Thermohaline circulation: a missing equation and its climate-change implications

Hsien-Wang Ou

Climate Dynamics

Observational, Theoretical and
Computational Research on the Climate
System

ISSN 0930-7575

Clim Dyn

DOI 10.1007/s00382-017-3632-y

Your article is protected by copyright and all rights are held exclusively by Springer-Verlag Berlin Heidelberg. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Thermohaline circulation: a missing equation and its climate-change implications

Hsien-Wang Ou^{1,2}

Received: 12 September 2016 / Accepted: 21 February 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract We formulate a box model of coupled ocean–atmosphere to examine the differential fields interactive with the thermohaline circulation (THC) and their response to global warming. We discern a robust convective bound on the atmospheric heat transport, which would divide the climate regime into warm and cold branches; but unlike the saline mode of previous box models, the cold state, if allowed, has the same-signed—though weaker—density contrast and THC as the present climate, which may explain its emergence from coupled general circulation models. We underscore the nondeterminacy of the THC due to random eddy shedding and apply the fluctuation theorem to constrain the shedding rate, thus closing the problem. The derivation reveals an ocean propelled toward the maximum entropy production (MEP) on millennial timescale (termed “MEP-adjustment”), the long timescale arising from the compounding effect of microscopic fluctuations in the shedding rate and their slight probability bias. Global warming may induce hysteresis between the two branches, like that seen in GCMs, but the cold transition is far more sensitive to the moistening than the heating effects as the latter would be countered by the hydrological feedback. The uni- or bi-modality of the current state—hence

whether the THC may recover after the cold transition—depends on the global-mean convective flux and may not be easily assessed due to its observed uncertainty.

Keywords Thermohaline circulation · Fluctuation theorem · Maximum entropy production · Ocean hysteresis · Global warming

1 Introduction

We define the thermohaline circulation (THC, all acronyms listed in “[Appendix A](#)”) as the mass exchange rate between warm and cold watermasses, which effectuates their net property flux. In coarse-grained general circulation models (GCMs), the exchange takes the form of the meridional overturning circulation (MOC) whereby the warm outflow across the subtropical front is returned via upwelling through the thermocline. In the actual ocean or fine-grained GCMs however, the warm outflow may be partly composed of eddy shedding, and if there were returning cold eddies, they would reduce the net outflow transport hence the MOC but not the THC. In other words, our THC is a generalization of the MOC transport to include the eddy exchange across the front.

Being the transport mechanism between the warm and cold watermasses, the THC would moderate their temperature and salinity difference, yet at the same time the THC is dynamically linked to the differential density¹ on which the

Lamont-Doherty Earth Observatory Contribution Number 8096.

Hsien-Wang Ou—Retired from Lamont-Doherty Earth Observatory of Columbia University.

✉ Hsien-Wang Ou
hsienou0905@gmail.com

¹ Department of Earth and Environmental Sciences, Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY 10964, USA

² 100 Overlook Terrace, Apt. 47, New York, NY 10040, USA

¹ It has been argued that the THC requires the wind work for its genesis on account of the Sandstrom theorem (Wunsch 2002), which however overlooks the dynamics that the thermal-induced pressure gradient can drive a flow. In addition, as discussed in Ou (2007), the varying wind work can be accommodated by changing thermocline depth (hence the potential energy) without impacting the THC to be derived here.

temperature and salinity differences exert opposite effects. Because of the latter, the ocean may exhibit bistable thermal and saline modes, as first illustrated by Stommel (1961) via an ocean box-model. Such bistability has provided a dynamical basis to the hysteresis seen in GCMs with its significant climate-change implications. But as an interpretative tool, ocean box-models suffer from two palpable shortfalls regarding its lack of physical closure and atmospheric coupling, as expounded below.

A significant gap in their closure is that the proportional constant linking the THC to the density contrast (called “admittance” here drawing its analogy from the electrostatics) is a free parameter, which renders the climate state undetermined as well. This nondeterminacy remains entrenched in coarse-grained GCMs as the admittance is embedded in the diapycnal diffusivity (Rahmstorf et al. 2005), which, over its probable observational range, may cause an order-of-magnitude change in the modelled THC (Bryan 1987; Dalan et al. 2005). The diapycnal diffusivity used in such GCMs thus is in effect an empirical parameter that is finely tuned to yield the observed state. This sensitivity necessarily impacts the GCM’s prognostic utility in climate-change studies, which, as we shall see, can be alleviated if eddy exchanges are properly included.

Observations show that the subtropical front is highly unstable to spawn copious eddies, which may facilitate the THC. Indeed, eddy census (Auer 1987) show that about 20 warm eddies are shed annually and, if one takes their mean depth and diameter to be 1 and 130 km, respectively, it amounts to a mass flux of O (10 Sv), which is commensurate with the observed THC. Since eddy shedding is a microscopic hence random process (Zaslavsky 1999), its rate is likely governed by probability laws of the nonequilibrium thermodynamics (NT); and it is the application of the NT to constrain the THC—thus fostering a physical closure—that particularly distinguishes our box model.

The other shortfall of the ocean box-models pertains to their saline mode, which features reversed density contrast and THC from the present, a state not yet corroborated by paleoclimate data. In contrast, the coupled GCMs (Manabe and Stouffer 1988) show a stable cold state with nominal—though weaker—density contrast and THC. Unlike the saline mode, this cold state can be identified with the glacial climate, whose dynamical basis however remains unknown. Through our coupled box-model, we show that the atmospheric coupling can be the progenitor of this cold state, which thus may fill this dynamical void.

With above closure and extension of the box model, we shall then apply it to examine the climate response to the global warming, defined as the rising global-mean temperature without implicating its origin. The GCMs have demonstrated a transition from the current warm state to the cold state with sufficient warming (Manabe and Stouffer 1993),

Fig. 1 The model configuration. Both ocean and atmosphere are composed of warm and cold boxes, which are symmetric about the equator and aligned at mid-latitudes. The absorbed solar flux (q') differentiates the SST of the ocean boxes (T'), which then differentiates the SAT of the atmospheric boxes (T_a') via the convective flux (q_c'). The moisture transport F_w differentiates the ocean salinity (S') and the THC (K) is subjected to random eddy shedding to be constrained by the fluctuation theorem. The primed variables are the cold-box deficits from global-means

followed by possible hysteresis. Given the practical implication of such dramatic climate change, several intercomparison modelling studies have been carried out (Gregory et al. 2005; Rahmstorf et al. 2005; Stouffer et al. 2006). Although the model outputs all show characteristic hysteresis, quantitative predictions of its thresholds vary widely, the source of which remains obscure because of the complexity of the GCMs. Our box model with its transparent physics may aid the synthesis of the GCM results.

For the organization of the paper, we formulate our coupled box-model in Sect. 2, followed by consideration of the NT in Sect. 3 to constrain the THC. In Sects. 4 and 5, we examine the THC response to prescribed freshwater perturbation and the global warming, respectively, which will be compared with the GCM results. We summarize the main findings in Sect. 6 and provide additional discussion in Sect. 7.

2 Model formulation

The box configuration is shown in Fig. 1, which can be a deductive outcome (Ou 2006) but is taken here as a minimal representation of the observed state. Both the troposphere and the ocean are divided into warm and cold boxes of homogeneous (eddy-averaged) property fields, which are symmetric about the equator and aligned at mid-latitudes. The internal boundaries can be identified in the ocean with the main thermocline outcropped as the subtropical front, and in the atmosphere with the polar front that separates the tropical and polar airmasses.

Since we are concerned with the differential fields interactive with the THC, we take the global-mean fields to be

known, so the differential forcing and property fields can be reduced to the cold-box deficits (except the density surplus, all are primed) from the global-means (over-barred); the reason we choose the cold boxes is because the cold water is bounded additionally by the freezing point. Retaining only the dominant energy pathways, the deficit in the absorbed shortwave (SW) flux (q') induces that in the sea-surface temperature (SST, T'), which in turn induces that in the surface-air temperature (SAT, T_a') via the convective flux (q_c').² The resulting atmospheric heat transport is accompanied by the moisture transport (F_w), which induces the salinity deficit (S'). The oceanic heat transport is via the THC (K), which is subjected to random eddy shedding across the subtropical front hence constrained by the NT.

For simplicity, the derivation proceeds in the non-dimensional form; all symbols and scale definitions (bracketed) are listed in appendices B and C, respectively, together with their “standard” values used in our model calculations. Sufficing for our purpose, we consider steady-state balances. For the ocean, the differential heat balance, expressed in the cold-box deficits, states

$$q' = q_c' + KT', \quad (1)$$

or the deficit in the absorbed solar flux (q') is balanced by that in the convective flux (q_c') and the heat carried in by the THC (K), with the convective flux given by

$$q_c' = T' - T_a'. \quad (2)$$

In terms of the poleward heat transport at mid-latitudes, (1) states that the total (lhs) is partitioned between the atmosphere and the ocean (the two rhs terms, respectively).

Since the atmospheric heat balance is already entailed in the last statement, it does not otherwise constrain the SAT, for which we shall invoke the maximum entropy production (MEP) to be discussed later (Sect. 3). Because of the (approximate) steady state, the atmospheric entropy production (EP, denoted σ_a) associated with the differential field equals the differential entropy flux exiting its lower boundary (see Footnote 4 for the neglect of the differential OLR hence the associated entropy flux). To derive this boundary flux, we begin with the local flux of q_s/T_a with q_s being the surface heat flux and T_a the SAT (in Kelvin). Denoting the global-means and deviations by overbars and primes respectively and noting that the temperature deviation is small compared with the global-mean temperature (in the absolute unit), the differential entropy flux is then approximately $\bar{T}_a^{-2} q_c' T_a'$ (see

Footnote 4 for the retention of only the convective flux q_c'). Applying the box approximation and nondimensionalizing, the differential atmospheric EP is then, expressed in the cold-box deficits,

$$\sigma_a = q_c' T_a'. \quad (3)$$

Substituting (2) into (3) and maximizing against T_a' yields.

$$T_a' = T'/2, \quad (4)$$

hence from (2),

$$q_c' = T'/2. \quad (5)$$

Expectedly, the deficit in the convective flux (same as the atmospheric heat transport) increases with the SST deficit; but since the actual convective flux may not be negative ($\bar{q}_c - q_c' \geq 0$), we have $q_c' \leq \bar{q}_c$ or the atmospheric heat transport is bounded above by the global-mean convective flux.

Combining this upper bound with (5), the atmospheric heat transport is then

$$q_c' = \min[T'/2, \bar{q}_c], \quad (6)$$

which is uniquely specified by the SST deficit. It is seen in particular that when the cold-water deficit T' exceeds $2\bar{q}_c$ (called “convective bound”), the atmospheric heat transport would saturate at $\bar{q}_c$ hence may no longer moderate the oceanic heat transport. This convective bound is a robust feature of a coupled ocean–atmosphere system, but not sufficiently recognized previously. It can be understood by considering an ocean that is increasingly differentiated from an initial state of uniform convective flux (hence equaling its global-mean), which would augment the atmospheric heat transport by redistributing the convective flux from the cold to warm surfaces. This would proceed until the convective flux over the cold surface is depleted, so the initial (global-mean) convective flux sets an upper bound on the atmospheric heat transport.

Substituting (6) into (1), we derive the following expressions of T' in terms of the (unknown) THC (K), separated by the convective bound:

$$T' = \frac{q'}{K + 1/2} \quad \text{for } T' \leq 2\bar{q}_c \quad (7)$$

$$= \frac{q' - \bar{q}_c}{K} \quad \text{for } T' \geq 2\bar{q}_c. \quad (8)$$

Setting $q' = 1$ by scale definition, we plot the above temperature in Fig. 2 with the horizontal dashed bars marking two different values of the convective bound for comparison. When combined with the THC constraint to be discussed later, this graph allows the determination of the climate state and will be referred as the “regime diagram”. Its use was pioneered by Stommel (1961), which turns out to

² The net surface longwave (LW) flux and the outgoing LW radiation (OLR) are relatively uniform compared with the convective flux hence neglected in the differential heating. We have also neglected the atmospheric absorption of the SW flux, which does not enter the surface heat balance to be considered.

Fig. 2 The regime diagram in which state variables are plotted against the THC. The convective bound (*horizontal dashed bars*, two examples are shown) divides the warm and cold branches marked by the opposite slope in the density curve. Microscopic fluctuations in the eddy shedding rate (shaded) pivot the admittance line (*thick dashed*) toward the MEP (*solid squares*), whose intersection with the density curve specifies the climate state (*solid ovals*). The warm MEP represents the current climate, which may be mono- or bi-stable depending on the convective bound

be highly effective as well in our explanation of the climate behavior.

As expected, the SST deficit T' increases (that is, the subpolar water gets colder) with decreasing THC, but its rising rate is slower below the convective bound—as the concurrent increase in the atmospheric heat transport (6) would decrease the oceanic heat transport (the total being fixed by q'). Above the convective bound, on the other hand, the atmospheric heat transport has saturated (6), so the fixed ocean heat transport yields a steeper rise in T' that is inverse in the THC.

The differential salt balance, expressed in the cold-box salinity deficit, states

$$F_w = KS', \quad (9)$$

or the freshwater input F_w to the cold box is countered by the salt carried in by the THC (rhs). As recounted in “Appendix D”, an explicit consideration of the hydrological cycle by Ou (2007), adjusted for the differing catchment- and ocean-areas of the cold box, yields a freshwater input of

$$F_w = \mu q_c', \quad (10)$$

where μ is a “moisture-content” parameter, a function of the global-mean SAT, as seen below. It is impractical to discuss all the steps in Ou (2007) leading to (10), which

however can be crudely understood as follows: at mid-latitudes, the moisture is transported by the same mass exchange as the atmospheric heat, so the two transports should be linked with a proportional constant μ measuring the moisture content of the air column. Because of the Clausius–Clapeyron equation, this moisture content is specified by the time-mean SAT in the frontal zone, which can be approximated by the global-mean SAT (Rahmstorf and Ganopolski 1999).

Substituting (6) into (10), we derive the following expressions of the salinity deficit:

$$S' = \frac{\mu T'}{2K} \quad \text{for } T' \leq 2\bar{q}_c \quad (11)$$

$$= \frac{\mu \bar{q}_c}{K} \quad \text{for } T' \geq 2\bar{q}_c, \quad (12)$$

which is plotted in Fig. 2 for $\mu \approx 0.3$ (see “Appendix D”). As expected, it increases (that is, the subpolar water gets fresher) with decreasing THC, but its rising rate contrasts sharply with that of the temperature deficit: below the convective bound, it is steeper because the moisture transport increases with the differential temperature (hence opposite the variation in the ocean heat transport); above the convective bound, on the other hand, the moisture transport has plateaued with the atmospheric heat transport, so the salinity deficit increases as the inverse of the THC, just as the temperature deficit.

This contrast manifests particularly distinctly in the density surplus ρ' of the cold box,

$$\rho' = T' - S', \quad (13)$$

plotted in Fig. 2. Being the distance between the T' and S' curves, the slope of ρ' curve is seen to change sign at the convective bound. This division into “warm” and “cold” branches (pertaining to the subpolar temperature itself) by the convective bound is a feature of the atmospheric coupling hence absent from ocean-only box models (Ruddick and Zhang 1996).

We have so far determined the state variables T' , S' and ρ' as functions of the THC (K), the latter thus needs to be constrained to specify the climate state. In box models, the THC is customarily assumed to be proportional to the cold-box density surplus (Stommel 1961; Marotzke and Stone 1995),

$$K = a\rho', \quad (14)$$

with an unknown admittance (a), a relation that is also supported by GCM calculations (Hughes and Weaver 1994). This is the straight “admittance line” (thick dashed) in the regime diagram whose slope is the inverse of the admittance and whose intersection with the density curve then specifies the climate state (ovals). The model closure thus is reduced to determining the admittance, as discussed next.

3 Admittance

As noted earlier (Sect. 1), the observed eddy shedding may account for a sizable portion of the THC, so for simplicity and without undue limitation (see later discussion in this section), we consider a THC that is comprised wholly of random shedding of warm eddies—compensated by returning cold eddies and/or upwelling. As such, the mass exchange rate is a product of the eddy shedding rate and eddy volume, and the latter would be proportional to the density contrast if the eddy radius is scaled by the deformation radius (Gill 1982). With this, we recover the relation (14) except the admittance is now identified with the eddy shedding rate—a measurable physical quantity.

Since eddies are microscopic constituents of the ocean, the eddy shedding is subjected to random fluctuations (see also Sect. 1). As such, we shall apply the fluctuation theorem (FT) to constrain the (expectant) shedding rate. The FT quantifies the probability of fluctuations in accordance with their EP (Evans et al. 1993; Wang et al. 2002) as the latter represents the number of compatible micro-paths on the phase space (Dewar 2005b). In its initial formulation, the EP pertains to that of the micro-paths (Evans et al. 1993; Wang et al. 2002), but Crooks (1999, see also Sevic et al. 2008) has generalized it to that between macro-states, which is more pertinent for our application—the only difference being that our macro-state is not externally driven but internally propelled.

The following derivation proceeds in two steps: we first derive the probability of the shedding-rate fluctuations from the FT; and based on this probability distribution, we then derive the (expectant) shedding rate. Given its short history, the FT is still under active development with varied interpretations; and then the FT has not been applied to the eddy-shedding problem (the closest is perhaps to the thermal plumes in the Rayleigh–Benard convection, see Ciliberto and Laroche 1998), so our derivation is admittedly heuristic, which hopefully can be strengthened by future advances.

Our box configuration offers a palpable simplification in the derivation since frontal eddies are either shed or not, so the shedding rate takes on binary fluctuations. The shedding events are independent (hence random) if they are separated by more than the decorrelation distance l , which can be taken to be several eddy diameters, and then since the events occur along the whole stretch of the front denoted by L , the shedding rate a is thus subjected to the fractional fluctuations of

$$(\ln a)' = a'/a = \pm \varepsilon, \quad (15)$$

where

$$\varepsilon \equiv l/L. \quad (16)$$

Given a mean eddy-diameter of 150 km (Auer 1987), one may take the decorrelation distance to be $O(300 \text{ km})$, so for a subtropical front of 3000 km length, ε is $O(0.1)$, a small number. The fluctuations in the admittance, by perturbing the climate state, would induce fluctuations in the EP (σ) as well. Applying the chain rule, the EP fluctuations corresponding to (15) is

$$(\ln \sigma)' = \pm \varepsilon \Gamma, \quad (17)$$

where

$$\Gamma \equiv \frac{d(\ln \sigma)}{d(\ln a)}, \quad (18)$$

involves only macro-variables (that is, it can be calculated from the regime diagram). As the fluctuations are time-reversible with odd-functioned EP (17), they satisfy this original requirement of the FT—although Crooks (1999) has proposed a less stringent condition of stochastic dynamics, which may also apply in our case.

Let p_+ and p_- denote the probability of the corresponding fluctuations (15), the FT states that, over an infinitesimal time interval δt ³ (Crooks 1999),

$$p_+/p_- = \exp(\varepsilon \Gamma \delta t). \quad (19)$$

With these fluctuations being the only possible outcomes, we have

$$p_+ + p_- = 1, \quad (20)$$

which, when subjected to (19), yields.

$$p_{\pm} \approx (1 \pm \varepsilon \Gamma \delta t)/2, \quad (21)$$

or the fluctuation with higher EP (that is, [17] is positive) is slightly favored. Since the macroscopic shedding rate is an expectant property, its change over the same time interval is, from (15) and (21),

$$\begin{aligned} \delta \ln a &= \varepsilon p_+ + (-\varepsilon) p_- \\ &= \varepsilon^2 \Gamma \delta t \end{aligned} \quad (22)$$

giving rise to a time-evolution equation

$$\frac{d(\ln a)}{dt} = \varepsilon^2 \Gamma. \quad (23)$$

This is the missing equation we are seeking, which would pivot the admittance line in Fig. 2 (shaded to indicate fluctuations) in accordance with the rhs of the existing climate, thus fully specifying its evolution. Since the above derivation is not rigorous, we shall regard (23) as a postulate, which in principle can be tested by fine-grained

³ The exponent in the FT (19) needs to be dimensionless, and with the macroscopic variables appearing in it already nondimensionalized, so should the time. Here we posit that the only relevant time-scale is the upper-ocean overturning time, to which our steady-state balances are also referenced.

GCMs that can simulate the random eddy shedding. Based on (23), we draw the following inferences.

1. The admittance is not a constant, but evolves toward higher EP, attaining stationarity when $\Gamma = 0$ or when the EP is maximized. This principle of maximum entropy production (MEP) has been widely applied in climate theories (see a review by Kleidon 2009) whose physical basis remains debated (Dewar 2005a; Grinstein and Linsker 2007); but if it were a deductive outcome of the FT, as posited here, it can be further justified.
2. The admittance evolves toward the MEP on a timescale that is ϵ^{-2} of the (decadal) upper-ocean overturning time (Ou 2012, his Sect. 2.3) hence millennial. We shall refer to this tendency as the “MEP-adjustment”, whose long timescale arises from the compounding effect of small fluctuations in the shedding rate and their slight probability bias. This slow tendency further justifies our construction of the regime diagram based on steady-state balances.
3. While atmospheric eddies are larger, ϵ remains small (low fraction of unity); and with the atmospheric mixing time measured in weeks, its MEP-adjustment time is still short compared with the upper-ocean overturning time, which justifies our earlier application of the MEP to the SAT (4).
4. Even if the THC has a deterministic laminar component, its remaining part is still subjected to fluctuations to be governed by (23); the deterministic component thus constitutes merely a lower bound, which does not otherwise impact the MEP-adjustment and/or the MEP state. This recognition justifies the simplifying assumption of an eddy-composed THC, and renders the THC and the resulting climate state insensitive to the observed diapycnal diffusivity, as alluded to earlier (Sect. 1).

To constrain the admittance, we shall next determine the MEP state of the ocean. Similar to (3), the oceanic EP is approximately

$$\sigma \approx (q' - q'_c)T', \quad (24)$$

with q'_c given by (6). Below the convective bound, the maximization of (24) yields $T' = q' = 1$ (by scale definition) hence, through (5), (1), (10) and (9), a “warm” MEP state of

$$(T', S', K) = (1, \mu, 1/2), \quad (25)$$

as indicated in the regime diagram. Above the convective bound on the other hand, q'_c is fixed, so the EP increases with T' (24) until the freezing point (T_f') beyond which the sea-ice must form to reduce q' in order to maintain the heat

balance (1); and since decreasing q' in (24) implies smaller EP, the EP thus has a maximum at the freezing point, which is the “cold” MEP state labelled in the regime diagram.

The warm MEP can be identified with the current climate, and for a crude comparison, we use the standard values listed in appendices B and C to yield a subpolar SST of 8 °C, salinity of 34.5 (hence a density ratio of 3.3) and THC of 16 Sv. All these are reasonable compared with the observed state (Peixoto and Oort 1992; Roemmich and Wunsch 1985) despite the extreme crudeness of our box model.

With the admittance line in Fig. 2 fixed by the current climate, it is seen that the case of higher convective-bound allows binary stable states (solid ovals), as well as an unstable (hence unrealized) saddle point (open oval). But differing from the saline mode of the ocean box-models (Rudick and Zhang 1996), the cold state has the same-signed density contrast and THC as—though much weaker than—the current climate. Such a state has emerged in coupled GCMs (Manabe and Stouffer 1988), and the convective bound uncovered here, which facilitates the cold branch, provides a plausible explanation.

4 Hysteresis

As we shall discuss in a forthcoming paper, the hysteresis between the two MEP states when subjected to the orbital forcing may account for the glacial cycles of late Pleistocene. For the present study, however, we are concerned with the hysteresis between the warm and cold branches when the freshwater input is perturbed by the global warming. Since both global warming and ocean response are rapid compared with the millennial MEP-adjustment, the admittance line may be regarded as fixed by the current climate.⁴ For the present study, therefore, there is no need to consider the time evolution of the admittance as entailed in (23).

Following parallel GCM approaches, we first consider in this section the case when the freshwater perturbation is prescribed, and then in the next section the more justified case when such perturbation is internally constrained by the global warming.

Physically, the freshwater perturbation would move the density curve in the regime diagram up and down, and when the warm branch detaches from the admittance line, there would be a transition to the cold branch and vice versa, causing the hysteresis. To illustrate this behavior,

⁴ This incidentally justifies the GCM practice of fixing the diapycnal diffusivity (a proxy of the admittance, see Sect. 1) after it is tuned for the current climate.

Fig. 3 Hysteresis loops when the current state (*upper circle*) is subjected to freshwater perturbation (w'). The inner loop (*dashed* in part) is for the lower convective bound, which is characterized by weakening and recovery of the THC within the positive perturbation due to the monostability. The outer loop (*solid*) is for the higher convective bound for which the bistability (two *circles* at zero perturbation) prevents the THC recovery except with negative perturbation

one simply adds a perturbation w' to the lhs of (10) and calculates the THC (“Appendix E”), and the result is plotted in Fig. 3 for the two convective bounds shown in the regime diagram.

As expected, the THC exhibits the hysteresis like that seen in Rahmstorf and Willebrand (1995) but, in contrast, the lower convective-bound (the inner loop, dashed in part) allows a partial THC shut-down without invoking multiple convection sites.⁵ As this case is seen from the regime diagram to be monostable, switching off the perturbation would cause the current state to recover. The case of higher convective-bound (the outer loop) on the other hand is bistable, so a negative perturbation is needed to close the hysteresis loop. For this example, the lower limb of the hysteresis loop has vanishing THC and—hence—density stratification (14), which however constitutes a singular limit when the model equations break down; but before reaching this limit, it is seen from the regime diagram that the cold box would encounter the freezing point, so there would be sea-ice formation to counter the freshwater input and affect the differential

solar heating and convective flux. While such feedbacks would render a nonzero THC, we deem the complication as beyond the scope of the present discussion.

Rahmstorf et al. (2005) noted considerable spread in the cold-transition threshold among different models and attributed it to the uncertain current state.⁶ As the latter depends on the admittance that is embodied in coarse-grained GCMs by the diapycnal diffusivity (Sect. 1), some model behavior can be interpreted via the regime diagram. A higher diapycnal diffusivity, for example, would flatten the admittance line to: (a) augment the THC, (b) raise the cold-transition threshold, and (c) favor the monostability—all are consistent with the GCM results (Schmittner and Weaver 2001; Prange et al. 2003; Dalan et al. 2005). With the current state as well as the admittance uniquely specified by the MEP, on the other hand, the (dimensionless) cold-transition threshold is a function only of μ (36), and with the latter set to 0.3 (see “Appendix D”), this threshold has a definitive value of 0.065. With the scale of the freshwater input estimated in “Appendix C”, the current freshwater input would be 0.43 Sv, which is commensurate with that used in Rahmstorf (1995, his Fig. 1); and the threshold perturbation is 0.18 Sv, which is within the range seen in Rahmstorf et al. (2005).

It is significant to note that even with the current state fixed by the MEP, it can still have different modalities (uni- or bi-) depending on the convective bound. As the two convective bounds used here differ by only $5 \text{ W} \cdot \text{m}^{-2}$ in the global-mean convective flux (that is, more than an order-of-magnitude smaller than its current value), they may not be easily discerned from observation. This sensitivity may explain why GCM runs exhibit both scenarios of permanent and temporary collapse of the THC (Rahmstorf and Ganopolski 1999). Other previous findings may also be explained via the regime diagram: allowing the atmospheric coupling, for example, would moderate the differential convective flux in (1) to augment the ocean heat transport, which in turn would raise the temperature curve hence the cold-transition threshold (Rahmstorf and Willebrand 1995); linking the atmospheric moisture- to heat-transport (10), on the other hand, would curl the density curve (in the warm branch) more sharply downward to promote the cold transition (Nakamura et al. 1994; Marozke and Stone 1995). With the above, it is seen that the regime diagram is highly effective in interpreting the GCM results.

⁵ Since the convections at different sites need not all cease at the same time, it allows partial shut-down of the THC. On the other hand, the direct effect of the local convection on the THC has been previously questioned as the THC depends primarily on the large-scale density contrast (14) (Maroztke and Scott 1999; Jungclaus et al. 2006).

⁶ Since the current state and the attendant freshwater input are uniquely determined in our model, there is no need or justification for aligning the warm transition among model runs by Rahmstorf et al. (2005).

5 Global-warming response

The additional ice-melt by global warming is of secondary importance compared with increasing moisture transport (Manabe and Stouffer 1994; Jungclaus et al. 2006), the latter being through the higher moisture content entailed in μ of (10) and (32), which we approximate as

$$\mu = \mu_0 + \mu_1 \delta \bar{T}_a, \quad (26)$$

where $\delta \bar{T}_a$ is the variation of the global-mean SAT (in °C), and subscripts “0” and “1” denote its current value and the perturbation per degree-warming, respectively. Their expressions are derived in “Appendix F” based on an explicit consideration of the hydrological cycle, from which we estimate the “moistening parameter” μ_1 to be 0.02. For the standard case, it amounts to 0.03 Sv increase in the freshwater input per degree-warming, similar to that found in GCM studies (Manabe and Stouffer 1994).

In addition to freshening the subpolar water, the global warming would reduce the ice albedo to increase the absorbed solar flux hence a smaller q' (recall that it is the cold-box deficit). To include this effect on the differential forcing, it is approximated as

$$q' = q_0' - q_1' \delta \bar{T}_a, \quad (27)$$

where the subscript usage follows that of (26); the “heating parameter” q_1' is derived in “Appendix G” and estimated to be 0.01 based on observations. For the standard case, it amounts to $1 \text{ W} \cdot \text{m}^{-2}$ decrease in the differential forcing per degree-warming. One should note that the direct radiative effect from rising CO_2 is spatially uniform hence does not affect the differential forcing.

The above effects of the global warming can be visualized in the regime diagram as follows: the decrease of the differential forcing (27) would lower the temperature hence the density-curves; and the increase of the moisture content (26) hence the freshwater input (10) would curl the density curve more sharply downward. Both thus would move the warm intersect (solid oval) downward along the admittance line and trigger the cold transition when the density curve becomes detached from the admittance line.

To distinguish the relative importance of the heating and moistening effect, partially coupled GCM runs have been carried out (Dixon et al. 1999; Mikolajewicz and Voss 2000; Gregory et al. 2005). The results are varied and difficult to interpret since they involve artificial turning-off of internal couplings. On the other hand, given the independent balances governing the moistening and heating parameters in our model (appendices F and G) and their considerable uncertainty, we are justified to examine the THC sensitivity to these parameters.

With the method described in “Appendix H”, we calculate and plot the hysteresis loop in Fig. 4 for the standard

Fig. 4 Same as Fig. 3 but induced by perturbation in the global-mean SAT. The darker and lighter shades are for $\pm 50\%$ variation in the heating parameter q_1' and $\pm 20\%$ variation in the moistening parameter μ_1 , respectively

case where shades around the cold transition represent, in lessening darkness, variation of $\pm 50\%$ for the heating parameter (q_1') and $\pm 20\%$ for the moistening parameter (μ_1). To the degree that the latter, even with smaller fractional variation, fully encloses the former implies that the cold transition is far more sensitive to the moistening than the heating parameters. Quantitatively, the cold transition occurs around 4.5°C warming, similar to that predicted by Manabe and Stouffer (1994) and Rahmstorf and Ganopolski (1999); it increases by only 1°C if the heating effect is completely turned off (that is, $q_1' = 0$), the latter thus plays only a minor role compared with the moistening effect.

This result is consistent with some GCM findings (Manabe and Stouffer 1994; Dixon et al. 1999), but at odds with others (Mikolajewicz and Voss 2000; Stocker et al. 2001; Gregory et al. 2005). Our model obviously may not address the inter-basin vapor transport contained in some GCMs, it nonetheless underscores an important but overlooked internal coupling: that is, the warming of the subpolar water would reduce the atmospheric heat- hence moisture-transport (10), thus countering the warming effect on the THC. Moreover, this countering accentuates with decreasing THC (9), so the density curve, while being lowered by heating, also becomes flatter to curtail the downward movement of the warm intersect. This is the reason why the differential heating is ineffective in our model to cause the cold transition, and it calls into question the justification of partial coupling when the freshwater input is fixed to isolate the heating effect.

Before the cold transition, one expects little change to the differential temperature, so it would not deter the global warming; after the cold transition, however, the much colder subpolar water would drag down the global-mean SAT (Rahmstorf and Ganopolski 1999) to propel the lower limb of the hysteresis loop. For the monostable case (the inner loop), the THC would recover by only about 3 °C lowering of this SAT—from its peak at the cold transition hence still 1.5 °C above the present; for the bistable case (the outer loop), on the other hand, the THC does not recover even if the global warming is eliminated by the cold transition. The uncertainty in the modality of the present climate (Sect. 4) thus may account for these differing scenarios in GCM runs (Manabe and Stouffer 1993; Rahmstorf and Ganopolski 1999; Vellinga and Wood 2002). Again, as discussed in Sect. 4, in a more refined model that considers the freezing-point bound on the water temperature, the THC would remain finite in the cold state, which does not alter the essential hysteresis loops.

6 Summary

We have formulated a box model of coupled ocean/atmosphere to examine its differential fields interactive with the THC, and their response to the global warming. Because of the non-negativity of the sea/air convective flux, we discern a robust “convective bound” on the atmospheric heat transport, which would divide the climate regime into warm and cold branches (pertaining to the subpolar temperature). Unlike the saline mode of the previous box models however, the cold state, if allowed, has the same-signed—though weaker—differential density and THC as the present, which thus may explain the emergence of such state from coupled GCMs.

We then underscore the nondeterminacy of the “admittance” linking the THC to the density contrast due to random eddy shedding and apply the fluctuation theorem of the nonequilibrium thermodynamics to constrain the shedding rate, thus closing the problem. The derivation shows that the ocean would be propelled toward the MEP on millennial timescale (termed “MEP-adjustment”), the long timescale arising from the compounding effect of small fluctuations in the shedding rate and their slight probability bias.

The above physical closure allows us to examine deductively the ocean response to the global warming, for which the admittance can be assumed unvarying hence fixed by the current MEP state. We first prescribe the freshwater perturbation, which yields similar hysteresis as seen in GCMs; but some differences in the model runs are readily explained via our regime diagram. We then consider the more relevant case when the perturbation in both

differential heating and freshwater input are internally constrained. The cold transition would occur with about 4.5°C rise in the global-mean SAT, similar to that seen in GCMs; but this threshold is far more sensitive to the moistening than the heating effect due to the hydrological feedback on the latter. The THC may or may not recover depending on whether the current climate is mono- or bi-stable, which cannot be easily assessed because of the observational uncertainty in the global-mean convective flux.

7 Discussion

Our MEP solution pertains to the coupled ocean/atmosphere devoid of the land influence hence may not be closely compared with the observed zonal-mean fields, which are strongly impacted by expansive landmass in high northern latitudes. The small heat capacity of the landmass unbounded by the freezing point strongly depresses the zonal-mean SAT (Jones et al. 1999; their plate 4) to augment its meridional contrast beyond that of the SST. This has contributed to the wide-spread use of the SAT in restoring the SST in ocean-only GCMs while in actuality the SAT is differentially heated by the SST to be of smaller meridional range absent the land influence, as deduced in our model.

Although the regime diagram is constructed from steady-state balances, some temporal behavior may nonetheless be inferred. First, as seen in Ou (2012), the positive feedback between subpolar salinity and THC renders a saline damping time much greater (centennial) than the thermal one (decadal); as such, the former, being limiting, would govern the time evolution around the hysteresis loop, including the warm/cold transitions—a deduction that is consistent with the GCM results (Manabe and Stouffer 1994; Rahmstorf and Ganopolski 1999). Second, if the freshwater perturbation is terminated during the cold transition, the latter would reverse course to return to the initial state without traversing the loop (see for example Dalan et al. 2005). To reconcile with their bimodality, Manabe and Stouffer (1999) have attributed the THC recovery in Schiller et al. (1997) to the above transient, but an alternative interpretation is that the initial state of Schiller et al. (1997) is in fact monostable and the hysteresis loop has already run its course. Third, as seen in Ou (2012), the disparate thermal and saline damping time would induce Atlantic multi-decadal variability, which would amplify when approaching the cold transition (the Hopf bifurcation, see Rahmstorf 1995) to increasingly mask the anthropogenic signal.

Our closure hinges on the warm outflow through the subtropical front that defines our THC hence it makes no assumption on where the THC is returned via upwelling,

which may in fact be concentrated in the Southern Ocean (Marshall and Speer 2012). Furthermore, it is this cross-frontal heat transport that we posit the importance of eddy shedding; that is, although the heat transport within the subtropics may be dominated by the gyre circulation—due in part to the subdued thermal distinction of eddies, it does not contravene the above eddy importance. In other words, these observational complications have little import on the deduced THC or the climate state of our box model.

The main advance in our closure is the application of the NT to constrain the eddy shedding rate hence the THC. To include the NT in GCMs remains challenging since they need to resolve the micro-scale to capture the random eddy shedding. Using a hydrodynamic GCM, Hurlburt and Hogan (2000) showed that only when the grid spacing is reduced to 1/64 degree would the eddy shedding approach its observed rate. As regard the cross-frontal heat transport, the coupled GCM runs of Griffies et al. (2015, their Fig. 9d–f) show that the eddy contribution (at 40°N) would approach the mean one when the grid spacing is reduced to 0.1°. But as seen in Hurlburt and Hogan (2000) and Bryan et al. (2007), the eddy/mean partition has not yet converged at 0.1°, so one expects further enhancement of the eddy importance with improving resolution. This eddy/mean partition notwithstanding, we see in Sect. 3 that so long as the eddy part is significant, the THC remains undetermined to be subjected to the MEP. In addition, since the EP is a global property, the eddy part may not be parameterized via local eddy diffusivity (see also Pierrehumbert 1991); there is thus little recourse in the GCM approach but to resolve the micro-scale.

Additional challenges to the GCM study of the climate change lie in the long time-integration needed and then such integration constitutes merely a single microscopic path, which may differ from the observed one as both are members of the ensemble, and only the ensemble mean as derived here is uniquely constrained. There is thus inherent limitation to the GCM prediction of the climate change due, in our context, to fluctuations in the eddy shedding—irrespective of its improved spatial resolution and detailed physics.

Appendix A: Acronyms

EP	Entropy production
FT	Fluctuation theorem
GCM	General circulation model
LW	Longwave
MEP	Maximum entropy production
MOC	Meridional overturning circulation
NT	Nonequilibrium thermodynamics

OLR	Outgoing longwave radiation
SAT	Surface-air temperature
SST	Sea-surface temperature
SW	Shortwave
THC	Thermohaline circulation

Appendix B: Symbols

a	Admittance
A	Surface area spanned by cold ocean-box ($= 1.9 \times 10^7 \text{ km}^2$)
Bo^{-1}	Inverse Bowen ratio over cold water ($= 1$)
C_d	Drag coefficient ($= 10^{-3}$)
$C_{p,a}$	Specific heat of surface air ($= 10^3 \text{ JKg}^{-1}\text{K}^{-1}$)
$c_{p,o}$	Specific heat of ocean ($= 4.2 \times 10^3 \text{ JKg}^{-1}\text{K}^{-1}$)
F_w	Freshwater transport to the cold box
g	Gravitational acceleration ($= 9.8 \text{ ms}^{-2}$)
K	Mass exchange rate between ocean boxes (THC)
l	Decorrelation distance of eddy shedding ($= 300 \text{ km}$)
L	Latent heat of vaporization ($2.26 \times 10^6 \text{ JKg}^{-1}$)
	Length of subtropical front ($= 3000 \text{ km}$)
q'	Cold-box deficit of absorbed solar flux
$\bar{q}_c$	Global-mean convective flux
q'_c	Cold-box deficit of convective flux
$\bar{S}$	Global-mean surface salinity ($= 35$)
$\bar{T}$	Global-mean SST ($= 15^\circ\text{C}$)
T'	Temperature deficit of cold ocean-box
$\bar{T}_a$	Global-mean SAT
T'_a	Temperature deficit of cold atmosphere-box
T'_f	Freezing temperature
$ u' $	Surface turbulent wind ($= 7 \text{ m/s}$)
α	Thermal expansion coefficient ($= 1.7 \times 10^{-4} \text{ }^\circ\text{C}^{-1}$)
α^*	Surface transfer coefficient ($\equiv C_d \rho_a C_{p,a} u' [1 + Bo^{-1}] = 14 \text{ Wm}^{-2} \text{ }^\circ\text{C}^{-1}$)
β	Saline contraction coefficient ($= 7.6 \times 10^{-4}$)
ρ'	Density surplus of cold ocean-box.
ρ_a	Surface air density ($= 1 \text{ Kgm}^{-3}$)
ρ_o	Ocean density ($= 10^3 \text{ Kgm}^{-3}$)
σ	Ocean entropy production rate.
σ_a	Atmosphere entropy production rate
μ	Moisture-content parameter ($= 0.3$)

Appendix C: Scale definitions

$$[q'] = 100 \text{ Wm}^{-2}$$

$$[T'] \equiv [q']/\alpha^* = 7.1^\circ\text{C}$$

$$[S'] = \alpha [T'] / \beta = 1.6$$

$$[\rho'] = \rho_o \alpha [T'] = 1.36 \text{ Kg m}^{-3}$$

$$[K] = (2\rho_o c_{p,o} [T'])^{-1} [q'] A = 31 \text{ Sv}$$

$$[a] = [K] / [\rho']$$

$$[F_w] = 2[K][S'] / \bar{S} = 2.8 \text{ Sv}$$

$$[\sigma] = 2\bar{T}^{-2} [q'] [T']$$

$$[\sigma_a] = 2\bar{T}_a^{-2} [q'] [T']$$

Appendix D: Moisture transport

By considering the hydrological cycle, Ou (2007) has linked the moisture transport at mid-latitudes to the atmospheric heat transport F_a^* (starred for dimensional variables). Adjusting this moisture transport for the freshwater input F_w^* to the cold box because of the differing catchment (A_c) and ocean (A) areas, we have

$$F_w^* = \mu^* F_a^* A_c / A, \quad (28)$$

where

$$\mu^* = \frac{1}{\rho_o L (1 + 1/Bo_m^{-1})}. \quad (29)$$

In the above, Bo_m^{-1} (the inverse “meridional” Bowen ratio) reflects the moisture content of the air column and is of the form

$$Bo_m^{-1} = \text{const.} \times e / T, \quad (30)$$

where T is the mean SAT in the frontal zone (approximated by the global-mean SAT) and e , its saturation vapor pressure, a relation that has been validated by observation (Ou 2007). When nondimensionalized, (28) becomes

$$F_w = \mu q_c', \quad (31)$$

where

$$\mu = \frac{c}{1 + 1/b} \quad (32)$$

with

$$b = Bo_m^{-1} \quad (33)$$

and

$$c = \frac{A_c}{A} \cdot \frac{C_{p,o} \bar{\beta} \bar{S}}{L \alpha}. \quad (34)$$

For the standard case, we use $b \approx 0.5$ (Ou 2007, his Fig. 4), and setting $A_c/A \approx 3$, we obtain $c \approx 0.87$, so $\mu \approx 0.3$. As seen in (25), μ yields the inverse density-ratio of the

current climate, which is observed to be about 0.3 (Tippins and Tomczak 2003), so our selection of A_c/A , given its high uncertainty, is partly to match this observation.

Appendix E: Response to freshwater perturbation

Subjected to the external freshwater perturbation w' , the salinity balance (9)–(10) states

$$\mu q_c' + w' = K S'. \quad (35)$$

Eliminating T' , S' and ρ' from (1), (6)–(14), we derive, for the warm branch,

$$w' = \frac{2K - \mu}{2K + 1} - \frac{K^2}{a}, \quad (36)$$

where the admittance a can be seen from (13), (14) and (25) to be given by $a = [2(1 - \mu)]^{-1}$ and, for the cold branch,

$$w' = 1 - (1 + \mu) \bar{q}_c - \frac{K^2}{a}. \quad (37)$$

These equations allow us to calculate, in a reverse manner, the perturbation w' given K , as plotted in Fig. 3.

Appendix F: Moistening parameter

With global warming, the inverse “meridional” Bowen ratio b of (33) and (30) would be perturbed as

$$\begin{aligned} \ln' b &= \ln' e - \ln' T \\ &= \left(\frac{L}{RT} - 1 \right) \ln' T \quad (\text{Clausius – Clapeyron equation}) \\ &= b_1 T', \end{aligned} \quad (38)$$

where

$$b_1 \equiv \left(\frac{L}{RT} - 1 \right) \cdot \frac{1}{T} \quad (39)$$

and the temperature can be approximated by the global-mean SAT (see the discussion following [10]). Setting the latter to 288°K yields $b_1 \approx 0.092$. From (32), we derive the perturbation in μ ,

$$\begin{aligned} \ln' \mu &= \frac{1}{1 + b} \ln' b \\ &\approx \frac{b_1}{1 + b_0} T' \end{aligned} \quad (40)$$

where b_0 is the unperturbed value of b . Keeping only the first-order terms in the perturbation and replacing T' with the symbol $\delta \bar{T}_a$, we arrive at

$$\mu \approx \mu_0 + \mu_1 \delta \bar{T}_a \quad (41)$$

where $\mu_0 = 0.3$ (“Appendix D”) and

$$\mu_1 = \frac{\mu_0 b_1}{1 + b_0}. \quad (42)$$

For the standard case, $\mu_1 \approx 0.02$.

Appendix G: Heating parameter

The cold-box deficit of the absorbed solar flux from its global-mean is

$$q^* = \bar{q}_s^* (1 - \bar{a}_l) - q_s^* (1 - a_l), \quad (43)$$

where stars indicate dimensional variables, q_s^* is the incident “solar” flux for the cold box and a_l is its surface albedo, both over-barred for their global-mean counterparts. As their representative values, we set $\bar{q}_s^* = 300 \text{ Wm}^{-2}$, $q_s^* = 200 \text{ Wm}^{-2}$, $\bar{a}_l = .3$ and $a_l = 0.45$ (Robock 1980, his Fig. 20) to yield $q^* = 100 \text{ Wm}^{-2}$, which sets the forcing scale $[q']$. A global warming of $\delta\bar{T}_a$ (in °C) would reduce the cold-box albedo by δa_l hence the global-mean albedo by $\delta\bar{a}_l = \delta a_l/2$ since the cold box spans half the global surface. Dividing (43) by $[q']$, the non-dimensionalized forcing is then

$$q' = q_0' - q_1' \delta\bar{T}_a, \quad (44)$$

where subscripts 0 and 1 denote its unperturbed value and its perturbation per degree-warming with $q_0' = 1$ and

$$q_1' = \frac{1}{[q']} \left(q_s^* - \frac{1}{2} \bar{q}_s^* \right) \frac{\delta a_l}{\delta\bar{T}_a}. \quad (45)$$

Based on Robock (1980, his Table 11 for B in high northern latitudes), the fractional ice cover decreases by about 0.03 per degree-warming, and taking the albedo difference between the ice-covered and ice-free surface to be 0.7, we estimate $\delta a_l / \delta\bar{T}_a \approx 0.03 \times 0.7 = 0.021$. Applying foregoing radiative fluxes to (45) yields $q_1' \approx 0.01$.

Appendix H: Response to global warming

Eliminating T' , S' and ρ' from (1), (6)–(14) yields, for the warm branch,

$$q' \left(1 - \frac{\mu}{2K} \right) = \frac{K}{a} \left(\frac{1}{2} + K \right). \quad (46)$$

Applying (41), (44) and retaining only linear terms in $\delta\bar{T}_a$ (the higher-order terms are more than an order-of-magnitude smaller) yields

$$\delta\bar{T}_a \approx C/B, \quad (47)$$

where

$$C = q_0' \left(1 - \frac{\mu_0}{2K} \right) - \frac{K}{a} \left(\frac{1}{2} + K \right), \quad (48)$$

$$B = q_1' \left(1 - \frac{\mu_0}{2K} \right) + \frac{q_0' \mu_1}{2K}, \quad (49)$$

which allows the calculation of $\delta\bar{T}_a$ given K . The counterpart to (46) for the cold branch is

$$q' - (1 + \mu) \bar{q}_c = K^2/a. \quad (50)$$

We have again (47) but with

$$C = q_0' - (1 + \mu_0) \bar{q}_c - K^2/a, \quad (51)$$

$$B = q_1' + \mu_1 \bar{q}_c. \quad (52)$$

References

- Auer SJ (1987) Five-year climatological survey of the Gulf Stream system and its associated rings. *J Geophys Res* 92:11709–11726
- Bryan F (1987) Parameter sensitivity of primitive equation ocean general circulation models. *J Phys Oceanogr* 17:970–985
- Bryan FO, Hecht MW, Smith RD (2007) Resolution convergence and sensitivity studies with North Atlantic circulation models. Part I: The western boundary current system. *Ocean Model* 16(3):141–159
- Ciliberto S, Laroche C (1998) An experimental test of the Gallavotti-Cohen fluctuation theorem. *J Phys IV France* 4:1998629. doi:10.1051/jp
- Crooks GE (1999) Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys Rev E* 60(3):2721–2726
- Dalan F, Stone P, Kamenkovich IV, Scott JR (2005) Sensitivity of the Ocean's Climate to Diapycnal Diffusivity in an EMIC. Part I: Equilibrium State. *J Clim* 18:2460–2481
- Dewar RC (2005a) Maximum entropy production and the fluctuation theorem. *J Phys A Math Gen* 38:L371–L381
- Dewar RC (2005b) Maximum entropy production and nonequilibrium statistical mechanics. In: Kleidon A, Lorenz RD (eds) *Nonequilibrium thermodynamics and the production of entropy: life, Earth, and beyond*. Springer, Heidelberg
- Dixon KW, Delworth TL, Spelman MJ, Stouffer RJ (1999) The influence of transient surface fluxes on North Atlantic overturning in a coupled GCM climate change experiment. *Geophys Res Lett* 26(17):2749–2752
- Evans DJ, Cohen EGD, Morriss GP (1993) Probability of second law violations in shearing steady flows. *Phys Rev Lett* 71:2401–2404
- Gill AE (1982) *Atmosphere–Ocean dynamics*. V30, Academic Press, USA
- Gregory JM, Dixon KW, Stouffer RRJ, Weaver AJ, Driesschaert E, Eby M, Fichefet T, Hasumi H, Hu A, Jungclaus JH, Kamenkovich IV, Levermann A, Montoya M, Murakami S, Nawrath S, Oka A, Sokolov AP, Thorpe RB (2005) A model intercomparison of changes in the Atlantic thermohaline circulation in response to increasing atmospheric CO₂ concentration. *Geophys Res Lett* 32:L12703. doi:10.1029/2005GL023209
- Griffies SM, Winton M, Anderson WG, Benson R, Delworth TL, Dufour CO, Dunne JP, Goddard P, Morrison AK, Rosati A, Wittenberg AT (2015) Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *J Clim* 28(3):952–977
- Grinstein G, Linsker R (2007) Comments on a derivation and application of the ‘maximum entropy production’ principle. *J Phys A Math Theor* 40:9717–9720
- Hughes TM, Weaver AJ (1994) Multiple equilibria of an asymmetric two-basin ocean model. *J Phys Oceanogr* 24(3):619–637
- Hurlburt HE, Hogan PJ (2000) Impact of 1/8 to 1/64 degree resolution on Gulf Stream model–data comparisons in basin-scale subtropical Atlantic Ocean models. *Dyn Atm Ocn* 32(3–4):283–329

- Jones PD, New M, Parker DE, Martin S, Rigor IG (1999) Surface air temperature and its changes over the past 150 years. *Rev Geophys* 37(2):173–199
- Jungclauss JH, Haak H, Esch M, Roeckner E, Marotzke J (2006) Will Greenland melting halt the thermohaline circulation? *Geophys Res Lett* 33:L17708. doi:10.1029/2006GL026815
- Kleidon A (2009) Non-equilibrium thermodynamics and maximum entropy production in the Earth system: applications and implications. *Naturwissenschaften* 96:653–677
- Manabe S, Stouffer RJ (1988) Two stable equilibria of a coupled ocean-atmosphere model. *J Clim* 1:841–866. doi:10.1126/science.1189250
- Manabe S, Stouffer RJ (1993) Century-scale effects of increased atmospheric CO₂ on the ocean-atmosphere system. *Nature* 364(6434):215–218
- Manabe S, Stouffer RJ (1994) Multiple-century response of a coupled ocean-atmosphere model to an increase of atmospheric carbon dioxide. *J Clim* 7(1):5–23
- Manabe S, Stouffer RJ (1999) Are two modes of thermohaline circulation stable? *Tellus* 51A:400–411
- Marotzke J, Scott JR (1999) Convective mixing and the thermohaline circulation. *J Phys Oceanogr* 29:2962–2970
- Marotzke J, Stone P (1995) Atmospheric transports, the thermohaline circulation, and flux adjustments in a simple coupled model. *J Phys Oceanogr* 25:1350–1364
- Marshall J, Speer K (2012) Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nature Geos* 5(3):171–180
- Mikolajewicz U, Voss R (2000) The role of the individual air-sea flux components in CO₂-induced changes of the ocean's circulation and climate. *Clim Dyn* 16(8):627–642
- Nakamura M, Stone PH, Marotzke J (1994) Destabilization of the thermohaline circulation by atmospheric eddy transports. *J Clim* 7:1870–1882
- Ou HW (2006) Meridional thermal field of a coupled ocean-atmosphere system: a conceptual model. *Tellus* 58A:404–415
- Ou HW (2007) Hydrological cycle and ocean stratification in a coupled climate system: a theoretical study. *Tellus* 59A:683–694
- Ou HW (2012) A minimal model of the Atlantic multidecadal variability: its genesis and predictability. *Clim Dyn* 38(3):775–794. doi:10.1007/s00382-011-1007-3
- Peixoto JP, Oort AH (1992) *Physics of climate*. Amer Inst Phys, New York, p 522
- Pierrehumbert RT (1991) Large-scale horizontal mixing in planetary atmospheres. *Phys Fluids* A3(5):1250–1260
- Prange M, Lohmann G, Paul A (2003) Influence of vertical mixing on the thermohaline hysteresis: analyses of an OGCM. *J Phys Oceanogr* 33:1707–1721
- Rahmstorf S (1995) Bifurcations of the Atlantic thermohaline circulation in response to changes in the hydrological cycle. *Nature* 378:145–149
- Rahmstorf S, Ganopolski A (1999) Long-term global warming scenarios computed with an efficient coupled climate model. *Clim Change* 43(2):353–367
- Rahmstorf S, Willebrand J (1995) The role of temperature feedback in stabilizing the thermohaline circulation. *J Phys Oceanogr* 25:787–805
- Rahmstorf S, Crucifix M, Ganopolski A, Goosse M, Kamenkovich I, Knutti R, Lohmann G, Marsh R, Mysak LA, Wang Z, Weaver AJ (2005) Thermohaline circulation hysteresis: a model intercomparison. *Geophys Res Lett* 32(23):L23605. doi:10.1029/2005GL023655
- Robock A (1980) The seasonal cycle of snow cover, sea ice and surface albedo. *Mon Wea Rev* 108(3):267–285
- Roemmich D, Wunsch C (1985) Two transatlantic sections: Meridional circulation and heat flux in the subtropical North Atlantic Ocean. *Deep Sea Res A Ocn Res Paps* 32(6):619–664
- Ruddick B, Zhang L (1996) Qualitative behavior and nonoscillation of Stommel's thermohaline box model. *J Clim* 9:2768–2777
- Schiller A, Mikolajewicz U, Voss R (1997) The stability of the North Atlantic thermohaline circulation in a coupled ocean-atmosphere general circulation model. *Clim Dyn* 13(5):325–347
- Schmittner A, Weaver AJ (2001) Dependence of multiple climate states on ocean mixing parameters. *Geophys Res Lett* 28:1027–1030
- Sevick EM, Prabhakar R, Williams SR, Searles DJ (2008) Fluctuation theorems. *Ann Rev Phys Chem* 59:603
- Stocker TF, Knutti R, Plattner GK (2001) The future of the thermohaline circulation—a perspective. In: Seidov D, Maslin M, Haupt BJ (eds) *The oceans and rapid climate change: past, present, and future*, vol 126. American Geophysical Union, Washington, pp 277–293
- Stommel H (1961) Thermohaline convection with two stable regimes of flow. *Tellus* 13:224–230
- Stouffer R, Yin J, Gregory J, Dixon K, Spelman M, Hurlin W, Weaver A, Eby M, Flato G, Hasumi H, Hu A, Jungclauss J, Kamenkovich I, Levermann A, Montoya M, Murakami S, Nawrath S, Oka A, Peltier W, Robitaille D, Sokolov A, Vettoretti G, Weber S (2006) Investigating the causes of the response of the Thermohaline Circulation to past and future climate changes. *J Clim* 19:1365–1387
- Tippins D, Tomczak M (2003) Meridional Turner angles and density compensation in the upper ocean. *Ocean Dyn* 53(4):332–342
- Vellinga M, Wood RA (2002) Global climatic impacts of a collapse of the Atlantic thermohaline circulation. *Clim Change* 54(3):251–267
- Wang GM, Sevick EM, Mittag E, Searles DJ, Evans DJ (2002) Experimental demonstration of violations of the Second Law of Thermodynamics for small systems and short time scales. *Phys Rev Lett* 89(5):050601/1–050601/4
- Wunsch C (2002) What is the thermohaline circulation? *Science* 298:1179–1180
- Zaslavsky G (1999) Chaotic dynamics and the origins of statistical laws. *Phys Today* 8:39–45