Ocean Gyres and Abrupt Change in the Thermohaline Circulation: A Conceptual Analysis

Hannah Longworth

School of Ocean and Earth Science, and Graduate School, Southampton Oceanography Centre, Southampton, United Kingdom

Jochem Marotzke

Max Planck Institute for Meteorology, Hamburg, Germany

Thomas F. Stocker

Climate and Environmental Physics, University of Berne, Berne, Switzerland

(Manuscript received 14 January 2004, in final form 28 August 2004)

ABSTRACT

The implications are investigated of representing ocean gyre circulations by a diffusion term in the Stommel and Rooth box models of the thermohaline circulation (THC) in one and two hemispheres, respectively. The approach includes mostly analytical solution and study of the bifurcation structure, but also numerical integration and feedback analysis. Sufficient diffusion (gyre strength) eliminates multiple equilibria from either model, highlighting the need for accurate gyre circulation strength in general circulation models (GCMs) when considering the potential for abrupt climate change associated with THC shutdown.

With diffusion, steady-state flow strength in the Rooth model depends on freshwater forcing (i.e., implied atmospheric water vapor transport) in both hemispheres, not only on that in the upwelling hemisphere, as in the nondiffusive case. With asymmetric freshwater forcing, two solutions (strong stable and weak unstable) are found with sinking in the hemisphere with stronger forcing and one solution with sinking in the other hemisphere. Under increased freshwater forcing the two solutions in the hemisphere with stronger forcing meet in a saddle-node bifurcation (if diffusion is sufficiently strong to prevent a subcritical Hopf bifurcation first), followed by flow reversal. Thus, the bifurcation structure with respect to freshwater forcing of the diffusive Rooth model of two-hemisphere THC is similar to that of the Stommel model of single-hemisphere THC, albeit with a very different dynamical interpretation. Gyre circulations stabilize high-latitude sinking in the Stommel model. In the Rooth model, gyre circulations only stabilize high-latitude sinking if the freshwater forcing is weaker in the sinking hemisphere than in the upwelling hemisphere, by an amount that increases with diffusion. The values of diffusion and freshwater forcing at which qualitative change in behavior occurs correspond to the range of the values used in and obtained with GCMs, suggesting that this analysis can provide a conceptual foundation for analyzing the stability of the interhemispheric THC, and also for the potential of the Atlantic THC to undergo abrupt change.

1. Introduction

Two international reports (Cubasch et al. 2001; National Research Council 2002) and several papers (e.g., Stocker 2000; Rahmstorf 2002; Clark et al. 2002; Alley et al. 2003) have recently reviewed the possibility that

Corresponding author address: Dr. Jochem Marotzke, Max Planck Institute for Meteorology, Bundesstrasse 53, 20146 Hamburg, Germany.
E-mail: marotzke@dkrz.de

abrupt climate change might be caused by a collapse of the Atlantic thermohaline circulation (THC). In particular, the possibility has been investigated that an increase in atmospheric greenhouse gas concentrations might create a THC collapse (e.g., Manabe and Stouffer 1993; Stocker and Schmittner 1997). This attention underscores the importance of understanding the processes that might lead to abrupt change in THC and climate so that a rational assessment is possible of the probability of such an event.

National Research Council (2002, p. 14) and Alley

et al. (2003) suggested two complementary definitions of "abrupt climate change." In addition to an impact-oriented definition that compares magnitude, extent, and speed of climate change with the ability of societies to cope with change, it was stated that "technically, an abrupt climate change occurs when the climate system is forced to cross some threshold, triggering a transition to a new state at a rate determined by the climate system itself and faster than the cause."

This paper is motivated by the need for a conceptual discussion of a fundamental point: Prior to trying to assess how far climate is from a threshold in, say, atmospheric CO_2 concentration and THC stability, we must ask ourselves whether such a threshold exists. Indeed, the model results shown in the Third Assessment Report by the Intergovernmental Panel on Climate Change (IPCC) mostly show a projected THC weakening (Cubasch et al. 2001), but the changes occur on the time scale of the forcing and, therefore, are not classified as "abrupt." However, the uncertainties of such estimates are large (e.g., Knutti et al. 2003). Here, we use two classical box models of the THC, in one hemisphere (Stommel 1961; Marotzke 1990, 2000) and in two hemispheres (Rooth 1982; Rahmstorf 1996; Scott et al. 1999), to explore the possibility of such a threshold existing in reality and the implications of this for reversibility and irreversibility of climate change.

In contrast to most previous papers, the work presented here contains a representation of wind-driven ocean gyres, which enter the model through a diffusive parameterization and whose influence is profound. This aspect has not been analyzed much in previous box model papers, although the stabilizing influence of wind-induced salinity transport on the THC has long been known. Marotzke (1990) and Oka et al. (2001) emphasized the Ekman salinity transport if a halocline was present at high latitudes, while Winton and Sarachik (1993) and Schiller et al. (1997) concentrated on the enhanced subpolar gyre salinity transport in a state of collapsed THC. Furthermore, diffusion in GCMs is known to influence the potential for multiple equilibria of the Atlantic THC (Manabe and Stouffer 1999; Schmittner and Weaver 2001; Prange et al. 2003).

Nevertheless, gyre salt transport has mostly either been omitted from box model studies or accounted for by defining atmospheric water vapor transport to be only the component that is not redistributed by wind-driven flow (Rahmstorf 1996; Titz et al. 2002b), thus preventing feedback between THC and gyre transport. One exception is Thual and McWilliams (1992), who included horizontal diffusion in a Stommel and a three-box model and vertical and horizontal diffusion in a 3 × 2 box model. However, altered scaling, which combined the diffusive and advective model fluxes, prevented an explicit exploration of the implications of diffusion. Notable recent developments are Shaffer and Olsen (2001) who generalized the coupled box model of Marotzke and Stone (1995) to include horizontal diffusion, as well as a vertical split in the oceanic boxes and vertical diffusion. Shaffer and Olsen (2001) found that high-latitude sinking was stabilized in their model by increasing horizontal or vertical mixing. Pasquero and Tziperman (2004) coupled a Stommel-type box model with an explicit horizontally oriented torus, mimicking gyre circulation, and concentrated on the variability caused by the interaction of wind- and buoyancy-driven flow.

Here, we use a much simpler rendition of a box model with a representation of gyre transports than in either Shaffer and Olsen (2001) or Pasquero and Tziperman (2004) so that we can explicitly present exact analytical solutions. The implications for abrupt climate change of incorporating gyre transports in the Stommel model have been broadly summarized before in National Research Council (2002) and Alley et al. (2003). However, no technical or mathematical detail was given in either reference. The larger part of the present paper is then taken up by investigating the role of gyre transports in the Rooth model. To our knowledge, interhemispheric THC dynamics have never been analyzed in this way. The layout of this paper is as follows. Section 2 deals with the Stommel model and section 3 deals with the Rooth model. Combined discussion and conclusions for both models are presented in section 4.

2. Gyres in the Stommel model

a. Model formulation and steady-state solutions

We use a modification to the original Stommel (1961) model, in the simplified variant amenable to analytical solution (Marotzke 1990, 2000). The geometry of the model is shown in Fig. 1. The North Atlantic high and low latitudes are represented by two well-mixed boxes, 1 and 2. Freshwater forcing is defined as the influence on salinity of the cycle comprising net evaporation at low latitudes, atmospheric water vapor transport to high latitudes, and net precipitation at high latitudes. We neglect the mass flux associated with net precipitation at high latitudes, P, and only consider the induced salinity changes by using a virtual surface salinity flux Φ (Bryan 1986):

$$\Phi = -S_0 P/H, \qquad (1)$$

where H is depth and S_0 a reference salinity (e.g., Marotzke and Willebrand 1991). Model constants are those of Marotzke (1996); see Table 1. We assume that
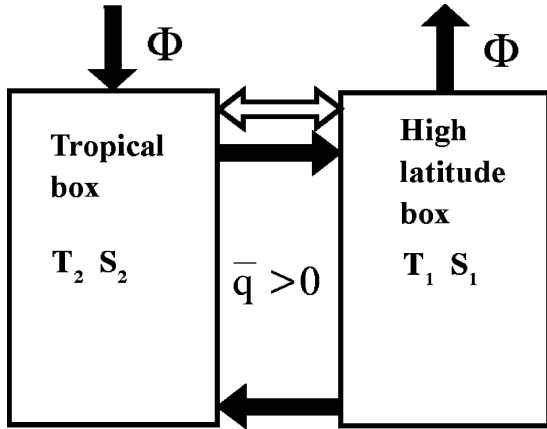
FIG. 1. The Stommel model with diffusion. Filled and unfilled arrows are the advective and diffusive flow components, respectively. Advective arrows reverse under flow reversal but diffusive arrows are unchanged.

the temperatures, T_1 and T_2, are completely prescribed. This is equivalent to assuming a restoring condition (Haney 1971) for heat flux with infinitely strong coupling and thus constitutes an extreme but physically interpretable limiting case of a coupled box model (Marotzke 1996). We have investigated the validity of assuming box temperatures to be fixed external parameters through numerical integration of a model variant with finite restoring strength and have not found any qualitative difference. For simplicity, and to limit the paper's length, we do not include these numerical results here.

The boxes are connected by pipes near the surface and the bottom; the pipes are assumed to have vanishing volume but are conduits for the flow. The thermohaline circulation strength is denoted by q (strictly speaking, q represents THC/volume; q has units of per second). We use the sign convention that $q > 0$ denotes poleward surface flow, implying equatorward bottom flow and, conceptually, sinking at high latitudes. This is the picture that we are used to when thinking about the North Atlantic THC. Conversely, $q < 0$ means equatorward surface flow and poleward bottom flow. We assume a very simple flow law for q, namely, that it

depends linearly on the density difference between high and low latitudes:

$$q = k(\rho_1 - \rho_2)/\rho_0 = k[\alpha(T_2 - T_1) - \beta(S_2 - S_1)], \quad (2)$$

where ρ_0 is a reference density and k is a hydraulic constant that contains all dynamics, that is, the connection between density and the flow field. Here α and β are respectively the constant thermal and haline expansion coefficients, given in Table 1; for simplicity, we employ a linear equation of state. Expressing THC dynamics through the intrahemispheric density gradient as in (2) is problematical, especially when thinking of the cross-hemispheric Atlantic THC. Welander (1986) represented the Atlantic THC by two back-to-back Stommel models, but it may be more appropriate to use the density difference between high northern and high southern latitudes as the driver determining flow strength, as first done in a box model by Rooth (1982). Rahmstorf (1996) and Marotzke (2000) have discussed this juxtaposition of Stommel's and Rooth's models, and the general circulation model results of Klinger and Marotzke (1999) gave support to viewing Rooth's model as dynamically superior. We will nevertheless proceed with the Stommel model here because it allows us to make a number of fundamental points in the simplest possible way. Section 3 will analyze the effect of gyre transports in the Rooth model.

Salt transports by the gyre circulation are parameterized by diffusive terms, where k_d (s^{-1}) is a diffusion coefficient analogous to the hydraulic constant. The salt conservation equations of the system are thus

$$\dot{S}_1 = -\Phi + |q|(S_2 - S_1) + k_d(S_2 - S_1), \quad (3)$$

$$\dot{S}_2 = \Phi - |q|(S_2 - S_1) - k_d(S_2 - S_1). \quad (4)$$

With the abbreviations for meridional differences of temperature and salinity,

$$T \equiv T_2 - T_1; \quad S \equiv S_2 - S_1, \quad (5)$$

taking the time derivative of (2) using (3)–(5) gives

$$\dot{q} = -2k\beta\Phi - 2(|q| + k_d)(q - k\alpha T). \quad (6)$$

When seeking analytical solutions for steady states, the modulus in (6) requires us to consider separately two cases. Equilibrium quantities are marked by an overbar and are readily verified as

Temperature dominance:

$$\bar{q} > 0, \quad \alpha T > \beta \bar{S}, \quad (7)$$

$$\bar{q}_{A/B} = \frac{1}{2}\left\{(k\alpha T - k_d) \pm \sqrt{(k\alpha T + k_d)^2 - 4k\beta\Phi}\right\}, \quad (8)$$

TABLE 1. Parameters of the Stommel model.

Parameter	Symbol	Value
Thermal expansion coefficient	α	1.5×10^{-4} K^{-1}
Haline expansion coefficient	β	8.0×10^{-4} psu^{-1}
Hydraulic constant	k	2×10^{-8} s^{-1}
Depth	H	5000 m
Reference salinity	S_o	35

where A and B denote the solutions. For a positive radicand, defined by

$$\frac{k\beta\Phi}{(k\alpha T + k_d)^2} < \frac{1}{4}, \tag{9}$$

and provided that condition (7), temperature dominance, is not violated, the model has two equilibrium solutions for poleward near-surface flow, as has the original Stommel model. The branch with weaker flow is unstable to infinitesimal perturbations, again as in the nondiffusive model. A systematic investigation of the existence of the various solutions is presented below.

Salinity dominance:

$$\bar{q} < 0, \quad \alpha T < \beta\bar{S}, \tag{10}$$

$$\bar{q}_C = \frac{1}{2}\left\{(k\alpha T + k_d) - \sqrt{(k\alpha T - k_d)^2 + 4k\beta\Phi}\right\}. \tag{11}$$

Notice that we must discard the positive root in (11); the radicand is greater than $k\alpha T - k_d$, so the positive root would imply $\bar{q} > 0$, violating the condition (10).

b. Conditions for existence of equilibria (bifurcation structure)

The solutions (8) and (11) are governed by two dimensionless parameters, the salinity flux, $E \equiv k\beta\Phi/(k\alpha T)^2$, and the diffusivity, $D \equiv k_d/(k\alpha T)$. Nondimensional equilibrium flow strength, $Q \equiv q/(k\alpha T)$ as a function of E and D, shows under what conditions multiple equilibria exist (Fig. 2). The characteristic Stommel bifurcation diagram with three flow solutions is recovered for $D = 0$ and $E < 0.25$. The middle solution B, on the branch that slopes from the lower left to the upper right, is unstable to infinitesimal perturbations. Diffusion increases the maximum freshwater forcing, E_1, for which the high-latitude sinking solutions A and B exist (e.g., from $E_1 = 0.25$ to 0.5625 for D of 0 and 0.5, respectively). From (9), we find

$$E_1 = \frac{1}{4}(1 + D)^2, \tag{12}$$

with the corresponding flow strength:

$$\bar{Q}_{A/B}(E_1) = \frac{1}{2}(1 + D). \tag{13}$$

The upper limit of freshwater forcing corresponds to the maximum salinity flux that the advective and diffusive transports can redistribute, preventing the salinity gradient from reversing the thermally driven flow. For growing diffusion, the range of validity of the unstable, weak solution B diminishes. It is readily shown from (8) that solution B does not exist if $D > 1$ or $E > 1$.
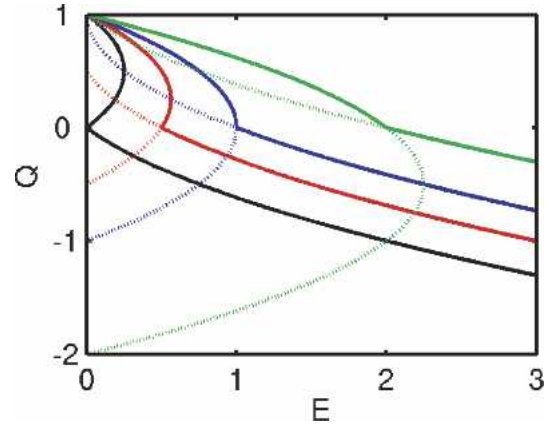

FIG. 2. Bifurcation structure of the Stommel model with diffusion. Black: $D = 0$; red: $D = 0.5$; blue: $D = 1$; green: $D = 2$. Solid lines: Permissible solutions, dotted lines: full solutions. Solution A marks the branch with $Q > 0$ and a negative slope; solution B the branch with $Q > 0$ and a positive slope; solution C the branch with $Q < 0$. Point E_1 occurs where the $Q > 0$ branch runs vertically; point E_2 occurs where the solution curves cut through $Q = 0$. See text for definitions.

A *lower* limit, E_2, on freshwater forcing for the occurrence of multiple equilibria is defined by the intersection of the weak positive and the negative flow solutions, B and C respectively, at $Q = 0$. One readily finds that the condition $E = E_2$ implies

$$E_2 = D. \tag{14}$$

Solutions B and C are invalid for $E < E_2$ since this would imply $\bar{Q}_B < 0$ and $\bar{Q}_C > 0$, respectively. Thus, only the strong positive solution, A, exists for $E < E_2$. Physically, this limit arises because diffusion reduces the meridional salinity difference but not the temperature difference. As the salinity difference brakes the THC strength in solution A but drives it in solution C, the former is strengthened by diffusion whereas the latter is weakened. Both statements are readily verified by differentiating (8) and (11) with respect to k_d.

In summary, multiple equilibria exist for $D < 1$ and

$$D \equiv E_2 < E < E_1 \equiv \frac{1}{4}(1 + D)^2. \tag{15}$$

Alternatively, (12) and (14) show that for a fixed $E < 1$, multiple equilibria exist for a diffusion (gyre strength) in the range of

$$2\sqrt{E} - 1 \equiv D_1 < D < D_2 \equiv E. \tag{16}$$

All three solutions merge at a critical diffusivity, D^c, and a critical freshwater flux, E^c:
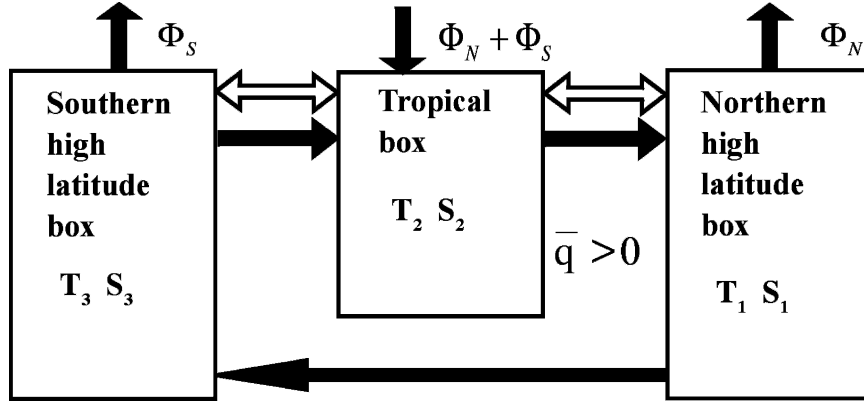
$$D^c = E^c = 1. \tag{17}$$

FIG. 3. As in Fig. 1 but for the Rooth model.

For $D > 1$, an increase of E leads to a direct transition from solution A to C, without the possibility of the intermediate unstable solution (see Fig. 2). Conversely, if $E > 1$, an increase of D leads to a direct transition from solution C to solution A. The condition (17) defines a *codimension-2 bifurcation* (Strogatz 1994, p. 70, "This fancy terminology essentially means that we have had to tune two parameters, here D and E, to achieve this type of bifurcation").

3. Gyres in the Rooth model

a. Model formulation

The version of Rooth's model (Fig. 3) used here is based on that of Scott et al. (1999). The boxes are well mixed and have equal volumes. Boxes 1 and 3 are the northern and southern high latitudes, respectively; box 2 is the tropical box. Flow between boxes is via insulated current pathways of negligible mass. The advective flow component sinks in one high-latitude box and upwells in the other with return surface flow via the tropical box. Northern sinking and southern upwelling correspond to $q > 0$; reversed if $q < 0$. Temperature and salinity of box i are T_i and S_i. Consistent with section 2, constant atmospheric water vapor transports from low to high latitudes are assumed, represented by equivalent salinity fluxes Φ_N and Φ_S, while box temperatures are governed by Newtonian relaxation at rate λ. Parameter values are those of Scott et al. (1999); see Table 2. Overturning strength is a linear function of the density difference between the two high-latitude boxes with a linear equation of state:

$$q = \frac{k}{\rho_0}(\rho_1 - \rho_3) = k[\alpha(T_3 - T_1) - \beta(S_3 - S_1)]. \quad (18)$$

Symbols are as those in section 2 except for k, which is a different hydraulic constant because similar strength flow now is created by a much smaller density difference. As in section 2, a diffusion term with coefficient k_d parameterizes heat and salt transport by oceanic subtropical gyres. Temperature and salinity of each box are governed by

$$\dot{T}_1 = \lambda(T_p - T_1) + k_d(T_2 - T_1)$$
$$+ \begin{cases} q(T_2 - T_1); & q \geq 0 \\ |q|(T_3 - T_1); & q < 0 \end{cases} \quad (19)$$

$$\dot{T}_2 = \lambda(T_e - T_2) + k_d(T_1 + T_3 - 2T_2)$$
$$+ \begin{cases} q(T_3 - T_2); & q \geq 0 \\ |q|(T_1 - T_2); & q < 0 \end{cases} \quad (20)$$

$$\dot{T}_3 = \lambda(T_p - T_3) + k_d(T_2 - T_3)$$
$$+ \begin{cases} q(T_1 - T_3); & q \geq 0 \\ |q|(T_2 - T_3); & q < 0 \end{cases} \quad (21)$$

$$\dot{S}_1 = -\Phi_N + k_d(S_2 - S_1) + \begin{cases} q(S_2 - S_1); & q \geq 0 \\ |q|(S_3 - S_1); & q < 0 \end{cases}$$
$$(22)$$

$$\dot{S}_2 = \Phi_N + \Phi_S + k_d(S_1 + S_3 - 2S_2)$$
$$+ \begin{cases} q(S_3 - S_2); & q \geq 0 \\ |q|(S_1 - S_2); & q < 0 \end{cases} \quad (23)$$

$$\dot{S}_3 = -\Phi_S + k_d(S_2 - S_3) + \begin{cases} q(S_1 - S_3); & q \geq 0 \\ |q|(S_2 - S_3); & q < 0 \end{cases}.$$
$$(24)$$

b. Numerical solutions

Integration of the Rooth model from the initial conditions $(T_1, T_2, T_3) = (1, 10, 0.1)$, and $(S_1, S_2, S_3) = (35, 35.7, 34.3)$ is performed under different ratios of

TABLE 2. Parameters of the Rooth model.

Constant	Symbol	Value
High-latitude thermal bath	T_p	0°C
Equatorial thermal bath	T_e	30°C
Newtonian relaxation constant	λ	12.9×10^{-10} s^{-1}
Hydraulic constant	k	1.5×10^{-6} s^{-1}
Depth	H	5000 m
Reference salinity	S_o	35 psu

northern to southern freshwater forcing, which is an additional degree of freedom compared to the Stommel model. Following Scott et al. (1999), we define cases I, II, and III as $\Phi_N = \Phi_S$ (case I), $\Phi_N = 0.5 \Phi_S$ (case II), and $\Phi_N = 1.5 \Phi_S$ (case III). Here Φ_S is fixed at 0.9×10^{-10} psu s^{-1}, corresponding to net precipitation at high southern latitudes of about 0.5 m yr^{-1}, while Φ_N takes the values 0.9×10^{-10}, 0.45×10^{-10}, and 1.35×10^{-10} psu s^{-1}, respectively. An equivalent treatment could be applied to the Southern Hemisphere flux, but the effects may instead be implied from symmetry. For reference, the equilibrium flow strength without diffusion is 1.51×10^{-10} s^{-1}, corresponding to 15 Sv (1Sv $\equiv$ 10^6 m^3 s^{-1}) in a box 5000 km $\times$ 4000 km $\times$ 5000 m.

Starting from a northern sinking state with zero diffusion, overturning weakens with increasing diffusion due to redistribution of salt dominating that of heat. In the Rooth model, the THC fundamentally, and conceptually, is thermally driven, as reflected by the high-latitude sinking (see National Research Council 2002, p. 103). The THC induces a pole-to-pole temperature gradient that opposes flow, with northern sinking, $\overline{T}_1 > \overline{T}_3$, since warm water is advected from box 2 to 1. For realistic choices of thermal and haline expansion coefficients, flow strength and direction are therefore largely "driven" by the pole-to-pole salinity difference (Scott et al. 1999), in contrast to the Stommel model in which salinity forcing brakes the thermally driven high-latitude sinking. As seen from Table 3, diffusion reduces $\overline{T}_1 - \overline{T}_3$ (the opposing temperature gradient), tending to strengthen northern sinking. The weakened overturning observed is therefore due to the corresponding decrease of $\overline{S}_1 - \overline{S}_3$, the driving pole-to-pole salinity gradient. In case III this decrease causes collapse of northern sinking; for some critical k_d between 5.0×10^{-11} and 5.1×10^{-11} s^{-1}, $\overline{q}$ changes from 6.77×10^{-11} s^{-1} to -1.54×10^{-10} s^{-1}. Likewise under increased diffusion the southern branch of case II undergoes a transition to northern sinking for some k_d between 1.7×10^{-11} and 1.8×10^{-11} s^{-1} ($\overline{q}$ changes from -7.19×10^{-11} s^{-1} to 1.39×10^{-10} s^{-1}).

An interpretation of the magnitude of the critical k_d can be given in terms of either a horizontal gyre strength or a horizontal diffusivity. As done before for flow q, using a box volume of $V = 10^{17}$ m^3 allows us to directly translate the critical k_d of case III, $k_d \approx 5 \times 10^{-11}$ s^{-1} into an equivalent horizontal gyre of 5 Sv, assuming it operates on the full meridional salinity difference. If, in contrast, the zonal salinity difference is only 25% of the meridional difference, the same effect is caused by a gyre of 20-Sv strength. With a length scale of 5000 km, the equivalent dimensional critical diffusivity of case III is around 10^3 m^2 s^{-1}, a standard value used in numerical ocean models. Variations of k_d on the order of 10^{-10} s^{-1} correspond to the ones investigated systematically in the two-dimensional model of Schmittner and Weaver (2001), where they were shown to cause qualitative changes in the solutions.

Yet another interpretation is obtained from calculating the implied diffusive freshwater transport, FW. Using again the critical k_d of case III gives

$$\text{FW} = \frac{k_d V \Delta S}{S_0} = 0.13 \text{ Sv.} \qquad (25)$$

TABLE 3. Steady-state flow characteristics of the Rooth model. Lines in italic are quasi-no-flow states.

|	k_d (s^{-1})	$\overline{q}$ (s^{-1})	$\overline{T}_1 - \overline{T}_3$ (°C)	$\overline{S}_2 - \overline{S}_1$	$\overline{S}_2 - \overline{S}_3$	$\overline{S}_1 - \overline{S}_3$
Case I	0	1.51×10^{-10}	2.52	0.5976	1.1952	0.5976
	1×10^{-11}	1.48×10^{-10}	2.34	0.5961	1.1527	0.5566
	1×10^{-10}	5.43×10^{-11}	0.85	0.5833	0.7886	0.2052
	1×10^{-8}	*4.74×10^{-22}*	*0.00*	*0.0090*	*0.0090*	*0.0000*
Case II	0	1.51×10^{-10}	2.52	0.2988	0.8964	0.5976
	1×10^{-11}	1.44×10^{-10}	2.38	0.2923	0.8579	0.5656
	1×10^{-10}	9.62×10^{-11}	1.40	0.2294	0.5712	0.3418
	1×10^{-8}	*5.27×10^{-12}*	*0.00*	*0.0045*	*0.0090*	*0.0045*
Case III	0	1.51×10^{-10}	2.52	0.8964	1.4939	0.5976
	1×10^{-11}	1.38×10^{-10}	2.30	0.9131	1.4601	0.5470
	1×10^{-10}	-1.24×10^{-10}	-1.71	0.8252	0.4019	-0.4234
	1×10^{-8}	*-5.27×10^{-12}*	*0.00*	*0.0135*	*0.0090*	*-0.0045*
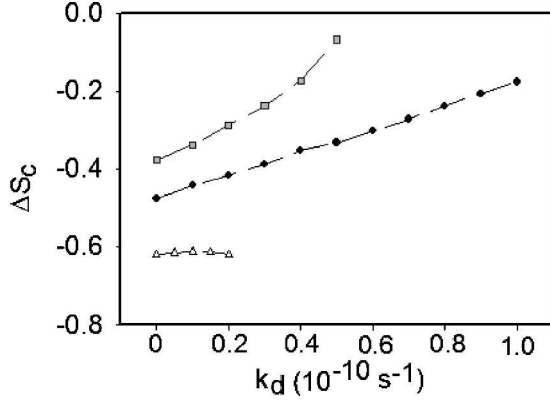
FIG. 4. Critical salinity perturbations of the Rooth model with diffusion. Solid circles: case I; open triangles: case II; gray squares: case III.

By comparison, in the low-resolution, single-hemisphere experiments of Marotzke (1990), a gyre freshwater transport of 0.05 Sv and a wind-driven shallow overturning freshwater transport of equal size were sufficient to reestablish a temporarily collapsed THC. Thus, the important transitions in box model behavior happen in a realistic parameter range, implying that any particular GCM realization might be on either side of a critical k_d.

Under symmetric freshwater forcing (case I), strong diffusion ($k_d \geq 1 \times 10^{-8}$ s^{-1}) eliminates the pole-to-pole density gradient, thus overturning tends to zero. With asymmetric freshwater forcing, however, a pole-to-pole salinity (but not temperature) gradient is maintained; note that the ratio $(\overline{S}_2 - \overline{S}_1):(\overline{S}_2 - \overline{S}_3)$ is the same as $\Phi_N: \Phi_S$ at equilibrium (see italic lines in Table 3). The resulting flow solution is termed a quasi-no-flow state; hence, flow would be zero were it not for the asymmetric forcing. Such quasi-no-flow states are dynamically different to those of Scott et al. (1999) and the no-flow state of Rooth (1982). Without diffusion the no-flow solution is a transient symmetric state, unstable to small perturbations (Rooth 1982), whereas the quasi-no-flow states induced by diffusion are stable to salinity perturbations up to an anomaly of −1.0 added to the sinking box. The direction of the quasi-no-flow overturning for cases II and III are also reversed compared to those with no diffusion; case III has positive overturning and case II has negative overturning in Fig. 5 of Scott et al. (1999).

We now numerically investigate numerically the stability of the northern sinking solution to instantaneous reductions in box-1 salinity. A critical salinity perturbation, ΔS_c, is defined as the minimum ΔS that, if removed from the sinking box and added to the tropical box, causes permanent flow reversal. Critical perturba-
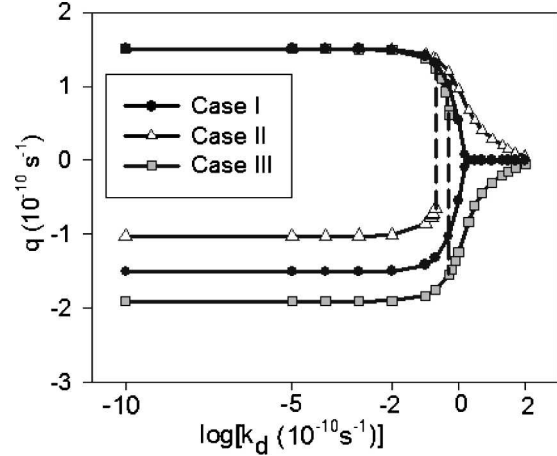

FIG. 5. Numerical solutions to Rooth's model with diffusion under variable temperatures. The lower branch solutions have been found with initial conditions implying initial southern sinking flow. Dashed vertical lines mark transitions to the opposite solution branch as k_d crosses a critical value.

tion magnitude decreases with diffusion for cases I and III (Fig. 4), so diffusion destabilizes northern sinking in these cases. The increased magnitude of ΔS_c for case II with $k_d = 2 \times 10^{-11}$ s^{-1} relative to $k_d = 0$ indicates weak stabilization; however, there is destabilization between $k_d = 0$ and 1×10^{-11} s^{-1}. Critical perturbations are only given at low diffusion for case II since the negative overturning solution is not found for $k_d > 2 \times 10^{-11}$ s^{-1} (refer to Fig. 5). That diffusion destabilizes appears surprising. However, Schmittner and Weaver (2001) also found destabilization of northern sinking under increasing diffusion and excess precipitation in the North Atlantic; an explanation is now provided through the analysis of a simplified setup, in which the temperatures in the Rooth model are fixed.

c. Analytical steady-state solutions

In steady state with $\overline{q} > 0$, it follows from (22) and (23) that

$$\overline{S}_2 - \overline{S}_1 = \frac{\Phi_N}{(\overline{q} + k_d)}, \quad (26)$$

$$\overline{S}_2 - \overline{S}_3 = \frac{\overline{q}(\Phi_N + \Phi_S) + k_d\Phi_S}{(\overline{q} + k_d)^2}, \quad (27)$$

$$\overline{S}_1 - \overline{S}_3 = \frac{k_d(\Phi_S - \Phi_N) + \overline{q}\Phi_S}{(\overline{q} + k_d)^2}, \quad (28)$$

and, with the flow law (18), that

$$\overline{q} = k\left\{\alpha(\overline{T}_3 - \overline{T}_1) + \frac{\overline{q}\beta\Phi_S - k_d\beta(\Phi_N - \Phi_S)}{(\overline{q} + k_d)^2}\right\}. \quad (29)$$

By setting $k_d = 0$ we can solve for $\overline{q}$ and recover the analytical solution of the nondiffusive Rooth model as derived by Rahmstorf (1996):

$$\overline{q} = \frac{k\alpha(\overline{T}_3 - \overline{T}_1)}{2} + \frac{1}{2}\sqrt{k^2\alpha^2(\overline{T}_3 - \overline{T}_1)^2 + 4k\beta\Phi_S}.$$

(30)

If $\overline{q} < 0$, the steady-state solution including diffusion is readily shown to satisfy

$$\overline{q} = k\left\{\alpha(\overline{T}_3 - \overline{T}_1) + \frac{\overline{q}\beta\Phi_N - k_d\beta(\Phi_N - \Phi_S)}{(k_d - \overline{q})^2}\right\}. \quad (31)$$

With fixed symmetric temperatures, $T_1 = T_3$, expansion of (29) and (31) yields the cubic equation

$$\overline{q}^3 \pm 2k_d\overline{q}^2 + \overline{q}(k_d^2 - k\beta\Phi_{S/N}) + k_dk\beta(\Phi_N - \Phi_S) = 0,$$

(32)

where the upper sign applies for $\overline{q} > 0$ and the lower sign for $\overline{q} < 0$, and $\Phi_{S/N}$ denotes Φ_S for $\overline{q} > 0$ and Φ_N for $\overline{q} < 0$. With diffusion and asymmetric freshwater forcing, (32) shows that steady-state flow strength is influenced by the freshwater flux in both the sinking and upwelling hemispheres. This stands in marked contrast to the nondiffusive steady state, (30), which shows dependence on freshwater flux in the upwelling hemisphere only; generally, the influence of the sinking hemisphere forcing increases with diffusion.

Figure 6 shows solutions to (32) as a function of diffusion under the three freshwater-forcing scenarios of section 3b ($\Phi_S = 0.9 \times 10^{-10}$ psu s^{-1}; case I: $\Phi_N = \Phi_S$, case II: $\Phi_N = 0.5\,\Phi_S$, case III: $\Phi_N = 1.5\,\Phi_S$). Numerical integration under fixed temperatures is used to determine flow stability. Note that the limit, $k_d \to 0$, is singular as no symmetric equilibrium solution, $\overline{q} = 0$, exists, regardless of freshwater forcing (Rooth 1982).

As shown in the appendix, for case I ($\Phi_N = \Phi_S$), there are positive, negative, and zero solutions to (32) if

$$k_d < \sqrt{k\beta\Phi}. \quad (33)$$

For the parameter choices of Table 2, $k_d < 10^{-10}$ s^{-1} is required, corresponding to a horizontal gyre of less than 10 Sv working on the full meridional salinity difference. For stronger diffusion, no valid nonzero solutions to (32) exist, leaving only the zero flow solution. Intersection of the three solutions occurs at $k_d = \sqrt{k\beta\Phi}$. Numerical integration finds the zero flow solution to be unstable for $k_d < \sqrt{k\beta\Phi}$ and stable for $k_d > \sqrt{k\beta\Phi}$, consistent with the numerical results of section 3b.

Under asymmetric forcing, we always find a stable strong flow solution with sinking in the hemisphere
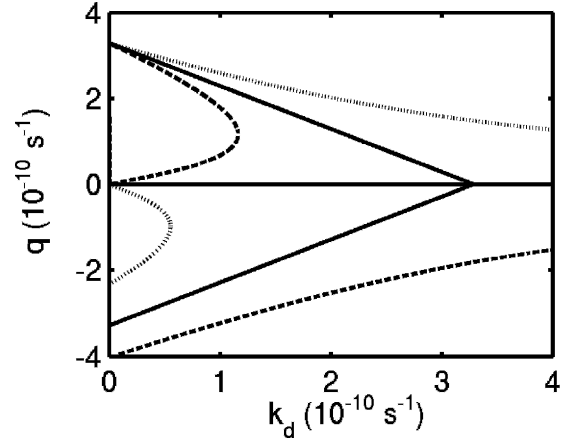

FIG. 6. Analytical flow solutions to the Rooth model as a function of diffusion. Temperature is symmetric about the equator. Solid line: case I; dotted line: case II; dashed line: case III.

with weaker freshwater forcing (northern for case II and southern for case III, Fig. 6). These solutions approach zero as diffusion tends to infinity, equivalent to the quasi-no-flow solutions of section 3b. For weak enough k_d we find, in addition, two steady states with sinking in the hemisphere with stronger freshwater forcing. Of these two solutions the stronger, stable, solution weakens with diffusion while the weaker, unstable, solution strengthens; both converge at a critical diffusion strength above which neither is found. Qualitatively, the fixed temperature analysis is in agreement with Fig. 5; quasi-no-flow and stable no-flow solutions at high diffusion, and collapse of northern (southern) sinking in case III (II) with increased diffusion.

The appendix shows that multiple equilibria exist in the Rooth model under the complementary conditions:

$\overline{q} > 0$:

$$0 < k_dk\beta(\Phi_N - \Phi_S)$$

$$< \frac{2}{27}\{(k_d^2 + 3k\beta\Phi_S)^{3/2} + k_d^3 - 9k_dk\beta\Phi_S\}, \quad (34)$$

$\overline{q} < 0$:

$$0 < k_dk\beta(\Phi_S - \Phi_N)$$

$$< \frac{2}{27}\{(k_d^2 + 3k\beta\Phi_N)^{3/2} + k_d^3 - 9k_dk\beta\Phi_N\}. \quad (35)$$

Additional equilibria, if they exist, exhibit sinking in the hemisphere with stronger freshwater forcing. Multiple equilibria are only possible below an upper limit of freshwater-forcing asymmetry, which is decreased by diffusion. A saddle-node bifurcation occurs at a critical $\Phi_N - \Phi_S$, where the stable and unstable solutions with

sinking in the dominant freshwater-forcing hemisphere meet. However, as shown in section 3e, a Hopf bifurcation may occur before the saddle node but this is not the case in Fig. 6.

d. Feedback analysis

To understand which processes stabilize or destabilize the steady states depicted in Fig. 6, we apply the feedback analysis detailed in Marotzke (1996). Temperatures are fixed and symmetric about the equator. Now,

$$q = k\beta(S_1 - S_3). \tag{36}$$

Let $S_1 = \bar{S}_1 + S_1'$ (likewise for S_3) and $q = \bar{q} + q'$, to obtain, by (36)

$$\dot{q}' = k\beta(\dot{S}_1' - \dot{S}_3'). \tag{37}$$

For $\bar{q} > 0$, (22) and (24) give, with $S_1' \ll \bar{S}_1$, $S_3' \ll \bar{S}_3$, $q' \ll \bar{q}$,

$$\frac{\dot{q}'}{k\beta} = \bar{q}(-2S_1' + S_2' + S_3') + q'(-2\bar{S}_1 + \bar{S}_2 + \bar{S}_3)$$
$$+ k_d(S_3' - S_1'). \tag{38}$$

With $q' = k\beta(S_1' - S_3')$, by salt conservation, $S_2' = -S_1' - S_3'$, and expressing the steady-state salinity differences in terms of the freshwater forcing, using (26) and (28) gives

$$\frac{\dot{q}'}{k\beta} = -3\bar{q}S_1' + q'\left(\frac{\Phi_N(1 + k_d/(k_d + \bar{q})) - \Phi_S}{(k_d + \bar{q})} - \frac{k_d}{k\beta}\right). \tag{39}$$

Linearizing the tendency equation (22) for S_1 and combining with (39) gives the system

$$\begin{pmatrix} \dot{S}_1' \\ \dot{q}' \end{pmatrix} = \mathbf{A} \begin{pmatrix} S_1' \\ q' \end{pmatrix} \tag{40}$$

$$\mathbf{A} \equiv \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \equiv \begin{pmatrix} -3(\bar{q} + k_d) & \left(\dfrac{(\bar{q} + k_d)}{k\beta} + \dfrac{\Phi_N}{(\bar{q} + k_d)}\right) \\ -3\bar{q}k\beta & \left(k\beta\dfrac{(\bar{q} + k_d)(\Phi_N - \Phi_S) + k_d\Phi_N}{(\bar{q} + k_d)^2} - k_d\right) \end{pmatrix}. \tag{41}$$

If $\bar{q} < 0$, the equivalent of (40) and (41) follows accordingly.

Element A_{11} in (41) shows the negative (stabilizing) mean flow feedback (Marotzke 1996) aided by diffusion, the tendency in S_1' is of opposite sign to the perturbation, due to this term. The off-diagonal terms likewise stabilize, as is seen from the odd number of negative couplings within the feedback loop between S_1' and q' (Kump et al. 1999). The coupling from S_1' to q', A_{21}, is negative and that from q' to S_1', A_{12}, is positive. Element A_{22} (and hence its associated feedback of q' directly upon itself, called here the q'-q' feedback) in (41) is of undetermined sign; its contributions are now analyzed in detail (Fig. 7).

Regardless of freshwater forcing, for nonzero flow solutions, the two terms in A_{22} work against each other (dashed and dotted lines in Fig. 7). The combined feedback for the strong northern sinking solution behaves as the salinity transport feedback in the nondiffusive model of Scott et al. (1999), neutral in case I, stabilizing in case II, and destabilizing in case III. The zero-flow solution has the q'-q' feedback undergo a transition from destabilizing to stabilizing at large enough diffusion.

The "pure diffusion" contribution to the q'-q' feedback is stabilizing (dotted lines in Fig. 7); therefore, changes to the q'-q' feedback stabilize if this dominates over the destabilizing change to the salinity transport feedback. With symmetric freshwater forcing, the salinity transport feedback increases from zero at the same rate as the diffusion feedback becomes negative (Fig. 7a) such that the q'-q' feedback is zero regardless of diffusion. This can also be seen by setting $\Phi_N = \Phi_S$ in A_{22} in (41) and applying the condition (A5) from the appendix for steady-state flow under symmetric forcing. Decrease of flow strength with diffusion (Fig. 6) weakens the mean flow feedback in (41) and thus destabilizes; by symmetry this is also true for the southern sinking solution, while the zero flow solution is stabilized (Fig. 7d). Diffusion significantly reduces the large positive salinity transport feedback of the latter. The net feedback acting on a flow anomaly (salinity transport and diffusive) becomes negative (i.e., the no-flow solution is stable) at the diffusion strength that eliminates the strong flow solutions.

Physically, a negative flow perturbation to high-latitude sinking freshens each high-latitude box. De-
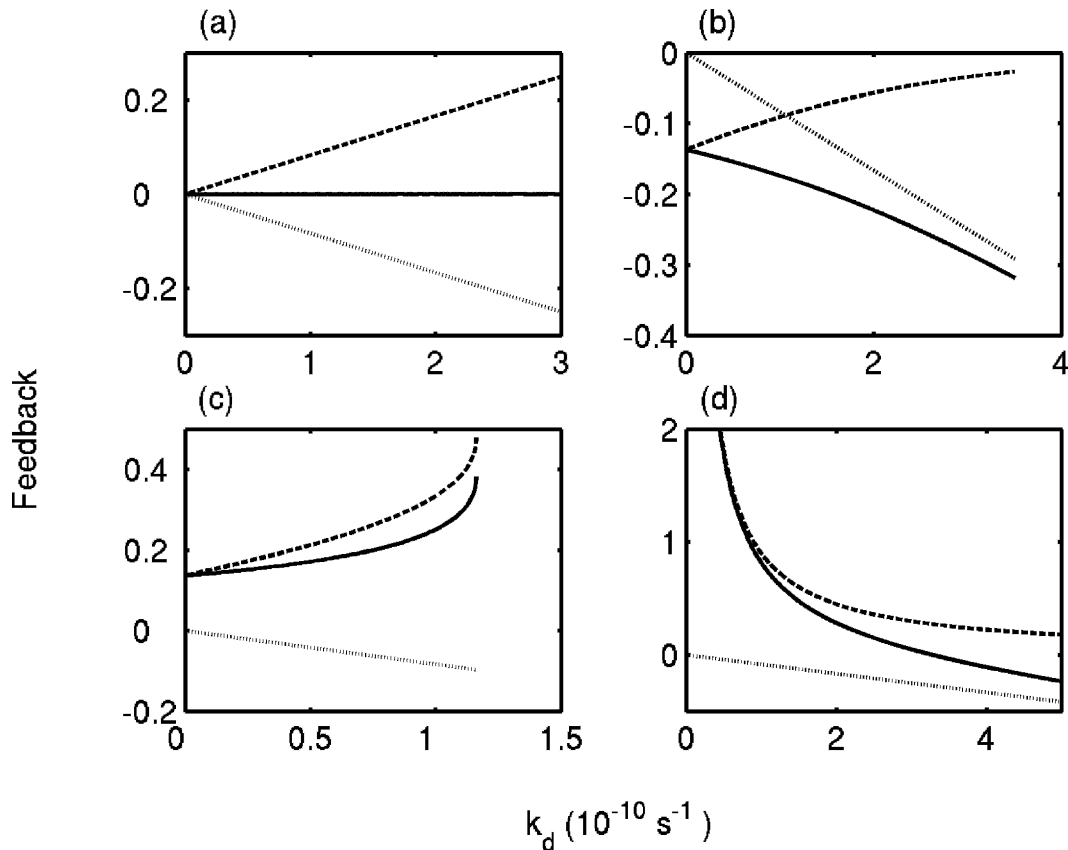
FIG. 7. Dependence of element A_{22} of (41) to diffusion in the Rooth model. (a) Case I: strong positive flow; (b) case II: strong positive flow; (c) case III: strong positive flow; (d) case I: zero flow solution. Solid: total of A_{22}; dashed: first term; dotted: second term.

pending on the steady-state salinity differences, this freshening is stronger in either the upwelling or the sinking hemisphere, or equal in both. In the latter case, the feedback is neutral. If the freshening is stronger in the upwelling hemisphere, the flow is reinvigorated, and the feedback is negative. If the freshening by reduced flow is stronger in the sinking hemisphere, we find positive feedback. With northern sinking, the criterion for neutral feedback is, by (26) and (28), that

$$\frac{\overline{S}_2 - \overline{S}_1}{\overline{S}_1 - \overline{S}_3} = \frac{(\overline{q} + k_d)\Phi_N}{(\overline{q} + k_d)\Phi_S - k_d\Phi_N} = 1. \quad (42)$$

Without diffusion, neutrality ensues under symmetric freshwater forcing. Diffusion makes the salinity advection feedback less stabilizing; (42) shows that neutrality requires $\Phi_N < \Phi_S$. This occurs because diffusion reduces the steady-state salinity differences. At equilibrium, the salinity difference in the upwelling hemisphere is larger than that of the sinking hemisphere, and so is decreased more by diffusion. This implies that

the pole-to-pole salinity gradient is decreased relative to the salinity difference in the sinking hemisphere. If the latter is already greater, implying positive salinity transport feedback, the strength of the feedback is thus increased. If the salinity transport feedback is negative, its magnitude decreases. In either case, destabilizing ensues. The same analysis with symmetric freshwater forcing gives a positive salinity transport feedback, in contrast to a neutral feedback with zero diffusion.

e. Hopf bifurcations

Above, Fig. 6 shows the dependence of the flow on a continuous change in k_d, for select cases of freshwater forcing. Now we continuously vary the asymmetry in the freshwater forcing, for some select cases of k_d. We are particularly interested in the linear stability of the steady states, which may be determined from the eigenvalues of the coefficient matrix of (41). The eigenvalues are complex over the parameter ranges considered; therefore, stable solutions result if real parts are negative, which is true if

$$3(\bar{q} + k_d)^3 + k_d(\bar{q} + k_d)^2 - k\beta(\Phi_N - \Phi_S)(\bar{q} + k_d)$$
$$- k_d k\beta\Phi_N > 0; \quad \bar{q} > 0, \tag{43}$$

$$3(k_d - \bar{q})^3 + k_d(k_d - \bar{q})^2 + k\beta(\Phi_N - \Phi_S)(k_d - \bar{q})$$
$$- k_d k\beta\Phi_S > 0; \quad \bar{q} < 0. \tag{44}$$

This can be seen from noting that the sum of the eigenvalues equals trace(**A**), the sum of the diagonal elements.

In a phase diagram showing steady-state flow strength against Northern Hemisphere freshwater forcing, Scott et al. (1999) had speculatively connected the stable equilibrium branches (northern and southern sinking) across the zero-flow line by a curve denoting unstable quasi equilibria ("quasi" because they could not even be in steady state for symmetric forcing). By adding even a small amount of diffusion, we can here make rigorous the connection in phase space between the northern sinking and southern sinking equilibria, and present a comprehensive bifurcation diagram (Fig. 8).

With weak diffusion, three equilibria are found for a relatively wide range of Φ_N. Above some critical value of Φ_N (here, about a tenfold asymmetry, Fig. 8), however, a saddle node bifurcation occurs and the northern sinking is no longer a solution. As expected, the weaker of the northern sinking solutions is unstable. In contrast to the Stommel model, however, the strong sinking solution is not always stable, undergoing a subcritical Hopf bifurcation. Stronger diffusion destabilizes the northern sinking further, for $\Phi_N > \Phi_S$, and the range of multiple equilibria shrinks. One consequence is that the Hopf bifurcation does not occur any more. The blue and the red curves in Fig. 8 are equivalent to gyre strengths of 10 Sv and 1 Sv, respectively, if operating on the full meridional salinity difference. They represent two different ways of destabilizing North Atlantic Deep Water formation under increased Northern Hemisphere freshwater flux forcing, which could happen with or without a phase of THC oscillations (e.g., Rahmstorf 1995). Even stronger diffusion means that only the hemisphere with the weaker freshwater flux shows sinking, consistent with the notion that diffusion suppresses the "weaker" equilibrium and supports a unique solution. Note that for all nonzero values of diffusion, the curves meet at $q = 0$ for symmetric freshwater flux forcing.

4. Discussion and conclusions

A diffusive parameterization of gyre circulations in the Stommel and Rooth box models of the THC mod-
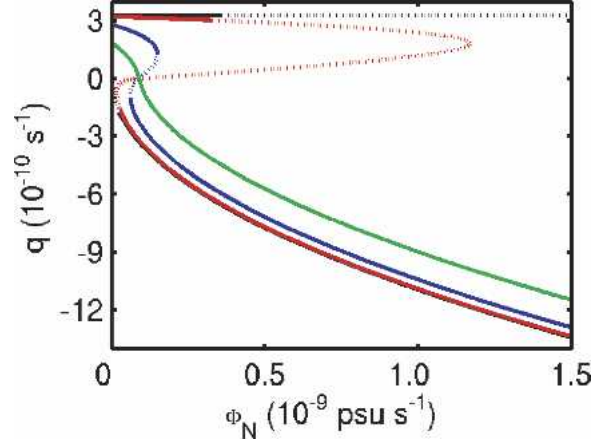

FIG. 8. Analytical flow solutions to the Rooth model as a function of Φ_N, $\Phi_S = 0.9 \times 10^{-10}$ psu s^{-1} (fixed). Temperature is symmetric about the equator. Black: $k_d = 0$; red: $k_d = 1 \times 10^{-11}$ s^{-1}; blue: $k_d = 1 \times 10^{-10}$ s^{-1}; green: $k_d = 4 \times 10^{-10}$ s^{-1}. Solid lines are stable solutions, dotted lines unstable; stability is determined by (43), (44), or numerical integration.

ifies the existence and stability of multiple flow solutions. We present a quantitative explanation of the results in National Research Council (2002) regarding the Stommel model. We argue that our results provide a conceptual foundation for analyzing complex numerical models and even the real world: Is the climate system best characterized by a THC interacting with "strong" or with "weak" gyre salinity transports? Only in the latter case would the THC be capable of undergoing abrupt change, with concomitant potential permanent change caused by temporary perturbations (e.g., hysteresis; Stocker and Wright 1991).

The nondiffusive Rooth model had not been amenable to bifurcation analysis since symmetric or nearly symmetric equilibria did not exist. In contrast, the diffusive Rooth model can be comprehensively analyzed for its bifurcation structure, under appropriate simplifications. These are fixed and symmetric temperatures—a reasonable approximation as salinity rather than temperature gradient changes dominate flow strength response to diffusion.

We have only analyzed uncoupled models of the THC here. However, previous experience with coupled box models suggests that atmospheric coupling does not change the basic bifurcation structures of either the Stommel or the Rooth model (Marotzke 1996; Scott et al. 1999). This statement must remain conjectural at this stage but forms the working hypothesis for this paper.

With asymmetric freshwater forcing the bifurcation structure of the diffusive Rooth model is qualitatively similar to that of the Stommel model, albeit with a very different dynamical interpretation. In the right range of

diffusion strength and asymmetry in freshwater forcing we find two equilibrium solutions with sinking in the hemisphere with stronger freshwater forcing, the stronger of which is stable while the weaker is unstable. We always find one stable equilibrium with sinking in the hemisphere with weaker freshwater forcing. In contrast to the Stommel model, the stable solutions have roughly the same flow strength, which is a better match with results from numerical models. In contrast to the nondiffusive Rooth model, all equilibria of the diffusive Rooth model depend on the freshwater forcing in both hemispheres.

With relatively weak gyre transport, an increase in the freshwater forcing in the sinking hemisphere leads to a subcritical Hopf bifurcation—oscillations set in with a growing amplitude, and the THC eventually switches to sinking in the opposite hemisphere (see also Scott et al. 1999; Titz et al. 2002a). Under stronger gyre transport, such a transition occurs via a saddle-node bifurcation, as in the Stommel model, without going through an oscillatory phase first. Thus, THC collapse could sometimes be accompanied by oscillations, sometimes not.

We propose that, where the Stommel bifurcation structure has previously been used to describe GCM hysteresis (e.g., Rahmstorf 1995), diffusive Rooth dynamics provide a more relevant model. The preference arises since expressing cross-hemispheric Atlantic THC dynamics through the intrahemispheric density gradient is problematical. While two back-to-back Stommel models have been used (Welander 1986), the density difference between high northern and southern latitudes is more appropriate as the driver determining flow strength (Stocker et al. 1992; Hughes and Weaver 1994; Rahmstorf 1996; Klinger and Marotzke 1999; Marotzke 2000; Thorpe et al. 2001). Rahmstorf (1996) found similarity between GCM THC hysteresis and that of the Rooth model but required an implied equatorward Southern Hemisphere atmospheric water vapor transport; that is, the gyre transport of freshwater exceeded that of the atmosphere (also employed in Titz et al. 2002b). This is inconsistent with the diffusive Rooth model since it would require Southern Hemisphere salinity differences in excess of 8 psu.

In the diffusive Rooth model, flow weakens with increased freshwater forcing in the sinking hemisphere. This stands in marked contrast to the nondiffusive Rooth model in which flow strength is a function of the upwelling hemisphere freshwater forcing only and increases with it (Rahmstorf 1996; Scott et al. 1999). Only weak evidence for the latter was found in a GCM study (Wang et al. 1999); the diffusive Rooth model may well be more realistic in suggesting a more even role of

Northern and Southern Hemisphere freshwater forcing.

Gyre circulations in the Stommel model stabilize high-latitude sinking, an expected consequence of poleward salt flux. In the diffusive Rooth model, however, only the sinking solution in the hemisphere with weaker freshwater forcing can be stabilized by diffusion. That diffusion destabilizes a steady state is only seemingly paradoxical. Strong diffusion eliminates multiple equilibria because the system behavior is more nearly linear. This change comes at the expense of the "weaker" equilibrium, the one with sinking in the hemisphere of stronger freshwater forcing.

The values of mixing and freshwater forcing at which qualitative change in behavior occurs in the Rooth model correspond to the range of the values used in and obtained with GCMs, suggesting that this analysis can provide a conceptual foundation for analyzing the stability of the interhemispheric THC. Notice that the heading "mixing" here subsumes a multitude of processes such as gyre transport, Ekman transport, isopycnic mixing, eddy-induced transport, or diapycnic mixing (i.e., anything that leads to a reduction in meridional salinity gradient). The potential for the THC to undergo an abrupt change is dependent on whether THC dynamics are characterized by strong or by weak mixing. It is possible that relatively subtle changes in a GCM decide which characterization applies.

APPENDIX

Bifurcation Structure of the Diffusive Rooth Model

The subsequent treatment (following that in Strogatz 1994, his section 3.6) shows that the Rooth model with diffusion and fixed temperatures symmetric about the equator can be characterized by an "imperfect bifurcation" structure. We analyze the steady-state condition (32) for the flow with fixed symmetric temperatures, $T_1 = T_3$:

$$\bar{q}^3 \pm 2k_d\bar{q}^2 + \bar{q}(k_d^2 - k\beta\Phi_{S/N}) + k_dk\beta(\Phi_N - \Phi_S) = 0.$$

(A1)

To account for asymmetry in the system introduced by $\Phi_N \neq \Phi_S$ we define an "imperfection parameter" (i.e.,

marking deviation from perfect symmetry: Strogatz 1994):

$$h \equiv k_d k \beta (\Phi_N - \Phi_S), \tag{A2}$$

and the abbreviation

$$y_\pm(\bar{q}) \equiv \bar{q}^3 \pm 2k_d \bar{q}^2 + \bar{q}(k_d^2 - k\beta\Phi_{S/N}), \tag{A3}$$

where the upper sign holds for positive argument, $\bar{q} > 0$, and the lower sign for negative argument, $\bar{q} < 0$. Hence, we can rewrite (A1) as

$$y_\pm(\bar{q}) + h = 0. \tag{A4}$$

With symmetric freshwater forcing, that is, $h = 0$, (A1) reduces to

$$\bar{q}[(\bar{q} \pm k_d)^2 - k\beta\Phi] = 0 \tag{A5}$$

with the solutions

$$\bar{q} = 0; \quad \bar{q} = -k_d + \sqrt{k\beta\Phi}; \quad \bar{q} = k_d - \sqrt{k\beta\Phi}, \tag{A6}$$

provided that

$$k_d < \sqrt{k\beta\Phi}. \tag{A7}$$

Condition (A7) defines the maximum diffusion strength for given Φ to support three flow solutions. For stronger diffusion, no valid nonzero solutions to (A5) exist, leaving only the zero flow solution. If (A7) is met, there are positive, negative, and zero solutions to (A5); hence three solutions are found as shown in Fig. 6 (solid lines). Intersection of the three solutions occurs at $k_d = \sqrt{k\beta\Phi}$, which is a cusp point with codimension-2 bifurcation; both k_d and Φ must be tuned to attain this point (Strogatz 1994, p. 70). Numerical integration finds the zero flow solution to be unstable for $k_d < \sqrt{k\beta\Phi}$ and stable for $k_d > \sqrt{k\beta\Phi}$, consistent with the numerical results of section 3b.

We now investigate the properties of the cubics, $y_\pm(\bar{q})$, (A3), to determine the range of h under which multiple solutions to (A1) are found. Since $y_\pm(\bar{q} \to \pm\infty) \to \pm\infty$; $y_\pm(0) = 0$, at least one intersection exists between y and $-h$, for any choice of h. For $h > 0$, $\Phi_N > \Phi_S$, this intersection occurs for $\bar{q} < 0$. The converse is true for $h < 0$. This means that one solution is guaranteed to exist, with sinking in the hemisphere with the weaker freshwater flux.

A necessary condition for multiple intersections is that $y(\bar{q})$ slopes downward at the origin. If that happens, $y_+(\bar{q})$ has a minimum and $y_-(\bar{q})$ has a maximum. For $h > 0$, $\Phi_N > \Phi_S$, there *could* then be two additional solutions to (A1), upper sign, with $\bar{q} > 0$. Analogously, for $h < 0$, $\Phi_N < \Phi_S$, there *could* be two additional solutions to (A1), lower sign, with $\bar{q} < 0$. Whether these additional solutions *do exist* depends on whether the

minimum or maximum is pronounced enough so that $y(\bar{q})$ intersects $-h$.

One finds for the derivative of $y_\pm(\bar{q})$, from (A3):

$$y'_\pm(\bar{q}) = 3\bar{q}^2 \pm 4k_d\bar{q} + (k_d^2 - k\beta\Phi_{S/N}), \tag{A8}$$

$$y'_\pm(0) = k_d^2 - k\beta\Phi_{S/N}. \tag{A9}$$

Thus, $y_+(\bar{q})$ has a minimum, at some $q_1 > 0$, if

$$k_d < \sqrt{k\beta\Phi_S} \tag{A10}$$

and $y_-(\bar{q})$ has a maximum, at some $\bar{q}_2 < 0$, if

$$k_d < \sqrt{k\beta\Phi_N}. \tag{A11}$$

The additional equilibria can only exist with sinking in the hemisphere with stronger freshwater flux forcing. Thus, of the two criteria, (A10) and (A11), the more stringent one defines the upper bound for k_d. We conclude that asymmetric freshwater flux forcing reduces the maximum strength in gyre transport under which multiple equilibria are possible in the Rooth model.

The location of the extrema in $y(\bar{q})$ are found from (A8) as

$$\bar{q}_1 = \frac{1}{3}\left(-2k_d + \sqrt{k_d^2 + 3k\beta\Phi_S}\right), \tag{A12}$$

which is positive if (A10) holds, and

$$\bar{q}_2 = \frac{1}{3}\left(2k_d - \sqrt{k_d^2 + 3k\beta\Phi_N}\right), \tag{A13}$$

which is negative if (A11) holds. Insertion into (A3) shows that multiple equilibria exist in the Rooth model under the complementary conditions:

$$\bar{q} > 0:$$
$$0 < k_d k \beta(\Phi_N - \Phi_S)$$
$$< \frac{2}{27}\{(k_d^2 + 3k\beta\Phi_S)^{3/2} + k_d^3 - 9k_d k\beta\Phi_S\}, \tag{A14}$$

$$\bar{q} < 0:$$
$$0 < k_d k \beta(\Phi_S - \Phi_N)$$
$$< \frac{2}{27}\{(k_d^2 + 3k\beta\Phi_N)^{3/2} + k_d^3 - 9k_d k\beta\Phi_N\}, \tag{A15}$$

which are (34) and (35) in the main text.

Insertion of the necessary conditions (A10) and (A11) into (A14) and (A15), respectively, shows that attainment of the bounds (A10) and (A11) implies $\Phi_N = \Phi_S$. Conversely, asymmetric freshwater forcing implies further reduced maximum k_d for which multiple equilibria are possible. Multiple equilibria are therefore only possible below an upper limit of freshwater-forcing asymmetry, which is decreased by diffusion. A

saddle-node bifurcation occurs at a critical $\Phi_N - \Phi_S$, where the stable and unstable solutions with sinking in the dominant freshwater-forcing hemisphere meet. However, as shown in section 3e, a Hopf bifurcation may occur before the saddle-node but this is not the case in Fig. 6.

REFERENCES

Alley, R. B., and Coauthors, 2003: Abrupt climate change. *Science,* **299,** 2005–2010.

Bryan, F., 1986: High-latitude salinity effects and interhemispheric thermohaline circulations. *Nature,* **323,** 301–304.

Clark, P. U., N. G. Pisias, T. F. Stocker, and A. J. Weaver, 2002: The role of the thermohaline circulation in abrupt climate change. *Nature,* **415,** 863–869.

Cubasch, U., and Coauthors, 2001: Projections of future climate change. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 525–582.

Haney, R. L., 1971: Surface thermal boundary condition for ocean circulation models. *J. Phys. Oceanogr.,* **1,** 241–248.

Hughes, T. M. C., and A. J. Weaver, 1994: Multiple equilibria of an asymmetric two-basin model. *J. Phys. Oceanogr.,* **24,** 619–637.

Klinger, B. A., and J. Marotzke, 1999: Behavior of double-hemisphere thermohaline flows in a single basin. *J. Phys. Oceanogr.,* **29,** 382–400.

Knutti, R., T. F. Stocker, F. Joos, and G.-K. Plattner, 2003: Probabilistic climate change projections using neural networks. *Climate Dyn.,* **21,** 257–272.

Kump, L. R., J. F. Kasting, and R. G. Crane, 1999: *The Earth System.* Prentice Hall, 351 pp.

Manabe, S., and R. J. Stouffer, 1993: Century-scale effects of increased atmospheric CO_2 on the ocean–atmosphere system. *Nature,* **364,** 215–218.

——, and ——, 1999: Are two modes of the thermohaline circulation stable? *Tellus,* **51A,** 400–411.

Marotzke, J., 1990: Instabilities and multiple equilibria of the thermohaline circulation. Ph.D. thesis, Berichte aus dem Institut für Meereskunde, 126 pp.

——, 1996: Analysis of thermohaline feedbacks. *Decadal Climate Variability: Dynamics and Predictability,* D. L. T. Anderson and J. Willebrand, Eds., Springer-Verlag, 333–378.

——, 2000: Abrupt climate change and thermohaline circulation: Mechanisms and predictability. *Proc. Natl. Acad. Sci. USA,* **97,** 1347–1350.

——, and J. Willebrand, 1991: Multiple equilibria of the global thermohaline circulation. *J. Phys. Oceanogr.,* **21,** 1372–1385.

——, and P. H. Stone, 1995: Atmospheric transports, the thermohaline circulation, and flux adjustments in a simple coupled model. *J. Phys. Oceanogr.,* **25,** 1350–1364.

National Research Council, 2002: *Abrupt Climate Change: Inevitable Surprises.* National Academy Press, 230 pp.

Oka, A., H. Hasumi, and N. Suginohara, 2001: Stabilization of thermohaline circulation by wind-driven and vertical diffusive salt transport. *Climate Dyn.,* **18,** 71–83.

Pasquero, C., and E. Tziperman, 2004: Effects of a wind-driven gyre on thermohaline circulation variability. *J. Phys. Oceanogr.,* **34,** 805–816.

Prange, M., G. Lohmann, and A. Paul, 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.

——, 1996: On the freshwater forcing and transport of the Atlantic thermohaline circulation. *Climate Dyn.,* **12,** 799–811.

——, 2002: Ocean circulation and climate during the past 120 000 years. *Nature,* **419,** 207–214.

Rooth, C., 1982: Hydrology and ocean circulation. *Progress in Oceanography,* Vol. 11, Pergamon Press, 131–149.

Schiller, A., U. Mikolajewicz, and R. Voss, 1997: The stability of the North Atlantic thermohaline circulation in a coupled ocean–atmosphere general circulation model. *Climate Dyn.,* **13,** 325–347.

Schmittner, A., and A. J. Weaver, 2001: Dependence of multiple climate states on ocean mixing parameters. *Geophys. Res. Lett.,* **28,** 1027–1030.

Scott, J. R., J. Marotzke, and P. H. Stone, 1999: Interhemispheric thermohaline circulation in a coupled box model. *J. Phys. Oceanogr.,* **29,** 351–365.

Shaffer, G., and S. M. Olsen, 2001: Sensitivity of the thermohaline circulation and climate to ocean exchanges in a simple coupled model. *Climate Dyn.,* **17,** 433–444.

Stocker, T. F., 2000: Past and future reorganization in the climate system. *Quat. Sci. Rev.,* **19,** 301–319.

——, and D. G. Wright, 1991: Rapid transitions of the ocean's deep circulation induced by changes in surface water fluxes. *Nature,* **351,** 729–732.

——, and A. Schmittner, 1997: Influence of CO_2 emission rates on the stability of the thermohaline circulation. *Nature,* **388,** 862–865.

——, D. G. Wright, and W. S. Broecker, 1992: The influence of high-latitude surface forcing on the global thermohaline circulation. *Paleoceanography,* **7,** 529–541.

Stommel, H., 1961: Thermohaline convection with two stable regimes of flow. *Tellus,* **13,** 224–230.

Strogatz, S. H., 1994: *Nonlinear Dynamics and Chaos.* Perseus, 498 pp.

Thorpe, R. B., J. M. Gregory, T. C. Johns, R. A. Wood, and J. F. B. Mitchell, 2001: Mechanisms determining the Atlantic thermohaline circulation response to greenhouse gas forcing in a non-flux-adjusted coupled climate model. *J. Climate,* **14,** 3102–3116.

Thual, O., and J. C. McWilliams, 1992: The catastrophe structure of thermohaline convection in a 2-dimensional fluid model and a comparison with low-order box models. *Geophys. Astrophys. Fluid Dyn.,* **64,** 67–95.

Titz, S., T. Kuhlbrodt, and U. Feudel, 2002a: Homoclinic bifurcation in an ocean circulation box model. *Int. J. Bifurcation Chaos,* **12,** 869–875.

——, ——, S. Rahmstorf, and U. Feudel, 2002b: On freshwater-dependent interhemispheric bifurcations in box models of the thermohaline circulation. *Tellus,* **54A,** 89–98.

Wang, X. L., P. H. Stone, and J. Marotzke, 1999: Global thermohaline circulation. Part I: Sensitivity to atmospheric moisture transport. *J. Climate,* **12,** 71–82.

Welander, P., 1986: Thermohaline effects in the ocean circulation and related simple models. *Large-Scale Transport Processes in Oceans and Atmosphere,* J. Willebrand and D. L. T. Anderson, Eds., NATO ASI Series, Vol. C190, D. Reidel, 163–200.

Winton, M., and E. S. Sarachik, 1993: Thermohaline oscillations induced by strong steady salinity forcing of ocean general circulation models. *J. Phys. Oceanogr.,* **23,** 1389–1410.