Effects of Noise on Thorpe Scales and Run Lengths

HELEN L. JOHNSON AND CHRIS GARRETT

Ocean Physics, School of Earth and Ocean Sciences, University of Victoria, British Columbia, Canada

(Manuscript received 15 October 2003, in final form 17 May 2004)

ABSTRACT

Estimating the diapycnal mixing rate from standard CTD data by identifying overturning regions in the water column (the Thorpe-scale approach) provides good spatial and temporal coverage but is sometimes limited by instrument noise. This noise leads to spurious density inversions that are difficult to distinguish from real turbulent overturns. Previous efforts to eliminate noise may have overcorrected and hence underestimated the level of mixing. Here idealized density profiles are used to identify the magnitude and characteristics of overturning regions arising entirely from instrument noise, in order to establish a standard against which CTD data can be compared. The key nondimensional parameters are 1) the amplitude of the noise scaled by the density change over the section of profile considered, and 2) the number of data points in the section of profile. In some cases the product of these, which is equal to the amplitude of the noise scaled by the average density difference between consecutive measurements, is more useful than the second parameter. The probability distribution of ''run length,'' a useful diagnostic, varies significantly across this parameter space. Reasons for this are discussed, and it is shown that CTD data very rarely lie in a region of parameter space where comparison with the probability density function (PDF) of run lengths for a random uncorrelated series, or its rms value $\sqrt{6}$, is appropriate. The distribution of Thorpe displacements arising entirely from instrument noise, as well as the Thorpe scale and the statistics of density inversions, is also discussed. Analysis of CTD data from the interfaces of the thermohaline staircase in the deep Canada Basin illustrates how the results can be applied in practice to help to distinguish between signal and noise in marginal regimes. Density inversions seen in these data are shown to be no different from those that would result from instrument noise.

1. Introduction

Recent numerical modeling studies have led to an increased awareness of the importance of small-scale ocean mixing for the large-scale ocean circulation (e.g., Welander 1986; Bryan 1987; Marotzke 1997). Since the processes that lead to mixing generally occur on scales that are too small to be resolved by numerical models their effects must be parameterized. Accurate measurement of the diapycnal mixing rate K_ν, especially its spatial and temporal variability, is essential for the assessment and improvement of mixing parameterizations. It is also important for a range of engineering and biological applications.

The diapycnal mixing rate can be determined using measurements of velocity and/or temperature microstructure (e.g., St. Laurent and Schmitt 1999). However, the instrumentation required is at present too expensive and specialized for the measurement of ocean mixing to be routine. As a result, spatial and temporal data coverage is sparse. One alternative is to estimate K_ν from the 10-m vertical shear (Gregg 1989). With allowance for a dependence on strain (Polzin et al. 1995) this approach has been shown to be accurate within a factor of 2 for a range of ocean environments (Gregg et al. 2003). However, it is based on a semiempirical relation, relying upon an inexact theory and assumptions about the background internal wave field that may not hold true everywhere (Kunze et al. 2002).

In this paper we revisit a third technique for estimating K_ν, which uses standard conductivity–temperature–depth (CTD) measurements to identify density inversions in the water column (Thorpe 1977). Such inversions signify overturning motions, which result in mixing. This seems to be more direct than the approach based on shear and strain and, because it uses standard oceanographic instrumentation, has the potential to provide good spatial and temporal coverage. The Thorpe scale L_T, a measure of the vertical scale of overturning eddies, is defined as the root-mean-square (rms) displacement of water parcels required to reorder a measured potential density profile such that it is gravitationally stable. It is thought to be related to the Ozmidov length scale, $L_O = (\epsilon/N^3)^{1/2}$, where ϵ is the rate of dis-

Corresponding author address: Dr. Helen L. Johnson, School of Earth and Ocean Sciences, University of Victoria, P.O. Box 3055, Victoria, BC V8W 3P6, Canada.
E-mail: helenj@uvic.ca

© 2004 American Meteorological Society

sipation of turbulent kinetic energy and N is the buoyancy frequency (Ozmidov 1965). Dillon (1982) and others (e.g., Crawford 1986) find that $L_O \approx 0.8L_T$. Since the vertical mixing rate is given by $K_v = \Gamma\epsilon/N^2$, where Γ is the mixing efficiency, which Oakey (1982) finds to be 0.2, K_v can be written in terms of the Thorpe scale as $K_v \approx 0.1NL_T^2$. This equation is plausible on dimensional grounds, although the coefficient of 0.1 is not firmly established (see, e.g., Baumert and Peters 2000). The expression for K_V is valid only in regions that are stably stratified on the large scale and where mixing is a result of mechanical energy input (rather than convective processes).

The Thorpe scale approach has several limitations. For example, ship motion often causes the CTD package to oscillate up and down during a cast, preventing it from sampling the undisturbed water column. Obtaining a density profile of the required accuracy from conductivity and temperature measurements also presents challenges. Despite this, Thorpe scales have been shown to give reasonable estimates of vertical mixing in regions where the stratification is high ($N > 0.01$ s^{-1}) and where mixing rates are larger than about 2×10^{-4} m^2 s^{-1} (e.g., Stansfield et al. 2001). When the stratification is high even small vertical displacements are likely to result in measurable density signals. Thorpe scales have recently been used to provide valuable insight in a range of different regions (e.g., Finnigan et al. 2002; Ferron et al. 1998).

As the sensitivity of CTD instruments improves, application of the Thorpe-scale technique in lower mixing environments more typical of the open ocean is also becoming more promising. However, instrument noise poses a serious problem, leading to spurious density inversions in measured CTD profiles that, if treated as physical overturns, result in an overestimate of K_v. There is an increasing need to better understand the effect of instrument noise on Thorpe-scale mixing estimates so that it can be more effectively eliminated.

A variety of approaches have been used to eliminate noise. Thorpe (1977) dealt with noise at the overturn detection stage, defining the Thorpe displacement L required of each water parcel during the reordering of the profile to be equal to zero unless the density of the sorted, gravitationally stable profile differs from that of the measured profile by more than a predetermined noise level. Ferron et al. (1998) improved on this, defining an intermediate density profile in which the density of two neighboring points only differs if the original difference in density between them is detectable above some predetermined noise threshold. Calculating Thorpe displacements by sorting this intermediate profile reduces the likelihood of noise-generated density inversions while maintaining the size of real overturns. A complete overturn is defined here as a region encompassing density data points that must be exchanged with each other (and no others) in order to achieve a stable profile. Alford and Pinkel (2000) take a different approach, requiring inversions to be present in both temperature and conductivity profiles for them to be recognized as real overturns. They state that, while they may not have caught *all* overturns, they are confident that no spurious ones have been introduced. This philosophy is typical of most people's treatment.

Galbraith and Kelley (1996) propose the most comprehensive series of tests for identifying overturns within a CTD profile and distinguishing these from density inversions created by instrument noise. As well as considering limits on density and vertical resolution, they suggest that the ''run length'' may be a useful diagnostic. This is a statistical measure defined by examining a series and grouping adjacent values of the same sign into ''runs.'' The number of data points in each run is the run length. Galbraith and Kelley calculate run lengths for the series of density differences between the measured density profile and the reordered, stably stratified profile (the Thorpe fluctuation series) and compare the distribution with that expected from a random uncorrelated series. We return to this later in the paper.

On the basis of the run length and other tests, Galbraith and Kelley (1996) reject overturning regions believed to be associated with instrument noise. (They also suggest means of recognizing systematic errors resulting from, for example, the mismatch in response time of temperature and conductivity sensors, and the thermal inertia of the conductivity cell itself. We do not deal with these systematic errors here.) By comparing the results obtained with and without applying the Galbraith and Kelley tests, Stansfield et al. (2001) find them to be robust for their datasets in that the tests seem to reject all spurious overturns created by adding noise to a sorted profile. However, Stansfield et al. (2001) comment that it is likely that the Galbraith and Kelley method does reject some true overturns. Thus it seems that, like other previous attempts to eliminate noise, the Galbraith and Kelley approach may underestimate the mixing. It therefore sets a conservative lower bound on the value of K_v. A natural upper bound is found by assuming that all density inversions present in the data are real overturns arising due to mixing processes.

Stansfield et al. (2001) show that in the Strait of Juan de Fuca, where mixing rates are of the order 5×10^{-4} m^2 s^{-1}, these upper and lower estimates of the Thorpe scale (and hence K_v) lie close together (using the Galbraith and Kelley approach reduces estimates of the Thorpe scale by less than 25%). However, in regions where the signal-to-noise ratio is low, these bounds on K_v may differ widely.

In a recent paper on the thermohaline structure of the deep Canada Basin in the Arctic Ocean, Timmermans et al. (2003) consider the level of turbulent mixing in the interfaces between well-mixed layers of a thermohaline staircase. Their aim is to establish whether the mixing in such interfaces is sufficient to support a heat flux through the staircase comparable with the geothermal flux from below. While they observe Thorpe scales of order 1 m, which is of the required magnitude, density differences within the overturns are small and both Thorpe scales and rms run lengths are close to those expected from a random series. Taking a Galbraith and Kelley approach therefore leads Timmermans et al. (2003) to conclude that the inversions seen within the data are likely to be the result of instrument noise and that K_v is, in fact, too small to support the required heat flux. However, some uncertainty persists, and the question is critical to understanding how the geothermal heat escapes the deep Canada Basin. This is a situation where there is perhaps some room for improvement on the tests proposed by Galbraith and Kelley (1996).

In this paper we investigate the effect of instrument noise on overturn statistics, Thorpe scales, and other relevant parameters. By identifying the magnitude and nature of overturning regions that may be expected to arise due to instrument noise, we establish more effective means of distinguishing between signal and noise. Our ultimate aim is to improve on the lower bound established by Galbraith and Kelley (1996) and hence narrow the gap between conservative and generous estimates of K_v in regions where the signal-to-noise ratio is low.

The remainder of the paper will be structured as follows. In section 2, idealized density profiles in which inversions arise entirely due to instrument noise are used to demonstrate that care must be taken when applying the run-length diagnostic. We identify the key nondimensional parameters and establish the behavior of the run length in this parameter space. Section 3 looks at the effect that noise has on the distribution of Thorpe displacements and on the Thorpe scale, and in section 4 the statistics of density inversions generated entirely by noise are investigated. To demonstrate how our improved understanding of the characteristics of noise may be applied in a low signal-to-noise regime, new CTD data from the deep Canada Basin are analyzed in section 5, in the light of the previous three sections, to establish whether the mixing in the thermohaline interfaces is distinguishable from instrument noise. The results are summarized and discussed in section 6.

2. Effect of noise on run length

When noise is added to a density profile, spurious density inversions are created. Galbraith and Kelley (1996) suggest that the run length may be a useful statistical measure in detecting such inversions and distinguishing them from real overturns. (The word *overturn* will be reserved throughout for structures arising due to real turbulent mixing. The term *density inversion* will be used to refer to all such structures, real and spurious, observed in a density profile.)

Before discussing the behavior of the run-length diagnostic in the presence of instrument noise, we require some definition of the series for which run lengths will be calculated. During Thorpe analysis the profile data are first sorted such that they increase monotonically in terms of potential density, and then the difference between the measured profile and the sorted profile is calculated. The series of density differences that results is a *Thorpe fluctuation* series. The corresponding *Thorpe displacement* series consists of the vertical displacements L required during the sorting of the profile, with each displacement referenced at the vertical position from which the water parcel originated in the measured profile. The sum of the Thorpe fluctuations or displacements over a complete overturn is zero.

The run length is defined by examining a series and grouping adjacent values of the same sign into ''runs.'' The number of data points in each run is the run length. For a random, *uncorrelated* series with an equal number of positive and negative values the probability of a run of length s is 2^{-s}. For real overturns one might expect there to be long runs in the Thorpe fluctuation and displacement series at the top of the density inversion, with long runs of the opposite sign at the bottom of the inversion. This will result in a higher probability of long runs than for a random series, something that Galbraith and Kelley (1996) regard as evidence of a real signal.

We note here that in a Thorpe displacement series each displacement is referenced at the vertical position from which the water parcel originated in the measured profile. If Thorpe displacements are indexed according to the position at which the moved parcel lies in the *sorted* profile, the Thorpe scale L_T will be unchanged, but the distribution of run lengths will be very different. This is because, since the measured profile is multivalued, consecutive values of L may be associated with parcels that originated in very different parts of the water column.

Galbraith and Kelley (1996) consider the probability distribution of run lengths calculated from the Thorpe fluctuation series of an entire profile at a typical CTD station. They then define a cutoff run length as the shortest run length that occurs at least twice as frequently as might be expected from a random series. Sections of a profile containing density inversions with an rms run length smaller than this cutoff are deemed to have arisen because of noise and are discounted. The choice of the factor 2 here is rather arbitrary. Instead, Timmermans et al. (2003) choose to compare their observations of rms run length, calculated from the Thorpe displacement series for a section of profile, with the rms run length that might be expected from a random series: $(\Sigma_1^\infty s^2 2^{-s})^{1/2} = \sqrt{6}$. For real overturns the probability of long runs will be $>2^{-s}$, and we might expect the rms run length to be $>\sqrt{6}$.

Here we demonstrate that, in the presence of random noise of an appreciable magnitude, a series of Thorpe fluctuations/displacements does not display the same run-length characteristics as a random series. Successive values in a finite Thorpe fluctuation (or displacement) series are correlated, and, as a consequence, long runs may arise *even in the absence of real overturns.* Very

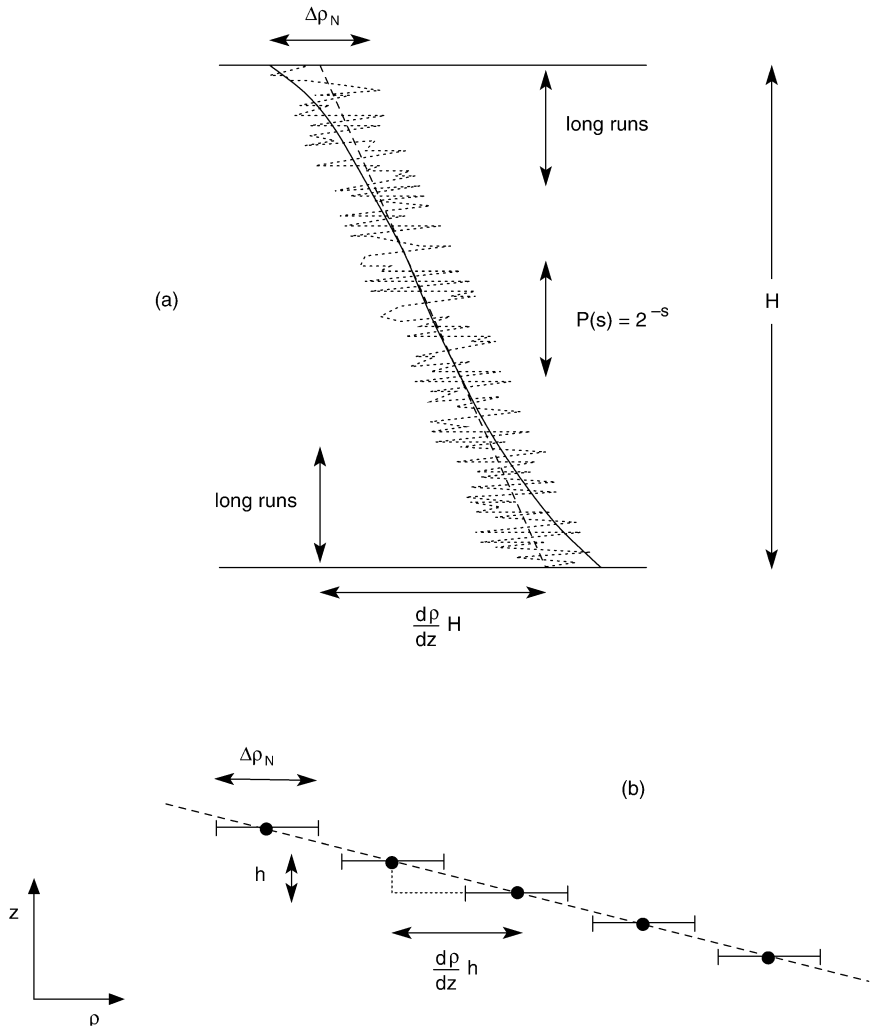

FIG. 1. Schematic to illustrate the two problems associated with using a series of Thorpe fluctuations as a basis for calculating run length. (a) The dashed line represents the original linear profile to which normally distributed random noise of amplitude $\Delta\rho_N$ is added, resulting in the dotted profile. The solid line shows the monotonically increasing sorted profile. The Thorpe fluctuation is the difference between the dotted and solid lines at each point. Note that the Thorpe fluctuation is more likely to be positive at the top of the profile and negative at the bottom, resulting in longer runs at these two extremes. (b) The problem of a low vertical sampling resolution is illustrated. The noise amplitude $\Delta\rho_N$ is smaller than the average density difference between consecutive measurements ($hd\rho/dz$), such that even a noisy profile will be stable, introducing a bias toward short runs.

short runs are also possible if the noise amplitude is comparable with the average density difference between measurements. Noise alone can therefore result in an rms run length that is either greater than or less than $\sqrt{6}$. Thus, comparing either the distribution of run lengths or the rms run length with that expected from a random uncorrelated series is not a reliable way of distinguishing between signal and noise.

Figure 1 illustrates the problem. Figure 1a shows a profile, varying linearly with depth, to which random uncorrelated noise has been added. The sorted profile will depend upon both the initial background gradient and the amplitude of the noise. The Thorpe fluctuation (which is equal to the difference between the dotted and solid lines at each depth) is more likely to be positive at the top of the profile and negative at the bottom, resulting in longer runs at these two extremes. Adding noise to a profile therefore results in a bias toward long runs at its top and bottom, as a direct result of the sorting procedure.

The proportion of the profile over which longer runs are to be expected depends upon the amplitude of the noise. In the limit of very large noise (compared with the background stratification), it is the noise amplitude alone that determines the gradient of the sorted profile with which the measured profile will be compared.

A second problem concerns the vertical sampling resolution. The run lengths calculated from a Thorpe fluctuation (or displacement) series obviously also depend upon the extent to which the noise causes the profile to become unstable. If the amplitude of the noise is small as compared with the average density difference between consecutive measurements, then the perturbed profile will be close to stable. This introduces a bias toward short runs (see Fig. 1b).

Consider a linear profile of length H and vertical sampling interval h, with stratification $d\rho/dz$, to which noise of a fixed distribution with standard deviation (equivalent to an rms amplitude) of $\Delta\rho_N$ is added. We choose to add normally distributed noise because this reflects the distribution of temperature values in the deep mixed layer below the thermohaline staircase structure of the Canada Basin in the Arctic Ocean, where instrument noise dominates the signal (see section 5). The run length varies as a function of the two nondimensional parameters in the system, which are

1) the amplitude of the noise, scaled by the density change over the section of profile considered:

$$Q = \frac{\Delta\rho_N}{(d\rho/dz)H}, \quad \text{and} \tag{1}$$

2) the number of points in the section of profile considered:

$$n = \frac{H}{h}. \tag{2}$$

This second parameter n is equivalent to the ratio B/Q, where B is a third parameter equal to the amplitude of the noise scaled by the average density difference between consecutive measurements:

$$B = \frac{\Delta\rho_N}{(d\rho/dz)h}. \tag{3}$$

We choose to use n instead of B to avoid the unphysical parameter space where $n < 1$. For $n > 1$ but small, it should be noted that lots of realizations, each with a different series of random noise, are required to define the run length distribution since the number of runs in each realization is small. Note also that $d\rho/dz$ here is the stratification of the original linear background profile. While it may seem more sensible to scale the noise with the stratification of the *sorted* profile, since this is easily obtained during the Thorpe analysis, this sorted stratification is not uniform over the length of the profile. In practice, for $Q < 0.3$ the two differ by very little except at the very top and bottom of the section of profile considered.

Figure 2 shows the probability distribution of the Thorpe fluctuation run length, for several combinations of the two parameters. The dashed line in each case shows the probability density function (PDF) of run length for a random uncorrelated series [$P(s) = 2^{-s}$, where s is the run length]. It is clear from the figure that increasing the amplitude of the noise increases the probability of long runs, while reducing the number of points in the profile reduces the probability of long runs (as well as limiting the maximum possible run length). Note that the range of values of Q shown stretches well beyond what we might expect from CTD data in order to demonstrate the behavior of the run-length diagnostic across the entire parameter space. Ideally, CTD profile data will have low levels of noise and a large number of points. If the density resolution is used as an estimate of the noise amplitude, then the Stansfield et al. (2001) datasets from the Strait of Juan de Fuca correspond to $Q \lesssim 0.005$ and $n \approx 1000$; B is equivalent to the inverse of the Stansfield et al. (2001) parameter R and is $O(1)$. Similarly, for the Galbraith and Kelley (1996) datasets, $Q \approx 0.0002$, $n \approx 3000$ and $B \approx 1$. On the other hand, for the interfaces within the thermohaline staircase discussed by Timmermans et al. (2003), $0.02 < Q < 0.3$, $20 < n < 200$, and $2 < B < 40$.

Figure 3a shows a plot of rms run length as a function of Q and n. The rms run length increases smoothly as a function of both parameters, with a rate of increase which drops off as each parameter increases. For small values of Q the rms run length asymptotes to the theoretical prediction for a random series ($\sqrt{6} = 2.45$) at very large n. This is more clear in Fig. 3b where the rms run length is plotted as a function of $\log_{10}Q$ and $\log_{10}n$ in order to emphasize the physically realistic low-amplitude noise regime (where both Q and B are small). The dashed lines show contours of $B(=nQ)$. In this limit it is clear that B dominates—the rms run length is determined entirely by the likelihood of noise to make the profile unstable given the average density difference between consecutive measurements. Only for $B \gtrsim 50$, which corresponds to a very limited region of parameter space, is the rms run length $\approx \sqrt{6}$.

Typical CTD data, such as those analyzed by Stansfield et al. (2001) and Galbraith and Kelley (1996), have $B \approx 1$ and lie in (and beyond) the upper-left corner highlighted in Fig. 3b. The rms run length that might be expected as a result of instrument noise is approximately 1.2, *significantly shorter* than the theoretical prediction of 2.45 for a random series. This suggests that, in rejecting density inversions based on comparison with a probability distribution that has an rms value of 2.45, the Galbraith and Kelley run-length criterion is likely to be too harsh. Some real overturns will typically be rejected as noise.

Except at small n the rms run length increases rapidly with parameter Q, becoming almost double the theoretical limit by the time the noise amplitude is equal to the background density difference over the length of the profile ($Q = 1$). For the Timmermans et al. (2003) deep Canada Basin data $0.02 < Q < 0.3$, $20 < n < 200$, and $2 < B < 40$. Figure 3b illustrates that we might

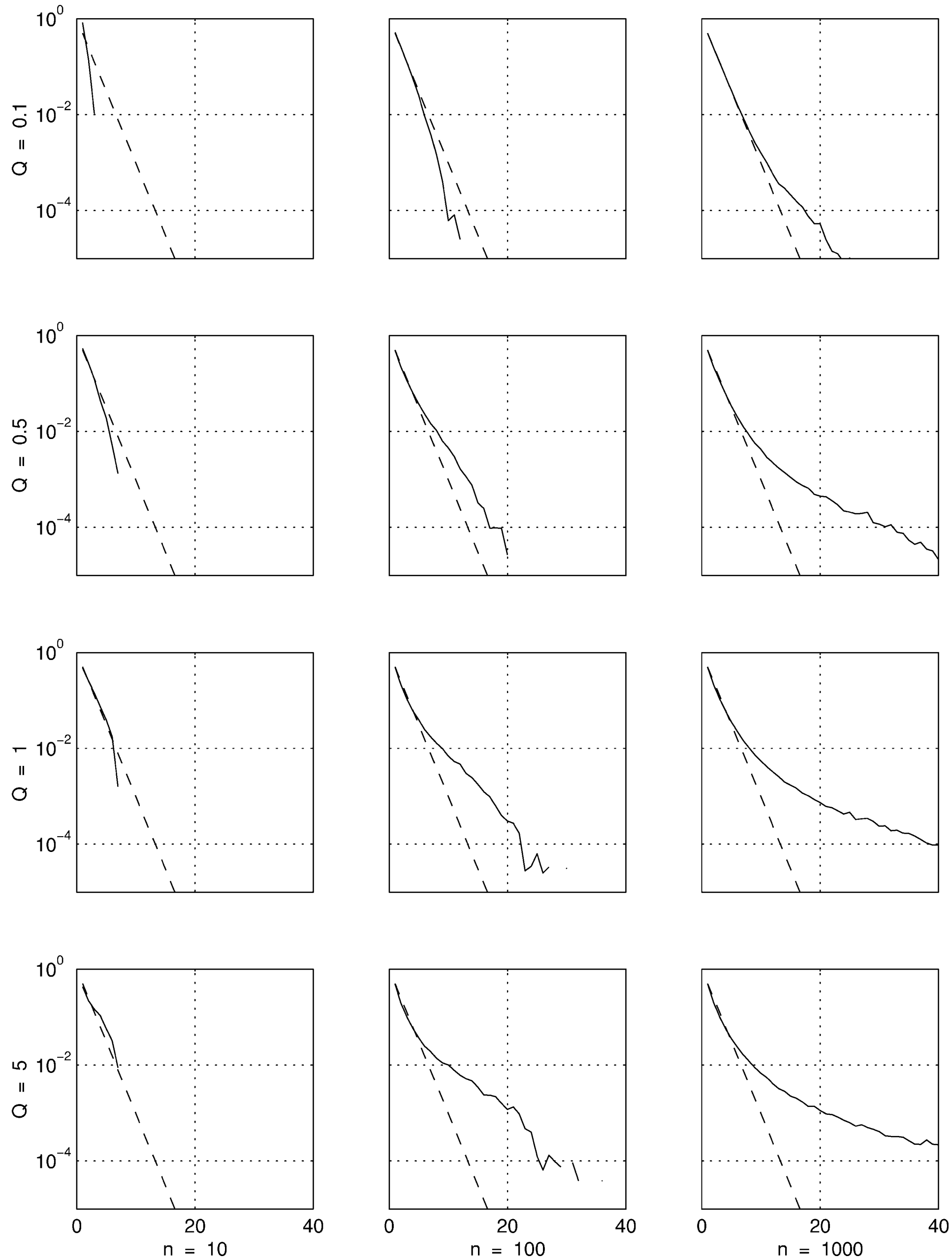

FIG. 2. Run-length probability distribution for the Thorpe fluctuations resulting from a linear density profile to which normally distributed random noise has been added. A range of parameters Q and n is shown. The dashed line in each case shows the PDF of run length for a random uncorrelated series [$P(s) = 2^{-s}$, where s is the run length]. One thousand realizations were performed in each case.

therefore expect rms run lengths both above and below the theoretical value of 2.45.

The arguments so far are based on normally distributed noise added to a linear background profile. If uniformly distributed noise is added instead, again with a standard deviation of $\Delta\rho_N$, then the sorted profile has less extreme values at its top and bottom, and the bias toward long runs is marginally reduced. This leads to a reduction of approximately 15% in the rms run length for high Q (and $B > 50$). A linear background stratification is chosen because, as well as being a reasonably good approximation for the Arctic interface profiles, this allows us to define the relevant parameters easily. While the degree to which a Thorpe fluctuation/displacement run-length series differs from a random series will undoubtedly change with the shape of the background pro-

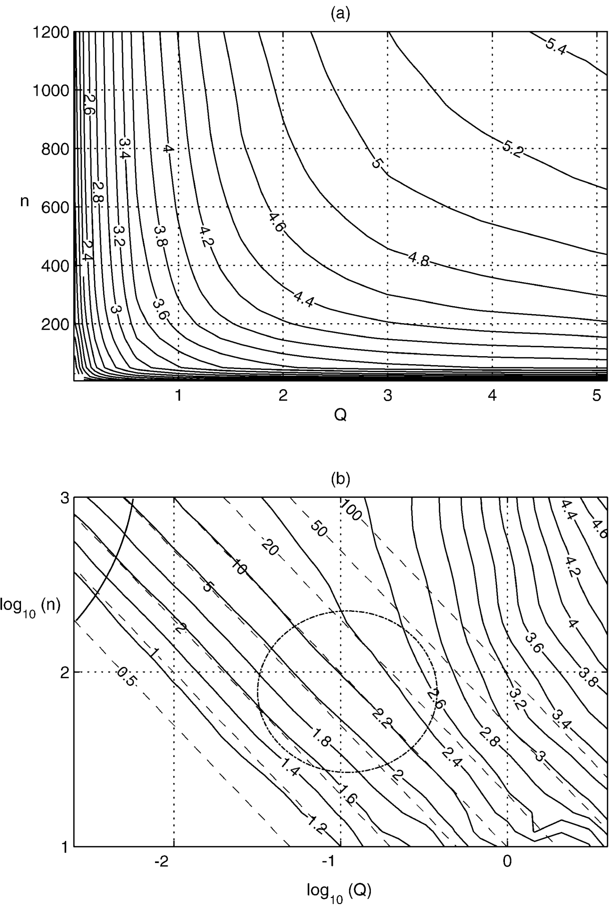

FIG. 3. Rms run length of the Thorpe fluctuations resulting from a linear density profile to which normally distributed random noise has been added (average of 1000 realizations). (a) Rms run length plotted as a function of parameters Q and n; (b) rms run length (solid lines) shown as a function of $\log_{10}Q$ and $\log_{10}n$ to emphasize the low-amplitude noise regime. The dashed lines are lines of constant B (0.5–100), the solid line in the upper-left corner indicates the region of parameter space in which typical CTD data lie, and the ellipse represents the region of parameter space relevant to the Timmermans et al. (2003) deep Canada Basin data.

file, the essence of our results will remain unaltered: comparing the run-length distribution with a random series is not a reliable indicator of noise.

In summary, Figs. 2 and 3 illustrate that it is not generally appropriate to compare the PDF of run lengths to $P(s) = 2^{-s}$, or the rms run length to $\sqrt{6}$, since in the presence of noise Thorpe fluctuation and displacement series do not behave as random uncorrelated series. The two figures do provide an alternative standard with which to compare run lengths in order to assess the likelihood of them arising because of instrument noise.

3. Effect of noise on the Thorpe scale

Here we consider the distribution of Thorpe displacements (and the resulting Thorpe scale) for inversions in a density profile generated entirely by noise. Our motivation is to further improve our understanding of the signature of instrument noise and establish other criteria to help distinguish a real signal from noisy data. Again, we consider a linearly stratified density profile of length H and vertical sampling interval h, with stratification $d\rho/dz$, to which normally distributed noise is added that has a standard deviation (rms amplitude) of $\Delta\rho_N$. All density inversions generated are spurious.

Figure 4 shows the probability distribution of Thorpe displacements as a function of parameter Q, with $n = 5000$ in all panels. The displacements are scaled by the total height of the profile, H, such that the maximum possible displacement is equal to 1. When $Q \lesssim 0.01$, the Thorpe displacements are approximately normally distributed, as one might expect. (When uniformly distributed noise is added instead, the PDF of Thorpe displacements is ‘‘top hat’’ shaped for small Q.) However, as the amplitude of the noise increases, the number of parcels displaced and the distance they move both increase; the PDF changes shape with $p(L/H)$ dropping off linearly with L/H for small displacements. In the noise-dominated limit (see panel 6 where $Q = 5$) the PDF is completely linear, consistent with a situation where every parcel has an equal probability of being displaced to every other point. In this case the probability of a displacement m, for $1 \le m \le n$, is

$$P(m) = \frac{2(n - m)}{n^2} \tag{4}$$

with a probability $1/n$ of zero displacement (Stansfield et al. 2001).

Plotted in Fig. 5a is the rms of these displacements, the Thorpe scale, as a function of Q and n. (Displacements of zero only represent a significant fraction of the distribution for very small B and were included in the calculation.) The Thorpe scale is now normalized by the amplitude of the noise:

$$L'_T = \frac{L_T}{[\Delta\rho_N/(d\rho/dz)]} = \frac{L_T}{\Delta\rho_N}\left(\frac{d\rho}{dz}\right). \tag{5}$$

For small values of Q, $L'_T = 1$; that is, the Thorpe scale is equal to the vertical displacement expected, given the background stratification and the amplitude of the noise. As Q increases, however, the normalized Thorpe scale drops off quickly. This is because the Thorpe scale becomes limited by the length of the profile.

From the definition of parameter Q in Eq. (1), the normalized Thorpe scale L'_T can also be expressed as

$$L'_T = \frac{L_T}{HQ}.$$

Shown in Fig. 5b is $L'_T \times Q$ (or the ratio L_T/H of the dimensional Thorpe scale to the height of the profile), again as a function of Q and n. Note that the range of parameter n differs from that in Fig. 5a. Except for $n \lesssim 10$, as Q increases L_T/H asymptotically approaches a value of approximately 0.41. This asymptotic value

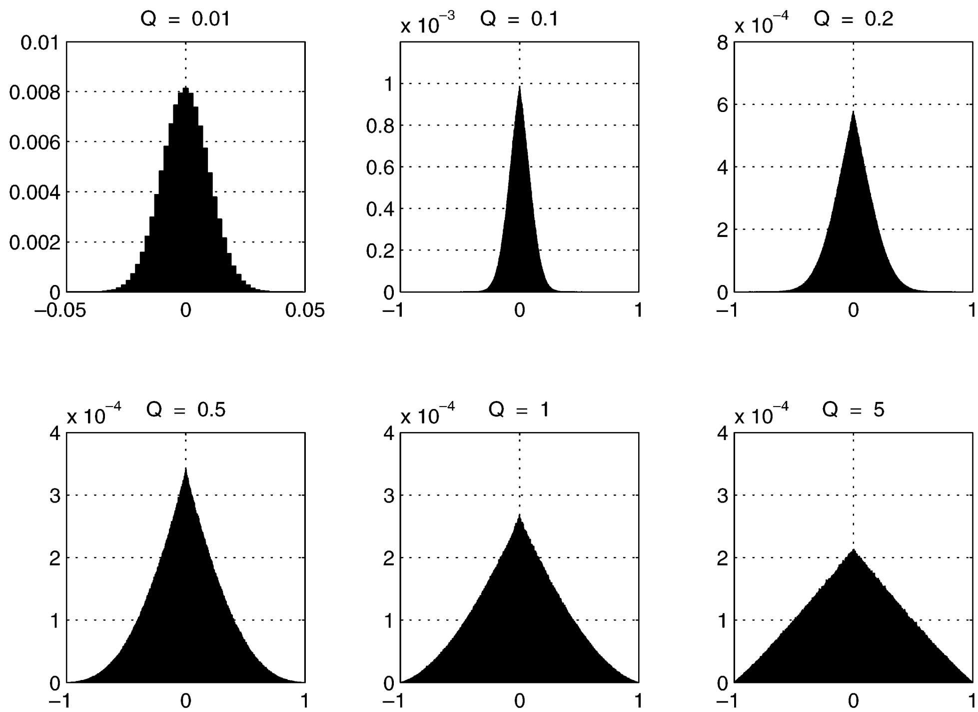

FIG. 4. Probability distribution of Thorpe displacements scaled by the height of the profile, L/H, for a range of values of the parameter Q; n is equal to 5000 throughout. Note that the limits on the axes change among panels.

occurs in the noise-dominated limit where each parcel has an equal probability of being displaced to every other point and is as we might expect from the linear probability distribution in Eq. (4):

$$\frac{L_T}{H} = \frac{1}{n}\left[\sum_{m=1}^{n} m^2 P(m)\right]^{1/2} = \left[\frac{2}{3}\frac{(n+1)\left(n+\frac{1}{2}\right)}{n^2} - \frac{1}{2}\frac{(n+1)^2}{n^2}\right]^{1/2}. \quad (6)$$

As $n \to \infty$:

$$\frac{L_T}{H} \to \left(\frac{2}{3} - \frac{1}{2}\right)^{1/2} = \frac{1}{\sqrt{6}} \approx 0.41.$$

When uniformly rather than normally distributed noise is added to the linear profile, there is almost no difference in the Thorpe scales that result.

Lorke and Wüest (2002) have recently suggested that the maximum displacement $L_{\max}$ required in the reordering of a profile is a more robust indicator of diapycnal mixing than the rms displacement L_T. They show that the two are related by a universal spectrum of Thorpe displacements and then argue that $L_{\max}$ is easier to resolve, since only the largest displacement need be detectable, whereas an estimate of the rms value, L_T, requires the entire distribution of displacements to be properly resolved [although, as Stansfield et al. (2001) show, L_T is dominated by the largest displacements in any case]. The ratio $L_T/L_{\max}$ expected to result from normally distributed random noise ranges from about 0.2 to 0.5 over a physically realistic range of parameters Q and n (not shown). If the ratio seen in data were to lie outside this range, it may provide a useful way of distinguishing between signal and noise. However, as Lorke and Wüest (2002) demonstrate using data from a range of stratifications and mixing levels, the ratio lies between 0.19 and 0.58, and so is of no help here.

4. Statistics of density inversions generated simply by noise

While the Thorpe scale itself presents an average picture, the size and number of overturns within a profile can provide important insight into the physical processes that lead to small-scale mixing. The Galbraith and Kelley approach of rejecting density inversions that fail a series of stringent tests also first requires the identification of individual overturning regions within a profile.

A ‘‘complete inversion’’ (either a real overturn or a spurious inversion generated by noise) is defined as the minimum extent of a region encompassing density data points that must be exchanged with each other (and no others) in order to achieve a stable profile. Figure 6 shows the number of complete inversions identified across the two-parameter space, together with the mean

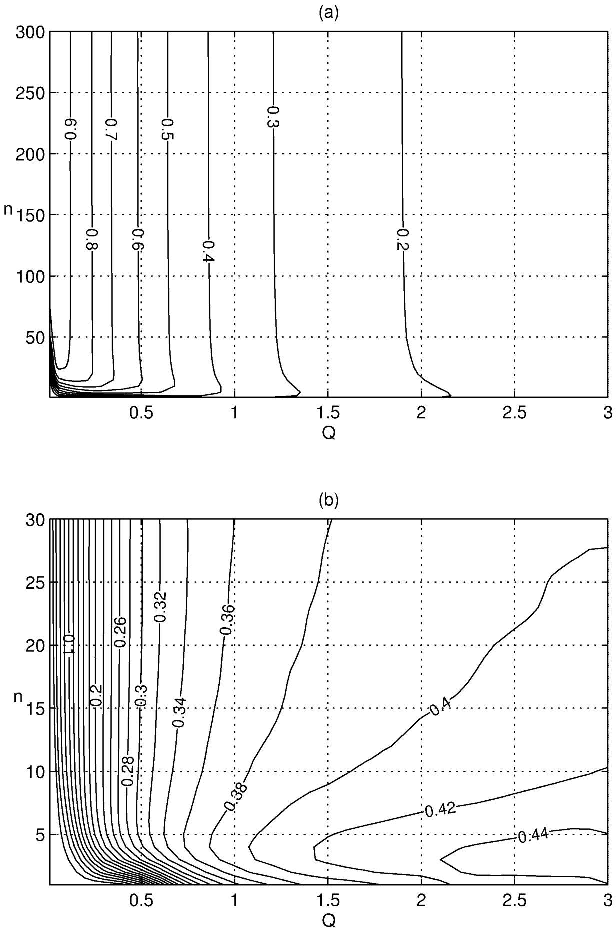

FIG. 5. (a) Thorpe scale, normalized by the amplitude of the noise, as a function of Q and n and (b) normalized Thorpe scale multiplied by Q. This is equivalent to the ratio L_T/H of the dimensional Thorpe scale to the length of the profile.

length of these inversions. It demonstrates that, when the vertical sampling resolution is sufficiently high, even small-amplitude noise results in the entire column overturning; that is, density inversions will overlap such that parcels must be exchanged throughout the whole water column to achieve stability, producing a single overturning region of length n. Consideration of the distribution of inversion lengths (not shown) reinforces this conclusion. The dashed line in Figs. 6a and 6b is the $B = 10$ contour. For $B < 10$ the tendency for noise to cause density inversions is reduced, and so there is less tendency for overlap.

In light of these results one might at first consider the existence of a single large overturn spanning most of the profile as evidence of noise. However, in active parts of the water column density inversions arising from real turbulent mixing can overlap significantly too (i.e., they are close enough together and have large enough density variations that water parcels must be exchanged between them in the sorting process). Therefore it is not clear at this stage whether this tendency for noise-generated density inversions to overlap can be used to distinguish between signal and noise.

5. Thermohaline staircase interfaces in the deep Canada Basin

The deep Canada Basin of the Arctic Ocean provides one example of a region in which mixing levels are low and the Thorpe scale approach is subject to problems arising from instrument noise. At depths between about 2400 and 2800 m there is a potentially double-diffusive feature known as a thermohaline staircase. It consists of convectively mixed homogeneous layers approximately 50 m in thickness separated by stably stratified interface regions in which the temperature changes

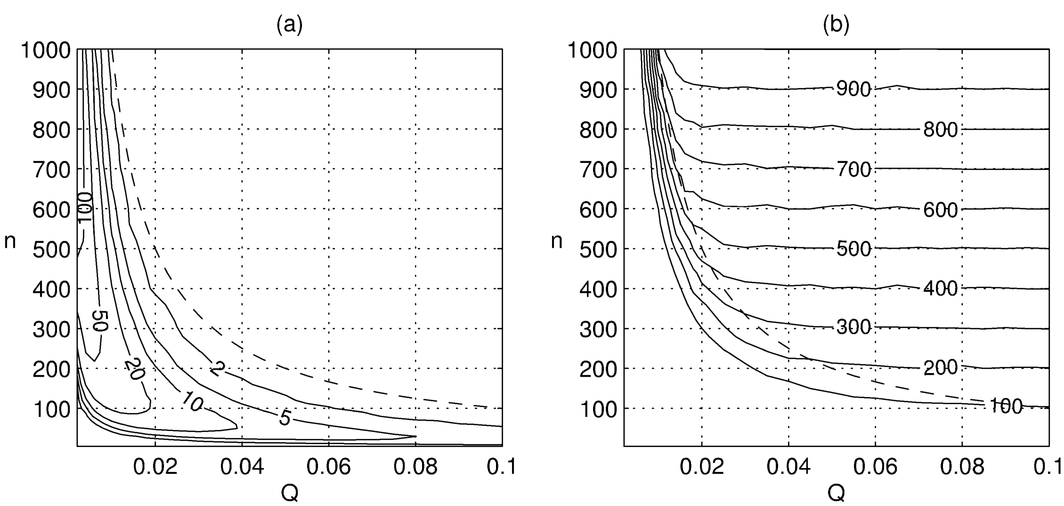

FIG. 6. Mean (a) number and (b) length of density inversions for the two-parameter space (calculated over 1000 realizations). The length of inversions is given in terms of the number of points. The dashed line in both (a) and (b) corresponds to $B = 10$.

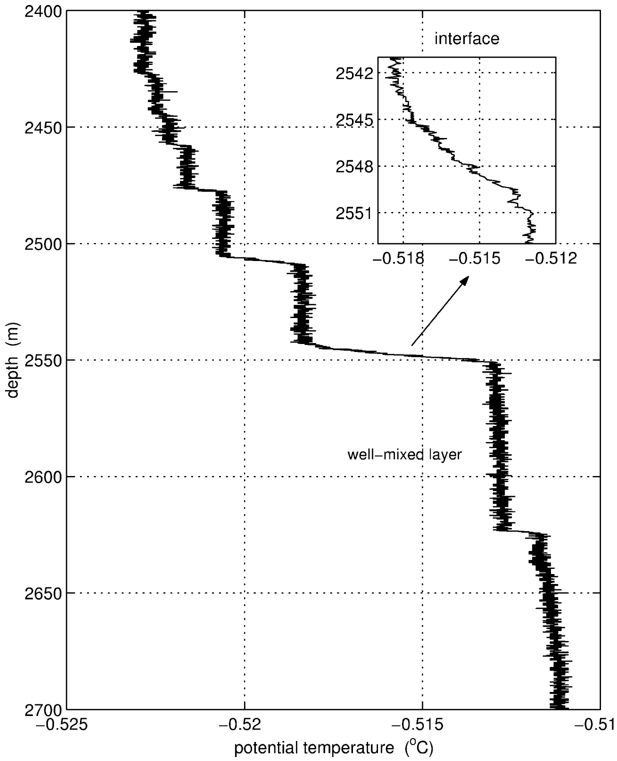

FIG. 7. Typical profile of potential temperature over the thermohaline staircase in the deep Canada Basin (Aug 2002). The inset shows a section of the profile within an interface between two well-mixed steps.

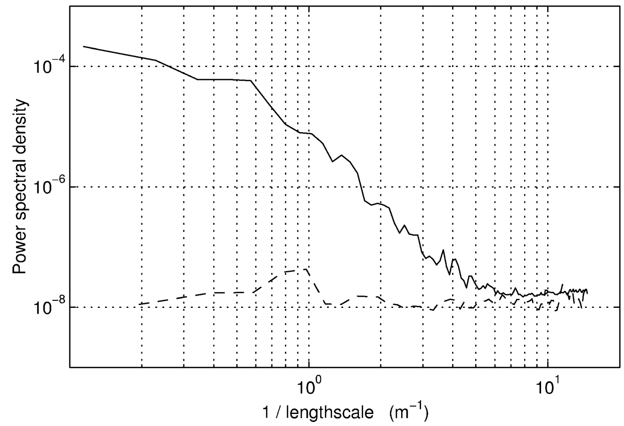

FIG. 8. Spectrum of θ for the profile shown in Fig. 7 from 50 to 3100 m (solid line), and from 2900 to 3100 m (dashed line).

sharply by about 0.005°C over a few meters. [A full discussion of this staircase can be found in Timmermans et al. (2003).]

Even when sampled with a high-resolution CTD and a vertical sampling interval of about 0.03 m the interface regions within the staircase have parameter values $0.02 < Q < 0.3$, $20 < n < 200$, and $2 < B < 40$. As such they represent a regime in which comparison with the revised standard for noise generated inversions developed here may be important. Measurement of shear-induced mixing within the interfaces is of interest because it tells us how much heat is being fluxed upward through the staircase from the geothermally heated layer below. If the density inversions in the interfaces are all treated as real overturns, then the resulting K_v is large enough for the entire geothermal heat flux to escape.

Timmermans et al. (2003) calculate Thorpe scales and run lengths within two interface profiles and compare the results with values obtained when random noise is added to the corresponding sorted interface profiles. They conclude that both Thorpe scales and run lengths are not significantly different from those that might be expected, given the instrument noise level. However, this conclusion is based on comparison with an rms run length of $\sqrt{6}$, which we have shown to be appropriate only over a very limited region of parameter space.

Here we revisit this issue using a larger volume of data collected in August 2002. Having established the effect of random instrument noise on run length, Thorpe scale, and density inversion statistics, we are in a position to perform a more detailed comparison and establish whether the structure seen in the interface profiles does indeed result from real mixing or merely from instrument noise. This section should be viewed as an example of how the understanding gained in the previous three sections can be applied practically to help us to distinguish between signal and noise in marginal regimes.

Figure 7 shows a typical profile of potential temperature (θ) over the well-mixed layers and interfaces that make up the thermohaline staircase. It is one of a series of CTD profiles taken in August 2002 and not included in the Timmermans et al. (2003) analysis. Following Timmermans et al. (2003) we analyze profiles of potential temperature here because the salinity data are not sufficiently accurate to use potential density. The difference in temperature across each interface is ≤0.005°C. The inset shows an enlarged view of the profile within one interface. Inversions on scales of about a meter are visible in the data.

The resolution of the Sea-Bird CTD SBE-911plus thermistor is stated by the manufacturer to be 0.0004°C, approximately 10% of the temperature difference across a typical interface. This means that any density inversions present in the interfaces will certainly have a large Q. However, to quantify the level of instrument noise associated with the data more accurately, the standard deviation of potential temperature variations in the deep mixed layer, below the thermohaline staircase, was calculated for each profile. Potential temperature measurements in this deep mixed layer were found to be approximately normally distributed with a standard deviation of 0.000 13°C, one-third of the manufacturer's quoted value. This value of 0.000 13°C was taken to be the standard deviation (or rms amplitude) of the instrument noise and to set $\Delta\rho_N$ in our parameter definitions.

The dashed line in Fig. 8 shows the spectrum of θ in the deep mixed layer below 2900 m, for the CTD profile

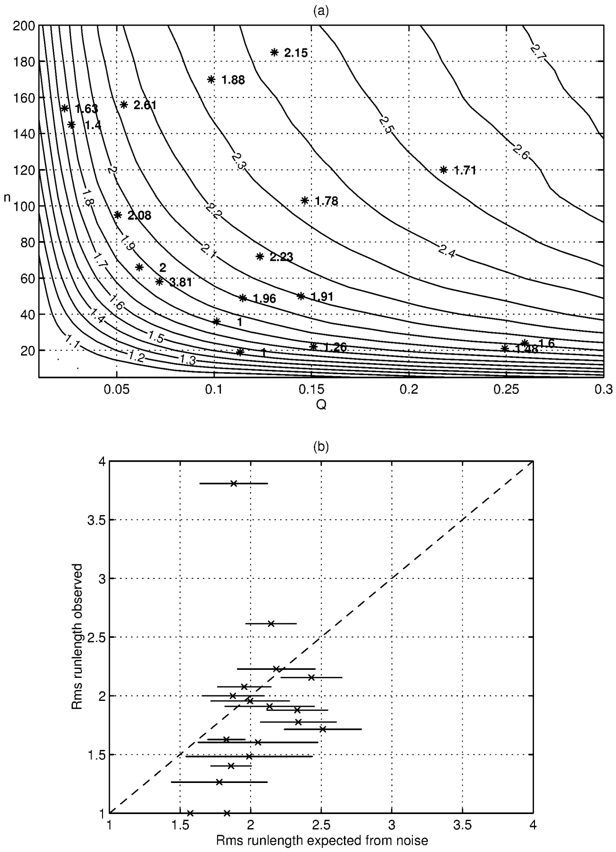

FIG. 9. (a) Rms run length (averaged over 1000 realizations) expected as a result of normally distributed random noise. Asterisks represent the position of an interface profile in the two-parameter space and are labeled with the measured rms run length. (b) Each of the measured rms run-length values plotted against the corresponding noise-induced value, for ease of comparison. The error bars indicate ±1 std dev.

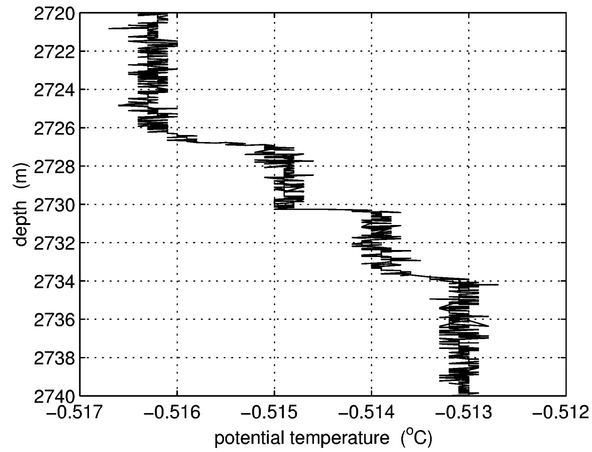

FIG. 10. Vertical profile of potential temperature over one of the interfaces that has an rms run length several standard deviations higher than that which might be expected from random noise. Note that the interface actually consists of three separate thinner interfaces, separated by well-mixed steps.

in Fig. 7. The spectrum here is white, and the variance of θ is $\sigma^2 = (1.6 \times 10^{-8})°C^2$, which agrees with our estimate for the standard deviation of $\sigma = 0.000\ 13°C$. The solid line in Fig. 8 shows the spectrum of θ over the whole profile from 50 to 3100 m. It demonstrates that on vertical scales smaller than about 15-cm variability in the profile is white and likely dominated by instrument noise. However, since the Thorpe displacements within an overturn can be many times larger than the Thorpe scale [and it is the largest displacements that contribute most (Stansfield et al. 2001)], this suggests that values of L_T as low as a few tens of centimeters ought to be discernible.

For the purposes of data analysis, an interface was defined as the approximately linearly varying section of profile between two well-mixed steps in potential temperature. Any reduction of the gradient in θ toward the ends of each interface was not included. Only data from the downcast of each profile were used, and these were processed such that only the first occurrence of each pressure remained, in order to minimize the likelihood of including turbulence generated by the CTD itself as it moved up and down through the water due to ship heave. This reduced the vertical sampling resolution to approximately 0.05 m in most cases.

It is impossible to calculate the Galbraith and Kelley (1996) cutoff run length for the interface profiles since for no run length is the probability of occurrence greater than their criterion of 2×2^{-s}. In fact, for runs of all lengths except 1 the probability of occurrence is $<2^{-s}$, and the probability distribution looks like the middle top panel in Fig. 2. This suggests that we are in a regime governed by the low vertical sampling resolution and, while Galbraith and Kelley would likely conclude that the mixing is too small to be detected above the noise threshold, it highlights the need for an extension of their approach.

Figure 9a shows the rms run length that we might expect to arise as a result of normally distributed noise over the relevant region of parameter space. The average of 1000 realizations is plotted. Superimposed are values of rms run length measured in 18 interfaces in the deep Canada Basin during August 2002. Figure 9b shows each of the measured rms run-length values plotted against the corresponding noise-induced value, for ease of comparison. The error bars indicate ±1 standard deviation.

In almost all of the 18 interfaces the rms run length lies within one or two standard deviations of the value that we would expect to arise in a profile constructed simply from random noise. Values are generally also slightly lower than those expected from noise. The only exceptions to this are the two interfaces with observed rms run lengths of 3.81 and 2.61. A closer look at the second of these, however, shows that θ does not vary continuously over the interface, which instead contains several smaller well-mixed steps (see Fig. 10). This may

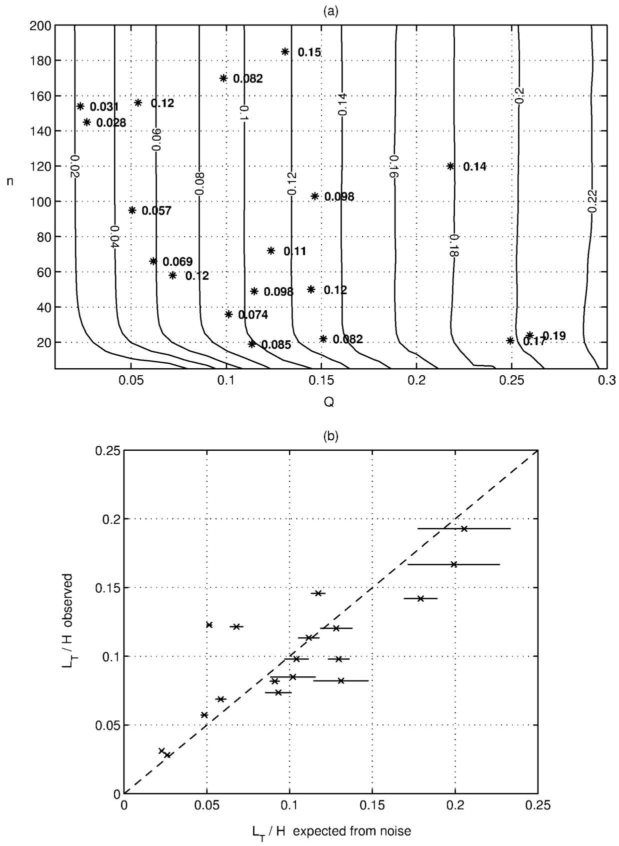

FIG. 11. (a) Dimensional Thorpe scale normalized by the height of the profile, as in Fig. 5. Superimposed are the Thorpe scales measured in each of the 18 interfaces. (b) Each of the measured values of L_T/H plotted against the corresponding noise-induced value, for ease of comparison. The error bars indicate ± 1 std dev.

be a result of the interface splitting discussed by Kelley (1988). If the individual interfaces within this short section of profile are considered in turn, their rms run lengths lie within the range expected from instrument noise. This leaves only one of the 18 interfaces with an rms run length that lies more than two standard deviations away from that expected due to random noise. Since for a normal distribution we expect 5% of cases to lie outside two standard deviations of the mean, there is no evidence from the run-length diagnostic that the inversions present in the interface profiles are the result of anything other than instrument noise.

It should be noted that there is some uncertainty associated with the positioning of the interface points in parameter space. This arises because of the uncertainty in noise level $\Delta\theta_N$, background gradient $d\theta/dz$, and interface profile length H and amounts to an error of about 10% in Q. Choosing the gradient of the sorted profile rather than the measured profile results in a negligible change in Q.

Figure 11a shows the Thorpe scale, normalized by the height of the profile, that we might expect to arise from normally distributed random noise, as in Fig. 5b. Superimposed are the values measured in the 18 interface profiles. In Fig. 11b each of the measured values of L_T/H is plotted against the corresponding noise-induced value, for ease of comparison. Again, the error bars indicate ± 1 standard deviation. Three of the interfaces have Thorpe scales significantly higher than might be expected to arise simply because of noise. Two of these correspond to the interfaces already identified above (recall that one of these appears to consist of several well-mixed steps rather than turbulent overturns). The third, with $L_T/H = 0.15$, has a run length that is well within the bounds expected from random noise. Again, we cannot regard these instances of elevated L_T as statistically significant.

There is one other concern. Predictions for the rms run length and Thorpe scale arising from noise, shown in Figs. 9 and 11 and elsewhere throughout the paper, are based on a distribution of temperature values rounded to the nearest 10^{-16}. However, the θ data used here were collected at a resolution of 0.0001°C, which is close to the noise level of 0.000 13°C. When the resolution of the artificially generated profiles (constructed from random noise added to a linear profile) is limited in the same way, the rms runlength changes slightly but not by enough to affect any of the conclusions arrived at here. Thorpe scales remain virtually unchanged.

While in some interfaces as many as 14 distinct inversions were detected, in most cases only 1 or 2 distinct inversions were apparent because of the overlap discussed in section 4. Parcels must be exchanged throughout the entire profile to achieve stability. Calculating the Thorpe scales of individual inversions within each profile (or neglecting the displacements of zero) results in only a small increase over the Thorpe scales shown in Fig. 11.

Since only 1 of 18 interfaces has both a Thorpe scale and an rms run length that cannot be explained (at the 95% confidence level) by noise, we conclude that the density inversions present in the interfaces between steps of the thermohaline staircase are indistinguishable from the effects of random noise. (The existence of real overturns in one profile suggests the possibility that the interfaces may be intermittently turbulent, but the heat flux through the staircase is certainly much less than the geothermal flux from below.) On the basis of our revised expectations for the characteristics of noise in a density profile we therefore conclude that the structure in the interface temperature profiles arises because of instrument noise.

6. Discussion and conclusions

There is currently a need for good spatial and temporal coverage in our measurements of the diapycnal mixing rate in order to address fundamental questions about the link between small-scale mixing and the global ocean circulation (and to parameterize the effects of small-scale processes in numerical climate models).

Since it is based on standard CTD instrumentation, the Thorpe scale approach is attractive, and establishing to what extent and in what regimes it can contribute is a pressing problem.

The Thorpe scale technique has several limitations. Here we have addressed only the problem of distinguishing a signal arising from turbulent mixing from the spurious density inversions generated by instrument noise. Our focus is on environments where the stratification and mixing levels are low since this is where existing conservative approaches most risk throwing out important ''babies'' with the ''bathwater'' (i.e., discarding signal as well as noise). In this paper we have identified the characteristics of density inversions arising entirely as a result of instrument noise. As such, we have established a standard against which CTD data can be compared for the purpose of distinguishing between signal and noise. Our key results are summarized here.

- The ''run length'' introduced by Galbraith and Kelly (1996) can be a useful diagnostic (see, e.g., Stansfield et al. 2001). However, the sorting procedure implicit in the definition of Thorpe fluctuation and displacement series causes them to behave differently from a random, uncorrelated series, even in the absence of real overturns. As a result, it is not generally appropriate to compare the PDF of run lengths with that expected from a random series [$P(s) = 2^{-s}$], or the rms run length with $\sqrt{6}$, in order to distinguish between signal and noise. Comparison with a modified PDF (or rms) is required.
- CTD data, including those analyzed by Galbraith and Kelley (1996), often lie in a regime where the noise amplitude is small in comparison with the density change over the whole section of profile considered, yet comparable to the average density difference between consecutive measurements. Expected values of the rms run length in this low-noise regime are less than 1.5, significantly shorter than the $\sqrt{6}$ expected from a random uncorrelated series. This implies that rejecting density inversions based on comparison with the value $\sqrt{6}$ is likely to result in the rejection of real turbulent overturns in many CTD datasets.
- The interfaces within the thermohaline staircase in the deep Canada Basin, previously discussed in detail by Timmermans et al. (2003), are in a regime where the amplitude of the noise is a significant fraction of the density change over the section of profile considered, yet still of the same order as the average density difference between measurements. As such, this region provides a useful test bed for our revised standard against which CTD data might be compared. From their original analysis Timmermans et al. (2003) concluded that inversions present in the data were likely to have arisen entirely due to instrument noise. A more detailed comparison presented here, based on the expectations for noise-generated density inversions and on new data, leads us to the same conclusion.

This example illustrates how the insight gained here can be made use of in distinguishing between signal and noise. However, further work is required in order to automate the overturn detection and noise rejection process. In this sense, the work presented here can only be supplementary to that of Galbraith and Kelley (1996), who propose a practical series of definitive tests, and merely extends their criteria such that, by comparing with a more accurate picture of the density inversions generated by noise, fewer real overturns are likely to be rejected. An obvious next step would be to use analytical (or empirical) relations for the noise-generated run length and Thorpe scale as functions of Q and n.

Improving our ability to distinguish between signal and noise may allow us to push the bounds of the Thorpe scale technique into regions where the stratification and/or amount of turbulent mixing is low. This includes a large fraction of the world's oceans. However, it is important to remember that there are also other problems associated with the Thorpe-scale approach, and in reality the noise issues dealt with here may not be the limiting factor. Rarely can we use inversions in potential temperature as a reliable indicator of turbulent overturns, because of the possibility of salinity-compensated intrusions. In most areas inversions in potential density must be identified. However, obtaining a reliable density profile from conductivity and temperature measurements is itself often difficult because of the mismatch in response times of the temperature and conductivity sensors. Another major issue is the likelihood of turbulence generated by the CTD package itself (particularly when it is included in a rosette and there is large ship motion), which may lead to very different forms of the noise distribution from that assumed here. Only measurements made by a free-falling CTD can really hope to represent the undisturbed water column. We should also emphasize here the statistical nature of any mixing estimate. Turbulence is nonstationary and inhomogeneous, and the Thorpe scale approach is not exempt from the need for large amounts of data in both space and time to give reliable estimates. (The Canada Basin data presented here are marginal in this respect.)

In light of these other limitations the revised approach suggested above is unlikely to make a big difference in the high-mixing environments where the Thorpe-scale technique has mostly been applied so far. Nevertheless, the work presented here represents a first step in establishing the degree to which the Thorpe-scale approach might contribute to global mixing measurements in the future.

Acknowledgments. We thank Fiona McLaughlin and Eddy Carmack of the Institute of Ocean Sciences, Sidney, British Columbia, as well as the captain and crew of the CCGS *Louis S. St.-Laurent*, for the 2002 Canada Basin data. Fruitful discussions with Mary-Louise Timmermans, Richard Dewey, Ann Gargett, Eric Kunze, Nathan Gillett, and Jim Moum are gratefully acknowledged. We also thank Dan Kelley and an anonymous reviewer for their insightful comments and suggestions on the manuscript. Financial support was provided by NSERC and ONR.

REFERENCES

Alford, M. H., and R. Pinkel, 2000: Observations of overturning in the thermocline: The context of ocean mixing. *J. Phys. Oceanogr.,* **30,** 805–832.

Baumert, H., and H. Peters, 2000: Second-order closures and length scales for weakly stratified turbulent shear flows. *J. Geophys. Res.,* **105,** 6453–6468.

Bryan, F. O., 1987: Parameter sensitivity of primitive equation ocean general circulation models. *J. Phys. Oceanogr.,* **17,** 970–985.

Crawford, W. R., 1986: A comparison of length scales and decay times of turbulence in stably stratified flows. *J. Phys. Oceanogr.,* **16,** 1847–1854.

Dillon, T. M., 1982: Vertical overturns: A comparison of Thorpe and Ozmidov scales. *J. Geophys. Res.,* **87,** 9601–9613.

Ferron, B., H. Mercier, K. Speer, A. Gargett, and K. Polzin, 1998: Mixing in the Romanche Fracture Zone. *J. Phys. Oceanogr.,* **28,** 1929–1945.

Finnigan, T. D., D. S. Luther, and R. Lukas, 2002: Observations of enhanced diapycnal mixing near the Hawaiian Ridge. *J. Phys. Oceanogr.,* **32,** 2988–3002.

Galbraith, P. S., and D. E. Kelley, 1996: Identifying overturns in CTD profiles. *J. Atmos. Oceanic Technol.,* **13,** 688–702.

Gregg, M. C., 1989: Scaling turbulent dissipation in the thermocline. *J. Geophys. Res.,* **94,** 9686–9698.

——, T. B. Sanford, and D. P. Winkel, 2003: Reduced mixing from the breaking of internal waves in equatorial waters. *Nature,* **422,** 513–515.

Kelley, D. E., 1988: Explaining effective diffusivities within diffusive oceanic staircases. *Small-Scale Turbulence and Mixing in the Ocean,* J. C. J. Nihoul and B. M. Jamart, Eds., Elsevier Science, 481–502.

Kunze, E., L. K. Rosenfeld, G. S. Carter, and M. C. Gregg, 2002: Internal waves in Monterey Submarine Canyon. *J. Phys. Oceanogr.,* **32,** 1890–1913.

Lorke, A., and A. Wüest, 2002: Probability density of displacement and overturning length scales under diverse stratification. *J. Geophys. Res.,* **107,** 3214, doi:10.1029/2001JC001154.

Marotzke, J., 1997: Boundary mixing and the dynamics of three-dimensional thermohaline circulations. *J. Phys. Oceanogr.,* **27,** 1713–1728.

Oakey, N. S., 1982: Determination of the rate of dissipation of turbulent energy from simultaneous temperature and velocity shear microstructure measurements. *J. Phys. Oceanogr.,* **12,** 256–271.

Ozmidov, R. V., 1965: On the turbulent exchange in a stably stratified ocean. *Izv. Acad. Sci. USSR, Atmos. Oceanic Phys.,* **1,** 861–871.

Polzin, K. L., J. M. Toole, and R. W. Schmitt, 1995: Finescale parameterization of turbulent dissipation. *J. Phys. Oceanogr.,* **25,** 306–328.

Stansfield, K., C. Garrett, and R. Dewey, 2001: The probability distribution of the Thorpe displacement within overturns in Juan de Fuca Strait. *J. Phys. Oceanogr.,* **31,** 3421–3434.

St. Laurent, L., and R. W. Schmitt, 1999: The contribution of salt fingers to vertical mixing in the North Atlantic Tracer Release Experiment. *J. Phys. Oceanogr.,* **29,** 1404–1424.

Thorpe, S. A., 1977: Turbulence and mixing in a Scottish loch. *Philos. Trans. Roy. Soc. London,* **286A,** 125–181.

Timmermans, M.-L., C. Garrett, and E. Carmack, 2003: the thermohaline structure and evolution of the deep waters in the Canada Basin, Arctic Ocean. *Deep-Sea Res.,* **50A,** 1305–1321.

Welander, P., 1986: Thermohaline effects in the ocean circulation and related simple models. *Large-Scale Transport Processes in the Oceans and Atmosphere,* J. Willebrand and D. L. T. Anderson, Eds., D. Reidel, 163–200.