Author's response to: Referee #2's comments on "Commonly used methods fail to detect known phase speeds of simulated signals of Sea Surface Height Anomalies" *by* Y. De-Leon and N. Paldor

Characterizing the properties of oceanic Rossby waves is central to understanding the role of the ocean in the climate as much of the response of the ocean to large-scale forcing is mediated by these waves. Indeed, this issue has attracted considerable attention across the ocean sciences, particularly since the advent of accurate altimetry measurements in 1992 when it became possible, in principle, to observe the signature of Rossby waves at the ocean surface, yet many aspects of such waves remain poorly understood. In particular, it has been found that observed phase speeds derived from altimetry data are systematically faster than the speeds suggested by the theory of Rossby waves. A number of explanations for the disagreement between observations and theory have been proposed, including the effects of the mean zonal flow and bottom topography or the fact that many of the westward-propagating features observed in the altimetry data are, in fact, eddies rather than Rossby waves.

The present study tests the ability of several methods to estimate the phase speed of Rossby waves on simulated data, and finds that such methods very often fail to estimate the true phase speed. The authors then conclude that this is the most likely reason for the differences between observed and theoretical phase speeds. The paper is well written, the figures are mostly adequate and clear, and the experiments designed to assess the skill of the various detection appear to have been conducted appropriately. Unfortunately, although the overall aim of the paper is worth pursuing, some aspects of the paper raise doubts and I do not believe that the results from the performed experiments support the authors' conclusion that "none of the methods is reliable for estimating the phase speed of Rossby waves in the real ocean". The authors are right in concluding that none of the methods is able to estimate the true phase speed in the simulated data, but this conclusion cannot be extrapolated to the observed data since, to the extent that I understand the issue, I don't think the simulated data provides an accurate representation of Rossby waves in the real ocean. In conclusion, I think that the manuscript requires substantial revisions and thus I cannot recommend it for publication as it stands. Details on my main concerns and other minor points are provided below.

Following the reviewer's comments, we will include in the revised manuscript oceanic phenomena other than Rossby waves in which the same radon Transform and 2d-FFT methods are employed. Our findings are relevant to all observations (e.g. near shore dynamics, eddy propagation) where propagation speeds are extracted from time-longitude diagram. Our choice of parameter ranges is drawn from the massive usage of the examined methods in the extraction of Rossby wave phase

speed from time-longitude diagrams of satellite observed SSHA signals.

Main points:

1. It is unclear to me from Section2.1 how exactly the simulated data are generated. The authors state that “The values of the phase speeds, C, are uniformly distributed in the 0 to -18 cm/s range”. Does that mean that each of the 20 or 50 modes is assigned a different phase speed within that range?

Yes, that’s exactly what we did. We will clarify it in the revised version of the manuscript.

Long Rossby waves in the ocean are approximately non-dispersive and so their phase speed is the same at all frequencies. Hence, assigning a different speed to every mode, if this is indeed what is done here, seems unjustified. Could you please clarify how exactly phase speed are ascribed to each mode? How do the results change if the same phase speed is used for all modes?

The emphasis is on “Long” while we include all wavenumbers, long and short, so the waves should be considered dispersive. In the case when all modes have the same phase speed, the 2D-FFT methods still fail in many cases (see the existing remark in the second paragraph of the Discussion) while the Radon transform methods will probably detect the phase speed correctly (we will add a note to this effect in the same paragraph).

Also, the range -18 to 0 cm /s contains some rather extreme values, do you get the same results if the speeds are generated from the range (-10, -2) cm/s?

The range of phase speed we employ is an “envelope” of observed values of Rossby waves. As per the reviewer’s suggestion we calculated the detection accuracy of the 4 methods in smaller ranges of frequency and phase speed and the conclusions from these results will be added to the revised version of the manuscript.

In low latitudes the phase speed of Rossby waves can easily exceed 15 cm/sec, and in high latitudes it is less than 1 cm/s. See e.g. Fig. 7 of Killworth et al. in the Journal of Physical Oceanography (1997), attached below. We will omit the words "in mid-latitudes" in the 2nd paragraph of section 2.1.

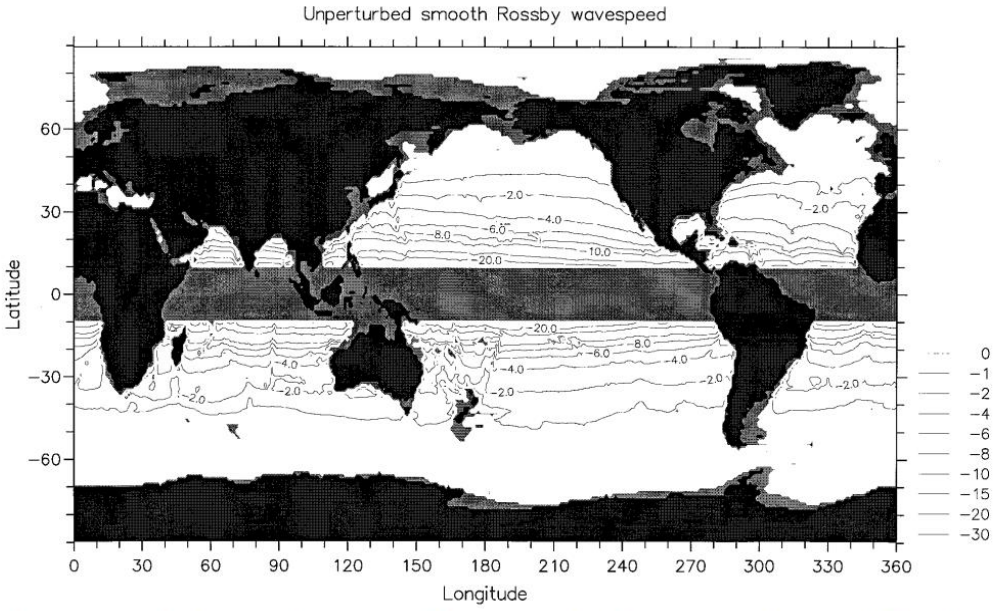

FIG. 7. The unperturbed fastest long planetary wave speed (in cm s^{-1}) (simply $-\beta/f^2$ times the square of the unperturbed internal wave speed shown in Fig. 6). Contour intervals are nonuniform: 30, 20, 15, 10, 8, 6, 4, 2, 1, and 0 cm s^{-1} westward. (The values 30 and 0 do not occur, but are added for comparison with later diagrams.) The values are masked within 10° of the equator, where equatorial, rather than long planetary wave, theory should hold.

2. On a similar comment, the theory of Rossby waves indicates that Rossby waves have a maximum frequency, which for the ocean is quite restrictive. For example, no baroclinic Rossby waves with periods shorter than 13 weeks are possible poleward of about 15o latitude. Here, the periods are taken from the range 5 to 200 weeks, which again seems to include some rather extreme values. Could you please provide a reference supporting such high frequencies for observed Rossby waves? How do the results change if you restrict the periods of the Rossby waves to, for example, the range 15 to 100 weeks?

Assume that a typical propagation speed is 5 cm/s and examine a wave with 5000 km wavelength. Then:

$$T = \frac{2\pi}{\omega} = \frac{2\pi}{kC} = \frac{\lambda}{C} \Rightarrow T = \frac{5 \cdot 10^6\ m}{0.05\ m/s} = 10^8 \text{ seconds} \approx 1150 \text{ days} = 165 \text{ weeks}$$

Considering the Nyquist frequency constraint our choice of longest period of 200 weeks does not seem to be an over-estimate for Rossby waves. Lower values of C and higher values of λ will yield longer periods.

The lower value of 5 weeks does not differ much from 13 weeks. However, as stated our response to comment point #1 above, we will include a brief description of the results for smaller ranges of both frequency and phase speed.

3. Theoretical phase speeds are not only different from observations, they are systematically slower. If the simulated data were an accurate representation of the

real ocean and the detection methods were really the issue here, then the authors should also find a systematic bias in the estimated phase speed. However, there is no mention of this in the paper. The authors only state that all methods fail to estimate the true phase speed of Rossby waves. Do you find any systematic biases? Could you please further elaborate on this?

Right, the observed speeds are always faster than the harmonic speeds but have no systematic bias compared to the Trapped wave's speeds. A clear example of this behavior is given in the comparison shown in Fig. 5 of De-Leon and Paldor, 2017b (reproduced below). The red curve is the Trapped wave speed and the Green curve – the Harmonic speed. Symbols are the observational speeds that are distributed systematically **above** the harmonic speed but with no obvious bias compared to the trapped wave speed.

In addition, we don't state "…that all methods fail to estimate the true phase speed of Rossby waves", but that they fail to estimate a dominant input phase speed regardless of its physical origin i.e. Rossby waves are an example.

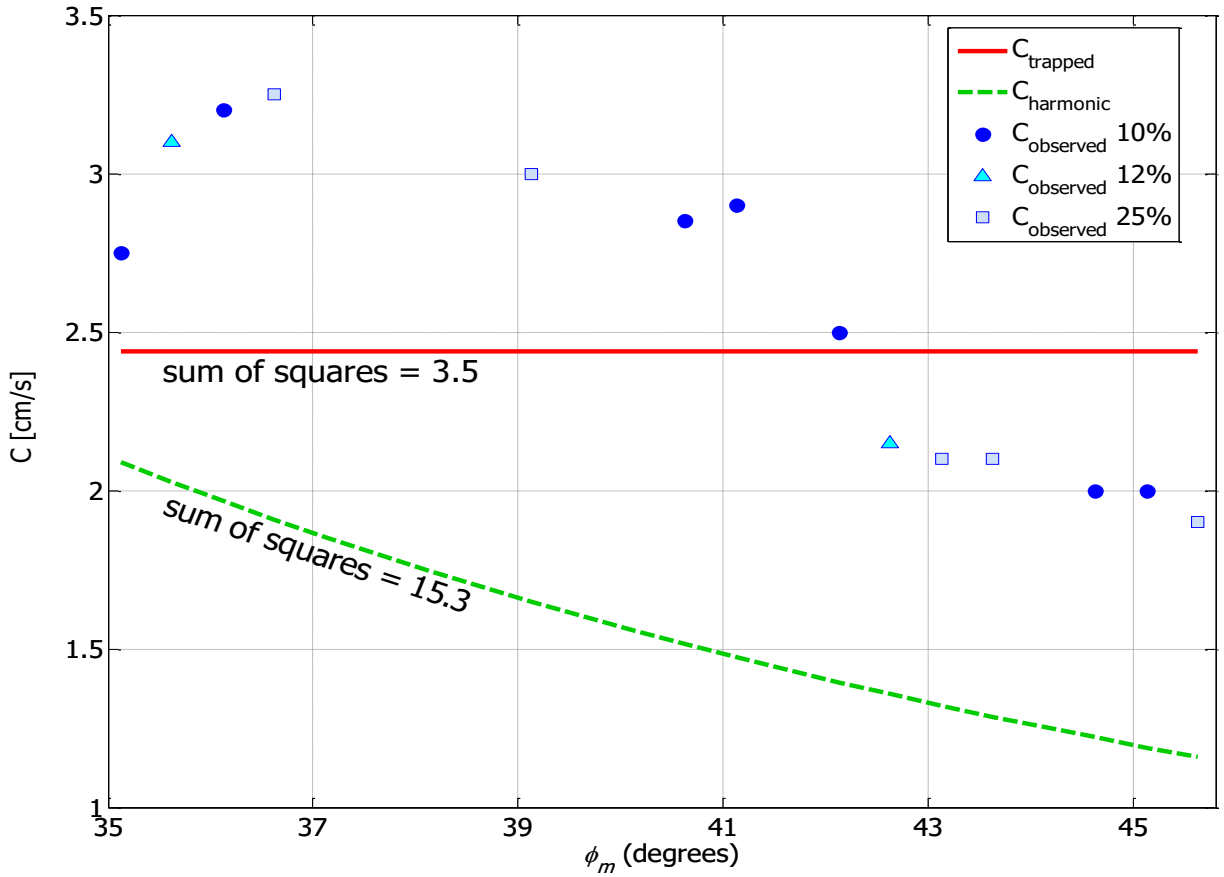

Figure 5. The observed phase speeds and the two theoretical phase speeds (trapped and harmonic) as a function of ϕ_m in intervals of 0.5° latitude. Blue dots denote latitudes where the estimates of at least two methods agreed by 10 % or less, triangles denote latitudes where such estimates agreed by 11 to 12 % and squares denote latitudes where the agreement is 25 %. No reliable estimates were obtained north of 35° S and in some more latitudes. The sum of squares of the distances in $(\text{cm s}^{-1})^2$ between trapped wave phase speeds and observed speeds (3.5) is much smaller than that of harmonic phase speeds (15.3).

4. In assessing the skill of the various methods, the authors assign a score of ½ if the dominant mode falls in one of its nearest neighbors. This seems to me like a rather arbitrary choice. Why not the second nearest neighbor or the third one? Can you estimate a "standard error" for the phase speed estimates based on the multiple realizations and assign a score of 1 when the true value is within one standard error and zero otherwise? This would be, in my view, a fairer metric for skill. Also, I think that 50 realizations is not sufficient and would suggest you use at least 100, if not 1000.

Indeed, our choice is arbitrary but so is any other choice. We will emphasize it in the revised text. The number of cases where the detected mode was 1-bin away from the dominant input mode (i.e. the score was ½) is very small in all signals we examined. As for the number of realizations, we didn't find significant difference between 25, 50 or 100 repeats.

Minor points:

Page 1. The spatiotemporal resolutions quoted here for the altimetry data refer to the grid size and time step of the altimetry gridded products rather than the scales that can actually be resolved by altimeters. Depending on latitude, the spatial separation between altimetry tracks can be of several hundred kilometers and altimeters take measurements over the same location once every 10 days at most. I think that some clarification is needed here, along with some references.

We define the grid in the same way it is defined by Aviso in their description of the altimetry gridded products they distribute to the community.

Page 1. "these features propagate..." What features? Please clarify. We removed this sentence.

Page 1. "Rossby waves that propagate westward" I suggest you remove "that propagate westward" as this seems redundant in this particular sentence. We removed this sentence.

Page 1. replace "diagrams at certain latitude" with "diagrams at a certain latitude". Done.

Page 2. "phase speed exceeds". We removed this sentence.

Page 2. I suggest "in the -18 to 0 cm/s range". Done.

Page 5. I suggest "None of the methods can identify a dominant input ..." Done.

Author's response to: Referee #1's comments on "Commonly used methods fail to detect known phase speeds of simulated signals of Sea Surface Height Anomalies" *by* Y. De-Leon and N. Paldor

Summary and recommendation:

The main aim of this paper is to challenge the reliability of the observational basis for the 'too-fast' Rossby waves evidenced by Chelton and Schlax (1996) based on 4 years of Topex-Poseidon satellite altimeter data. The authors derive their conclusion from showing that it is possible to construct a synthetic Rossby wave signal composed of 20 to 50 sine waves with random known speeds, which standard techniques such as the Radon and Fourier transforms fail to identify accurately. In a previous study, Paldor et al. had showed such techniques to work well for a synthetic signal composed of three basic waves only, so the difficulties experienced by the Radon and Fourier transforms in this paper appear to result from the increase in many more basic waves in the synthetic signal constructed. As to the motivation for the present study, Nathan Paldor's group has been working on the 'too-fast' Rossby wave issue for many years, promoting the view that the observed phase speed enhancement results from latitudinal trapping due to Earth's curvature. So far, however, Paldor's group appear to have found it difficult to vindicate their theory from observations; but rather than concluding that the problem might rest with their theory, as others theoreticians may have done, the present study proposes that the blame should lie with the observations and the kind of techniques used to analyse them instead, not their theory.

We appreciate the concise summary the reviewer has written about Paldor's work in the last decade but neither the theoretical work itself nor the reviewer's summary have anything to do with the work under review that examines the applicability of Radon Transform and 2D-FFT methods to time-longitude (Hovmöller) diagrams. We share the reviewer's frustration with the minute impact that a higher-order theory that **consistently** accounts for the latitudinal variation of Coriolis parameter (instead of the traditional paradigm that "f is constant though its derivative is non-zero") had in planar GFD (not only spherical as the reviewer erroneously claims!).

Since no additional assumptions or approximations are employed in the Trapped wave theory (in comparison to the Harmonic traditional theory), and only higher order terms are consistently included, we see no basis for the claim: "… that the problem might rest with their theory". The reviewer is invited to refute the Trapped wave theory in another forum.

As far as presentation is concerned, the paper is clearly written, and the analysis appears to be competently done. However, as a contribution to the general issue of what satellite altimeter data actually tell us about westward propagation in the ocean and about the usefulness/validity of the standard Rossby wave theory, this

study appears to be very biased in its approach and therefore of very little scientific value, clearly failing to meet the required standards for publication. This is unfortunate, because I otherwise find Paldor's work on the rigorous analysis of the waves supported by the shallow water equations to useful and valuable. As far as I understand the issue, their work appears to be essentially concerned with refining the standard flat-bottom, no mean flow, linear theory of the shallow-water waves on the sphere, and has therefore no bearing with real Rossby waves, which theoretical advances over the past 50 years have clearly showed to be strongly affected by both the background mean flow and topography. The rationale for my assessment is contained in the following remarks and observations.

Again, the current work does not deal with the consistent wave theory of Rossby waves (on a sphere or a plane) but with methods for extracting propagation speeds from slopes of contour levels on time-longitude (Hovmöller) diagrams. In our view, the reviewer's assessments: 1) that the paper is "clearly written" and 2) that the analysis is "competently done" along with the prevalent usage of these methods in recent (see the response below to main point #1) interpretations of various oceanic observations should render the paper suitable for publication in Ocean Science.

Main points

1. The authors fail to mention that the reliability of Chelton and Schlax (1996)'s conclusions has already challenged by Dudley Chelton himself and his collaborators in Chelton et al. (2011), in which the authors argue that westward propagation in the oceans is dominated by meso-scale eddies rather than linear Rossby waves in contrast to what CS96 had previously assumed. Since then, how to disentangle the meso-scale eddy field from the background Rossby wave field has been a major challenge that only a few authors have tried to tackle. Since we know that meso-scale eddies tend to have an equatorward or poleward drift depending on whether they are cyclonic or anti-cyclonic, it is clear that determining their propagation characteristics cannot be easily achieved from the use of Hoevmuller diagrams in longitude/time, which is why eddy tracking algorithms have been developed. Since we don't really know to what extent the propagation speed of eddies differs from that the more linear background Rossby wave field, it seems clear that there is some degree of uncertainty about how CS96's results should be interpreted. In any case, it is clear from Chelton et al. (2011) that there is no observational basis for their synthetic signal.

We changed the focus of the paper from satellite derived SSHA signals to propagation speeds derived from time-longitude diagrams (but we cannot ignore the simple fact that the Radon transform and 2D FFT methods were heavily employed in SSHA signals derived from satellites). Both the Hovmöller diagrams and the methods employed to interpret them were used in recent years (last 5-6 years) and not only prior to 2011. Additional such references will be included in a revised version of the manuscript.

2. Theoretical developments prompted by Chelton and Schlax (1996) have clearly revealed that the background mean flow and bottom topography have a major impact on the propagation and vertical structures of Rossby waves, and hence that the standard theory can never be a satisfactory description of actual Rossby wave propagation regardless of what satellite altimeter data actually tell us. Indeed, Aoki et al. (2009) and Hunt et al. (2012) have both convincingly established that the standard theory cannot account for the features of simulated Rossby waves propagation, which can only be satisfactorily explained when both the mean flow and bottom topography are accounted for. Flat bottom, no mean flow, modes are completely unable to capture the vertical structure of simulated Rossby wave variability. Irrespective of what the observations tell us, I believe it is pretty clear that the authors' approach cannot tell us anything about actual Rossby waves.

Again – the manuscript does **not** deal with the theory of Rossby waves (be it Trapped or Harmonic)! We only examine the accuracy of the methods used to extract propagation speeds from time-longitude diagrams. Indeed, the manuscript does not "tell us anything about actual Rossby waves" and the reviewer's comment belongs somewhere else and not in a review of the issue our paper addresses.

3. Contrary to what this paper and previous ones assert, theoretical studies of the standard theory based on the WKB approximation are able to account for both the trapping of the Rossby waves as well as for Earth curvature, and it is misleading to refer to such theories as harmonic theories. In WKB theory, one will typically express the pressure anomaly in the form

$$p = A(x,y,z,t)e^{i\Sigma(x,y,t)}$$
$$k = \nabla\Sigma \quad , \quad \omega = -\frac{\partial\Sigma}{\partial \mathrm{t}}$$

In such an approach, the amplitude is slowly varying, and will in general decay with latitude, thus capturing the trapped wave behaviour emphasised by the authors. The function Σ is a rapidly varying phase function, allowing to define a local wave vector and frequency. Note that a single WKB wave mode is able to represent the observed beta-refraction and a latitudinally varying phase speed.
In contrast, the basic wave mode considered by Paldor's group is separable in latitude, and typically chosen of the form

$$p = A(y)e^{i(kx-\omega t)}$$

Arguably, if the term 'harmonic mode' needs to be used, it seems more appropriate to the modes considered by Paldor's group, since it is clearly what they chose for the temporal and zonal dependence of their mode. As a result, such a mode does not capture the beta-refraction pattern described by Shopf et al. (1981) for instance, raising the question of how useful this kind of mode is to describe mid-latitude Rossby waves.

Though this point has nothing to do with the sermon of our paper we agree with the reviewer. The difference between the Trapped wave theory and the Harmonic wave theory is precisely the form of A(y) in the last expression for *p*: In the traditional, Harmonic, theory the variation of *p* is sinusoidal so these waves (that spread over the entire latitudinal domain) are named Harmonic while in the Trapped wave theory A(y) has the form of Airy function whose maximum is located near the equatorward boundary (southern in the northern hemisphere). In mid-latitudes Schopf's theory employs the usual "f is constant though its derivative does not vanish" while in his equatorial ray theory the frequency is y-dependent so the concept of separation of variables, that underlies the form of *p* is entirely lost (d/dy should include the latitudinal derivative of the frequency). Again – we emphasize that these (interesting) issues have nothing to do with the sermon of the present work!

Commonly used methods fail to detect known propagation speeds of simulated signals from time-longitude (Hovmöller) diagrams

Yair De-Leon[1], Nathan Paldor [1]
[1]Fredy and Nadine Herrmann Institute of Earth Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem, 9190401, Israel

Correspondence to: Nathan Paldor (nathan.paldor@huji.ac.il)

Abstract. This work examines the accuracy and validity of two variants of Radon transform and two variants of the Two-Dimensional Fast Fourier Transform (2D FFT) that have been previously used for estimating the propagation speed of oceanic signals such as Sea Surface Height Anomalies (SSHA) derived from satellite borne altimeters based on time-longitude (Hovmöller) diagrams. The examination employs numerically simulated signals made up of 20 or 50 modes where one, randomly selected, mode has a larger amplitude than the uniform amplitude of the other modes. Since the dominant input mode is ab-initio known, we can clearly define "success" in detecting its phase/propagation speed. We show that all previously employed variants fail to detect the phase speed of the dominant input mode even when its amplitude is 5 times larger than all other modes and that they successfully detect the phase speed of the dominant input mode only when its amplitude is 10 times (or more) larger than the other modes. This requirement is an unrealistic limitation on oceanic observations such as SSHA. In addition, three of the variant methods "detect" a dominant mode even when all modes have the exact same amplitude. The accuracy with which the four methods identify a dominant input mode **decreases** with the increase in the number of modes in the signal. Our findings are relevant to the reliability of phase speed estimates of SSHA observations and the reported "too fast" phase speed of baroclinic Rossby waves in the ocean.

1 Introduction

Time-longitude (Hovmöller) diagrams at a given latitude of an oceanic variable, $\eta(x,t)$, that represents e.g. temperature, sea surface height, chlorophyll, etc., obtained for example by satellite observations are often used for estimating the propagation rate of the oceanic variable. The rate at which $\eta(x,t)$ propagates is determined by the inverse slopes of same-amplitude contours. These slopes are calculated by applying various methods, employed in image processing and detailed below in Sect. 2.2, to the raw data or to processed data (e.g. Polito and Liu, 2003 who separated the data into tiles of different periods). The methods examined here were used in many oceanic sub-areas such as: The propagation speed of Rossby waves using SSHA (e.g. Tulloch et al., 2009) or data other than SSHA (Belonenko et al., 2018; Xie et al., 2016); Nearshore wave dynamics (Almar et al., 2014); Eddy detection (Abernathey and Marshall, 2013; Oliveira and Polito, 2018) and Intraseasonal variability from mooring (Hu et al. 2018) to name a few.

Deleted: hase speeds of simulated signals of Sea Surface Hei
Formatted
Deleted: SSHA signals made up of 20 or 50 modes where one,
Formatted
Deleted: measurement
Deleted: of
Formatted
Deleted: made
Formatted
Formatted
Formatted
Deleted: e.g.
Deleted: S
Formatted
Deleted: can be
Formatted
Deleted: such
Formatted
Deleted: measurements
Formatted
Deleted:
Formatted
Deleted: of Sea Surface Height Anomalies (SSHA) were routin
Deleted: ¶
Formatted
Deleted: SSHA
Formatted
Deleted: westward
Formatted
Deleted: can be derived using time-longitude (Hovmöller)
Deleted: are proportional to the propagation speed of these
Deleted: were
Formatted
Deleted: distributed by AVISO (supported by CNES; see
Formatted
Deleted: se
Formatted
Deleted: studied regarding t
Formatted
Deleted: rate
Formatted
Deleted: , n
Formatted
Deleted: ,
Formatted
Deleted:
Formatted

De-Leon and Paldor (2017a) examined the accuracy of various methods in estimating the phase speed of waves by applying these methods to an artificially generated signal made of 3 sine functions (modes) with known phase speeds and amplitudes, compounded by large-amplitude random white noise. All methods have successfully filtered out the high amplitude white-noise from the 3-harmonic signal and accurately detected the main mode (some of them also detected the secondary modes). However, such a signal is too synthetic/ideal and cannot be compared to real oceanic observations that include tens, if not hundreds, of modes with different frequencies and propagation speeds and not just 3 modes compounded by white noise.

In this short study we simulate "oceanic observations" and examine whether the methods detect a single dominant propagation speed out of many (20 or 50) speeds. In Sect. 2 we provide details on the generation of "observed" η signal, the methods for evaluating the propagation speed of the signal and the tests we apply to the signals. The results are shown in Sect. 3 and discussed in Sect. 4.

2 "Data" and methods

2.1 Generating the simulated "observations"

The η signals (where η is any oceanic variable) used here are generated numerically by summing up N purely propagating sine functions (modes, hereafter) of the form $\sin(kx-\omega t)$ where k is the zonal wavenumber, x is longitude, t is time and $\omega=kC$ is the frequency (where C is the zonal propagation speed). The number of participating modes, N, is taken to be either 20 or 50 and N-1 of these modes have an amplitude of 1 while the amplitude of the additional N^{th}, randomly chosen, mode is either larger than or equal to 1. The sum of all N modes constitutes the η-signal which is analyzed by the methods described in 2.2.

The "spatial domain" (x) is chosen between "longitudes" 70-130° on a Cartesian grid with a 1/4° resolution, and the "time" (t) duration is 20 years (1044 weeks) with temporal resolution of one datum per week, similar to publicly distributed products by e.g. Aviso. The values of the propagation speeds, C, are uniformly distributed in the -18 to 0 cm s^{-1} range (i.e. each of the 20 or 50 modes is assigned a different propagation speed within that range), which is typical for baroclinic Rossby waves in the ocean (see e.g. Fig. 7 of Killworth et al., 1997; Barron et al., 2009). The values of the frequencies, ω, are selected randomly so that the period, $2\pi/\omega$, falls in the range between 5 and 200 weeks while the values of the zonal wavenumbers, k, equal ω/C. The resulting signal was low-pass filtered by applying a 5-week-running-average at each grid point to eliminate short term variations.

The η signal made up of the filtered signal (i.e. the sum of N pure sine waves) at a given latitude, is plotted as a function of longitude and time (Hovmöller diagram). When a single dominant mode exists that has a certain propagation speed the pattern on this diagram is a straight line whose slope is the inverse of the dominant propagation speed (since the abscissa is longitude and the ordinate is time). An example of a time-longitude diagram of an artificial signal is shown in Fig. 1a (for signal with dominant input mode's amplitude of 1.5) where the slope of the solid blue line corresponds to the propagation

Deleted: The phase speed estimates from observations were compared with theoretical predictions of the classical harmonic theory of Rossby wave (e.g. Pedlosky, 1987; Gill, 1982) and it w found that the observed phase speed exceed the theoretical one in many locations in the world ocean (e.g. Chelton and Schlax, 199 Osychny and Cornillon, 2004). Explanations for these discrepanc were suggested in previous studies (see the introduction of De-Le and Paldor, 2017b for more details) in order to bridge that gap. O of the suggestions is the trapped wave theory (Paldor, Rubin and Mariano, 2007; Paldor and Sigalov, 2008; Gildor et al., 2016) wh the waves' phase speed is higher than that of harmonic waves. However, this trapped wave theory could be applied and compare oceanic observations only in the presence of a wall, while there a only a few such places in the world (the southern coast of Austral the only place where we succeeded to do that, see De-Leon and Paldor, 2017b). An attempt to define a virtual boundary based on linear fit of observed estimates have not yielded satisfactorily fit between the trapped wave theory and observations.¶
A possible reason for this lack of success in bridging the gap betw observations and theory is that the observed estimates are not reliable. …

Formatted: Highlight

Formatted: Highlight

Deleted: phase

Deleted: phase

Deleted: phase

Formatted: Not Highlight

Deleted: phase

Formatted: Not Highlight

Deleted: sea surface height "observations"

Deleted: phase

Formatted: Not Highlight

Formatted: Not Highlight

Deleted: hase

Deleted: $-18\ cm\ s^{-1}$

Formatted: Superscript

Deleted: to -18 $\frac{cm}{s}$

Formatted: Not Highlight

Formatted: Not Highlight

Deleted: in mid-latitudes

Deleted: (see e.g. Chelton and Schlax, 1996)

Formatted: Not Highlight

Deleted: SSHA

Deleted: AKA the

Deleted: phase

Deleted: phase

Deleted: phase

speed of the dominant input mode. The challenge is to estimate the dominant speeds using different methods and examine their success in detecting the propagation speed of the known dominant input mode. The methods examined here are detailed in the next subsection.

2.2 Methods for estimating the "observed" propagation speed

Four variants of methods have been employed for identifying the preferred direction of the same-amplitude contours on the Hovmöller diagram, each variant relates a certain measure of the power/intensity in a mode to its propagation speed. The first method is the Radon transform, used by e.g. Chelton and Schlax (1996); Chelton et al. (2003) and Tulloch et al. (2009) for analysing satellite observations of the ocean. In this method, one calculates the sum of the amplitudes along lines inclined at an angle θ and displaced a distance s from the origin. Then, the sum of squares of the values of these sums along all lines having the same angle is calculated and the angle at which this sum of squares is maximal, is the best estimate for the orientation of the lines on the image. The dominant propagation speed of the signal is then proportional to the tangent of this angle of maximum sum-of-squares. In the second method, which is a variant of the Radon transform, the variance of the amplitudes is calculated along every angle θ instead of the sum of amplitudes. This method was applied e.g. by Polito and Liu (2003 to local auto-correlation of segments of Hovmöller diagram) and Barron et al. (2009). Another, independent, method commonly used (e.g. Zang and Wunsch, 1999; Osychny and Cornillon, 2004) is the two-dimensional Fast Fourier Transform (2D FFT). An example of (ω, k) diagram, obtained by applying 2D FFT to the signal of Fig. 1a is shown in Fig. 1b. Here we use two variants of the 2D FFT method: In the first variant one sweeps over the 2D FFT spectra to find the direction in (ω, k) plane with maximum "energy" (this is the third method) while in the second variant one finds the maximal amplitude of the 2D FFT (i.e. one of the bright points in Fig. 1b) and calculates the ratio ω/k where ω and k are the frequency and zonal wavenumber of the maximal amplitude (this is the fourth method). Detailed description of these methods and the interpretation of observed signals are found in De-Leon and Paldor (2017a).

Each of the four variants of the methods can yield an estimate of the dominant propagation speed of a signal, based on the extremum of a graph that relates the calculated measure of a mode's intensity to its propagation speed. An estimation of the propagation speed based on a local extremum of this graph is accepted when this extremum is narrow and isolated compared to other local extrema. When the normalized amplitude (the term normalized amplitudes is used for the calculated amplitudes divided by the maximal amplitude in the domain) of a distinct peak is 1 while the (normalized) amplitudes of all other peaks are smaller than 0.8, this mode is considered the dominant mode. In the variance method, where the extrema are minima, the dominant mode is accepted when its amplitude is 0.0 while the normalized amplitudes of all other minima are larger than 0.2. The results shown in Fig. 2 demonstrate the emergence of a "rejected" peak that does not differ significantly from other peaks (2d) and "accepted" peaks whose intensities are significantly larger than those of other peaks (2a, 2c) or "accepted" trough that is significantly smaller than the other troughs (2b).

2.3 Examining the accuracy of dominant mode detection

Deleted: phase
Deleted: phase
Deleted: phase
Deleted: analyzing
Deleted: phase
Deleted: used
Deleted: on a
Deleted: applied to
Deleted: piecewise
Deleted: phase
Deleted: phase
Deleted: phase

Two types of tests are applied to these "observations" to examine the accuracy of the various methods in assessing the existence of a dominant propagation speed (i.e. mode) in the signal.

The first is a true-positive/false-negative test in which there is a dominant mode in the "observed" signal and a given method indicates whether this dominant mode exists (true-positive; TP) or not (false-negative; FN). In our case, one of the sine functions (this is the dominant input mode) is chosen randomly and its amplitude is set to be larger than 1. We check for each method if (at all) it identifies a dominant mode and if so, if it matches the propagation speed of this (larger amplitude) input mode. We divide the interval of propagation speeds into N (=20 or 50) bins of equi-distant values and the determination of the success of the methods in identifying a dominant mode is as follows: if the dominant mode found by the method falls in the expected bin of the dominant input mode - we score it by 1 ("TP"). If it is found in one of its next neighbors, it is scored by 1/2. A score of 0 ("FN") is assigned when the method cannot find any dominant mode and when it detects a dominant mode more than 1 bin away from the correct bin. For each of 5 values of dominant mode's amplitudes: 1.5, 2, 2.5, 5 and 10 and for 2 values of N (20 or 50) we repeat this procedure 50 times (i.e. for 50 different signals), sum up the scores and calculate the percentage of success in identifying the dominant input mode in the signals by TP/(TP+FN)*100.

The second is a false-positive/true-negative test in which no dominant mode exists in the "observed" signal and a given method indicates that there **exists** a dominant mode (false-positive; FP) or not (true-negative; TN). In our case, this is done by generating a signal in which all modes have identical amplitudes (=1) and checking whether a method erroneously detects a certain propagation speed as dominant. If a dominant mode is detected, we score it by 1 ("FP"), if a peak is detected but is too wide we score it by 1/2 and if there is no dominant mode (i.e. no distinct peak or more than one peak) we score it by 0 ("TN"). We repeat this procedure 50 times for each of N=20 or N=50, sum up the scores and calculate the percentage of erroneous detection of dominant mode in the signals by FP/(FP+TN)*100.

3 Results

An example of false determination of the dominant mode is shown in Fig. 2 for the signal shown in Fig. 1a. Figure 2a shows the distribution of the sum-of-squares of the Radon transform versus C (black markers), normalized such that the maximum value equals 1. Also plotted are the dashed black vertical lines located at the N values of the uniformly distributed propagation speeds, C, where the solid blue line is located at the C-value of the dominant input mode's propagation speed. These dashed black and solid blue vertical lines are also shown in panels (b)-(d) of Fig. 2. Clearly, the dominant propagation speed calculated by the Radon transform (where the black curve attains its maximum) does not match the propagation speed of the dominant input mode. Figure 2b shows the (normalized) mean of variances as a function of C (black markers), and here, too, the calculated propagation speed (the curve's minimum point) does not agree with the dominant input mode's speed. Figure 2c shows the (normalized) distribution of the sum-of-squares of the spectral coefficients (2D FFT-amplitudes) along different ω/k lines (sweeping) as a function of C (black markers) where a distinct peak exists but it's located far from

Deleted: phase
Deleted: hase
Deleted: phase
Deleted: -
Deleted: phase
Deleted: erroneous
Formatted: Highlight
Deleted: phase
Deleted: phase
Deleted: phase
Deleted: phase
Deleted: phase

the dominant input mode's propagation speed. Figure 2d shows the 20 highest (normalized) 2D FFT amplitudes of the (ω, k) diagram of Fig. 1b. There are many peaks with no clear single maximum and the dominant input mode has one of the lowest amplitudes. For this signal, none of the methods identified correctly the dominant input mode.

The statistics of success of each method in detecting the dominant input mode for the TP/FN test is shown in Fig. 3 for dominant input mode's amplitude of 2.5, 5 and 10. The results for amplitudes of 1.5 and 2 are not shown as the success rate of all methods at these amplitudes is between 10% and 30% only. In each case we repeated the procedure 50 times (i.e. for 50 signals) for N=20 and then for N=50, summed the scores and calculated the percentage of success by TP/(TP+FN)*100. The conclusions from these results are: 1) in order to identify the dominant input mode with more than 70% certainty, its amplitude should be larger than 5. 2) No method has clear advantage over the other methods. 3) Clearly, as N increases the dominant mode's amplitude has to increase, too, for a successful identification (so as to ensure that the ratio between the dominant mode's amplitude and the sum of all amplitudes is similar for different values of N since, e.g., 2.5/20>2.5/50). 4) The amplitude at which successful detection occurs decreases with the decrease in the ranges of propagation speeds and periods. Thus, for propagation speeds in the range of -10 to -2 cm s^{-1} and periods between 15 and 100 weeks, an amplitude of 5 is successfully detected in over 90% of the cases, while an amplitude of 2 yields poor results (results not shown).

The statistics of erroneous detection of a dominant mode (the FP/TN test where the amplitudes of all input modes equal 1 i.e. there is no dominant input mode) is shown in Fig. 4 for each of the methods (here again we have generated 50 signals for N=20 and for N=50 and calculated the percentage of erroneous detection by FP/(FP+TN)*100). The 2D FFT maxima method is the only method for which the percentage of error is smaller than 20% while other 3 methods err in at least 50% of the cases (i.e. they identify a single clear peak in one of the propagation speeds). As N increases this erroneous detection percentage decreases slightly in the 2 Radon variants but increases slightly in the 2D FFT sweeping method.

Deleted: phase

Deleted: When the

Deleted: become narrower, the amplitude for which successful detection occurs decreases

Deleted: 2

Deleted: 10

Deleted: yields more

Deleted: than

Deleted: successful detection

Deleted: phase

4 Discussion

None of the methods can identify a dominant input mode unless its amplitude is significantly larger than the others (by a factor larger than 5! in the present study's ranges of propagation speeds and periods) and most of them (except the 2D FFT maxima) erroneously detect a dominant mode when there is no such input mode. Though the 2D FFT maxima method does not falsely detect dominant mode when it does not exist, its performance in detecting dominant input mode when it exists is not satisfactory. For realistic signals of the ocean we don't know that there is a dominant mode with sufficiently large amplitude so none of the methods is reliable for estimating the propagation speed of e.g. Rossby waves.

When the values of ω and k are chosen in a range corresponding to the resolution limit and the Nyquist frequency, the success of 2D FFT in identifying the dominant input mode's propagation speed increases significantly, compared to the case where the values of C are determined in advance, ω is chosen randomly in a particular range and k is set as a result of ω/C so aliasing can occur. For that reason, even if the signal includes only one mode (i.e. one C), and both ω and k are chosen in the latter manner, there can be a wrong identification by the 2D FFT (but the Radon transform identifies it correctly). In the

Formatted: Font: (Default) Times New Roman, 10 pt, Complex Script Font: Times New Roman, 12 pt, (Complex) Arabic (Saudi Arabia), English (United Kingdom)

Deleted: All methods cannot identify a dominant input

Deleted: for

Deleted: phase

Deleted: phase

Deleted: , and therefore this method does not succeed by more 70% even if the amplitude of the dominant input mode is 10 times that of the other amplitudes

Formatted: Not Highlight

ocean we don't know ab-initio which wave numbers and frequencies exist so we cannot filter them out of the signal; hence aliasing can occur, and the percentage of success in detecting the real dominant mode is expected to decrease further.

The erroneous identification of the Radon and variance methods can be partially attributed to the non-linear relation between the angle θ and the propagation speed, C, which is proportional to $\tan(\theta)$ so the equi-distant values of C are converted to θ-values that are very close to one another. Figure 5 shows the distribution of the sum of squares of the Radon transform versus θ for the signal shown in Fig. 1a (while the distribution of the sum of squares of the Radon transform versus C for that signal is shown in Fig. 2a). It is clear from this figure that the peak is located in the vicinity of θ values corresponding to many C values. However, the performance of the Radon variants improves with the increase of the dominant input mode's amplitude, so the non-linear relation between the angle and propagation speed is not the only reason for the mismatch.

As N (the number of modes) increases (it is impossible to establish a-priori a bound on the number of modes in the ocean), the dominant input mode's amplitude should be larger in order to be separately identified from modes with similar characteristics. In addition, the width of the bins becomes narrower as N increases so fewer results can be evaluated as success. Also, in narrower ranges of propagation speeds and periods and for the same number of bins, all methods correctly detect the dominant input mode at lower threshold amplitudes (results not shown).

The weakness of the methods in identifying the dominant mode points to the difficulty in comparison between theories and observations of baroclinic Rossby waves in the ocean and this difficulty might explain the lack of "continuity" of propagation speed estimates between adjacent latitudes in one (or more) methods. It can also explain why a validation of the higher order trapped wave theory (where the β term is treated consistently) has been confirmed by observations only in the Indian Ocean south of Australia (De-Leon and Paldor, 2017b) and not in other parts of the world ocean. In contrast to the harmonic theory whose propagation speed estimates are always slower than the observed speeds, the trapped wave theory differs from observations sporadically (see Fig 5 of De-Leon and Paldor, 2017b).

Competing interests. The authors declare that they have no conflict of interest.

References

Abernathey, R. P., and Marshall, J.: Global surface eddy diffusivities derived from satellite altimetry, *J. Geophys. Res. Oceans*, **118**: 901– 916, doi:10.1002/jgrc.20066, 2013.

Almar, R., Rodrigo Cienfuegos, H. M., Bonneton, P., Tissier, M., and Ruessink,G.: On the use of the Radon Transform in studying nearshore wave dynamics, *Coastal Engineering*, **92**, 24-30, doi: 10.1016/j.coastaleng.2014.06.008, 2014.

Barron, C. N., Kara, A. B., and Jacobs, G. A.: Objective estimates of westward Rossby wave and eddy propagation from sea surface height analyses. *J. Geophys. Res.* **114**: C03013, doi: 10.1029/2008JC005044, 2009.

Deleted: phase

Deleted: where

Deleted: phase

Deleted: when the

Deleted: is widened

Deleted: the

Deleted: 's amplitude

Deleted: should be larger in order to correctly identify the dominant modelower threshold amplitude

Deleted: /algorithms

Deleted: phase

Formatted: Font: Bold, Complex Script Font: Bold

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Bold, Complex Script Font: Bold

Belonenko, T. V., Bashmachnikov, I. L. and Kubryakov A. A.: Horizontal advection of temperature and salinity by Rossby waves in the North Pacific, International *J. Rem. Sensing*, **39:8**, 2177-2188,DOI: 10.1080/01431161.2017.1420932, 2018.

Chelton, D. B. and Schlax, M. G.: Global observations of oceanic Rossby waves. *Science* **272**: 234-238, doi: 10.1126/science.272.5259.234, 1996.

Chelton, D. B., Schlax, M. G., Lyman, J. M., and Johnson, G.C.: Equatorially trapped Rossby waves in the presence of meridionally sheared baroclinic flow in the Pacific Ocean. *Prog. Oceanogr*. **56**: 323-380, doi: 10.1016/S0079-6611(03)00008-9, 2003.

De-Leon, Y., and Paldor, N.: An accurate procedure for estimating the phase speed of ocean waves from observations by satellite borne altimeters. *Acta Astronautica*, **137**, 504–511 http://dx.doi.org/10.1016/j.actaastro.2016.11.016, 2017a.

De-Leon, Y., and Paldor, N: Trapped planetary (Rossby) waves observed in the Indian Ocean by satellite borne altimeters. *Ocean Science*. **13**. 483-494, doi: 10.5194/os-13-483-2017, 2017b.

Hu, S., Sprintall, J., Guan, C., Sun, B., Wang, F., Yang, G., Jia, F., Wang, J., Hu D. and Chai, F.: Spatiotemporal features of intraseasonal oceanic variability in the Philippine Sea from mooring observations and numerical simulations. *J. Geophys. Res.: Oceans*, **123**, 4874– 4887. 2018. https://doi.org/10.1029/2017JC013653

Killworth, P. D., Chelton, D. B., and de Szoeke, R. A.: The speed of observed and theoretical long extratropical planetary waves. *J. Phys. Oceanogr.* **27**: 1946–1966, doi: 10.1175/1520-0485(1997)027<1946:TSOOAT>2.0.CO;2, 1997.

Oliveira, F.S.C., and Polito, P.S: Mesoscale eddy detection in satellite imagery of the oceans using the Radon transform, *Prog. in Oceanogr.*, **167**, 150-163, doi: 10.1016/j.pocean.2018.08.003, 2018.

Osychny, V., and Cornillon, P.: Properties of Rossby waves in the North Atlantic estimated from satellite data. *J. Phys. Oceanogr*. **34**: 61-76, doi: 10.1175/1520-0485(2004)034<0061:PORWIT>2.0.CO;2, 2004.

Polito, P. S., and Liu, W. T.: Global characterization of Rossby waves at several spectral bands. *J. Geophys. Res*. **108**(C1), 3018, doi:10.1029/2000JC000607, 2003.

Tulloch, R., Marshall, J., and Smith, K. S.: Interpretation of the propagation of surface altimetric observations in terms of planetary waves and geostrophic turbulence. *J. Geophys. Res*. **114**: C02005, doi: 10.1029/2008JC005055, 2009.

Xie, L., Zheng Q., Tian J., Zhang S., Feng. Y., and Li, X.: Cruise Observation of Rossby Waves with Finite Wavelengths Propagating from the Pacific to the South China Sea. J. Phys. Oceanogr., 46, 2897–2913,https://doi.org/10.1175/JPO-D-16-0071.1. 2016:

Zang, X., and Wunsch, C.: The observed dispersion relationship for North Pacific Rossby wave motions. *J. Phys. Oceanogr*. **29**: 2183 – 2190, doi: 10.1175/1520-0485(1999)029<2183:TODRFN>2.0.CO;2, 1999.

Formatted: Font: (Default) Times New Roman, Font color: Auto, Complex Script Font: +Body (Times New Roman), Pattern: Clear

Formatted: Font: (Default) Times New Roman, Font color: Auto, Complex Script Font: +Body (Times New Roman), Pattern: Clear

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Default Paragraph Font, Font: (Default) Times New Roman, Italic, Font color: Auto, Complex Script Font: +Body (Times New Roman), Italic, Pattern: Clear

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Default Paragraph Font, Font: (Default) Times New Roman, Italic, Font color: Auto, Complex Script Font: +Body (Times New Roman), Italic, Pattern: Clear

Formatted: Complex Script Font: +Body (Times New Rom 12 pt

Formatted: Font: (Default) Times New Roman, Font color: Auto, Complex Script Font: +Body (Times New Roman), Pattern: Clear

Formatted

Formatted: Font: Italic, Complex Script Font: Italic

Formatted

Formatted: Font: Italic, Complex Script Font: Italic

Formatted

Formatted

Formatted

Formatted

Field Code Changed

Formatted: Not Highlight

Formatted

Formatted: Not Highlight

Formatted

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Deleted: Gildor, H., Paldor, N., and Ben-Shushan, S.: Numerica

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Indent: First line: 0 cm

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Bold, Complex Script Font: Bold

Formatted: Complex Script Font: +Body (Times New Rom

Deleted: Paldor, N., Rubin, S., and Mariano, A. J.: A consister

Formatted

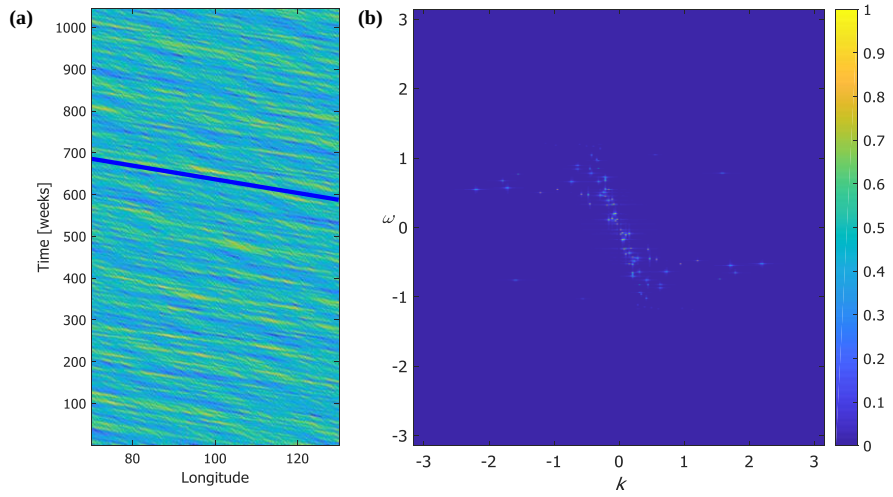

Figure 1: (a) An example of an artificial "observed" signal. (b) The associated (ω, k) diagram obtained by applying 2D FFT to the signal of panel (a). The solid blue line in panel (a) corresponds to the randomly chosen dominant input mode's propagation speed.

Deleted: phase

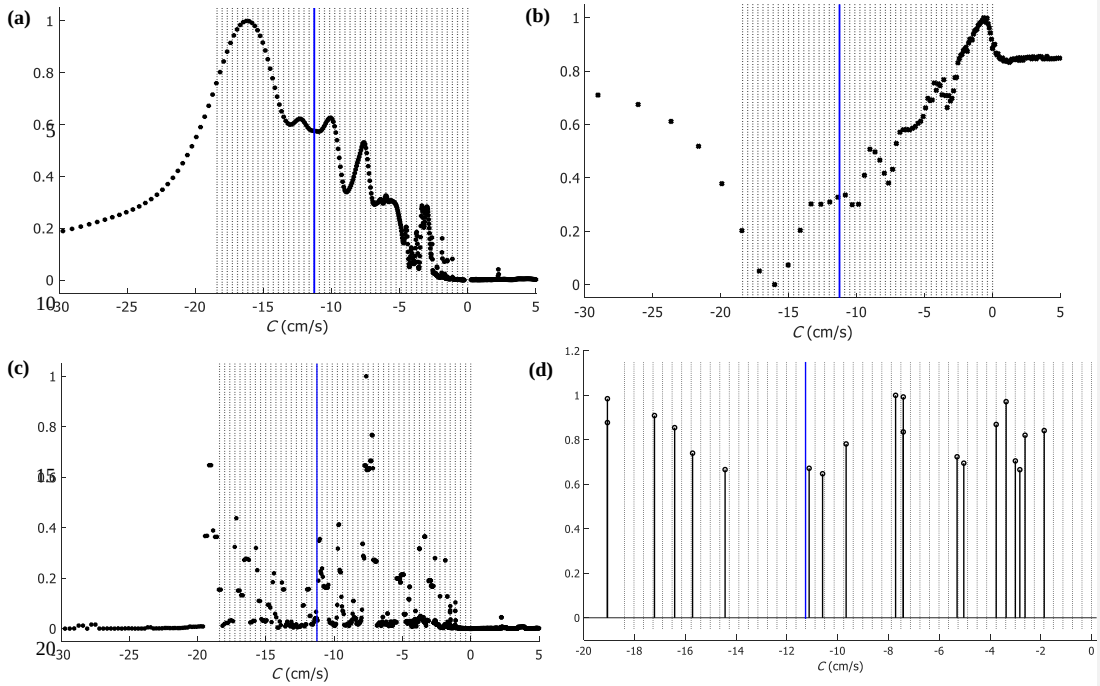

Figure 2: An application of the 4 methods to the artificially generated signal shown in Fig. 1a (panel a: Radon Transform, panel b: Variance, panel c: 2D FFT sweeping, panel d: 2D FFT maximal amplitude). Blue lines correspond to the dominant input mode propagation speed and dashed black lines correspond to the input modes' propagation speeds.

Deleted: phase

Deleted: phase

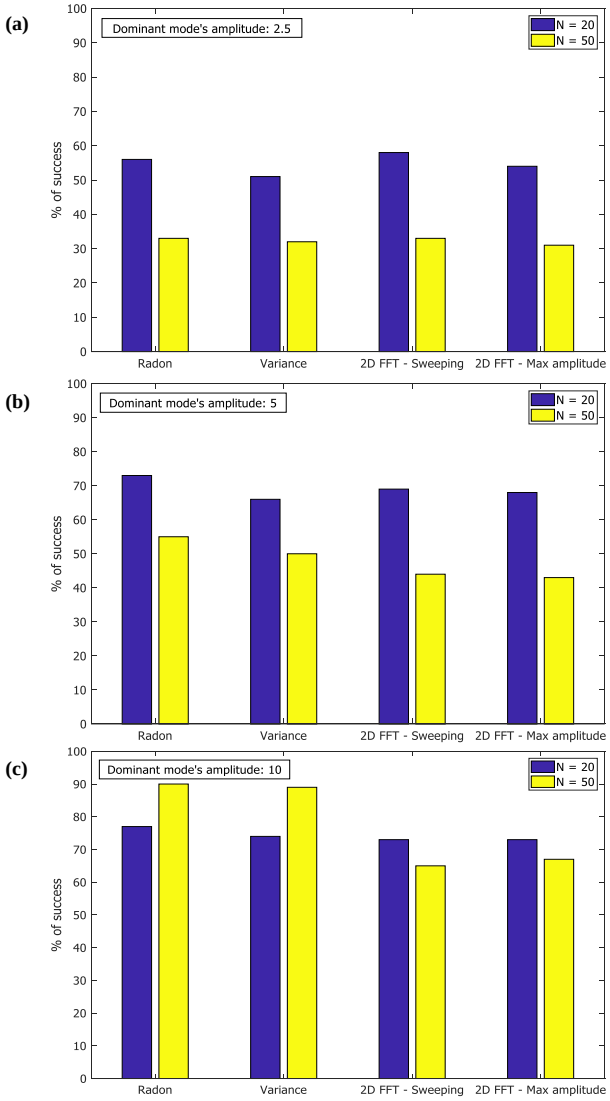

Figure 3: The percentage of success of each variant of the methods for dominant mode's amplitude of 2.5 (panel a), 5 (panel b) and 10 (panel c) for both N=20 and N=50.

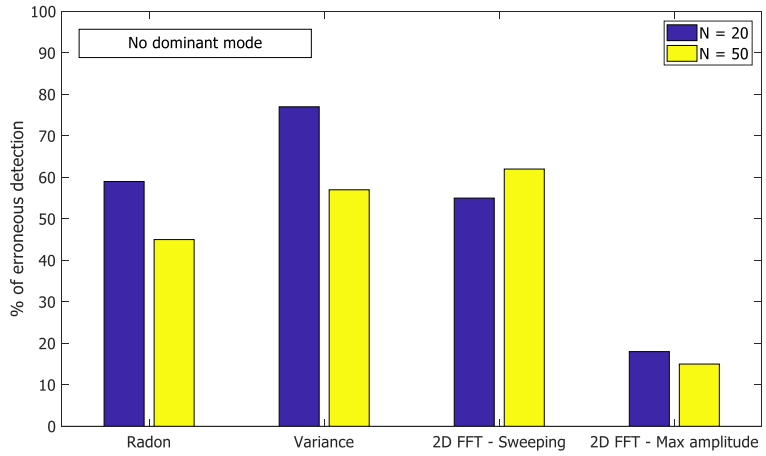

Figure 4: The percentage of error detection of each variant of the methods where there is no dominant input amplitude.

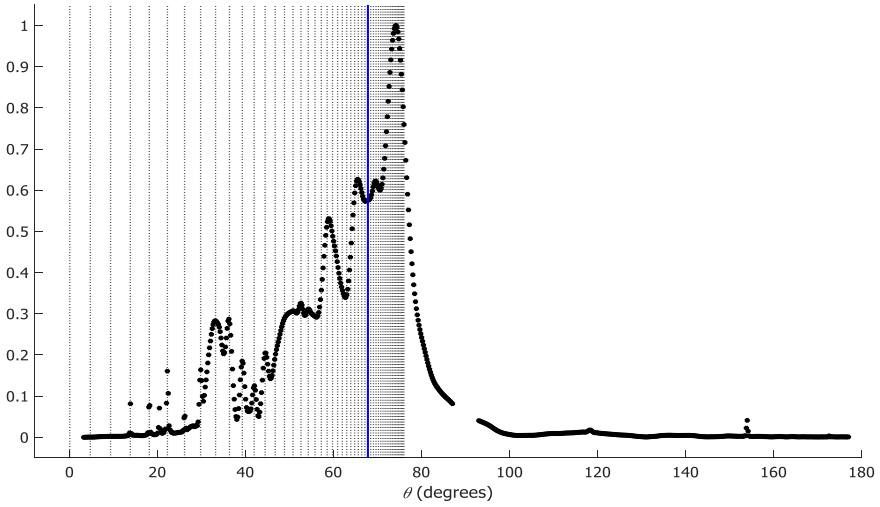

Figure 5: The distribution of the sum-of-squares of the Radon transform versus the Radon angle θ. Clearly, the distribution of the corresponding input propagation speeds (vertical dashed black lines) as a function of θ is not uniform due to the nonlinear relation between C and θ.

Deleted: phase