perform to its full potential.

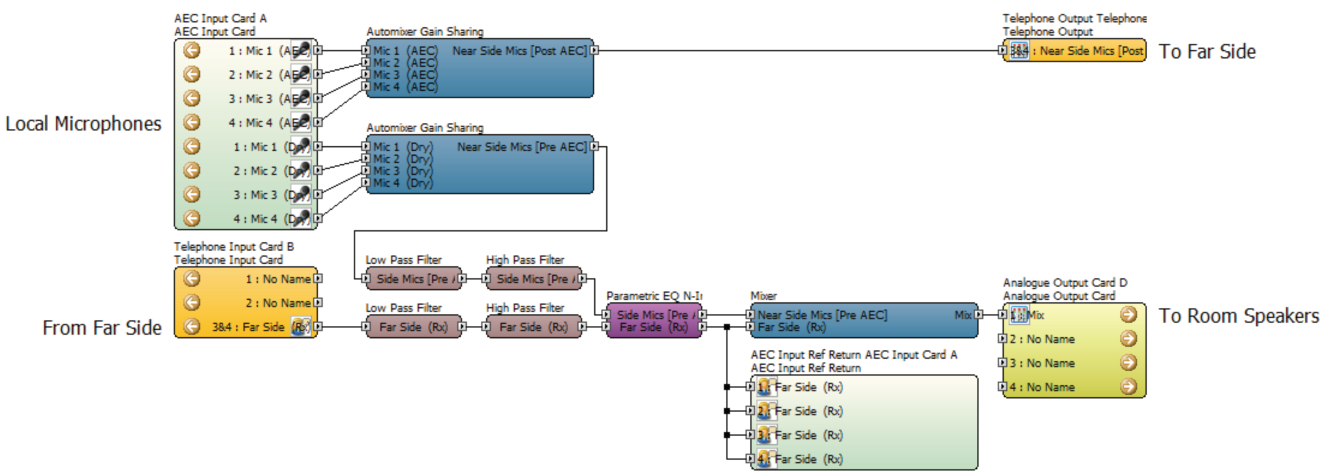


If the Reference is moved to the same location as in the previous 'No Reinforcement' example, it will satisfy the rule of placing the Reference 'as close as possible' to the speaker output, but in doing so the Reference will be fed with a mix of both the near side and far side signals. Since the Reference signal is the 'signal we want to remove from the input audio path' then this means that the AEC algorithm will cancel the speaker's voice coming into the AEC Input Card. Since the input microphone signal path is being fed to the far side as well as the local speakers then the far side will not be able to hear the speaker either. (Typically what happens is that only portions of the speaker's voice gets cancelled because of the VAD state and it causes the voice to distort and sound bad to both the far side and the near side)

How to use Acoustic Echo Cancellation



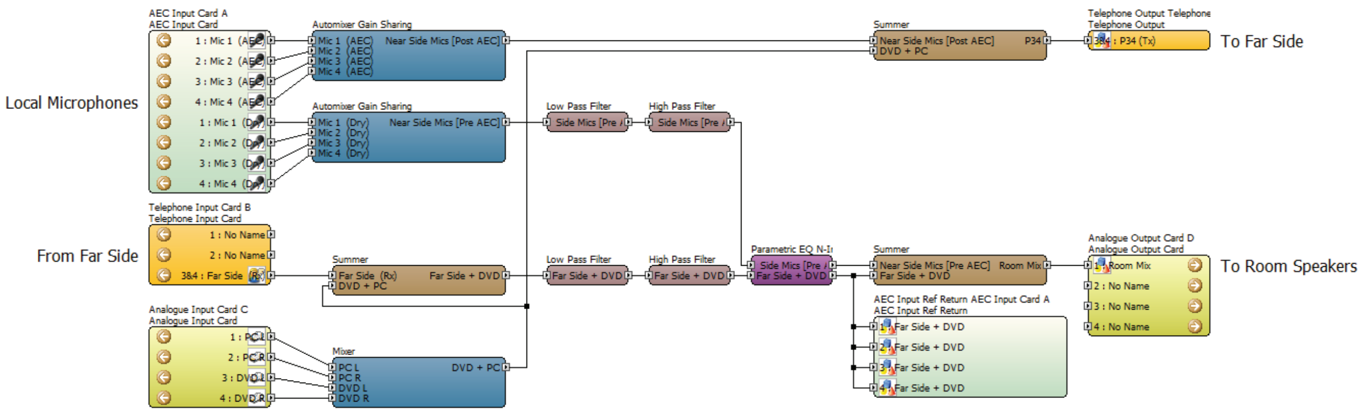
To solve this dilemma utilize an N-input parametric EQ and another set of High/Low Pass objects in order to provide the Reference the same signal as the room speaker. This means that you only Reference (remove) the far side signal while still feeding a mix of both near side and far side audio to the room speakers. **It is very important to make sure the same settings are maintained in both signal paths. In particular, care must be taken that any non-linear processing (such as compression or limiting) that happens to the speaker output signal, also happens to the Reference signal.** BSS Audio recommend using the 'copy parameter values' feature to ensure the settings are identical.



How to use Acoustic Echo Cancellation

Example: Local Media Distribution

Add to the previous Local Sound Reinforcement design by adding a local DVD player and a local PC audio input for presentations that the near side would like to share with the far side as well as having the near side speakers distribute this material.



This design requires that the local media inputs (DVD/PC) are sent directly to the far side via a 'Telephone Hybrid' to provide a high quality audio signal. For the near side to hear the local media inputs through their speakers, Reference (remove) the far side signal and local media signal from the microphone input signal path to prevent the far side from hearing their own voice (echo) and from getting a lower quality local media signal - which can be interpreted as an echo as well. Remember that you are already directly sending the local media signal to the far side so you need to make sure that the local media signal is not allowed to travel to the far side via the microphone input signal path as well.

Once again, it should be emphasized that the ERL meter needs to be in the 'green' zone while the local Media sources are playing during the conference or the AEC algorithm will not optimally remove the echoes.

This page intentionally blank