

Quantitative processing of broadband data as implemented in a scientific splitbeam echosounder

Lars Nonboe Andersen,^{1, a)} Dezhang Chu,² Nils Olav Handegard,³ Harald Heimvoll,¹ Rolf Korneliussen,³ Gavin J. Macaulay,³ Egil Ona,³ Ruben Patel,⁴ and Geir Pedersen³

¹*Kongsberg Maritime AS, Strandpromenaden 50, 3191, Horten, Norway*

²*Fishery Resource Analysis and Monitoring Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Blvd. E. Seattle, WA, 98112, USA*

³*Marine Ecosystem Acoustics, Institute of Marine Research, Bergen, 5001, Norway*

⁴*Codelab, Bergen, Norway*

The use of quantitative broadband echosounders for biological studies and surveys offers considerable advantages over narrowband echosounders. These include improved spectral-based target identification and significantly increased ability to resolve individual targets. Biological studies and surveys typically require accurate measures of backscatter strength and we present here a systematic and comprehensive explanation of how to derive quantitative estimates of target strength and volume backscattering, as a function of frequency from broadband echosounder signals.

^{a)}lars.nonboe.andersen@km.kongsberg.com

I. INTRODUCTION

Active acoustics is an efficient method for remote sensing of marine ecosystems. These methods have been used to cover a wide range of spatial and temporal scales, ranging from tens of km (?) to small scale behaviour patterns (?), and can provide information on key forage species, including fish and zooplankton, linking primary producers and top predators as well as other taxonomic groups important for the ecosystem functioning (?). As early as 1935, Oscars Sund observed the distribution of spawning cod in the Lofoten area using a single beam echosounder (?). The method was further developed to map the abundance of fish, driven by the need for fisheries management (?). More recently, the methods have been deployed on a wide range of platforms, including observatories, autonomous underwater vehicles (?), uncrewed surface vehicles (?) and vessels of opportunity, observing a wide range of ecosystem processes across different spatial and temporal scales (?).

Several different configurations of acoustic instruments are available, including multibeam sonars (?), synthetic aperture sonar, omnidirectional sonars (?), acoustic imaging sonars (?) as well as single- and multibeam echosounders (?). Narrowband single beam echosounders have been used extensively in fisheries management and ecosystem monitoring, and methodology for using these systems are well developed. The ability for accurate calibration of the system has made these especially useful for providing data for fisheries management (?).

Recent commercially available single beam echosounders produce pulses with a wide and continuous frequency range (broadband pulses), as opposed to conventional single frequency systems. This provides significantly better along-beam resolution, a higher signal to noise ratio than narrowband pulses (??), and improved frequency resolution for backscatter categorization (?).

Broad banded echosounders have been pursued actively for ecology in particularly based on the promise of improved categorization of species or group of species. Using high frequency broadband acoustics (150-600 kHz) Lavery et al (?) showed that broadband helped reduce the ambiguities in interpretation of acoustic scattering from zooplankton and oceanic microstructure. (?) used low frequency broadband pulses (1-6 kHz) with frequency content covering the resonance of swimbladdered fish, demonstrating both the spectral resolution and ability to classify size classes of fish within fish assemblages. (?) used broadband, with ground truthing, to identify that the composition of two Sound Scattering Layers (SSL) detected in the Bay of Biscay in springtime was characterized and identifying fine scale heterogeneity within the SSLs. While unsuccessful in discriminating between several species of

fishes with large swimbladders during the Alaska pollock survey, (?) showed that broadband may be helpful in characterizing smaller fishes with swimbladder and euphausiids. (?) was able to effectively discriminate three monospecific aggregations of species (hake, anchovy and krill) using broadband data (45-170 kHz), while attempts to classify using narrowband multifrequency failed.

The increased spatial resolution is another aspect of broadband acoustics that are actively being used both for improved categorization and increased ability to resolve targets near each other or near boundaries such as the seafloor. (?) explored different broadband pulse shapes to increase successful resolution of single targets in proximity as well as near boundaries in tank experiments. By applying high frequency broadband pulses to fish-like artificial targets, (?) demonstrated the potential for acoustic sizing of individually resolved fishes in a controlled ex situ environment. Utilizing the increased spatial resolution (?) were able to isolate single fishes and discriminate successfully between average frequency responses of walleye pollock and pointhead flounder.

An equally interesting outcome from the increased resolution and broad frequency bands is the ability to infer behavioural information on both single individuals, groups of individuals, and interaction between individuals or groups. (?) was able to extract orientation information about krill from using a combination of high frequency broadband acoustics and acoustic modelling. Fine-scale observations of single targets facilitated the study of interactions between prey and predators, capelin and cod, in the Barents sea ((?)).

There have been several scientific broadband echosounder systems developed for laboratory use (???), some prototype or custom-made systems (???????) and some commercially available systems (?????).

The theoretical foundation for signal processing comes from radar applications and is earlier documented for broadband echosounders (?). To tap the full potential of broadband echosounder data, further data processing methods are being developed (??). Examples include adapting pulses to improve target separation as well as further parameters from the individual targets. Different send pulses can be envisioned as well as various methods for acoustic target classification. When translating equations into code, several choices must be made. Some of these are well founded in signal processing literature whereas others are of a more practical and ad-hoc nature. The latter is typically missing in conventional papers, and this makes it difficult to test the implementation of the signal processing methods in new echo sounders and post processing software.

The objective of this paper and associated code is to present a systematic and comprehensive description of the data processing steps for calibrated broadband echosounder data. The steps include pulse compression, target strength as a function of frequency $[TS(f)]$, and volume backscattering strength as a function of frequency $[Sv(f)]$. The intention is that the code will be used to validate implementations in various relevant data processing software, provide a starting point for further developing signal processing, and serve as a learning resource for broadbanded acoustic data processing. We also note that the processing equations and methodology presented in this paper are similar to version 1.12.4 and earlier of the Simrad EK80 software.

II. SIGNAL FLOW AND INITIAL PROCESSING

A. Accompanying code

The code accompanying this paper is written in Python and is available as supplementary materials. All processing steps with respective figures in the paper can be reproduced by running the main scripts (TSf.py and Sv.py). Without loss of generality, we use the Simrad EK80 as an example since it is currently the most used broadband echosounder in the marine ecosystem acoustics field. By presenting the design goals, implementation details, and recommended procedures and processing required to obtain quantitative broadband data, the authors hope to encourage and facilitate the realistic use of broadband signals in marine ecosystem acoustics.

Our presentation uses nomenclature and approaches that are commonly used for narrowband echosounder systems, which were derived from radar processing (?). In particular, the expressions for target strength (TS) and volume backscattering strength (S_v) (?) are presented in a similar manner for broadband signals as for narrowband signals.

B. System overview

A basic echosounder system consists of a transducer, a transceiver, and a computer program that controls the operation of the transceiver and records received signals. During transmission the program defines the signals which are created as electric signals in the transceiver, converted to acoustic signals by the transducer and transmitted into the water. The acoustic signals propagate through the water, are reflected or scattered by objects in

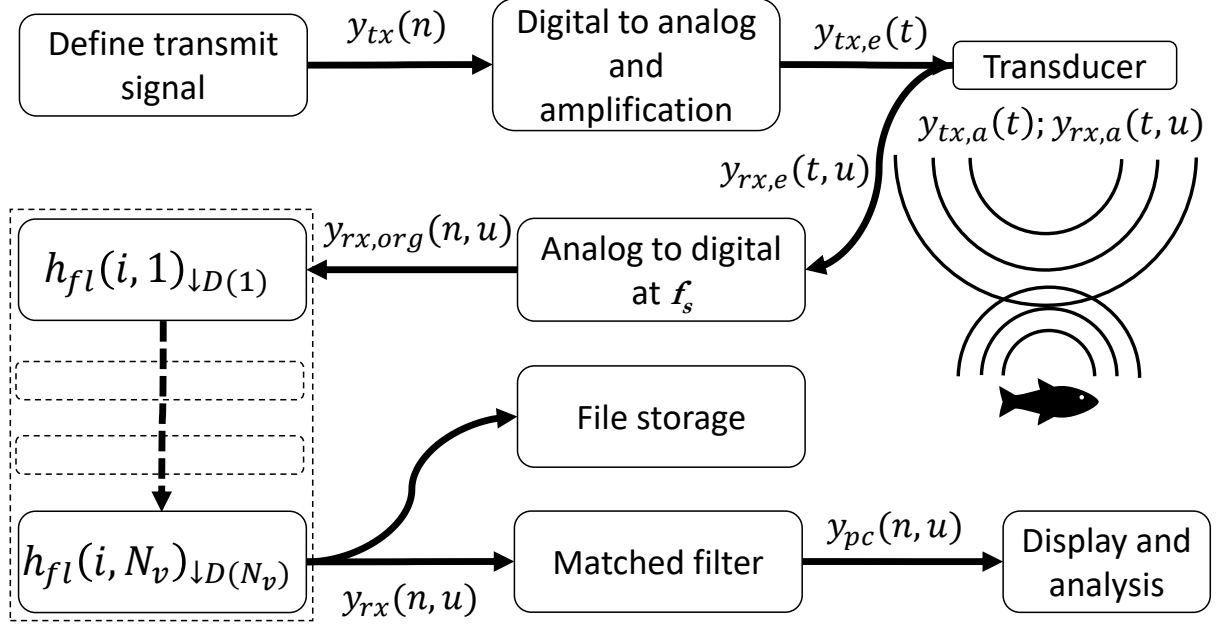


Figure 1. Signal and data flow in the Simrad EK80 system. An echosounder ping starts with the definition of a transmit signal (upper left) and ends with file storage (lower middle) and display and analysis (lower right). Note that all signals are complex-valued time-series.

the water, and propagate back to the transducer. During reception the transducer converts the received acoustic signals to electric signals, which are received, pre-amplified, filtered, digitized, and processed in the transceiver, and then transferred to the controlling program for further data processing and storage (Fig. 1). Many types of transmit signals are feasible - this paper considers only the transmission of linear frequency modulated signals (also known as linear chirps).

C. Signal generation

The controlling computer program generates a short-duration digital transmit signal (a ping), $y_{tx}(n)$, where n is a sample index in the discrete time domain. Typical BB pulses are linear upswep pulses windowed by an envelope function. The generated signal $y_{tx,e}(t)$ is converted to an analogue electric signal and amplified by the transceiver to obtain analogue signal $y_{tx,e}(t)$. The analogue and amplified signal is passed on to the transducer to generate the transmitted acoustic signal $y_{tx,a}(t)$ in the water column, where t is the time for the (analog) signal. For a split-aperture echosounder system, there are typically three or four

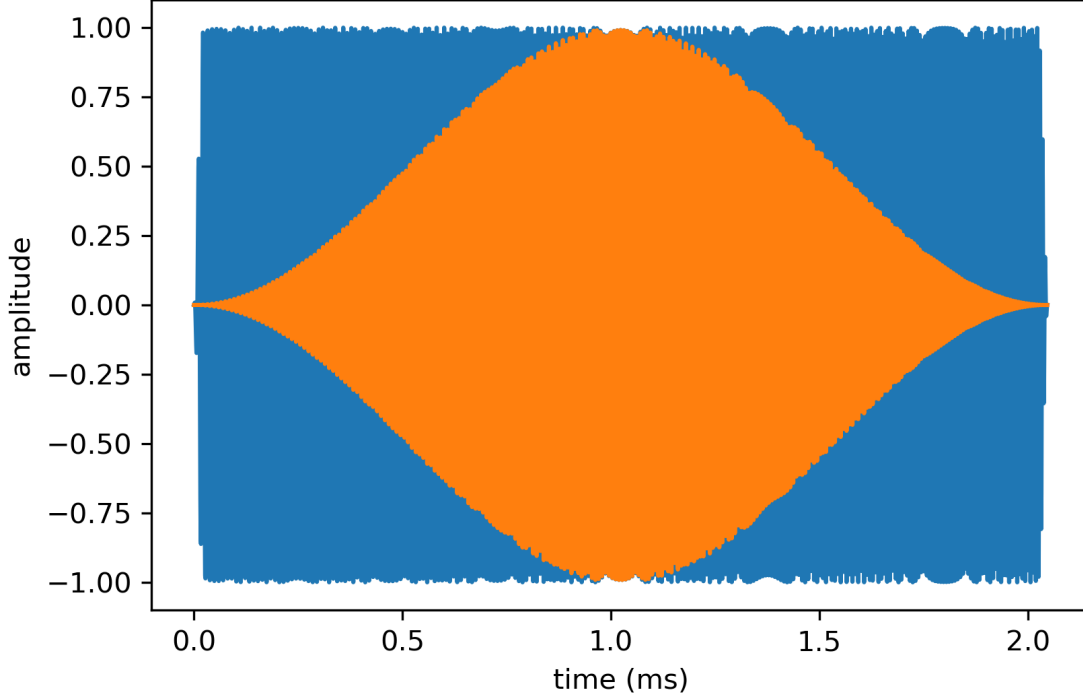


Figure 2. Ideal and enveloped transmit pulse between 34 kHz to 45 kHz with a slope of 0.057, as implemented in the EK80. The orange curve illustrates the effect of setting the slope to 0.5.

channels to allow estimation of the angle of arriving echoes, and the signal is typically transmitted with equal power across the channels.

In the example a linear sweep enveloped by a Hann function is implemented (Fig. 2), where the parameters are the initial frequency, the final frequency, the pulse duration, the sampling rate, and the proportion of the signal that is tapered in each end, respectively. A tapering of 0 and .5 indicates no tapering and tapering across the whole signal, respectively.

D. Signal reception

The returning acoustic signal, $y_{\text{rx},a}(t)$, is received by each transducer sector, u , and converted to an analog electric signal, $y_{\text{rx},e}(t, u)$, in the transducer and received by corresponding receiver channels, u , in the transceiver. The received electric signal, $y_{\text{rx},e}(t, u)$, from each

channel, u , is pre-amplified, filtered by an analog anti-aliasing filter, and digitized in the transceiver at a frequency of f_s , creating the digital signal, $y_{\text{rx,org}}(n, u)$.

To remove noise and reduce the quantity of data, the sampled signal from each channel is filtered and decimated in multiple stages, v , using complex bandpass filters, $h_{\text{fl}}(i, v)$, and decimation factors, $D(v)$. The individual filter coefficients for each filter and decimation stage are indexed by i . The output signal from each channel, u , from each filter and decimation stage, v , is then given by:

$$y_{\text{rx}}(n, u, v) = (y_{\text{rx}}(n, u, v-1) * h_{\text{fl}}(i, v))_{\downarrow D(v)}, v = 1, \dots, N_v, \quad (1)$$

where $y_{\text{rx}}(n, u, 0)$ is set to $y_{\text{rx,org}}(n, u)$, being the signal before decimation, $*$ indicates convolution and N_v is the total number of filter stages. The output signal from the final filter and decimation stage, $y_{\text{rx}}(n, u, N_v)$, is shortened to $y_{\text{rx}}(n, u)$ for convenience. For the output signal, $y_{\text{rx}}(n, u)$, the decimated sampling rate, $f_{s,\text{dec}}$, is given by:

$$f_{s,\text{dec}} = f_s \prod_{v=1}^{N_v} \frac{1}{D(v)}. \quad (2)$$

The characteristics of the bandpass filter and decimation factors are chosen with regard to the desired operating bandwidth, noise suppression levels, impulse response duration, and other common filter characteristics, with the aim of maintaining sufficient information in the data. In the example $N_v = 2$. The frequency responses of the filters are shown in Figure 3, and the corresponding filter coefficients and decimation factors are found in the test data set.

The original sample data $y_{\text{rx,org}}(n, u)$ are not available in the EK80 data files. Instead, the resulting filtered and decimated complex samples from each transducer channel $y_{\text{rx}}(n, u)$ are stored in the data files. The data are recorded together with additional data such as from position and motion sensors and system configuration data in raw data files for display and analysis by processing software. The filter parameters are in this paper used to generate the transmit pulse for the matched filter, see below.

E. Pulse compression

To increase signal-to-noise ratio and resolution along the acoustic beam a matched filter may be applied to the raw data samples (?). This technique is also known as pulse compression (?). One approach for a matched filter is to use a normalized version of the

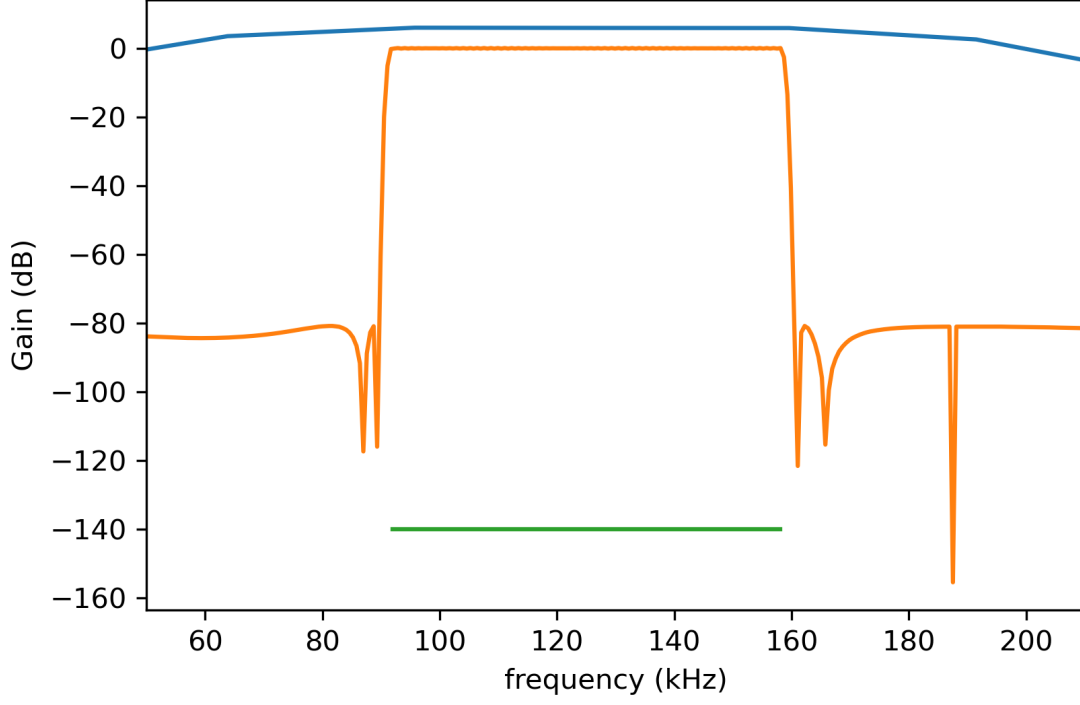


Figure 3. The frequency (power) response of the filters in our test set.

ideal transmit signal as the replicate signal, filtered and decimated using the same filters and decimation factors as applied in Eq. 1. The normalized ideal transmit signal, $\tilde{y}_{\text{tx}}(n)$, is given by:

$$\tilde{y}_{\text{tx}}(n) = \frac{y_{\text{tx}}(n)}{\max(y_{\text{tx}}(n))} \quad (3)$$

where \max is the maximum value of $y_{\text{tx}}(n)$. The filtered and decimated output signal, $\tilde{y}_{\text{tx}}(n, v)$, from each filter stage, v , using the normalized ideal transmit signal, $\tilde{y}_{\text{tx}}(n)$, as the input signal, is given by:

$$\tilde{y}_{\text{tx}}(n, v) = [\tilde{y}_{\text{tx}}(n, v-1) * h_{\text{fl}}(i, v)]_{\downarrow D(v)}, v = 1, \dots, N_v, \quad (4)$$

where $\tilde{y}_{\text{tx}}(n, 0)$ is set to $\tilde{y}_{\text{tx}}(n)$ and \downarrow indicates decimation by the factor $D(v)$. The output signal from the final filter and decimation stage, $\tilde{y}_{\text{tx}}(n, N_v)$, is used as the matched filter and is denoted as $y_{\text{mf}}(n)$ (Fig. 4).

The auto correlation function of the matched filter signal and the effective pulse duration, defined as the pulse duration at transmit power $p_{\text{tx},e}$ which produces the same energy as the

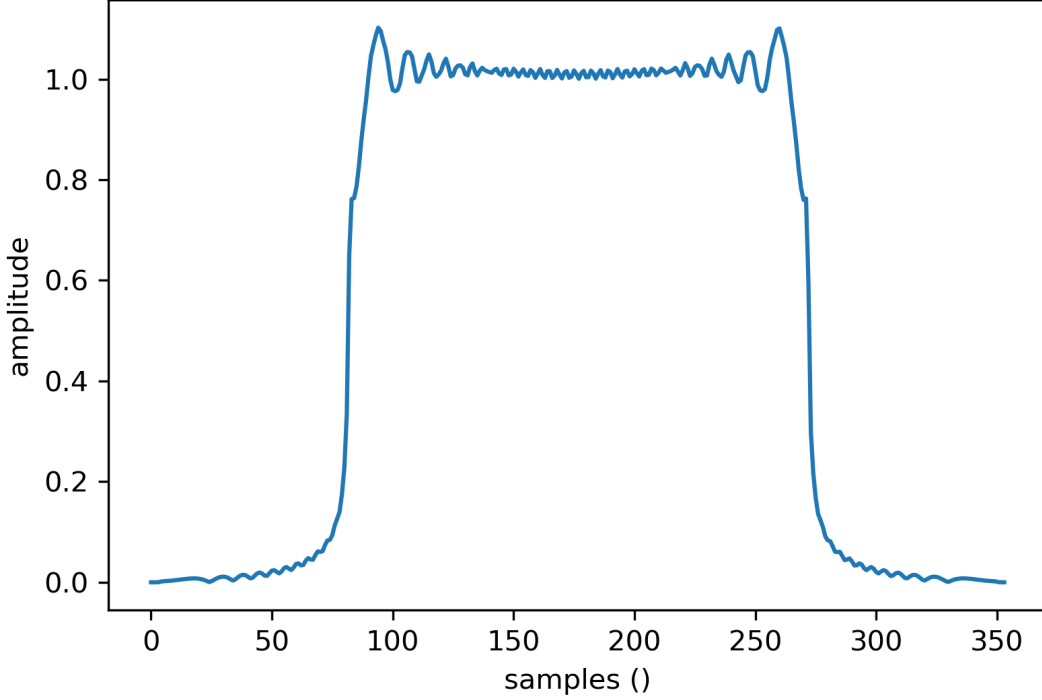


Figure 4. The absolute value of the filtered and decimated output signal $y_{mf}(n)$ used by the pulse compression.

actual transmitted pulse, will be used in later processing steps and are defined as

$$y_{mf,auto}(n) = \frac{y_{mf}(n) * y_{mf}^*(-n)}{\|y_{mf}\|_2^2} \quad (5)$$

and

$$\tau_{eff} = \frac{\sum p_{tx,auto}(n)}{\max(p_{tx,auto}(n))f_{s,dec}}, \quad (6)$$

where

$$p_{tx,auto}(n) = |y_{mf,auto}(n)|^2$$

is the square of the absolute value of the matched filter autocorrelation function, and the summation is calculated over a duration of twice the nominal pulse duration, 2τ . For an ideal system, i.e., no tapering at the rising and trailing edges of the transmitted signal, the effective pulse duration is the same as the transmit pulse duration.

To perform pulse compression the received signal, $y_{rx}(n, u)$, is convolved with a complex conjugated and time-reversed version of the matched filter signal with the matched filter

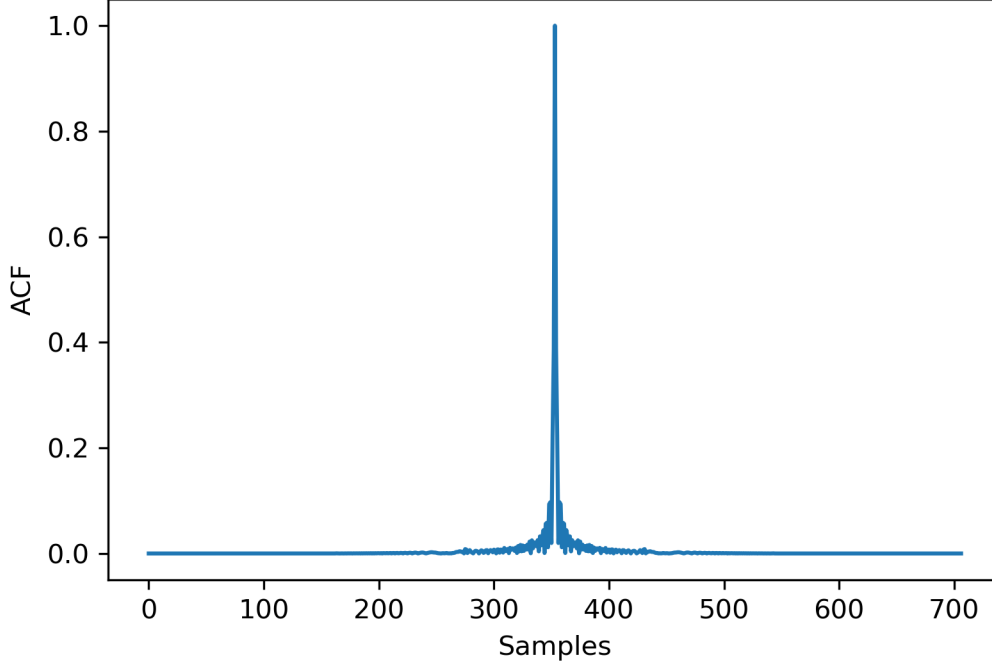


Figure 5. The autocorrelation function $p_{\text{tx,auto}}$ from the example. The corresponding $\tau_{\text{eff}} = 0.1\text{ms}$ and $\tau = 2\text{ms}$.

signal, and here also normalized with the l^2 -norm of the matched filter to maintain received signal power. The pulse compressed signal, $y_{\text{pc}}(n, u)$, then becomes

$$y_{\text{pc}}(n, u) = \frac{y_{\text{rx}}(n, u) * y_{\text{mf}}^*(-n)}{\|y_{\text{mf}}\|_2^2}, \quad (7)$$

where $\|y_{\text{mf}}\|$ indicates the l^2 -norm of y_{mf} , also known as the Euclidean norm. The received power samples are then used to estimate target strength and volume backscattering strength. For estimating received power samples, the mean signal, $y_{\text{pc}}(n)$, over all transducer sectors, N_{u} , will be used:

$$y_{\text{pc}}(n) = \frac{1}{N_{\text{u}}} \sum_{u=1}^{N_{\text{u}}} y_{\text{pc}}(n, u). \quad (8)$$

Compensation of echo strength for position in the acoustic beam requires an estimate of the echo arrival angle. This is obtained using the split-aperture method (?), which for broadband pulses can be implemented with the angle values contained in the complex-valued $y_{\text{pc}}(n)$ data, in combination with knowledge of transducer sector geometry. The principle is demonstrated with a transducer that is divided into four quadrants (Fig. 6). In this example

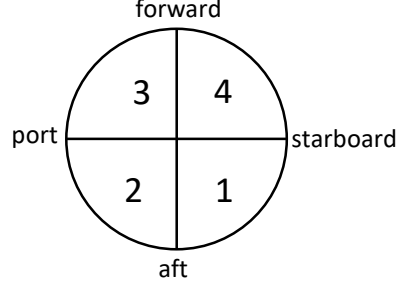


Figure 6. Transducer divided into four quadrants. The labels are directions often used when a transducer is mounted on a ship.

the summed signals from four halves (1+2, 2+3, 3+4, 4+1) are calculated as:

$$y_{\text{pc,fore}}(n) = \frac{1}{2} (y_{\text{pc}}(n, 3) + y_{\text{pc}}(n, 4)) , \quad (9)$$

$$y_{\text{pc,aft}}(n) = \frac{1}{2} (y_{\text{pc}}(n, 1) + y_{\text{pc}}(n, 2)) , \quad (10)$$

$$y_{\text{pc,star}}(n) = \frac{1}{2} (y_{\text{pc}}(n, 1) + y_{\text{pc}}(n, 4)) , \quad (11)$$

$$y_{\text{pc,port}}(n) = \frac{1}{2} (y_{\text{pc}}(n, 2) + y_{\text{pc}}(n, 3)) , \quad (12)$$

where fore, aft, star(board), and port indicate the relevant transducer halves.

F. Power and angle samples

The transceiver measures voltage over a load, $z_{\text{rx,e}}$, connected in series with the transducer impedance, $z_{\text{td,e}}$. When calculating various acoustic properties a system gain parameter will be used which assumes a matched receiver load. The total received power, $p_{\text{rx,e}}(n)$, from all transducer sectors for a matched receiver load (Fig. 7) is given by:

$$p_{\text{rx,e}}(n) = N_{\text{u}} \left(\frac{|y_{\text{pc}}(n)|}{2\sqrt{2}} \right)^2 \left(\frac{|z_{\text{rx,e}} + z_{\text{td,e}}|}{z_{\text{rx,e}}} \right)^2 \frac{1}{|z_{\text{td,e}}|} . \quad (13)$$

Forward/aft and port/starboard phase angles of target echoes are estimated by combining the transducer half signals thus:

$$y_{\theta}(n) = y_{\text{pc,fore}}(n) y_{\text{pc,aft}}^*(n) , \quad (14)$$

$$y_{\phi}(n) = y_{\text{pc,star}}(n) y_{\text{pc,port}}^*(n) , \quad (15)$$

where $y_{\theta}(n)$ is the electrical angle along the minor axis of the transducer (positive in the forward direction when ship-mounted) and $y_{\phi}(n)$ the electrical angle along the major axis

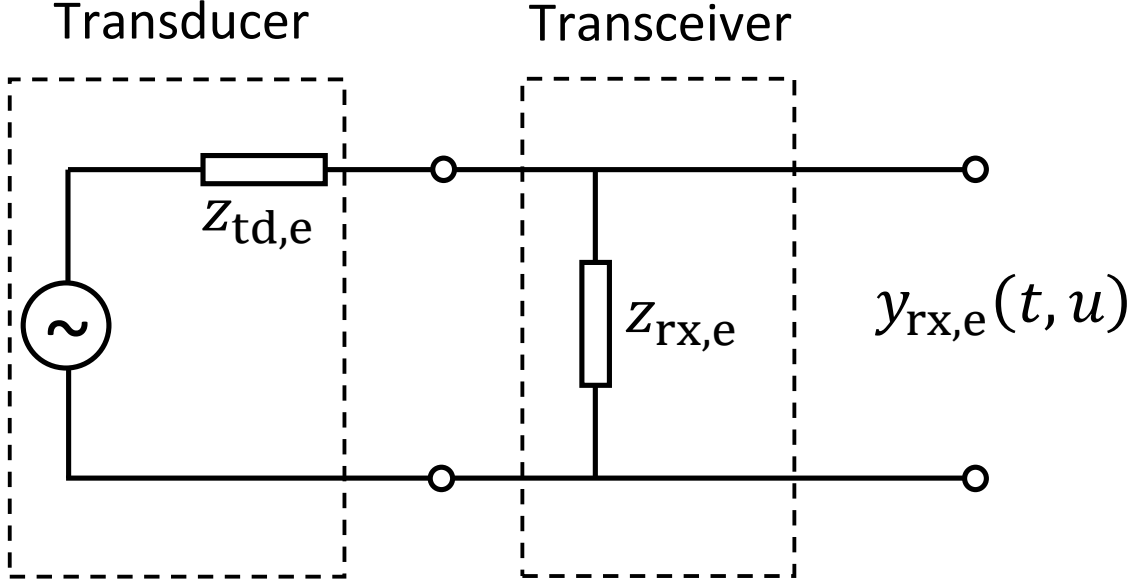


Figure 7. Equivalent circuit diagram of transducer/transceiver with system impedances.

of the transducer (positive to starboard when ship-mounted), where complex signals are represented in the form $e^{j2\pi ft}$, where $j = \sqrt{-1}$. The physical echo arrival angles (θ and ϕ) are then given by:

$$\theta(n) = \arcsin \left(\frac{\arctan2(\Im(y_\theta(n)), \Re(y_\theta(n)))}{\gamma_\theta} \right) \quad (16)$$

$$\phi(n) = \arcsin \left(\frac{\arctan2(\Im(y_\phi(n)), \Re(y_\phi(n)))}{\gamma_\phi} \right). \quad (17)$$

where γ_θ and γ_ϕ are constants that convert from phase angles to physical echo arrival angles and are derived from the transducer geometry and f_c the centre frequency of the chirp pulse (?). The inverse sine is indicated by \arcsin , the four quadrant inverse tangent which returns values in the interval $[-\pi, \pi]$ inclusive is indicated by $\arctan2$, the real part of a complex number by \Re and the imaginary part by \Im . As a mnemonic, the horizontal line in the symbol used for the forward/aft direction, θ , represents the pivot axis for the alongship angles and the near-vertical line in the ϕ symbol indicates the pivot axis for port/starboard angles.

III. TARGET STRENGTH

To illustrate each step of the process to obtain $TS(m)$ we use a single ping from a data file where a single metallic calibration sphere with 35 mm diameter (WC35) is suspended approximately 6 m below a 120 kHz transducer (fig. 9).

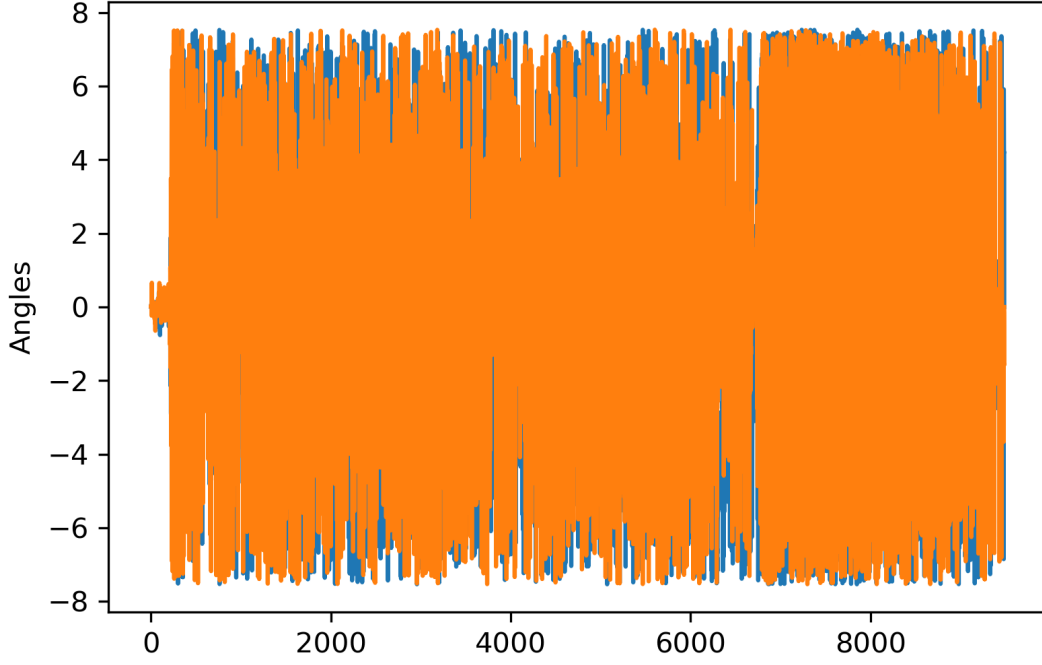


Figure 8. Physical angles θ and ϕ .

Echoes from single targets are often characterised by their TS, which is related to the differential backscattering cross section, σ_{bs} , via

$$TS = 10 \log_{10} \left(\frac{\sigma_{bs}}{r_0^2} \right), \quad (18)$$

where \log_{10} is the logarithm with base 10 and r_0 is 1 m.

The power-budget equation (i.e., sonar equation) for a single target (Formulation D, ?) at frequency f is:

$$TS(f) = 10 \log_{10}(P_{rx,e,t}(f)) + 40 \log_{10}(r) + 2\alpha(f)r - 10 \log_{10} \left(\frac{p_{tx,e}\lambda^2(f)g^2(\theta_t, \phi_t, f)}{16\pi^2} \right), \quad (19)$$

where $P_{rx,e,t}(f)$ is the Fourier transform of the received electric power in a matched load for a signal from a single target at frequency f , r_t is the range to the target, α the acoustic absorption at frequency f , $p_{tx,e}$ the transmitted electric power, λ the acoustic wavelength, and g the transducer gain incorporating both the on axis gain $g_0(f) = g(0, 0, f)$ and the beam pattern based on the estimated target bearing (θ_t, ϕ_t) .

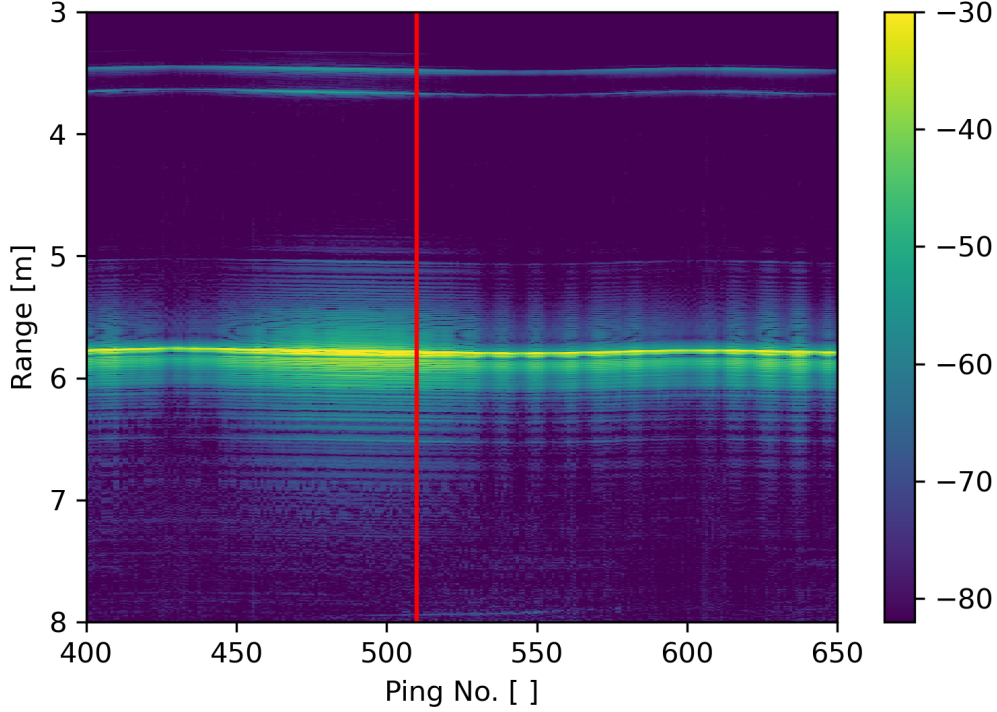


Figure 9. S_v as a function of frequency (m) and range (n) for a echogram raw data file. The red vertical line indicates ping to be processed in this section for $TS(m)$. Single target calibration sphere is seen at approximately 6 m range and the seafloor at approximately 12 m.

The point scattering strength, $S_p(n)$, is estimated by applying Eq. 19 to the received digitized power samples using the on-axis gain value with f set to the centre frequency of the broadband pulse, f_c :

$$S_p(n) = 10 \log_{10}(p_{rx,e}(n)) + 40 \log_{10}(r(n)) + 2\alpha(f_c)r(n) - 10 \log_{10} \left(\frac{p_{tx,e}\lambda^2(f_c)g_0^2(f_c)}{16\pi^2} \right), \quad (20)$$

noting that $S_p(n)$ is an average over frequency of all echoes from single or multiple targets received at sample n .

Based on the point scattering strength samples and the phase angle samples, single targets can be detected, and range and bearing to the single targets can be estimated. This is typically achieved through a single echo detection algorithm (SED). Here we will assume that the samples from the pulse compressed data $y_{pc}(n)$ originating from single target already have been identified, noting that the number of samples after the detected target may be higher than those before the peak to include scattering processes for beyond ideal

point targets. The alongship angle θ_n , athwartship angle ϕ_n and range r_n at the *peak* power $p_{rx,e}(n)$ within the detected target is used as estimates for θ_t , ϕ_t and r_t , respectively (Fig. 11). A simple pseudo SED algorithm is implemented in the code for illustrative purposes.

From the autocorrelation function of the matched filter signal, $y_{mf,auto}(n)$, the equivalent number of samples around the peak are extracted to create the reduced autocorrelation signal of the matched filter signal, $y_{mf,auto,red}(n)$. Depending on the target scattering characteristics and the distance to any adjacent single targets, the number of samples around the peak echo level in $y_{pc,t}(n)$ that contain the majority of the echo energy can be more or less than the total number of samples around the peak of $y_{mf,auto}(n)$. If the number of samples around the target is more than the total number of samples around the peak of $y_{mf,auto}(n)$ all samples around the peak of $y_{mf,auto}(n)$ are used. If the number of samples around the target is less than the total number of samples around the peak of $y_{mf,auto}(n)$, this lower number is used to create a reduced autocorrelation signal, $y_{mf,auto,red}(n)$.

The discrete Fourier transforms of the target signal, $Y_{pc,t}(m)$, and the reduced autocorrelation signal, $Y_{mf,auto,red}(m)$, are given by:

$$Y_{pc,t}(m) = \text{DFT}_{N_{\text{DFT}}}(y_{pc,t}(n)), \quad (21)$$

$$Y_{mf,auto,red}(m) = \text{DFT}_{N_{\text{DFT}}}(y_{mf,auto,red}(n)), \quad (22)$$

where DFT indicates the Fourier transform of length N_{DFT} and m the sample index in the frequency domain. The nomalized discrete Fourier transform of the target signal, $\tilde{Y}_{pc,t}(m)$, (Fig. 11) is then calculated by:

$$\tilde{Y}_{pc,t}(m) = \frac{Y_{pc,t}(m)}{Y_{mf,auto,red}(m)}. \quad (23)$$

Assuming, as a first approximation, that the impedances of the transceiver and transducer are independent of frequency, the received power into a matched load, $P_{rx,e,t}(m)$, is then estimated by:

$$P_{rx,e,t}(m) = N_u \left(\frac{|\tilde{Y}_{pc,t}(m)|}{2\sqrt{2}} \right)^2 \left(\frac{|z_{rx,e} + z_{td,e}|}{|z_{rx,e}|} \right)^2 \frac{1}{|z_{td,e}|}, \quad (24)$$

noting that any variation of impedance with frequency will be reflected in the g_0 obtained from the calibration process.

Target strength can then be estimated using Eq. 19:

$$\text{TS}(f) = 10 \log_{10}(P_{rx,e,t}(m)) + 40 \log_{10}(r_t) + 2\alpha(f)r_t - 10 \log_{10} \left(\frac{p_{tx,e}\lambda^2 g^2(\theta_t, \phi_t, f)}{16\pi^2} \right) \quad (25)$$

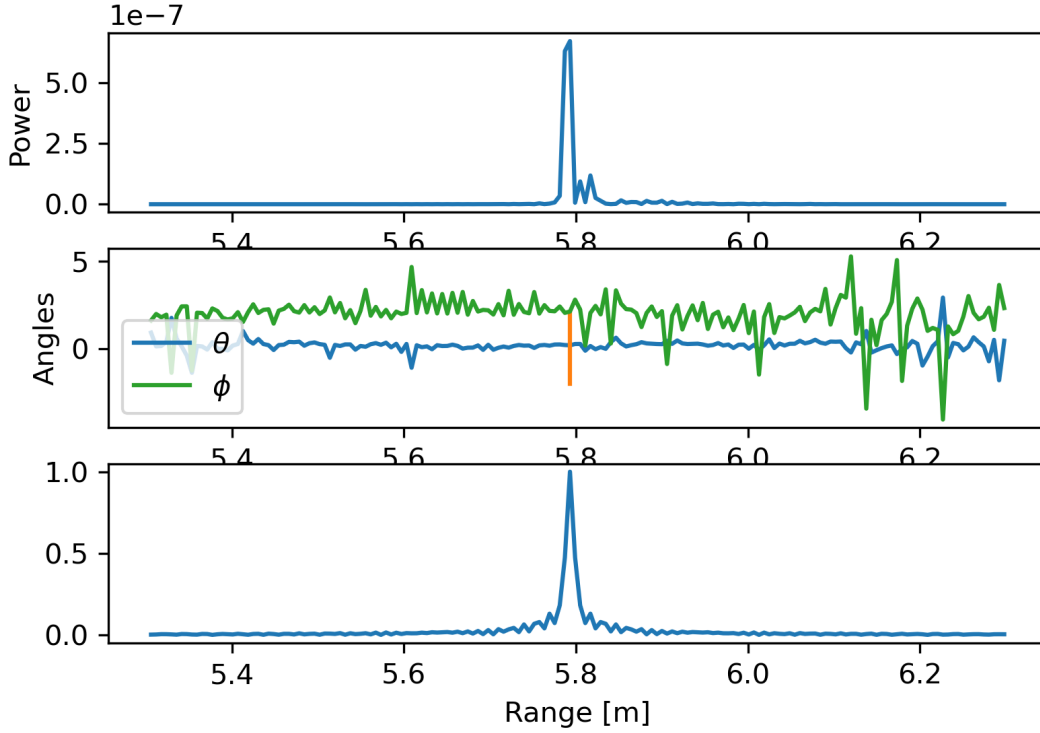


Figure 10. The $p_{\text{rx},e}(n)$ power (upper) and split beam angles (θ_t and ϕ_t) (middle) for the single target. The vertical line corresponds to the range r_t for the single target. The $y_{\text{mf,auto,red}}(n)$ (lower) is the autocorrelation function of the transmit signal reduced to the length of the target signal and aligned with the peak power of the target.

where the sample index m corresponding to frequency f can be estimated using

$$m = \lfloor \frac{f}{f_{\text{s,dec}}} N_{\text{DFT}} \rfloor \mod N_{\text{DFT}} \quad (26)$$

A frequency modulated pulse scattered by a metallic sphere will exhibit frequencies at which very little energy is returned due to destructive interference (?). This is visible in the TS (Fig. 11) and agrees well with theoretical estimates of the backscatter from spheres (?). The amplitude of the backscatter signal also clearly shows these nulls (Fig. 11), which are readily visible here due to the use of a linear chirp where time through the pulse corresponds to specific frequencies.

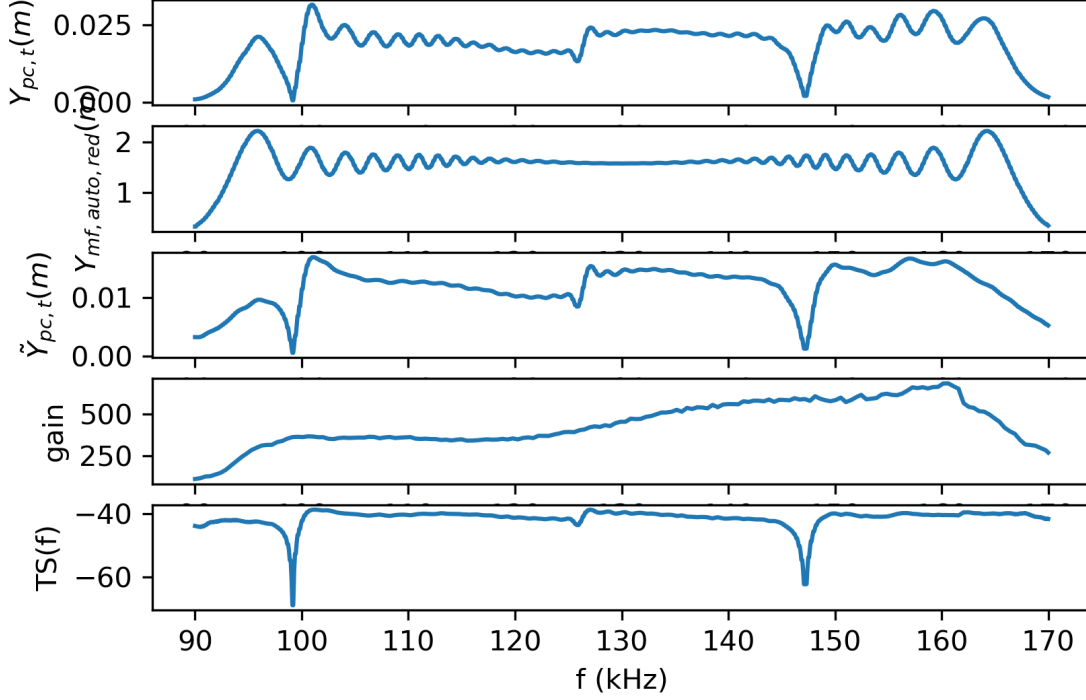


Figure 11. The discrete Fourier transform of the target signal $Y_{pc,t}(m)$ (a) and the reduced auto correlation signal $Y_{mf,auto,red}(m)$ (b), the normalized discrete Fourier transform of the target signal $\tilde{Y}_{pc,t}(m)$ (c) and the estimated TS(f) (d).

IV. VOLUME BACKSCATTERING STRENGTH

As a metallic sphere is a rather simple and ideal scatterer, we use S_v from a school of non-swimbladdered fish (fig. 12) collected with a 120 kHz centre frequency transducer to illustrate the processes to obtain $S_v(m)$.

Echoes from multiple scatterers can be quantified using volume backscattering strength, S_v , being the density of backscattering cross sections, and is given by:

$$S_v = 10 \log_{10} \frac{\sum \sigma_{bs}}{V}. \quad (27)$$

where V is the volume occupied by the scattering targets. The power-budget equation for multiple targets is then:

$$S_v(f) = 10 \log_{10}(P_{rx,e,v}(f)) + 20 \log_{10}(r_c) + 2\alpha(f)r_c - 10 \log_{10} \left(\frac{p_{tx,e} \lambda^2 c t_w \psi(f) g_0^2(f)}{32\pi^2} \right), \quad (28)$$

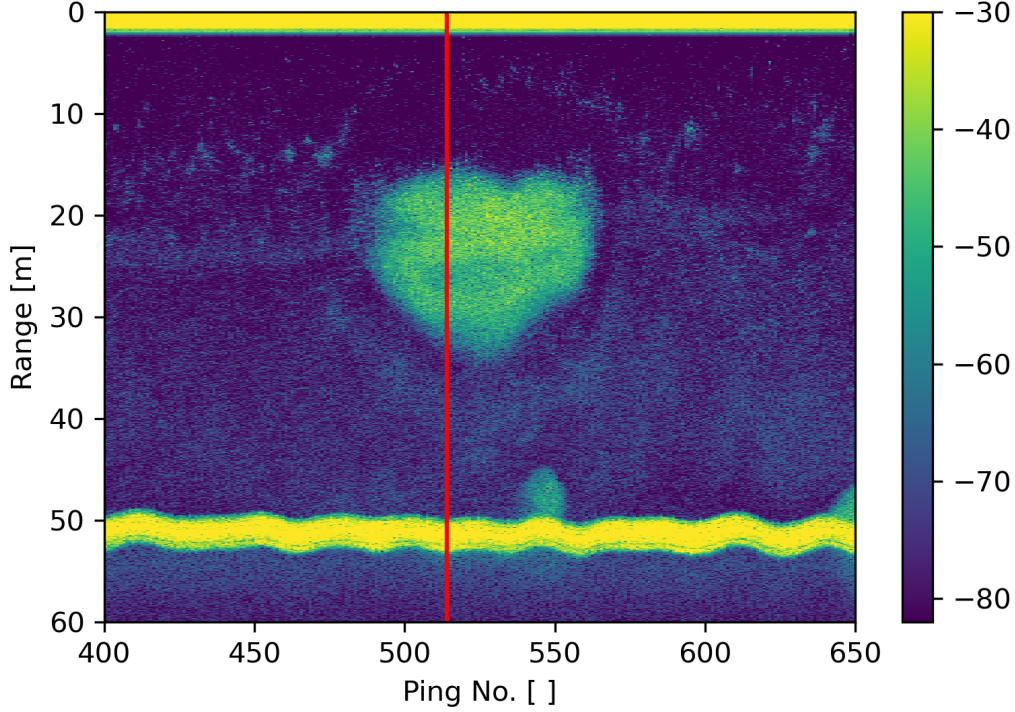


Figure 12. S_v as a function of frequency (m) and range (n) for a echogram raw data file. The red vertical line indicates ping to be processed in this section.

where $P_{rx,e,v}(f)$ is the received electric power in a matched load for the signal from a volume at frequency f , c the sound speed, t_w the duration of the time window, excluding the zero-padded portion if applied, used for evaluating the frequency spectrum, r_c is the range to the centre of the range volume covered by t_w , and ψ is the two-way equivalent beam angle. The two-way equivalent beam angle is a function of frequency that is derived from an empirical estimate of ψ at the nominal frequency, f_n :

$$\psi(f) = \psi(f_n) \left(\frac{f_n}{f} \right)^2. \quad (29)$$

Volume backscattering samples compressed over the operational frequency band are estimated by applying Eq. 28 to the received digitized power samples using the on-axis gain value with f set to the centre frequency of the broadband pulse, f_c :

$$S_v(n) = 10 \log_{10}(p_{rx,e}(n)) + 20 \log_{10}(r_c(n)) + 2\alpha(f_c)r_c(n) - 10 \log_{10} \left(\frac{p_{tx,e}\lambda^2(f_c)c\tau_{eff}\psi(f_c)g_0^2(f_c)}{32\pi^2} \right), \quad (30)$$

noting that $S_v(n)$ is an average over frequency of all echoes received at sample n . In this case, the time window, t_w , is the effective pulse duration, τ_{eff} , resulting from pulse compression.

Compensation of spherical spreading loss requires compensation of received power by a factor of r_c^2 , and hence compensation of amplitude by a factor of r_c :

$$y_{\text{pc,s}}(n) = y_{\text{pc}}(n)r_c(n). \quad (31)$$

where $y_{\text{pc,s}}(n)$ is the pulse compressed signal compensated for spherical spreading. A discrete Fourier transform is performed on the range compensated pulse compressed sample data using a normalized sliding Hann window, $w(i)$. The duration, t_w , of the sliding window is chosen as a compromise between along-beam range resolution and frequency resolution. We suggest that it be at least twice the pulse duration and for computational efficiency reasons should result in a number of samples, N_w , which is a power of 2.

The normalised Hann window, \tilde{w} , is given by:

$$\tilde{w}(i) = \frac{w(i)}{\left(\frac{\|w\|_2}{\sqrt{N_w}}\right)}, i = \frac{-N_w}{2}, \dots, \frac{N_w}{2} \quad (32)$$

and the discrete Fourier transform of the windowed data, $Y_{\text{pc,v}}(m)$, is then obtained from:

$$Y_{\text{pc,v}}(m) = \text{DFT}_{N_{\text{DFT}}} \left(\tilde{w}(i) \left(y_{\text{pc,s}}(i+n) \left[u\left(i + \frac{N_w}{2}\right) - u\left(i - \frac{N_w}{2}\right) \right] \right) \right), \quad (33)$$

where $u(i)$ is the step function and n is the sample data index for the centre of the sliding window. The discrete Fourier transform of the auto correlation function of the matched filter signal, $Y_{\text{mf,auto}}(m)$, also needs to be evaluated at the same frequencies:

$$Y_{\text{mf,auto}}(m) = \text{DFT}_{N_{\text{DFT}}}(y_{\text{mf,auto}}(n)). \quad (34)$$

The normalized discrete Fourier transform of the windowed data, $\tilde{Y}_{\text{pc,v}}(m)$, is then given by:

$$\tilde{Y}_{\text{pc,v}}(m) = \frac{Y_{\text{pc,v}}(m)}{Y_{\text{mf,auto}}(m)}, \quad (35)$$

and received power into a matched load, $P_{\text{rx,e,v}}(m)$, is estimated from:

$$P_{\text{rx,e,v}}(m) = N_u \left(\frac{|\tilde{Y}_{\text{pc,v}}(m)|}{2\sqrt{2}} \right)^2 \left(\frac{|z_{\text{rx,e}} + z_{\text{td,e}}|}{|z_{\text{rx,e}}|} \right)^2 \frac{1}{|z_{\text{td,e}}|}. \quad (36)$$

Finally, the discretized estimate of $S_v(f)$, $S_v(m)$, is given by:

$$S_v(f) = 10 \log_{10}(P_{\text{rx,e,v}}(m)) + 2\alpha(f)r_c - 10 \log_{10} \left(\frac{p_{\text{tx,e}}\lambda^2 c t_w \psi(f) g_0^2(f)}{32\pi^2} \right). \quad (37)$$

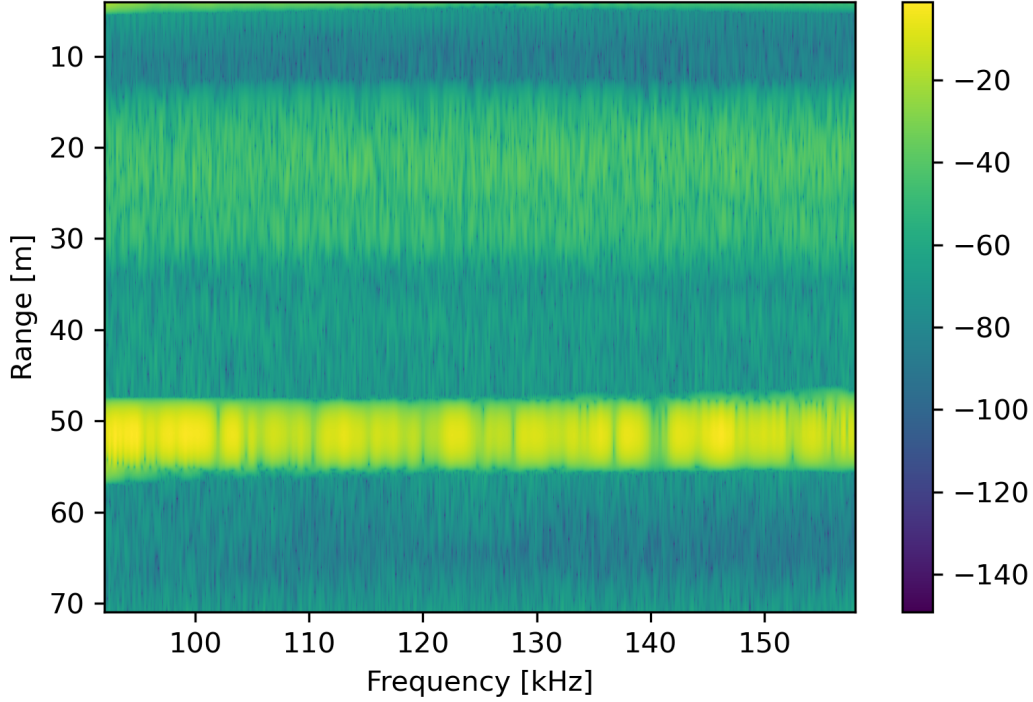


Figure 13. Sv as a function of frequency (m) and range (n) for a single ping

where the sample index m corresponding to frequency f can be estimated using

$$m = \lfloor \frac{f}{f_{s,dec}} N_{DFT} \rfloor \mod N_{DFT} \quad (38)$$

By selecting a set of centre samples n , Sv values can be presented as a function of range (n) and frequency (m) for each ping. The range for centre samples n could be chosen to half the window length or any other grid that the user prefer the data presented to be in. This can be useful when combining the $Sv(m)$ across a range of transducers. In our example we have simply chosen n to be on the same order of the range resolution (fig. 13).

For acoustic abundance estimation and classification purposes it is common to integrate the Sv over a range, in our example we integrate Sv from 15 to 34 m covering a school of non-swimbladdered fish (fig. 12). We note that we normally require averaging over several pings to obtain an unbiased estimate of Sv for echo integration purposes. Here we present the $Sv(m)$ over the layer for our example ping for illustrative purposes (fig. 14). Even though this is for a single ping, we still observe a positive slope of the frequency response indicative of non-swimbladdered fish.

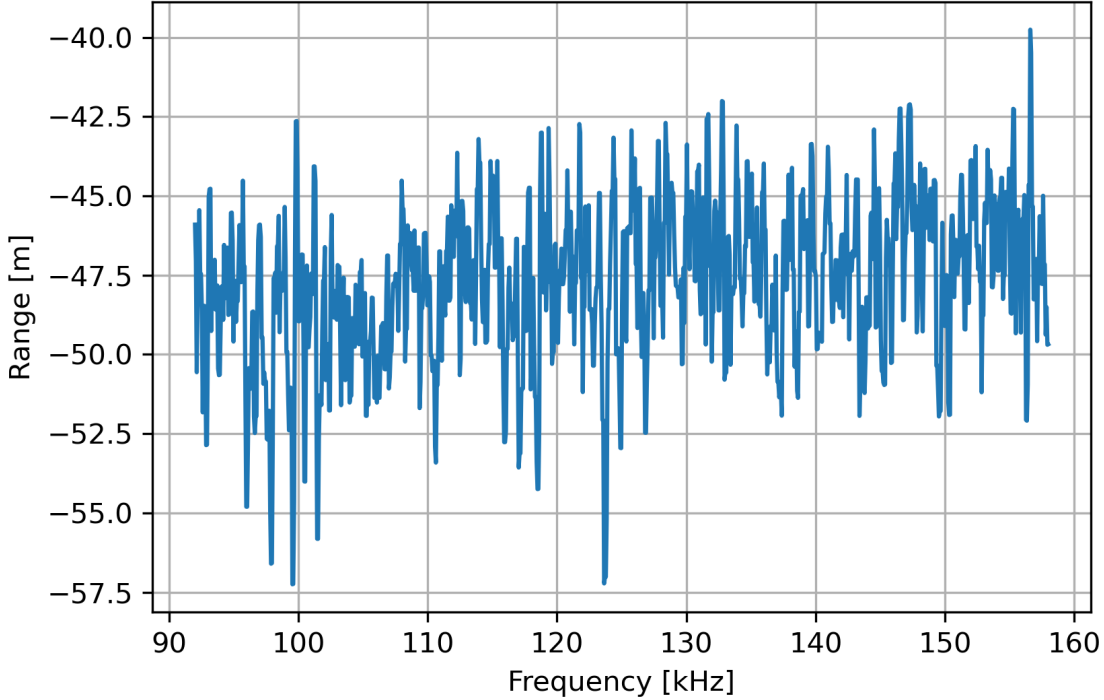


Figure 14. S_v as a function of frequency (m) averaged over a range interval covering a fish school

The trend for increasing S_v with frequency is well-known (?) and is consistent with the trend observed in this example. In contrast to data from isolated scatterers, such as metallic spheres, the benefit of pulse compression on the backscatter from an object that generates many overlapping echoes is not immediately obvious (fig. 12).

V. DISCUSSION

The use of broadband signals in fisheries acoustics is a developing field, and our contribution represents a comprehensive description of data processing steps. The inclusion of a corresponding code ensures a complete description. One could also envision that signal processing steps designed for specific purposes beyond these standard parameters can be developed, and our code may then serve as a starting point for further development. We also envision that the paper and code will be used both for educational purposes and for understanding the standard signal processing.

The contribution includes all the minor details required for the processing steps to work in practice. This includes steps well founded in the literature as well as practical and more ad-hoc choices, and a few examples of such choices follows. How to handle gaps in the calibration data, the choice of transmit pulse including tapering, calculation of efficient pulse duration, and decimation factors and filtering are examples of such choices. When convolving the received signal with the transmit pulse, there are various approaches to handle edge cases, and, in our implementation, we simply chose to remove the edge cases. We have also assumed a four-sector transducer, and the code must be adapted to other beam configurations if other configurations are needed. When estimating $TS(f)$ and $S_v(f)$ the resolution and accuracy will depend on the length N_{DFT} of the Fourier transform and the actual choice will be a compromise between accuracy and computational speed. Our objective is not to provide an evaluation of all these choices, but to serve as a reference implementation when benchmarking the choices.

$TS(f)$ is a common metric for studying single targets, used to extract features for single individuals. These features includes size, target classification and is used to extract behaviour through tracking. In our implementation, the target strength by frequency calculation assumes that a single target has successfully been identified. For this to work in proactive a robust single target detector is required. Several single echo detection algorithms exist, and various single echo detection algorithms may be required depending on situation. Typical SED algorithms are based on traditional single frequency pulses, and utilizing the information in the broadband data improved SEDs may be envisioned, but this is outside the scope of this paper.

$S_v(f)$ is a key parameter for echo integration. To estimate $S_v(f)$ a Fourier transform is used, repeatedly applied via a sliding window in range. The recommended size of the window is two times the pulse length, and is chosen as a compromise between spatial and frequency resolution. Since the duration of the sliding window can cause the spreading loss compensation to differ between the beginning and end of the window, the compensation for spreading loss is performed on the pulse compressed time domain data before the transform. Absorption loss compensation is also range dependent (and frequency dependent), but is insignificant for typical operating frequencies at short ranges and between the beginning and the end of the window at longer ranges. The compensation for absorption loss is therefore performed after applying the discrete Fourier transform. The choice of window also allow the data to be cast on a predefined range-frequency grid, which can be used to fit data across

transducers to an n-dimensional tensor typically employed by machine learning frameworks like Keras, Pytorch, and Tensor flow.

The formulation presented in this paper results in several frequency dependent parameters, such as transducer gain, two-way equivalent beam angle, and the water absorption coefficient, that are required to quantitatively estimate $TS(f)$ and $S_v(f)$ from received broadband signals. Methods to estimate these are not within the scope of this paper, but common practise is to use the conventional sphere backscatter calibration methodology (?) slightly enhanced for broadband (??). We note that these methods do not provide an operational method to estimate τ_{eff} or $\Phi(f)$, especially for ship-mounted transducers, and that empirical measurements of these parameters are necessary to fully calibrate both narrowband and broadband echosounders.

A set of equations for calculating calibrated, frequency-dependent, target strength and volume backscatter from broadband echosounder signals have been presented. Providing an open-source resource for anyone interested in learning and developing the techniques further.

VI. CONCLUSION

A set of equations for calculating calibrated, frequency-dependent, target strength and volume backscatter from broadband echosounder signals have been presented, with reference to the Simrad EK80 echosounder.

VII. DATA AVAILABILITY STATEMENT

The data associated with this article are available on request from the authors.