Separating overlapping echolocation: An updated method for estimating the number of echolocating animals in high background noise levels

Rebecca A. Hamilton, Josefin Starkhammar, Stefanie K. Gazda, et al.

Citation: The Journal of the Acoustical Society of America **150**, 709 (2021); doi: 10.1121/10.0005756
View online: https://doi.org/10.1121/10.0005756
View Table of Contents: https://asa.scitation.org/toc/jas/150/2
Published by the Acoustical Society of America

JASA ARTICLE

Separating overlapping echolocation: An updated method for estimating the number of echolocating animals in high background noise levels

Rebecca A. Hamilton,[1,a)] Josefin Starkhammar,[2,b)] Stefanie K. Gazda,[3] and Richard C. Connor[1]
[1]*Biology Department, University of Massachusetts Dartmouth, North Dartmouth, Massachusetts 02747, USA*
[2]*Department of Biomedical Engineering, Lund University, Lund, Sweden*
[3]*Department of Biology, University of Florida, Gainesville, Florida 32611, USA*

ABSTRACT:
Much can be learned by investigating the click trains of odontocetes, including estimating the number of vocalizing animals and comparing the acoustic behavior of different individuals. Analyzing such information gathered from groups of echolocating animals in a natural environment is complicated by two main factors: overlapping echolocation produced by multiple animals at the same time, and varying levels of background noise. Starkhammar *et al.* [(2011a). Biol. Lett. **7**(6), 836–839] described an algorithm that measures and compares the frequency spectra of individual clicks to identify groups of clicks produced by different individuals. This study presents an update to this click group separation algorithm that improves performance by comparing multiple click characteristics. There is a focus on reducing error when high background noise levels cause false click detection and recordings are of a limited frequency bandwidth, making the method applicable to a wide range of existing datasets. This method was successfully tested on recordings of free-swimming foraging dolphins with both low and high natural background noise levels. The algorithm can be adjusted via user-set parameters for application to recordings with varying sampling parameters and to species of varying click characteristics, allowing for estimates of the number of echolocating animals in free-swimming groups.
 https://doi.org/10.1121/10.0005756
(Received 25 March 2021; revised 7 June 2021; accepted 9 July 2021; published online 2 August 2021)
[Editor: James F. Lynch] Pages: 709–717

I. INTRODUCTION

Toothed whales and dolphins (odontocetes) use two broad categories of vocalizations: whistles and clicks (Janik, 2009; Tyack, 2019). Whistles are narrowband and frequency-modulated vocalizations used for social communication in most odontocete species (Morisaka, 2012; Tyack, 2019). These tonal vocalizations have been best studied in bottlenose dolphins (*Tursiops spp.*), which have typical whistle durations of 0.06–5.4 s and fundamental frequencies up to 41 kHz (Hiley *et al.*, 2017; Kaplan and Reiss, 2017; Jones *et al.*, 2020). Clicks can be modified for use in both navigation (i.e., echolocation clicks) and communication (e.g., burst pulse signals) (Arranz *et al.*, 2016).

The echolocation clicks of most toothed whales (most members of the family Delphinidae and all members of families Iniidae, Lipotidae, Monodontidae, Platanistidae, and Ziphiidae) are broadband, with central frequencies ranging between 33 and 102 kHz (Au, 2018; Fenton *et al.*, 2014). Specialists at hunting squid at long ranges, the clicks of sperm whales (family Physeteridae) are also broadband, but are particularly low in central frequency (15 kHz) and commonly documented as multipulsed (Møhl *et al.*, 2003; Fenton *et al.*, 2014). Some species produce relatively narrowband high frequency clicks (members of the families Phocoenidae, Pontoporiidae, and Kogiidae, the genus *Cephalorhynchus*, as well as *Lagenorhynchus cruciger* and *Lagenorhynchus australis*) with central frequencies between 123 and 142 kHz (Morisaka and Connor, 2007; Kyhn *et al.*, 2013; Fenton *et al.*, 2014; Jensen *et al.*, 2018; Galatius *et al.*, 2019). All echolocation clicks are short duration (<300 μs), directional, and are emitted in trains with interclick intervals (ICIs) up to 2 s depending on the distance of the target object from the animal, usually with a "terminal buzz" of decreasing ICI upon target approach (Fenton *et al.*, 2014; Au, 2018; Ridgway *et al.*, 2015; Miller *et al.*, 2004). As with whistles, echolocation clicks have been best studied in bottlenose dolphins (*Tursiops spp.*), which have typical click train durations of 0.1 to 5.0 s (Lilly and Miller, 1961; Finneran *et al.*, 2014; Jones *et al.*, 2020; Starkhammar *et al.*, 2011a; Starkhammar *et al.*, 2019).

Other than differences in central frequency, certain differences in characteristics of echolocation clicks (such as rises, notches, and peaks) between families have been used to build computational analysis tools that can distinguish

[a)]Electronic mail: rebeccahamilton64@gmail.com, ORCID: 0000-0002-0739-3746.
[b)]ORCID: 0000-0003-0514-7397.

clicks generated by odontocetes of different species (Roch *et al.*, 2011, Klinck and Mellinger, 2011; Palmer *et al.*, 2017; Yang *et al.*, 2020; Soldevilla *et al.*, 2008). However, using echolocation to distinguish between individuals of the same species, when clicks have the same species-specific characteristics, must rely on other measures.

Studies investigating echolocation production at the individual level can be used to estimate the number of vocalizing individuals and analyze behavior within groups, including answering questions related to cognition. For example, captive common bottlenose dolphins (*Tursiops truncatus*) have been shown to "eavesdrop" on the echolocation of others (Xitco and Roitblat, 1996), but examples of this in their wild counterparts are rare [Götz *et al.* (2006) and Gregg *et al.* (2007), but see Alcázar-Treviño *et al.* (2021)]. Additionally, analysis of echolocation can reveal differences in acoustic behavior and individual roles during group behaviors. Based on inter-click-intervals (ICI), Benoit-Bird and Au (2009) proposed that spinner dolphins use echolocation to determine the location of conspecifics during cooperative foraging. Studies of groups are hindered by the difficulty of analyzing overlapping echolocation in recordings of groups of free-swimming animals. For example, the individual ICI becomes difficult to assess when more than one animal echolocates concurrently. It is also particularly challenging when source localization is not possible, as in cases where animals are spatially close to each other or when an insufficient number of synchronized hydrophones are available. Several studies have attempted to obtain estimates for the number of echolocating animals from a single recording by comparing various click characteristics in order to separate overlapping echolocation click trains (Halkias and Ellis, 2006; Zaugg *et al.*, 2013, Bahl *et al.*, 2003, Baggenstoss, 2011). However, algorithms applied to the unusual click characteristics of sperm whales, the subject of many of these studies, may not prove useful for similar studies in other toothed whales. Additionally, many of these methods are computationally taxing, extracting a high number of parameter measures from each click [28 features in the case of Baggenstoss (2011)]. One study on beluga whales used inter-click interval as a single parameter to separate overlapping click trains, but this method proved unreliable when burst pulse signals were present in the data due to the rapid increase in ICI (Le Bot *et al.*, 2015).

A computationally inexpensive MATLAB-based click group separation (CGS) program (Starkhammar *et al.*, 2011b) was created to separate overlapping echolocation based on a different single metric: correlation of frequency spectra between individual clicks. With this CGS algorithm, the number of distinct echolocation click groups can be determined from recordings of multiple animals taken from a single hydrophone. The number of echolocation click groups detected can be used as a proxy for estimating the number of animals echolocating in a given recording. However, this program was developed for a controlled experimental setting: dolphins echolocated on targets positioned behind or in front of a 2D array of hydrophones in a relatively low background noise environment. The animals were swimming freely and in spontaneous groups inside a natural pen, and the hydrophone array was mounted on the netting surrounding the pen. Additionally, the recordings were hardware triggered and set to record a narrow time window around each click in a series with a high click rate (burst-mode sampling). Excluding the silence between clicks enabled simultaneous full sampling on 47 separate hydrophone channels while still keeping file sizes manageable during long periods of burst-mode recordings of echolocation activity. The original CGS algorithm, however, only utilized the output from one of the channels; the channel that recorded the highest amplitude during a click train sequence. When the algorithm was tested on continuous recordings from a shallow, noisy environment (recorded in another setting and with another hardware altogether), the program struggled to accurately separate click groups. Especially problematic were broadband, short duration sounds similar in frequency spectra to dolphin clicks, such as those produced by snapping shrimp, which were falsely identified by the program as echolocation clicks (Figs. S1, S2, S3 in supplementary material[1]). Such sounds artificially overestimated the number of click groups in the program output. The method also relied on the assumption that the frequency spectra varies gradually from click to click within a single click train. However, this assumption will hold only if successive clicks are recorded in the same region of the echolocation beam. Since the spectral content of the clicks varies across the cross section of the echolocation beam it is clear that the accuracy of an algorithm reliant upon the spectral content of successive clicks will also be dependent on the movement of the animal relative to the recording hydrophone. It will be less accurate if the animal is scanning rapidly across the hydrophone than if the hydrophone records the click train in the same position within the echolocation beam for every click. Thus, in cases where the animal is travelling fast or moving its head side to side while scanning, additional acoustic click information less impacted by the relative axis to the recording device is likely to improve the CGS method. One study on bottlenose dolphins showed a promising method for using both inter-click interval and click-to-click amplitude to separate overlapping echolocation (Lepper *et al.*, 2005).

This paper details significant modifications to the original CGS algorithm in order to improve results from under sampled recordings in environments of relatively high background noise levels. The accuracy of this new algorithm was tested using acoustic clips with known numbers of free-swimming echolocating animals in an environment with high background noise. The resulting algorithm is also compared with the original algorithm using the same data set as presented in the original publication.

II. METHODS

A. Data collection

Data used in this study came from two different sources. The first source of data were also used in the original

https://doi.org/10.1121/10.0005756

test of the CGS algorithm (Starkhammar *et al.*, 2011b). This was a recording of 19 free-swimming and spontaneously echolocating bottlenose dolphins (*Tursiops truncatus*) in a large open-sea pen with a maximum depth of 5 m, a sandy sea floor, and minimal reflective surfaces, thereby having very low background noise levels. Recordings were made at a sampling rate of 1 MHz and 12-bit resolution, with recordings of 150 μs time windows, plus a pre-trig window of 40 μs, designed to capture only clicks and not the silence between them, thereby limiting recording of samples representing silence, and thus also minimizing the data file size during long recording sessions. This hardware-based file size minimizing recording method was crucial for enabling hour-long simultaneous sampling of the 47-element hydrophone array with 1 MHz sampling rate. The full setup of this data collection and data parameters was described in Starkhammar *et al.* (2009) and Starkhammar *et al.* (2011b). Only the channel that recorded the strongest acoustic amplitude level during any given click train was used. One segment of recorded echolocation that was manually determined to contain overlaying click trains from multiple animals was selected as an example dataset. The manual classification was based on the visually apparent abrupt changes in amplitude and ICIs. *A priori* information on the number of acoustically active animals within the group of 19 individuals present is not available for this dataset. Therefore, additional data in which the number of echolocating animals could be known were also used.

The second source of data were acoustic recordings made of single free-swimming bottlenose dolphins in the inshore areas along the west coast of Florida, within 35 km south of the Cedar Keys. The environment consists of mostly shallow seagrass beds and oyster bars among small islands and inlets with an average depth of 1 m. This area is normally characterized by high biological background noise levels (average 181 +/− 7.4 dB re: 1 μPa over random sampling of 50 five-second samples), as is typical for shallow warm waters worldwide (Potter *et al.*, 1997). Underwater recordings were made during boat-based surveys from May to August 2018, using an HTI-96-MIN hydrophone (flat frequency response: 0.002–30 kHz +/− 1 dB) kept 0.5 or 0.25 m below the surface, depending on available depth. The sensitivity of the hydrophones including preamplification was −173 dB re 1 V/μPa. This dataset was sampled with the continuous sampling mode, as opposed to the first dataset, which was collected with burst-mode sampling. Recordings were made at a sampling rate of 96 kHz on a multitrack recorder with in-built anti-aliasing filter eliminating frequencies over 48 kHz (Tascam DR-680MKII). Data of this sampling rate were available because whistle vocalizations, which have fundamental frequencies only up to 41 kHz (Hiley *et al.*, 2017; Kaplan and Reiss, 2017), were the original focus of this data-collection protocol. We recognize that this frequency range excludes the higher frequencies of echolocation clicks but believe this is a strength rather than a weakness of the method, making it equally applicable to datasets originally designed for studies of other vocalizations. This means there is no requirement for dedicated fieldwork and data collection for the specific purpose of answering questions related to overlapping echolocation, which may not be possible for many research operations.

Whenever possible, the boat engine was turned off to reduce anthropogenic background noise. Concurrent synced video was captured using a handheld video camera (Sony Handycam HDR-SR11). All audio recordings were visualized as spectrograms in Adobe Audition v6.0 (Adobe Systems, down-sampled to 48 kHz, FFT length: 1024, Blackmann-Harris window). Clips of audio during active foraging by single animals were manually isolated for analysis if they met three criteria for audio quality assurance: (1) the event occurred while the focal animal was within 100 m of the research boat, (2) the boat engine was off, and (3) dolphin vocalizations were visible on the spectrogram over background noise.

During these recordings, visual observations and review of video footage confirmed that only the focal animal was present within 100 m of the hydrophone. Therefore, there was confidence that only the focal animal's echolocation was captured on these acoustic recordings. Echolocation was present in all audio clips extracted from foraging events. A lower limit frequency cutoff of 16 kHz was introduced in Adobe Audition to partially reduce low-frequency background noise, thus restricting subsequent analysis to the frequency band from 16 to 48 kHz.

B. Original algorithm

A classical threshold method was used to recognize clicks within an audio clip: any sample with an amplitude over the user-provided amplitude trigger level was analyzed. This parameter was set by the user depending on the amplitude of clicks over general background noise for each individual clip. All samples within a given time window, placed around a trig-event with the trig-event positioned in the middle of that time window, were analyzed as one single click. The total length of the time windows where 150 μs (150 samples) for the first, burst-mode sampled, dataset and 802 μs (77 samples) for the continuously sampled second dataset. For each subsequent click in an audio clip, the original algorithm compared the frequency spectral correlation between each click and the proceeding click, calculated according to

$$C_k = \frac{\sum_i U_k(f_i) \cdot U_{k+j}(f_i)}{\max\left[\sum_i U_k^2(f_i), \sum_i U_{k+j}^2(f_i)\right]}, \quad (1)$$

where $k = 1, 2, 3, \ldots, N-1$; $N =$ number of measured clicks in the echolocation sequence; $i = 1, 2, 3, \ldots, M$; $M =$ number of frequency components in the frequency spectra; $C_k =$ the correlation coefficient of the kth and $k+1$th clicks; $f_i =$ the ith frequency component in the spectrum; $U_k(f_i) =$ the

relative amplitude of the ith frequency component in the spectrum of the kth click; and $j = \text{lag} = 1$.

For clicks with spectral frequency of a similarity exceeding the user-defined spectral resolution threshold, clicks would be grouped together into click groups. The code restricted stepwise comparisons to 30 proceeding clicks to search for a reasonable match between the click being considered and previous click groups. Given how the algorithm works, "click groups" are defined as groups of clicks produced in sequence by the same animal rather than being defined by the any relative timing of clicks.

C. Algorithm modification to increase CGS algorithm accuracy

In addition to frequency spectra, two new properties of each click were measured: peak-to-peak amplitude and inter-click interval. The use of these metrics for comparison assumed that across consecutive clicks from the same animal, these characteristics change gradually rather than suddenly. Specifically, the new program used a decision-tree, classifying each click based on the results of specific comparisons (Fig. 1). Again, the method may compare the click with up to 30 proceeding clicks in order to group clicks together, preventing the number of click groups being overestimated due to interruption.

By introducing ICI as a comparison metric in the method, a new limitation was introduced. In cases where echolocation was continuous but contained clicks that fell below the amplitude trigger level set for that clip (as might happen if the dolphin momentarily turned off-axis from the hydrophone), clicks would be grouped by the program as separate groups based on a large inter-click interval. To reduce this overestimation of click group estimates, a parameter (trigger level ICI tolerance) was added that removes ICI as a characteristic for comparison for any clicks that had an amplitude within a set amount above the trigger level for the clip.

Because the algorithm was being tested on both data collected with burst mode sampling as well as continuous sampling, a Window Taper Parameter was added to suppress signals in the start and end of each click sampling window. This is needed in the rare event that two clicks from different animals coincide in time within one single time window or if other temporal noise sources interrupt the data. Such coinciding signals would cause the signal snippet to not end on an amplitude equal to zero. This would in its turn introduce false frequency components in the fast Fourier transform (FFT). Therefore, each data snippet (time window containing data) was windowed using the Tukey window function in MATLAB (The Mathworks, Natick, USA). The tapering parameter used as input to the Tukey window function was set to 0.6 in the burst-mode sampled dataset and 0.3 in the continuously sampled dataset. The different values of the tapering parameters here reflect adjustments to the two different sample modes and sample parameters used for recording the two different datasets. All adjustable algorithm parameters, their descriptions, units, and values are listed in Table I.

Despite removing all sound below 16 kHz from the continuously sampled data, there were still frequent intense background noises (such as snapping shrimp) incorrectly identified by the CGS program as echolocation clicks since

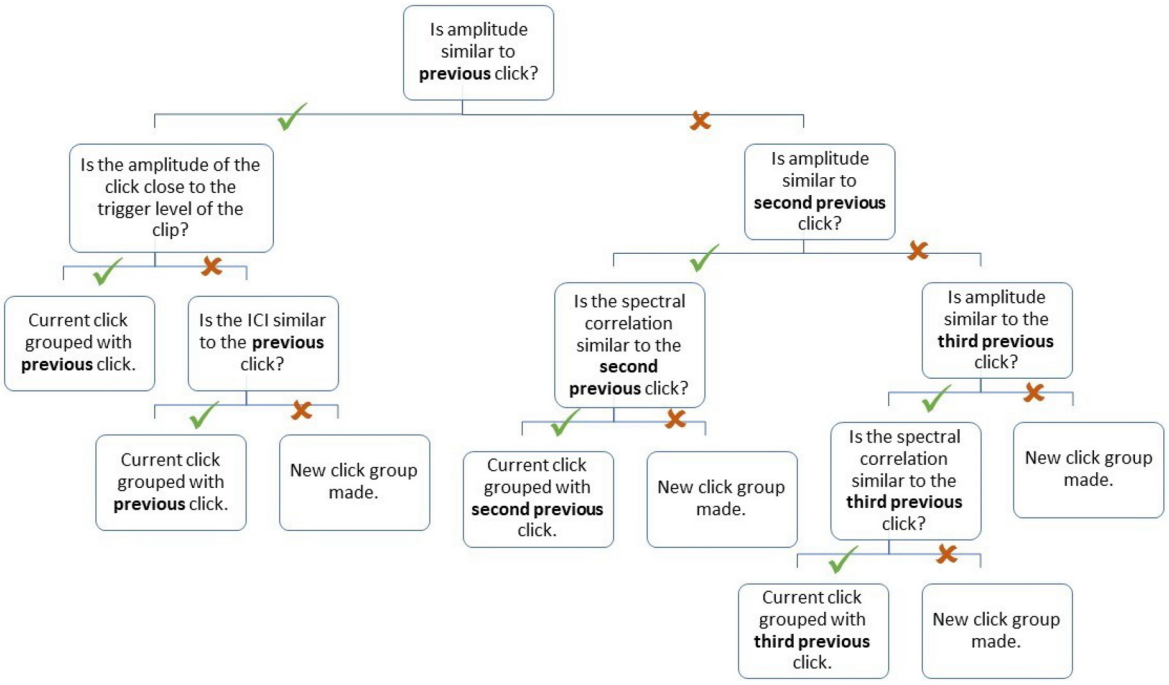

FIG. 1. (Color online) Decision tree used by custom MATLAB program to classify clicks into different click groups by comparing characteristics of each click with those of the proceeding clicks. The program would compare with up to 30 previous clicks to determine if there was a reasonable match based on click characteristics.

https://doi.org/10.1121/10.0005756

TABLE I. Parameter descriptions and values used in the final version of click group separation program as determined by manual fine-tuning. Two parameters (trigger level and LPRL) were set per data clip (section of interest within an audio recording) depending on SNR.

Parameter	Description	Units	Value
Trigger level	Minimum amplitude of clicks included in analysis, set per clip	Pascals/Volts	0.447 to 446.79/0.001 to 1
Peak-to-peak amplitude threshold	Maximum allowed amplitude difference between clicks grouped together. Calculated as a function of the chosen trigger level.	Pascals/Volts	72.2(Trigger Level^2) + 1.2(Trigger Level) + 0.12
Inter-click interval threshold	Maximum allowed ICI difference between clicks grouped together	Number of samples	10 000
Trigger level ICI tolerance	Minimum amplitude difference between trigger level and click for which ICI would be used to determine click similarity	Pascals/Volts	10% of Trigger Level
Spectral resolution threshold	Minimum correlation value of frequency spectrum content between clicks grouped together	Proportion	0.9
Low percentage removal limit	All detected peaks belonging to relatively unpopulated click groups (below this percentage value) are removed for data analysis	Percentage	0% to 12%
Window taper parameter	Tapering parameter of the click windows, preventing false frequency components in the analysis	Unitless	0.6 (burst-mode data) 0.3 (continuous data)
Previous click comparison number	The number of clicks earlier in the sequence for which the algorithm compares parameters	Number of clicks	30

all signal levels over the given threshold were included in the analysis. An additional procedure was added to allow for more accurate click group separation given such noise. This procedure used an added parameter, the low percentage removal limit (LPRL). The LPRL procedure assumes that background noise peaks that are falsely classified as new clicks trains account for only a small percentage of all detected click groups since they are unlikely to be consistent in ICI, amplitude, and frequency spectra. Additionally, they should be visually distinguishable from echolocation clicks on spectrograms by an experienced human observer. This was the case here in all clips used to test the program. When originally running the program, the LPRL parameter would be set to 0%. A manual operator compared the program output to the spectrogram of the same clip, thereby identifying any background peaks incorrectly identified as click groups by the program. The LPRL was then set manually to a percentage 1% higher than the output percentage of total clicks for any mis-identified click groups consisting of background peaks. When run again, the program runs normally in a first iteration, then re-runs while ignoring any "clicks" belonging to click groups representing less than the set percentage value (see Figs. S1, S2, and S3 in the supplementary material,[1] for example). This procedure to remove likely noise peaks helps to restore real click trains' true ICI values and peak-to-peak amplitude threshold values. Since these values are used when successive clicks are compared in the process of grouping the clicks together with earlier click train groups, the re-run of the algorithm will group click groups more accurately.

D. Algorithm evaluation

Manual fine-tuning of appropriate algorithm parameter values was done using data from both sources. On the burst-mode sampled data clip, the number of actively echolocating animals was not known (<19), so performance could only be judged by visually comparing result outputs when using different parameter values, comparing the output to that of the original algorithm version described in Starkhammar *et al.* (2011b) and to what could be expected from visually inspecting the peak-to-peak amplitude and ICI differences between clicks in the data. Then clips from the continuously sampled dataset, known to contain exactly one echolocating animal, were used to further determine the code's performance and adjust set parameters as needed (N = 9, 3 different individuals). These clips ranged from 1 to 11 s in duration, with a mean of 5.11 s. Finally, combinations of single-animal echolocation clips of different individuals were artificially merged using ADOBE AUDITION v6.0 (Adobe Systems) to create a second set of clips with overlapping and/or interrupting echolocation of exactly two animals (N = 11). Using these two sets of clips known to contain the echolocation of exactly 1 and 2 animals, respectively, allowed for finalizing parameter values for optimal performance of the algorithm. Only the trigger level and LPRL values required adjustment per data clip, depending on the background noise level of each (see Table I for all final parameter values).

To test the accuracy of the final version of the algorithm with the determined parameter values, the algorithm was tested on an additional nine merged clips containing echolocation of two different individuals and on a set of merged clips containing echolocation of three different individuals (N = 12). In order to test the hypothesis that the number of click groups as determined by the algorithm would significantly differ between clips known to contain one, two, and three animals, a generalized linear regression with Poisson distribution was used. Clips used in both parameter finetuning and in testing the final parameters were combined for

use in this model testing in order to achieve a high enough sample size to satisfy model assumptions.

III. RESULTS

Using the new version of the CGS algorithm, the burst-mode sampled data clip had 163 detected clicks, the same number as in Starkhammar *et al.* (2011b). The algorithms differed in the grouping of clicks. Visual inspection of the clicks indicates that the new algorithm made better matching decisions than the original version due to observable differences in ICI and peak-to-peak differences being more easily explained in the new version (Fig. 2).

In the continuously sampled dataset, each of the nine clips known to contain one echolocating individual (three different individuals) had a mean of 120.11 clicks detected (range: 9–401, SD: 118.26), grouped into a mean of 1.67 click groups (range: 1–3, SD: 0.71), meaning that the number of click groups was correctly estimated in four cases (44%), overestimated in five cases (55%), and never underestimated. In the 12 clips known to contain two animals echolocating used in parameter finetuning, the algorithm resulted in a correct output for the number of click groups in five cases (45.5%), overestimated in six cases (54.5%), and never underestimated. In examining both parameter finetuning and testing data clips known to contain two different animals echolocating ($N = 20$), there were an average of 171.4 clicks detected (range: 42–446, SD: 107.14) and grouped into a mean of 3.15 click groups (range: 2–6, SD: 1.272). In 12 clips known to contain three different animals echolocating, an average of 220.92 clicks were detected (range: 82–498, SD: 113.53) and grouped into a mean of 3.92 click groups (range: 3–5, SD: 0.669). Statistical analysis revealed significant differences in the number of click groups detected between the groups of clips containing single, two, or three echolocating animals (glm: single animal versus two animals: $Z = 2.216$, $p < 0.05$; single animal versus three animals: $Z = 2.881$, $p < 0.05$; two animals versus 3 animals: $Z = 0.218$, $p > 0.05$; Fig. 3 and Tables S1a, S1b in the supplementary material[1]). The adjustable parameter values used and click group results for each data files are given in Table S2.[1]

IV. DISCUSSION

The results of testing the updated CGS algorithm showed the program's ability to perform well on continuous

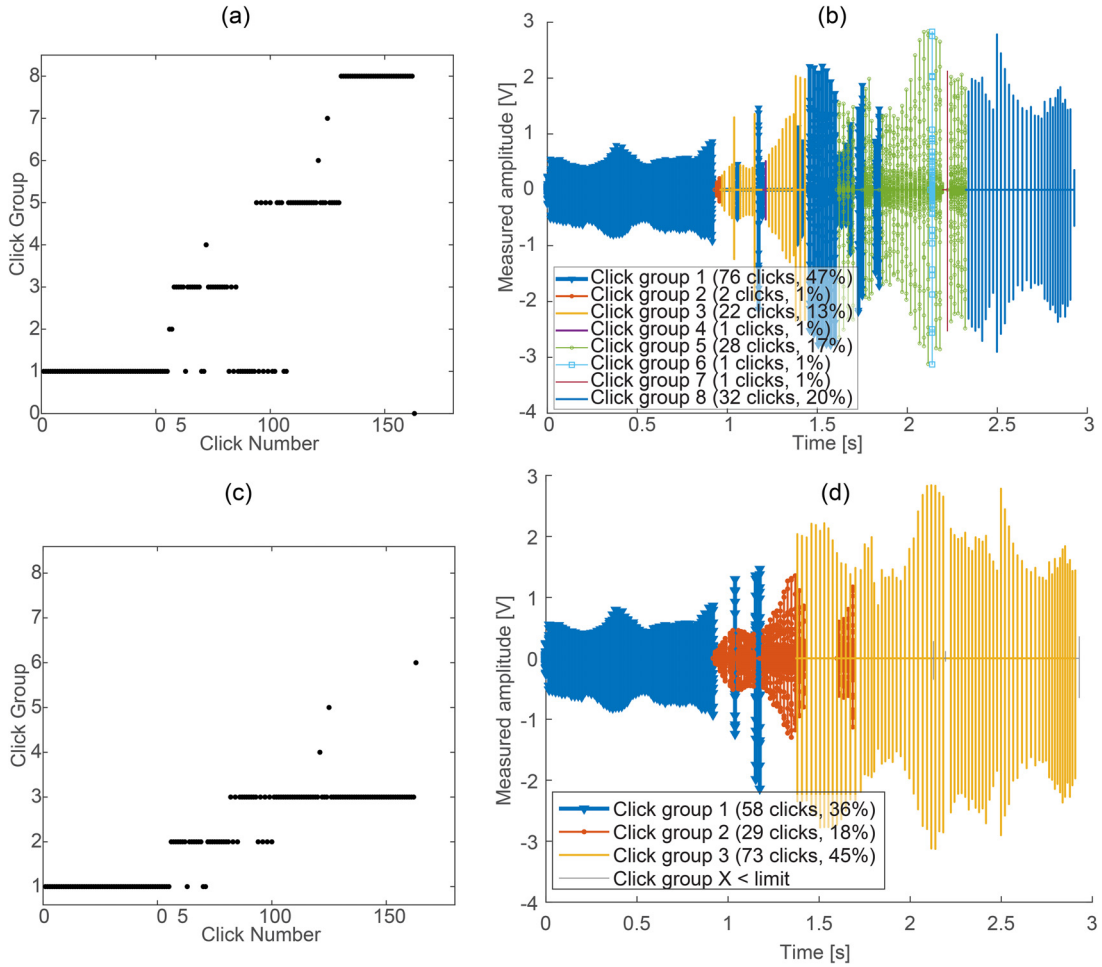

FIG. 2. (Color online) Automatic assignment of clicks to click groups from a burst-mode sampled recording of a group of free-swimming dolphins, for which the number of echolocating animals was not known *a priori*. Click groups are shown by color and line differences. (a) and (b) show the grouping made by the original version of the algorithm used in Starkhammar *et al.* (2011b) while (c) and (d) shows the grouping made by the updated algorithm.

https://doi.org/10.1121/10.0005756

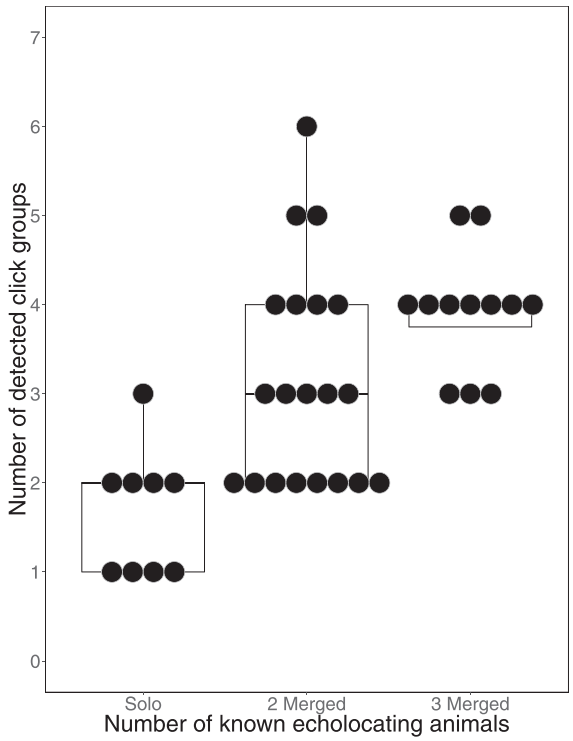

FIG. 3. Number of click groups detected in clips known to contain the echolocation of one, two, and three animals. Each dot represents a single data point. Dots were spread out horizontally to prevent them from overlapping. Asterisks indicate significant differences (**$p < 0.01$, *$p < 0.05$).

single hydrophone recordings of free-swimming dolphins. Using additional click parameters amplitude and inter-click-interval in combination with frequency spectral correlation allowed more clicks to be correctly grouped together, even in the case of relatively high background noise levels and a limited recording bandwidth in the continuously recorded data.

Use of the LPRL parameter eliminated the problem of background noises being classified as new click groups. It also allowed for correct grouping of clicks in trains together in cases when they were initially classified as separate due to such interrupting background noise. The results in Fig. 3 show that the algorithm never underestimated the number of echolocating animals, meaning the algorithm can be used reliably to determine the maximum possible number of animals echolocating in a clip.

The main data used in this study to design and test the method was collected at a sampling rate of 96 kHz since the original purpose of the data collection was to capture whistles. Despite not having captured the higher frequencies of echolocation clicks, the CGS algorithm performed well on this data. Therefore, we are confident that the CGS method can be successfully applied to *Tursiops spp.* acoustic datasets already in existence with similar recording parameters without requiring dedicated data collection for applying this method, which may not be feasible for many research operations.

We caution users of the CGS algorithm to be aware that the final version has one major limitation known to affect output results. This occurs if a continuous echolocation fades momentarily on the recording, as would happen when an animal temporarily moves off-axis from the hydrophone. In this case, resulting time-gaps between detected clicks of larger than 10 000 samples could cause classification of clicks into different click groups. The trigger level ICI tolerance parameter is designed to prevent this from occurring, but it will still be possible in cases where click amplitude is high enough above trigger level to not be caught by this parameter. We offer that this limitation can be overcome by manually removing classification of different click groups caused solely by ICI in these cases. However, caution should be taken in such an approach as this adjustment instead could introduce a false underestimation of the number of click groups. Users should decide if such manual adjustment should be used in each case, which will be determined by whether the user is more interested in estimating the number of click trains or the number of echolocating animals, and additionally by how they define a "click train." This limitation favors overestimating the number of click groups rather than underestimating. This allows for confidence that the number of click groups detected represents the maximum number of echolocating animals in a clip. In instances where the output shows a single click group, users can have confidence that only a single animal is echolocating in that given recording. Determining whether a single animal or multiple animals are echolocating in a recording may prove to be the most useful distinction in the study of acoustic role partitioning within groups, such as detecting the use of "eavesdropping" by group members on the echolocation of a single individual. This could be determined if, for example, the CGS method output shows only one animal in a group echolocating and other individuals change their behavior without echolocating themselves.

This updated CGS method, with flexible user-set parameters and protocols for limiting the influence of background peaks on results, should make this method applicable to analysis of datasets from a wide range of recording conditions, background noise levels, and studied broad-band echolocating species. Given the reliable performance of the algorithm on the current dataset with relatively high background noise, we expect the algorithm to perform equally as well on any data with equal or lower levels of background noise.

The updates made to CGS algorithm described here greatly improved performance of the algorithm on both burst-mode and continuous recordings of multiple echolocating animals. We expect that the algorithm will perform equally well on echolocation of other broad-band echolocating toothed whales with similar frequency range, amplitude, and ICI. The only assumption that needs to be met by any echolocation being analyzed by the CGS algorithm is that characteristics of clicks produced by the same animal will change gradually over sequential clicks and will, thereby, be more similar to each other than clicks produced by other animals in the same recording. While further testing will be necessary to determine if this method is reliable when used

with clicks modified for communication such as pops or burst pulses (Smolker and Connor, 1996; Arranz *et al.*, 2016; Ridgway *et al.*, 2015), or with odontocetes known to produce echolocation clicks of varying characteristics (i.e., sperm whales or harbor porpoises). For harbor porpoises and the 12 other narrow-band high frequency signaling species, there is less variation in the frequency spectra of clicks between individuals (Kyhn *et al.*, 2013). Although testing would be required, we expect this would present a challenge for the current CGS algorithm. In this case, we suggest edits to allow the code to create click groups based only on amplitude and inter-click-interval. We hope that such flexibility of the CGS algorithm will make it useful for researching all odontocetes, regardless of species or context. These include determining when and how free-swimming animals may utilize "eavesdropping" (Xitco and Roitblat, 1996; Götz *et al.*, 2006; Gregg *et al.*, 2007) or how individuals may partition acoustic roles while performing cooperative behaviors, under both experimental and natural conditions [e.g., King *et al.* (2020)].

ACKNOWLEDGMENTS

Fieldwork would not have been possible without the generosity of the Cedar Key Dolphin Project board members, volunteers, and donors. This study was funded by a National Geographic Society Early Career Grant, The Swedish Research Council (Grant No. 2015-04340), and a Louis M. Herman Research Scholarship. All fieldwork was collected under NMFS scientific research permits 14450 and 21938.

[1]See supplementary material at https://www.scitation.org/doi/suppl/10.1121/10.0005756 for further details on algorithm performance, model outputs, and raw data.

Alcázar-Treviño, J., Johnson, M., Arranz, P., Warren, V. E., Pérez-González, C. J., Marques, T., Madsen, P. T., and Aguilar de Soto, N. (**2021**). "Deep-diving beaked whales dive together but forage apart," Proc. R. Soc. B **288**(288), 20201905.

Arranz, P., DeRuiter, S. L., Stimpert, A. K., Neves, S., Friedlaender, A. S., Goldbogen, J. A., Visser, F., Calambokidis, J., Southall, B. L., and Tyack, P. L. (**2016**). "Discrimination of fast click-series produced by tagged Risso's dolphins (*Grampus griseus*) for echolocation or communication," J. Exp. Biol. **219**, 2898–2907.

Au, W. W. (**2018**). "Echolocation," in *Encyclopedia of Marine Mammals*, edited by B. Würsig, J. G. M. Thewissen, and K. M. Kovacs, 3rd ed. (Associated Press, New York), pp. 289–299.

Baggenstoss, P. M. (**2011**). "Separation of sperm whale click-trains for multipath rejection," J. Acoust. Soc. Am. **129**(6), 3598–3609.

Bahl, R., Ura, T., and Fukuchi, T. (**2003**). "Techniques for segregation and classification of several vocalizing sperm whales for AUV-based localization applications," in *Oceans 2003. Celebrating the past… Teaming toward the Future*, IEEE Cat. No. 03CH37492, IEEE, pp. 457.

Benoit-Bird, K. J., and Au, W. W. (**2009**). "Cooperative prey herding by the pelagic dolphin, Stenella longirostris," J. Acoust. Soc. Am. **125**(1), 125–137.

Fenton, B. M., Jensen, F. H., Kalko, E. K., and Tyack, P. L. (**2014**). "Sonar signals of bats and toothed whales," in *Biosonar* (Springer, Berlin), pp. 11–59.

Finneran, J. J., Branstetter, B. K., Houser, D. S., Moore, P. W., Mulsow, J., Martin, C., and Perisho, S. (**2014**). "High-resolution measurement of a bottlenose dolphin's (*Tursiops truncatus*) biosonar transmission beam pattern in the horizontal plane," J. Acoust. Soc. Am. **136**(4), 2025–2038.

Galatius, A., Olsen, M. T., Steeman, M. E., Racicot, R. A., Bradshaw, C. D., Kyhn, L. A., and Miller, L. A. (**2019**). "Raising your voice: Evolution of narrow-band high-frequency signals in toothed whales (Odontoceti)," Biol. J. Linnean Soc. **126**(2), 213–224.

Götz, T., Verfuss, U. K., and Schnitzler, H. (**2006**). "'Eavesdropping' in wild rough-toothed dolphins (*Steno bredanensis*)?," Biol. Lett. **2**(1), 5–7.

Gregg, J. D., Dudzinski, K. M., and Smith, H. V. (**2007**). "Do dolphins eavesdrop on the echolocation signals of conspecifics?," Int. J. Comp. Psychol. **20**(1), 65–88.

Halkias, X. C., and Ellis, D. P. (**2006**). "Estimating the number of marine mammals using recordings of clicks from one microphone," in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, IEEE.

Hiley, H. M., Perry, S., Hartley, S., and King, S. L. (**2017**). "What's occurring? Ultrasonic signature whistle use in Welsh bottlenose dolphins (*Tursiops truncatus*)," Bioacoustics **26**(1), 25–35.

Janik, V. M. (**2009**). "Acoustic communication in delphinids," Adv. Study Behav. **40**, 123–157.

Jensen, F. H., Johnson, M., Ladegaard, M., Wisniewska, D. M., and Madsen, P. T. (**2018**). "Narrow acoustic field of view drives frequency scaling in toothed whale biosonar," Curr. Biol. **28**(23), 3878–3885.

Jones, B., Zapetis, M., Samuelson, M. M., and Ridgway, S. (**2020**). "Sounds produced by bottlenose dolphins (Tursiops): A review of the defining characteristics and acoustic criteria of the dolphin vocal repertoire," Bioacoustics **29**, 399–440.

Kaplan, J. D., and Reiss, D. (**2017**). "Whistle acoustic parameters in Atlantic spotted dolphins (*Stenella frontalis*) and bottlenose dolphins (*Tursiops truncatus*) in two locations in the Bahamas and comparisons with other populations," Aquatic Mammals **43**(4), 364–377.

King, S. L., Guarino, E., Donegan, K., McMullen, C., and Jaakkola, K. (**2020**). "Bottlenose dolphins can communicate with vocal signals to solve a cooperative task," Royal Soc. Open Sci. **8**(3), 202073.

Klinck, H., and Mellinger, D. K. (**2011**). "The energy ratio mapping algorithm: A tool to improve the energy-based detection of odontocete echolocation clicks," J. Acoust. Soc. Am. **129**(4), 1807–1812.

Kyhn, L. A., Tougaard, J., Beedholm, K., Jensen, F. H., Ashe, E., Williams, R., and Madsen, P. T. (**2013**). "Clicking in a killer whale habitat: Narrow-band, high-frequency biosonar clicks of harbour porpoise (*Phocoena phocoena*) and Dall's porpoise (*Phocoenoides dalli*)," PloS One **8**(5), e63763–12.

Le Bot, O., Mars, J., Gervaise, C., and Simard, Y. (**2015**). "Rhythmic analysis for click train detection and source separation with examples on beluga whales," Appl. Acoust. **95**, 37–49.

Lepper, P., Dumortier, N., Dudzinski, K., and Datta, S. (**2005**). "Separation of complex excholocation signal trains from multiple bio-sonar sources," *Proceedings of the International Conference Underwater Acoustic Measurements: Technologies &Results*.

Lilly, J. C., and Miller, A. M. (**1961**). "Sounds emitted by the bottlenose dolphin," Science **133**(3465), 1689–1693.

Miller, P. J., Johnson, M. P., and Tyack, P. L. (**2004**). "Sperm whale behaviour indicates the use of echolocation click buzzes 'creaks' in prey capture," Proc. R. Soc. London Ser. B: Biol. Sci. **271**(1554), 2239–2247.

Møhl, B., Wahlberg, M., Madsen, P. T., Heerfordt, A., and Lund, A. (**2003**). "The monopulsed nature of sperm whale clicks," J. Acoust. Soc. Am. **114**(2), 1143–1154.

Morisaka, T., and Connor, R. (**2007**). "Predation by killer whales (*Orcinus orca*) and the evolution of whistle loss and narrow-band high frequency clicks in odontocetes," J. Evol. Biol. **20**(4), 1439–1458.

Morisaka, T. (**2012**). "Evolution of communication sounds in odontocetes: A review," Int. J. Comp. Psychol. **25**(1), 1–20.

Palmer, K. J., Brookes, K., and Rendell, L. (**2017**). "Categorizing click trains to increase taxonomic precision in echolocation click loggers," J. Acoust. Soc. Am. **142**(2), 863–877.

Potter, J. R., Wei, L. T., and Chitre, M. (**1997**). "High-frequency ambient noise in warm shallow waters," Sea Surf. Sound **97**, 21–25.

Ridgway, S., Samuelson Dibble, D., Van Alstyne, K., and Price, D. (**2015**). "On doing two things at once: Dolphin brain and nose coordinate sonar clicks, buzzes and emotional squeals with social sounds during fish capture," J. Exp. Biol. **218**(Pt 24), 3987–3995.

Roch, M. A., Klinck, H., Baumann-Pickering, S., Mellinger, D. K., Qui, S., Soldevilla, M. S., and Hildebrand, J. A. (**2011**). "Classification of

echolocation clicks from odontocetes in the Southern California Bight," J. Acoust. Soc. Am. **129**(1), 467–475.
Smolker, R., and Connor, R. C. (**1996**). "Pop'goes the dolphin: A vocalization male bottlenose dolphins produce during consortships," Behaviour **133**(9-10), 643–662.
Soldevilla, M. S., Henderson, E. E., Campbell, G. S., Wiggins, S. M., Hildebrand, J. A., and Roch, M. A. (**2008**). "Classification of Risso's and Pacific white-sided dolphins using spectral properties of echolocation clicks," J. Acoust. Soc. Am. **124**(1), 609–624.
Starkhammar, J., Amundin, M., Nilsson, J., Jansson, T., Kuczaj, S. A., Almqvist, M., and Persson, H. W. (**2009**). "47-channel burst-mode recording hydrophone system enabling measurements of the dynamic echolocation behavior of free-swimming dolphins," J. Acoust. Soc. Am. **126**(3), 959–962.
Starkhammar, J., Moore, P. W., Talmadge, L., and Houser, D. S. (**2011a**). "Frequency-dependent variation in the two-dimensional beam pattern of an echolocating dolphin," Biol. Lett. **7**(6), 836–839.
Starkhammar, J., Nilsson, J., Amundin, M., Kuczaj, S. A., Almqvist, M., and Persson, H. W. (**2011b**). "Separating overlapping click trains originating from multiple individuals in echolocation recordings," J. Acoust. Soc. Am. **129**(1), 458–466.
Starkhammar, J., Reinhold, I., Moore, P. W., Houser, D. S., and Sandsten, M. (**2019**). "Detailed analysis of two detected overlaying transient components within the echolocation beam of a bottlenose dolphin (*Tursiops truncatus*)," J. Acoust. Soc. Am. **145**(4), 2138–2148.
Tyack, P. (**2019**). "Communication by sound and by visual, tactile, and chemical sensing," in Ethology *and Behavioral* Ecology *of* Odontocetes (Springer, Berlin), pp. 25–50.
Xitco, M. J., and Roitblat, H. L. (**1996**). "Object recognition through eavesdropping: Passive echolocation in bottlenose dolphins," Anim. Learn. Behav. **24**(4), 355–365.
Yang, L., Sharpe, M., Temple, A. J., Jiddawi, N., Xu, X., and Berggren, P. (**2020**). "Description and classification of echolocation clicks of Indian Ocean humpback (*Sousa plumbea*) and Indo-Pacific bottlenose (*Tursiops aduncus*) dolphins from Menai Bay, Zanzibar, East Africa," PloS One **15**(3), e0230319.
Zaugg, S., van der Schaar, M., Houégnigan, L., and André, M. (**2013**). "Extraction of pulse repetition intervals from sperm whale click trains for ocean acoustic data mining," J. Acoust. Soc. Am. **133**(2), 902–911.