

# A Southern Ocean Mechanism for the Interhemispheric Coupling and Phasing of the Bipolar Seesaw

ANDREW F. THOMPSON

*California Institute of Technology, Pasadena, California*

SOPHIA K. HINES

*Lamont–Doherty Earth Observatory, Palisades, New York*

JESS F. ADKINS

*California Institute of Technology, Pasadena, California*

(Manuscript received 20 September 2018, in final form 19 April 2019)

## ABSTRACT

The last glacial period is punctuated by abrupt changes in Northern Hemisphere temperatures that are known as Dansgaard–Oeschger (DO) events. A striking and largely unexplained feature of DO events is an interhemispheric asymmetry characterized by cooling in Antarctica during periods of warming in Greenland and vice versa—the bipolar seesaw. Methane-synchronized ice core records indicate that the Southern Hemisphere lags the Northern Hemisphere by approximately 200 years. Here, we propose a mechanism that produces observed features of both the bipolar seesaw and the phasing of DO events. The spatial pattern of sea ice formation and melt in the Southern Ocean imposes a rigid constraint on where water masses are modified: waters are made denser near the coast where ice forms and waters are made lighter farther north where ice melts. This pattern, coupled to the tilt of density surfaces across the Southern Ocean and the stratification of the ocean basins, produces two modes of overturning corresponding to different bipolar seesaw states. We present evolution equations for a simplified ocean model that describes the transient adjustment of the basin stratification, the Southern Ocean surface density distribution, and the overturning strength as the ocean moves between these states in response to perturbations in North Atlantic Deep Water formation, which we take as a proxy for Greenland temperatures. Transitions between different overturning states occur over a multicentennial time scale, which is qualitatively consistent with the observed Southern Hemisphere lag. The volume of deep density layers varies inversely with the overturning strength, leading to significant changes in residence times. Evidence of these dynamics in more realistic circulation models is discussed.

## 1. Introduction

Ice core records from Greenland suggest that Dansgaard–Oeschger (DO) events dominated climate variability in the Northern Hemisphere (NH) throughout the last glacial period (Dansgaard et al. 1993; Blunier and Brook 2001). The life cycle of a DO event is characterized by an abrupt warming that occurs over a few decades, weak cooling over a period of hundreds to thousands of years, and then an abrupt drop back to cool conditions [see reviews by Landais et al. (2015) and Pedro et al. (2018)] (Fig. 1). Throughout the last glacial period, ice cores from

Antarctica show a coupled, but differing response in the Southern Hemisphere (SH). Transitions from cold to warm periods in the north are linked to transitions from warm to cool conditions in the south, a feature known as the bipolar seesaw. Recent analysis of methane records from high-accumulation ice cores in both hemispheres indicate that changes between the two polar regions were not only asymmetric but also asynchronous; the SH lags the NH response by roughly 200 years (Buizert et al. 2015a) (Fig. 1c). This time scale is longer than either the abrupt warming or cooling transitions in the north, but shorter than the full DO cycle, and thus is not a linear response to these changes. The multicentennial time scale suggests a key role for ocean circulation in the interhemispheric propagation of this climate signal (Gordon 1986; Broecker 1991).

---

*Corresponding author:* Andrew F. Thompson, andrewt@caltech.edu

DOI: 10.1175/JCLI-D-18-0621.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) ([www.ametsoc.org/PUBSReuseLicenses](https://www.ametsoc.org/PUBSReuseLicenses)).

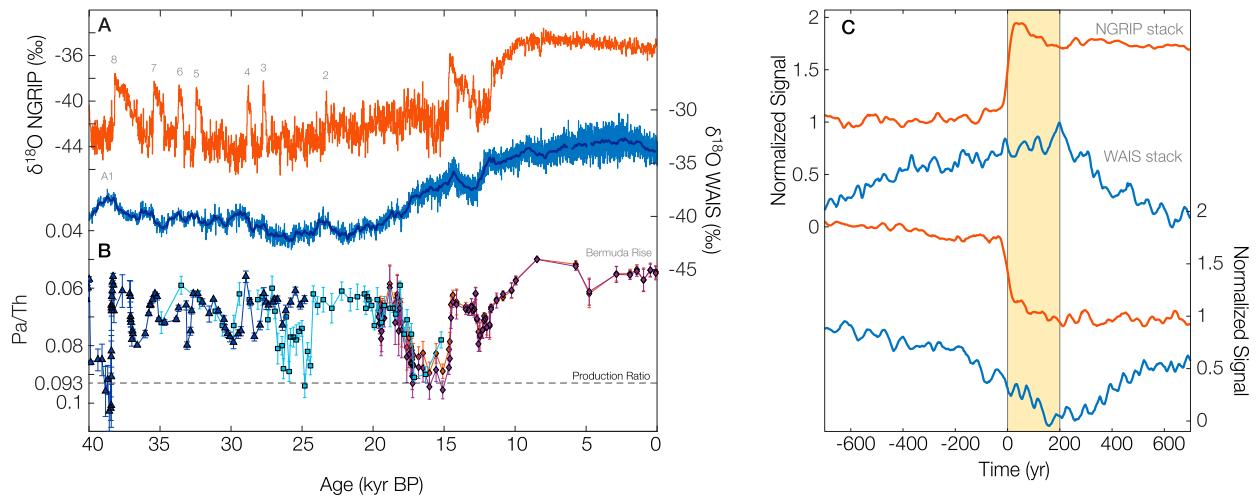


FIG. 1. Climate variability through the end of the last glacial. (a) Ice core  $\delta^{18}\text{O}$  records from NGRIP in Greenland (red, on the GICC05 time scale; Andersen et al. 2004) and the WAIS Divide in Antarctica (blue, on the WD2014 time scale; Buizert et al. 2015a) with DO and Antarctic Isotope Maximum events labeled. (b) Protactinium–thorium records from the Bermuda Rise showing variation in AMOC strength [red and purple diamonds from core OCE326-GGC5, 33°42'N, 57°k35'W, 4550 m (McManus et al. 2004); blue squares from Ocean Drilling Program core 1063, 33°41'N, 57°37'W, 4584 m (Lippold et al. 2009); dark blue triangles from core KNR191-CDH19, 33°41'N, 57°35'W, 4541 m (Henry et al. 2016)]. Values near the production ratio (dashed line) indicate weaker overturning. (c) Stacked records from NGRIP (red) and WAIS (blue) showing north–south timing at NH warming (top of panel) and cooling (bottom of panel) events (Buizert et al. 2015a). The signal is normalized so that NGRIP and WAIS records fit on the same axis; time 0 marks the midpoint of the NH rise or fall. See section 2c for a description of the data processing.

Early theories for the bipolar seesaw focused on two aspects of the ocean's role in the climate system: 1) the transport of heat in the Atlantic Ocean basin (Crowley 1992) and 2) the storage of heat in the abyssal ocean (Stocker and Johnsen 2003). An enhanced northward heat flux in the upper ocean, related to an intensification of North Atlantic Deep Water (NADW) formation, has been proposed as a mechanism that cools the SH (Crowley 1992). However, this theory does not account for the upwelling of NADW along density surfaces in the Southern Ocean (Talley 2013; Armour et al. 2016), nor does it consider coincident changes in the abyssal stratification and circulation. The differing rates of warming and cooling in each hemisphere during the DO life cycle (Fig. 1c) have led to the suggestion that an oceanic “thermal reservoir” could explain smoother transitions in temperature observed in Antarctic ice cores. In other words, heat storage, principally in the Southern Ocean, could filter the more abrupt changes in NADW and delay the onset of Antarctic warming after a northern warming event (Stocker and Johnsen 2003; Barker et al. 2009). In these studies, the time scale associated with this filtering process was empirically, rather than mechanistically, determined. A recent assessment of the heat storage mechanism by Pedro et al. (2018) supports the existence of this reservoir, but proposes that the storage occurs north of the Southern Ocean. Nevertheless, a single time scale neglects

increasing evidence that both the ocean's stratification and overturning strength change across DO events (Lynch-Stieglitz 2017), both of which will modify the efficiency of the ocean's thermal storage and thus the filtering time scale. Any link between NADW formation and Antarctic surface temperatures during DO cycles would imply a change in how heat is distributed within or transported across the Southern Ocean. Yet, the models discussed above do not consider leading-order dynamics that govern the adjustment of the Southern Ocean stratification and overturning: eddy generation, eddy transport, and surface wind and buoyancy forcing (Johnson and Bryden 1989; Marshall and Radko 2003; Lumpkin and Speer 2007; Marshall and Speer 2012; Munday et al. 2013).

The response of the overturning circulation to changes in high-latitude forcing has been explored in idealized but insightful models that focus on water mass transformation, processes that modify the density, or buoyancy,<sup>1</sup> of a fluid parcel. Closure of the overturning circulation requires that the loss of buoyancy at high-latitude sites of deep- and bottom-water formation must be balanced by a gain of buoyancy, either through direct

<sup>1</sup> Throughout, buoyancy  $b$  is defined to be linearly proportional to density:  $b \equiv g(1 - \rho/\rho_0)$ , where  $\rho$  is density,  $\rho_0$  is a reference density, and  $g$  is gravity.

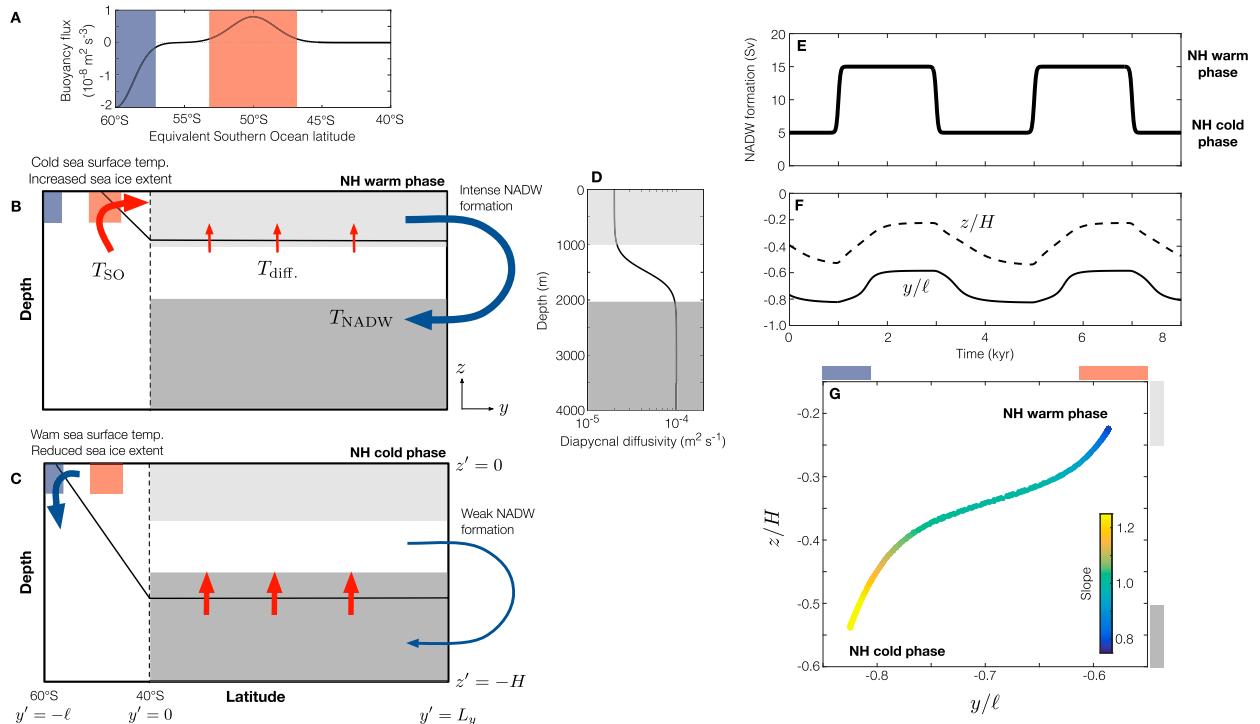


FIG. 2. The DO mechanism for a two-layer (two density class) ocean. (a) Prescribed surface buoyancy flux  $F_b(y)$  in the Southern Ocean; blue (negative) and red (positive) regions correspond to sites of sea ice formation and melt, respectively. (b) Schematic of NH warm phase: NADW formation is strong, diffusive upwelling is weak, and water is transformed from denser to lighter waters in the Southern Ocean. A larger fraction of the surface Southern Ocean is covered with denser, colder water. (c) Schematic of an NH cold phase: NADW formation is weak, the stratification increases at depth, which intensifies diffusive upwelling, and water is transformed from lighter to denser waters in the Southern Ocean. The surface Southern Ocean is exposed to lighter, warmer surface waters. (d) Prescribed diapycnal diffusivity  $\kappa(z)$  in the ocean basins; light- and dark-gray regions indicate weaker and strong mixing, respectively. (e) Time-dependent forcing  $T_{\text{NADW}}$  and (f) the temporal evolution of the interface depth in the basin  $z$  and the outcrop position  $y$  in the Southern Ocean. (g) Phase space of the layer-interface depth  $z$  and Southern Ocean outcrop position  $y$ ; the color indicates the steepness of the isopycnal slope  $s_p = -z/y$ , normalized by  $s_0 = \tau/(\rho_0 K)$ . Blue, red, and gray bars correspond to those appearing in (a) and (d).

surface buoyancy forcing where deep and bottom waters outcrop in the Southern Ocean, or through a vertical transfer of buoyancy into the interior by mixing across density surfaces at lower latitudes (Walín 1982; Lumpkin and Speer 2007). The steady, *global* circulation that is consistent with transporting water from sites of buoyancy loss to buoyancy gain has come to be known as the “residual” circulation (Marshall and Radko 2003). The Southern Ocean plays a unique role in the residual circulation: while the formation of NADW transforms water from lighter to heavier density classes and interior mixing transforms water from heavier to lighter density classes (at least in a zonal average), transformation in the Southern Ocean may induce a volume flux to either lighter or heavier density classes. The sense and rate of Southern Ocean transformation depends on both the surface buoyancy flux and the surface distribution of density, or buoyancy (Marshall 1997; Abernathy et al. 2016) [Fig. 2; Eq. (2)]. Furthermore, the spatial distribution of

surface buoyancy forcing in the Southern Ocean is constrained by patterns of sea ice growth and melt (Pellichero et al. 2018), a feature that is key to the new DO mechanism presented here.

Previous idealized circulation studies, often simplifying the ocean to a zonally averaged, two-layer system, have shown that rearrangement of the global stratification and overturning is typically associated with changes in water mass transformation in the Southern Ocean (Gnanadesikan 1999). These studies have also provided insight into how the ocean may occupy multiple steady states (Johnson et al. 2007; Cimattoribus et al. 2014), exhibit hysteresis (Wolfe and Cessi 2015; Hines et al. 2019), and have internal modes of variability on time scales comparable to the DO cycle (Johnson et al. 2007). Internal oscillations that arise in some of these models have been attributed to the advection of salt anomalies between hemispheres in the upper branch of the overturning circulation (Johnson et al. 2007; Wolfe and Cessi

2015). These models provide a well-tested framework for our approach to understanding the ocean's evolution through DO events.

Previous models have shown that the Southern Ocean can transition between a region of positive or negative water mass transformation (net production of lighter or denser waters, respectively) (Gnanadesikan 1999; Shakespeare and Hogg 2012; Marshall and Zanna 2014). However, this behavior has not been coupled to, or constrained by, the surface buoyancy flux distribution. Antarctic sea ice is formed in polynyas near the coast and, through brine rejection, produces denser waters over the continental shelf. Through the combined effect of surface winds and the surface ocean circulation, sea ice is carried equatorward and melts at lower latitudes and into lighter density classes (Abernathey et al. 2016; Haumann et al. 2016). This geographic constraint on the buoyancy flux profile is an emergent property of the climate system with a robust dipole structure even as sea ice extent varies (Ferrari et al. 2014; Jansen 2017). Changes in water mass transformation in the Southern Ocean may arise from modifications to the surface buoyancy flux, displacement of the surface outcrop location of density classes, or both (Nikurashin and Vallis 2011, 2012; Sun et al. 2016; Jansen 2017). Here, we present a model where the surface buoyancy distribution in the Southern Ocean evolves in time, as determined by the surface outcrop position of different density classes. This meridional buoyancy distribution is related to Southern Ocean sea surface temperatures and is taken here as a proxy for temperatures recorded in Antarctic ice cores. This proxy is supported by the recent study by Markle et al. (2017), which showed that the moisture source for the West Antarctic Ice Sheet (WAIS) Divide ice core comes from a band of latitudes with a width of roughly  $20^\circ$  centered at  $50^\circ\text{S}$ , a region that describes the Antarctic Circumpolar Current (ACC).

The recent review of the bipolar seesaw by Pedro et al. (2018) provides a useful summary of the current mechanistic understanding for the interhemispheric signal propagation of DO events and its related time scales. This review is firmly couched in the "heat reservoir" literature; indeed, the purpose of the review is specifically to test the mechanism of Stocker and Johnsen (2003). We instead propose that an understanding of the bipolar seesaw, and in particular the Southern Ocean's role in this climate feature, is better understood by appealing to changes to the ocean's density structure and the overturning circulation, both of which depend on water mass transformation rates, as opposed to ocean heat content. Transitions in overturning circulation states are necessarily tied to changes in water mass transformation, which in the Southern Ocean, depends on eddy processes and

interactions with sea ice that are directly related to the ocean's density structure. In other words, a dynamical understanding of the global ocean circulation state, not just local heat transports, is required to explain DO events. We acknowledge that a full explanation of the bipolar seesaw also requires an assessment of coupling to the atmosphere and cryosphere (Wunsch 2006; Markle et al. 2017; Pedro et al. 2018). Here, we only focus on the ocean component, but we identify a new mechanism, the link between overturning states and Southern Ocean surface densities and temperatures (section 2), that explains multiple observed aspects of the DO phenomenon.

In this study we make use of a simple ocean model that highlights the key dynamics of our proposed bipolar seesaw mechanism. Critical to the model development is a simplified evolution equation for the outcrop position of density surfaces in the Southern Ocean, which, to our knowledge, has not previously been proposed. This, along with a second equation describing the evolution of the stratification in the basin (e.g., Nikurashin and Vallis 2011), allows consideration of the global ocean's transient response to perturbations in surface forcing. In section 2a, we provide a conceptual description of the DO cycle using a zonally averaged, two-layer representation of the ocean. This idealization is designed to facilitate understanding of water mass transformation processes that give rise to the DO cycle. This section is meant to be understood without the detailed descriptions of water mass transformation parameterizations that are presented in section 2b. Note that a two-layer representation of the ocean can only support a single overturning cell. It has, however, been suggested that the ocean may transition between two-cell and single-cell states, the latter circulating through multiple ocean basins (Ferrari et al. 2014; Thompson et al. 2016, hereinafter T16). This behavior is lost when abstracting to a zonal average or to two layers. Therefore, in section 3 we also present simulations with four layers and two separate basins, an Atlantic and an Indo-Pacific, that provide a more complete picture of the restructuring of overturning pathways and changes to deep-ocean residence times, both of which have a significant impact on oceanic tracer distributions. In section 4 we discuss limitations of the model and provide supporting evidence from a comprehensive climate model that key elements of this DO mechanism are active. We conclude with a summary of results in section 5.

## 2. Model description

### a. Conceptual model of bipolar states in a two-layer ocean

To highlight the importance of an evolving surface buoyancy distribution in the Southern Ocean, we begin

by considering a two-layer, zonally symmetric ocean (Gnanadesikan 1999), in which a single interface separates two layers with uniform density.<sup>2</sup> The upper density class loses volume through the formation of NADW in the NH, represented as a volume flux  $T_{\text{NADW}}$ , and gains volume through upwelling from the lower density class as a result of turbulent mixing, typically represented as a diffusive volume flux  $T_{\text{diff}}$ . In steady state, water mass transformation in the Southern Ocean  $T_{\text{SO}}$ , which arises from the surface buoyancy forcing, must balance the residual volume flux when  $T_{\text{NADW}}$  and  $T_{\text{diff}}$  are summed (Figs. 2b,c). In this reduced model, the strength of these transformation processes, and thus the residual overturning, is determined from two properties: the depth  $z$  of the interface outside of the Southern Ocean (assumed to be uniform in the northern basins) and the outcrop position  $y$  of the interface in the Southern Ocean (Figs. 2b,c). The ratio  $-z/y$  represents the slope  $s_p$  of the interface across an ACC-like channel and serves to couple processes occurring in the Southern Ocean with low-latitude interior ocean dynamics.

A limitation of this zonally averaged approach is that it cannot resolve stratification adjustments associated with wave processes. Waves are responsible for the propagation of changes along boundaries (Kelvin waves) and ultimately for the adjustment of the interior of ocean gyres (Rossby waves) (Kawase 1987; Johnson and Marshall 2002). However, both Rossby and Kelvin waves are adiabatic motions—they do not contribute to water mass modification and therefore cannot produce a change in the volume of a given density class. Our model uses a single depth for the low-latitude stratification, which is consistent with observations that density surfaces are essentially flat in the deep ocean at low- and midlatitudes. Strong deviations from these flat isopycnals are not possible without generating strong, deep geostrophically balanced flows, although western boundary currents may be more sensitive to wave-driven changes (Johnson and Marshall 2002). Zonal structure in the low-latitude stratification, especially at the northern boundary of the ACC, may give rise to additional time scales that are not captured in this model. Zonally symmetric transient studies have been employed to study the global response of the overturning circulation in other recent studies—for example, Marshall and Zanna (2014) and Jansen and Nadeau (2019).

To probe both steady-state and transient responses of the ocean to changes in the strength of NADW formation, we introduce evolution equations for  $z$  and  $y$ .

Following the arguments above, the volume of water in the upper layer, and hence the zonally averaged depth of the layer interface in the northern basins, evolve according to

$$\dot{V} = -Az\dot{z} = \sum T_V = (T_{\text{diff}} + T_{\text{SO}} - T_{\text{NADW}}). \quad (1)$$

Here,  $\dot{X} \equiv dX/dt$ , where  $X$  is the indicated variable, and  $V$  and  $A$  are the volume and horizontal area of the upper density class;  $A$  is assumed to be constant here for simplicity, although in the ocean  $A$  is a function of depth, which may also affect the overturning structure (Ferrari et al. 2016; de Lavergne et al. 2017). The volume  $V$  increases (decreases), and thus  $z$  deepens (shoals), if the net volume flux  $\sum T_V$  into the upper layer is positive (negative). The different components of the total volume flux into and out of the upper density layer are shown schematically in Figs. 2b and 2c. The diffusive flux requires a representation of the strength of diffusive mixing with depth, typically given as a vertically varying eddy diffusivity  $\kappa(z)$ . We prescribe a vertical profile for  $\kappa$  in section 2b that is bottom intensified, which is guided by direct microstructure observations (Polzin et al. 1997; Waterhouse et al. 2014) (Fig. 2d). Last,  $T_{\text{SO}}$  is determined from a combination of the surface buoyancy flux and the surface buoyancy distribution, both determined from the single outcrop position in the ACC (Fig. 2a). The buoyancy flux distribution is prescribed, but the outcrop position may evolve in time. The dependence of  $T_{\text{diff}}$  and  $T_{\text{SO}}$  on  $z$  and  $y$  have been described in previous studies (Gnanadesikan 1999; Nikurashin and Vallis 2011, T16), but are given for completeness in section 2b. Systems that represent the stratification in two or more separate basins, for example, an Atlantic and an Indo-Pacific, require an additional contribution on the right hand side of Eq. (1),  $T_{\text{zonal}}$ , that accounts for the adiabatic convergence and divergence of mass due to zonal flow in the ACC [see Eq. (11)].

The outcrop position of the interface  $y$  adjusts as a result of three physical processes: the Southern Ocean surface wind stress, transport by mesoscale eddies, and the surface buoyancy flux. This evolution may be represented as

$$\dot{y} = \dot{y}_{\text{wind}} + \dot{y}_{\text{eddy}} + \dot{y}_{\text{buoy}} = \frac{1}{h_m} \left( -\frac{\tau}{\rho_0 f} + K_S s_p - \frac{F_b}{\partial_y b_s} \right), \quad (2)$$

where  $h_m$  is the mixed layer in the Southern Ocean, assumed to be uniform in space and time for simplicity, and  $y'$  is used to distinguish the meridional coordinate from the outcrop position  $y$ . Equation (2) is a transient version of the steady residual-mean balance presented

<sup>2</sup> The interface can also be considered as a single density surface separating a range of densities that are lighter above and heavier below; see further discussion in T16.

by Marshall and Radko (2003). The three processes that modify  $y$  are parameterized on the right-hand side of Eq. (2). Displacement outcrop positions by Ekman transport depend on the surface wind stress  $\tau$  ( $\text{N m}^{-2}$ ) and the Coriolis frequency  $f$  ( $\text{s}^{-1}$ ). Relaxation by eddies occurs through a Gent and McWilliams (1990) parameterization:  $K$  is the isopycnal eddy diffusivity ( $\text{m}^2 \text{s}^{-1}$ ), and  $s_\rho = -z/y$  (negative; dimensionless) is the isopycnal slope at the base of the mixed layer. The surface buoyancy distribution can change in response to surface forcing:  $F_b$  ( $\text{m}^2 \text{s}^{-3}$ ) and  $\partial_y b_s$  ( $\text{s}^{-2}$ ) are the surface buoyancy flux and the meridional surface buoyancy gradient, respectively, with the latter assumed to be positive. A westerly wind stress (assumed to be uniform here) advects the outcrop position equatorward ( $\dot{y} > 0$ ) in a surface Ekman layer and tends to increase the slope  $s_\rho$  of the interface spanning the ACC. Eddies, generated through baroclinic instability, counter the tilting of  $s_\rho$  and advect the outcrop position poleward ( $\dot{y} < 0$ ). A positive surface buoyancy flux, corresponding to a supply of buoyancy to the ocean, leads to  $\dot{y} < 0$ ; a negative surface buoyancy flux extracts buoyancy from the ocean and leads to  $\dot{y} > 0$ . Steady-state conditions ( $\dot{y} = \dot{z} = 0$ ) occur when the wind and eddy contributions to  $\dot{y}$  are balanced by the buoyancy flux term; this is the steady, residual circulation described by Marshall and Radko (2003). Only the last term in Eq. (2) exchanges volume between density classes since the displacement of the outcrop position by winds and eddies is adiabatic. The introduction of Eq. (2) allows for a mechanistic representation of the temporal evolution of the steeply tilted density surfaces in the ACC that ventilate the deep ocean (Toggweiler and Samuels 1995) and is key to our description of the bipolar seesaw and the 200-yr lag that characterize DO events.

This two-layer ocean produces two end-member behaviors that are consistent with observed bipolar seesaw characteristics (Figs. 2b,c). Periods of intense NADW formation are linked to warm conditions in the NH as warm, near-surface waters (i.e., upper layer) are drawn toward Arctic latitudes. In the modern ocean, NADW sinks to a depth of roughly 3000 m and makes up a significant portion of the weakly stratified abyss (i.e., the lower layer). The removal of mass from the upper layer and its injection in the lower layer shoals the interface  $z$  at low latitudes, typically to depths where  $\kappa$  is reduced, weakening diffusive transformation. Thus, to close the overturning, a net transformation from heavier (deeper) to lighter (shallower) density classes must occur in the Southern Ocean, which requires that the interface outcrops in a region where the surface buoyancy flux is positive,  $F_b > 0$ , or within a region of sea ice melt (Fig. 2b). During this NH warm phase, the interface

outcrops farther to the north of the ACC channel, such that warmer upper-ocean waters occupy a relatively small fraction of the Southern Ocean. We propose that this would lead to a colder mean Southern Ocean surface temperature, increased sea ice extent, and a cooler Antarctic climate (Fig. 2b). If the formation of NADW weakens, less NADW is injected into the deep ocean and the volume of the upper ocean increases as a result of a transient convergence of mass; both of these processes cause the interface  $z$  to deepen (Johnson et al. 2007). This lowering of the interface enhances low-latitude diffusive transformation and deep-water formation must occur at both poles to compensate. Southern deep-water formation, in this model, requires that the outcrop position  $y$  now reside in a region where  $F_b < 0$ , or within a region of sea ice formation (Fig. 2c). During this NH cold phase, the interface outcrops near the Antarctic coast, such that warmer upper-ocean waters occupy a larger fraction of the Southern Ocean, which we propose would lead to a warmer mean Southern Ocean surface temperature, reduced sea ice extent, and a warmer Antarctic climate (Fig. 2c). Both states predict oppositely signed surface temperature anomalies in the NH and SH—a key feature of the bipolar seesaw.

### b. Detailed model description

In practice, to integrate Eqs. (1) and (2) forward in time, the water mass transformation rates  $T_{\text{diff}}$ ,  $T_{\text{SO}}$ , and  $T_{\text{zonal}}$  need to be determined. We build on the steady-state model described in T16 and Hines et al. (2019), which follows similar models by Gnanadesikan (1999); Shakespeare and Hogg (2012); Marshall and Zanna (2014), among others. Introducing additional layers leads to more interfaces that follow the same evolution equations. Below, we show results from a four-layer model that has a similar architecture to T16, roughly representing Intermediate Water, Upper Circumpolar Deep Water, Lower Circumpolar Deep Water, and Bottom Water. The density, or buoyancy, of each layer is fixed. As in T16, different basins, representing the Atlantic and Indo-Pacific, can be included; in these configurations, the zonally averaged stratification, equivalent to the thickness of the layers, is solved in each basin. In all cases, the depth–latitude plane is divided into a northern basin region, where isopycnals are horizontal, and an ACC region, where isopycnals slope and outcrop at the surface. In the ACC region, mass may be transferred zonally and adiabatically between Atlantic and Indo-Pacific basins, when included.

The model is forced at the surface of the Southern Ocean by an imposed wind stress,  $\tau = \text{constant} = 0.1 \text{ N m}^{-2}$  in this study, as well as an imposed surface buoyancy flux. The other external forcing parameter is an imposed formation rate of NADW in the Atlantic basin  $T_{\text{NADW}}$ ; model

TABLE 1. Parameters for the isopycnal box model used to generate water mass transformation rates described in section 2b and used to produce the solutions in Figs. 4–7. The parameter  $T_{\text{NADW}}$  is taken as the primary forcing parameter; its temporal variations are described in section 2b. Exploration and discussion of the model sensitivity to these parameters appear in T16 and Hines et al. (2019). The temperature and buoyancy values  $T_1, \dots, T_4$  and  $b_1, \dots, b_4$  are representative of modern values.

Parameter	Value	Description
$L_y$	$1 \times 10^7$ m	Meridional basin length
$\ell$	$2 \times 10^6$ m	Meridional ACC length
$L_A$	$6 \times 10^6$ m	Width of Atlantic basin
$L_P$	$12 \times 10^6$ m	Width of Pacific basin
$H$	4000 m	Ocean depth
$h_m$	100 m	Depth of Southern Ocean mixed layer
$\tau$	$0.1 \text{ N m}^{-2}$	Reference wind stress
$K$	$1000 \text{ m}^2 \text{ s}^{-1}$	ACC along-isopycnal diffusivity
$F$	$-1 \times 10^{-4} \text{ s}^{-1}$	Coriolis parameter in the ACC
$\rho_0$	$1025 \text{ kg m}^{-3}$	Reference seawater density
$b_1, b_2, b_3, b_4$	$(5, 1, -1, -2.5) \times 10^{-3} \text{ m s}^{-2}$	Buoyancy of layers 1–4 (corresponding to AAIW, UCDW/PDW, LCDW/NADW, and AABW density classes)
$T_1, T_2, T_3, T_4$	$(10, 6, 2, 0)^\circ\text{C}$	Temperature of layers 1–4 (used for the Southern Ocean equivalent surface temperature calculation in Fig. 4f)
$F_0$	$2 \times 10^{-8} \text{ m}^2 \text{ s}^{-3}$	Reference surface buoyancy flux
$\ell_N, \ell_S$	$(2, 1.6) \times 10^5$ m	Widths of the sea ice melting and formation regions, respectively
$y_{\text{si}}$	$-1 \times 10^6$ m	Meridional position of sea ice edge in Southern Ocean
$\gamma_n$	0.4	Relative strength of buoyancy flux between sea ice melt and formation regions
$\kappa_0$	$1 \times 10^{-4} \text{ m}^2 \text{ s}^{-1}$	Mean diapycnal diffusivity
$\Delta\kappa$	$8 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$	Change in diapycnal diffusivity
$z_0$	1600 m	Depth of transition in diapycnal diffusivity
$d$	300 m	Thickness of transition in diapycnal diffusivity
$U_0$	$0.03 \text{ m s}^{-1}$	Mean (barotropic) zonal velocity
$T_{\text{NADW}}$	$0\text{--}20 \times 10^6 \text{ m}^3 \text{ s}^{-1}$	Forcing parameter: NADW formation rate (varied in experiments)

parameters are provided in Table 1. In the four-layer experiment,  $T_{\text{NADW}}$  determines the volume transport between layers 1 and 3. Hines et al. (2019) have explored the sensitivity of the ocean's steady-state configuration to changes in the density of newly formed NADW, in addition to its formation rate. The experiments presented in this study explore the response of the ocean's stratification, ventilation and overturning to perturbations in  $T_{\text{NADW}}$ . Taking  $T_{\text{NADW}}$  as an external forcing, the model evolves  $z$  and  $y$  for each layer interface in each basin. Steady state is achieved when the net volume flux into each layer is zero. Together, the depth and outcrop position of each interface defines a slope, which is assumed to be constant across the Southern Ocean. A discussion of more realistic isopycnal curvature in the Southern Ocean is presented in T16.

We now briefly describe parameterizations used to determine water mass transformation rates between density layers. A more detailed discussion of these relationships appears in T16.

### 1) DIFFUSIVE UPWELLING

Water mass modification at low latitudes is dominated by diffusive upwelling (Nikurashin and Vallis 2011). Zonally averaged transformation is computed in this

model, although the actual transformation may be localized over steep, sloping, bottom bathymetry (Ferrari et al. 2016). The transformation depends on the strength of the mixing, represented by a diapycnal diffusivity  $\kappa$ , and the stratification following

$$T_{\text{diff}} = \kappa(z)A/\Delta z, \quad (3)$$

where  $\Delta z$  is the thickness of the layer above the interface and  $A$  is the horizontal area of the interface. The value for  $\kappa$  is determined from the interface depth at each time step. More accurate representations of the second-order vertical derivative associated with this diffusive process could be applied in simulations that resolve a greater number of density classes, (e.g., Marshall and Zanna 2014). The choice of  $\Delta z$  as the density layer above a given interface tends to stabilize the model by increasing the diffusive flux into thinner layers. The depth dependence of  $\kappa$  is given by

$$\kappa(z) = \kappa_0 - \frac{\Delta\kappa}{2} \tanh\left(\frac{z - z_0}{d}\right), \quad (4)$$

where  $\kappa_0$  is a background diapycnal diffusivity,  $\Delta\kappa$  is the change in diapycnal diffusivity with depth, which occurs over a transition zone of thickness  $d$  centered at a depth  $z_0$

(Fig. 2d). Although this profile is idealized, it is motivated by the enhancement of the diapycnal diffusivity in the deep ocean over rough bottom topography (Waterhouse et al. 2014). Values for parameters are provided in Table 1.

## 2) SOUTHERN OCEAN BUOYANCY FLUXES AND RESIDUAL TRANSPORT

The outcropping of isopycnals at the surface of the Southern Ocean exposes density surfaces to direct surface buoyancy forcing resulting in water mass modification. The transformation rate at the outcropping location depends on both the local strength of the surface buoyancy forcing  $F_b$  as well as the local meridional gradient of the surface buoyancy distribution  $b_s = b(z' = 0)$ :

$$T_{\text{SO}} = \frac{F_b(y')}{\partial b_s / \partial y'} L, \quad (5)$$

where  $L$  is the zonal extent of the basin, which may be different for the Atlantic and Indo-Pacific. Because of the truncated model resolution, the relationship Eq. (5) is calculated as

$$T_{\text{SO}} = F_b(\Delta y / \Delta b) L, \quad (6)$$

where  $\Delta b = b_n - b_{n+1}$  and  $\Delta y = (y_{n-1} - y_{n+1})/2$ , with  $n$  representing layer number. The value of  $F_b$  is taken at the layer-interface outcrop position at each time step.

The surface buoyancy flux has contributions from heating/cooling as well as the freshwater cycle, involving evaporation, precipitation, glacial melt, and sea ice formation/melting. Sea ice patterns play the leading role in setting the Southern Ocean distribution of the surface buoyancy flux (Abernathey et al. 2016). Regions of  $F_b < 0$ , related to sea ice formation, are concentrated around the Antarctic continent, particularly in the Weddell and Ross Seas, known sites of bottom-water formation. Regions of  $F_b > 0$ , related to sea ice melt, are located farther to the north in a band between 50° and 60°S. North of the sea ice edge, direct heating from the atmosphere provides an additional heat flux, but its contribution to the overall buoyancy flux budget for the Southern Ocean is smaller than that of sea ice (Abernathey et al. 2016; Sun et al. 2016), and is not explicitly included here. In our model we define the buoyancy flux distribution as

$$F_b(y) = \gamma_N F_0 \exp \left[ - \left( \frac{y - y_{\text{si}}}{\ell_N} \right)^2 \right] - F_0 \exp \left[ - \left( \frac{y + \ell}{\ell_S} \right)^2 \right], \quad -\ell < y < 0, \quad (7)$$

shown in Fig. 2a. Here,  $F_0$  is a constant buoyancy flux,  $y_{\text{si}}$  is the meridional position of the sea ice edge in the

Southern Ocean (kept fixed), and  $\ell_N$  and  $\ell_S$  are the meridional widths of the positive and negative buoyancy forcing regions associated with sea ice melt and sea ice formation, respectively. The prefactor  $\gamma_N$  sets the relative strengths of buoyancy forcing in the sea ice melting and formation regions. The buoyancy flux is assumed to be the same in Atlantic and Pacific basins. An exploration of the model sensitivity to a zonally varying surface buoyancy flux, in steady state, is presented in T16, while the sensitivity of steady-state solutions to  $y_{\text{si}}$  is explored in Hines et al. (2019). Parameter values are presented in Table 1. Values were chosen to ensure that the surface buoyancy has zero spatial mean. However, this is not an essential constraint on the model since additional buoyancy fluxes are implied from the formation of NADW in the NH.

## 3) ZONAL TRANSPORT

Zonal flow in the Southern Ocean has both barotropic (depth independent) and baroclinic (vertically sheared) components; the latter is determined by the slopes of isopycnals spanning the ACC. T16 solved for both components of zonal flow in order to calculate the zonal volume convergence within each density layer in the Southern Ocean. Here, in the four-layer model, we only consider the barotropic component of the zonal flow, which we assume is constant, noting that the zonal transport of the ACC has been shown to be relatively insensitive to changes in surface forcing (Hallberg and Gnanadesikan 2006).

With this assumption, the zonal convergence  $T_{\text{zonal}}$  in the Southern Ocean is only dependent on the difference in the depth-latitude area of a given density class between the Atlantic and Indo-Pacific basins, or

$$T_{\text{zonal}} = -U_0(\Delta A_{\text{ACC}}), \quad (8)$$

where  $U_0$  is the barotropic zonal velocity and  $A_{\text{ACC}}$  is the areal coverage of a density class in the region  $-\ell < y < 0$ . The area  $A_{\text{ACC}}$  is determined geometrically from the values of  $z$  and  $y$  of the bounding interfaces. The sign convention is such that mass diverges from the basin with a larger value of  $A_{\text{ACC}}$  and converges in the basin with a smaller value of  $A_{\text{ACC}}$ .

## 4) SIMULATIONS

As our focus is on the response to a change in NADW formation rate, and not the trigger of DO events, we prescribe a time-dependent  $T_{\text{NADW}}$ . Both NH warm and cold phases of the bipolar seesaw are simulated using a 2000-yr, 10-Sv-amplitude ( $1 \text{ Sv} \equiv 10^6 \text{ m}^3 \text{ s}^{-1}$ ), top-hat oscillation (Fig. 2e). This choice is motivated by the widely held view that rapid changes in Greenland temperature are closely associated with changes in the strength of NADW formation, and is a reasonable

starting point for understanding the global response to changes of this type. The lag time scale determined from the ice core (section 1 of Buizert et al. 2015a),  $\sim 200$  years, is not prescribed by this forcing. We keep  $F_b$  fixed, but coupling with sea ice dynamics that allows for a temporally changing surface buoyancy flux distribution is an important future direction to explore, and is discussed in more detail in section 4.

The NADW forcing for the main suite of simulations (e.g., Figs. 2 and 4) is given by

$$T_{\text{NADW}}(t) = 5 + 5 \left[ \tanh\left(\frac{t-1000}{40}\right) - \tanh\left(\frac{t-3000}{40}\right) \right], \quad 0 < t < 4000 \quad (9)$$

where  $t$  is time in years and  $T_{\text{NADW}}$  is given in units of Sverdrups. This forcing is repeated every 4000 yr.

We introduce a  $T_{\text{NADW}}$  time series with noise in section 3 that uses the same base forcing as given in Eq. (9) but adds a noise component that was designed to have a red frequency spectrum ( $\omega^{-3}$ ) with a random phase, applied across a frequency range corresponding to periods between 200 and 2000 yr. Simulations that employed a noise component with a shallower spectral slope, emphasizing higher-frequency variability, did not qualitatively change the results presented here. The motivation for adding noise was that the initial forcing time series produces correlations at long lags because the system has no variations outside the abrupt shift (see section 3b). This is unrepresentative of the climate system. However, the noise is not meant to represent specific physical processes.

### c. Observations

We follow a similar procedure to Buizert et al. (2015a) to determine the phasing, or lag, of the bipolar seesaw. Buizert et al. (2015a) stacked the North Greenland Ice Core Project (NGRIP) and WAIS records at abrupt DO transitions (Fig. 1c). To do this, the authors first identified the abrupt DO transitions in the NGRIP  $\delta^{18}\text{O}$  record, defined as the midpoint between the values before and after the transition. Individual NGRIP DO events were then interpolated onto a regular temporal grid spanning 1200 years before and 1200 years after the transition at 1-yr resolution; the transition is labeled  $t = 0$ . WAIS events are aligned to the midpoint of the corresponding WAIS methane transition and interpolated in a similar fashion. Once all records are on the same age scale, they were averaged to create the stack (Fig. 1c).

The millennial-scale variability in the NGRIP and WAIS  $\delta^{18}\text{O}$  records was isolated using a fourth-order Butterworth filter (500–10 000-yr window) on the methane-synchronized portion of the records (31.2–68 kyr), as in Buizert et al.

(2015a). The filtered WAIS  $\delta^{18}\text{O}$  record was then differentiated and interpolated onto the same age scale as the NGRIP  $\delta^{18}\text{O}$  record, at 20-yr resolution. As noted in Buizert et al. (2015a,b), there is better correspondence between the NGRIP  $\delta^{18}\text{O}$  record and the U/Th-dated Hulu Cave  $\delta^{18}\text{O}$  record if the Greenland Ice Core Chronology 2005 (GICC05) age scale is multiplied by a factor of 1.0063. This factored version of the GICC05 age scale is used to create the methane-synchronized WAIS Divide deep ice core 2014 (WD2014) chronology. Following these observations, we used the NGRIP  $\delta^{18}\text{O}$  record on GICC05  $\times$  1.0063 and the WAIS  $\delta^{18}\text{O}$  record on WD2014. As can be seen in Fig. 3, this age scale generates close correspondence between NGRIP  $\delta^{18}\text{O}$  and WAIS methane, and was used in the calculation of the lagged correlation (Fig. 4i).

## 3. Bipolar seesaw mechanism and time scales

### a. DO evolution

We begin by summarizing the adjustment of the low-latitude stratification and the Southern Ocean outcrop position, using Eqs. (1) and (2), respectively, in the two-layer ocean configuration with a temporally varying  $T_{\text{NADW}}$ , as in Eq. (9). As the NADW formation rate  $T_{\text{NADW}}$  intensifies (weakens), the interface shoals (deepens), and the outcrop position moves equatorward (poleward) (Fig. 2f), consistent with the two phases of the bipolar seesaw (section 2a). For this particular set of model parameters, the depth of the interface in the northern basin varies over  $-0.6 < z/H < -0.2$  while the outcrop position in the Southern Ocean varies over  $\sim 25\%$  of the channel,  $-0.82 < y/\ell < -0.58$  (Fig. 2f). The magnitude of these displacements are a function of the model parameters (Hines et al. 2019). By examining the phase space  $[z/H, y/\ell]$ , the transition between states is associated with a change in Southern Ocean outcropping between positive and negative buoyancy forcing regions (Fig. 2g). Similarly, the interface depth moves across the  $\kappa$  transition zone between the two states. During strong  $T_{\text{NADW}}$  states, diffusive transformation alone is insufficient to close the overturning, so lighter waters are formed in the Southern Ocean. As  $T_{\text{NADW}}$  weakens, diffusive transformation upwells more water than is produced in the NH, and the system transitions to a state where deep water is made in the SH as well. In the real ocean, a volume balance resulting from the three transformation processes must occur within each density class, which is consistent with our understanding that different ranges of density become lighter and denser in the Southern Ocean.

We now extend the two-layer model to a somewhat more realistic representation of the ocean by simulating the evolution of four density classes in both an Atlantic

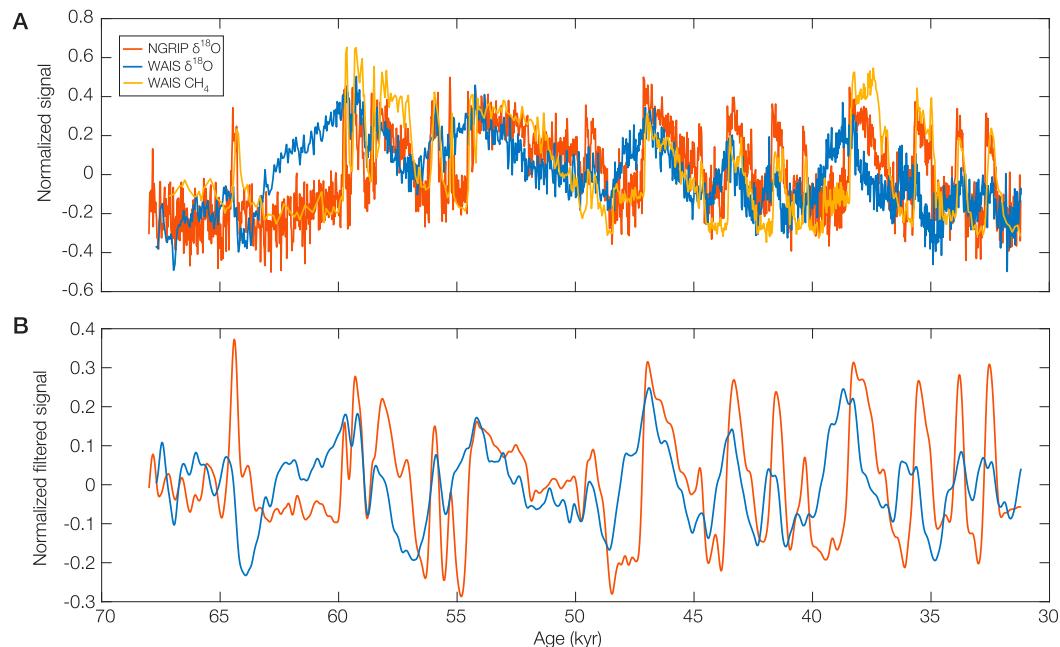


FIG. 3. (a) NGRIP  $\delta^{18}\text{O}$  on GICC05  $\times 1.0063$  (red), WAIS  $\delta^{18}\text{O}$  on WD2014 (blue), and WAIS methane on WD2014 (yellow). All records are normalized so that they fit on the same y axis. (b) Filtered  $\delta^{18}\text{O}$  records from NGRIP (red) and WAIS (blue). The filter is a fourth-order Butterworth filter (500–10 000-yr window), applied with zero lag.

and an Indo-Pacific basin. The additional layers and basin more accurately represent changes in the overturning structure and pathways. In the two-layer model, the strong  $T_{\text{NADW}}$  regime behaves similarly to the upper branch of the modern overturning circulation, where NADW formation is principally balanced by lightening in the Southern Ocean (Wolfe and Cessi 2011). During the weak  $T_{\text{NADW}}$  regime, diffusive upwelling balances deep-water sources, which is similar to the lower limb of the modern overturning circulation. By including four layers, both upper and lower overturning cells can be represented. It has also been suggested that these two cells have been isolated during glacial periods (Ferrari et al. 2014), but are coupled through zonal interbasin exchange in the modern ocean (Talley 2013); by including two basins, transitions between these regimes can also be resolved. For these results, Eq. (1) is modified to

$$\begin{aligned} \dot{V}^A &= (-A\dot{z})^A = \sum T_V^A \\ &= (T_{\text{diff}}^A + T_{\text{SO}}^A - T_{\text{NADW}} + T_{\text{zonal}}), \quad \text{Atlantic basin,} \end{aligned} \quad (10)$$

and

$$\begin{aligned} \dot{V}^P &= (-A\dot{z})^P = \sum T_V^P \\ &= (T_{\text{diff}}^P + T_{\text{SO}}^P - T_{\text{zonal}}), \quad \text{Pacific basin,} \end{aligned} \quad (11)$$

where  $T_{\text{zonal}}$  is defined in Eq. (8) and accounts for exchange between the Indo-Pacific and Atlantic (indicated by superscripts  $P$  and  $A$ , respectively) through the ACC.

The same top-hat  $T_{\text{NADW}}$  forcing used in the two-layer model, now a volume transport between layers 1 and 3, is applied only in the Atlantic basin of the NH (Fig. 4a). As before, periods of stronger NADW formation (NH warm phase) correspond to an intensified stratification of the upper ocean (layers 1 and 2) and a reduced stratification at depth to accommodate the injection of NADW in layer 3 (Fig. 4b). The outcrop positions of isopycnals in the ACC channel also reflect this shallow stratification, with shallow density classes outcropping in regions where the surface buoyancy flux is positive,  $F_b > 0$  (Fig. 4c). Conversely, during periods of reduced NADW formation (NH cool phase), abyssal layers are thin and isopycnals outcrop farther to the south in the ACC, where the surface buoyancy flux is negative,  $F_b < 0$ . The associated diapycnal water mass transformations ( $T_{\text{diff}}$  and  $T_{\text{SO}}$ ) for the four-layer model are shown in Fig. 5. Note that  $T_{\text{diff}}$  in the upper ocean and in the abyssal ocean (light- and dark-blue lines, respectively) are inversely related (Fig. 5b); this feature is discussed in section 3c. The greater variability of  $T_{\text{diff}}$  between layers 3 and 2 is due to the combined changes in stratification and mixing intensity as the interface moves through the  $\kappa$  transition zone. In the Southern Ocean,

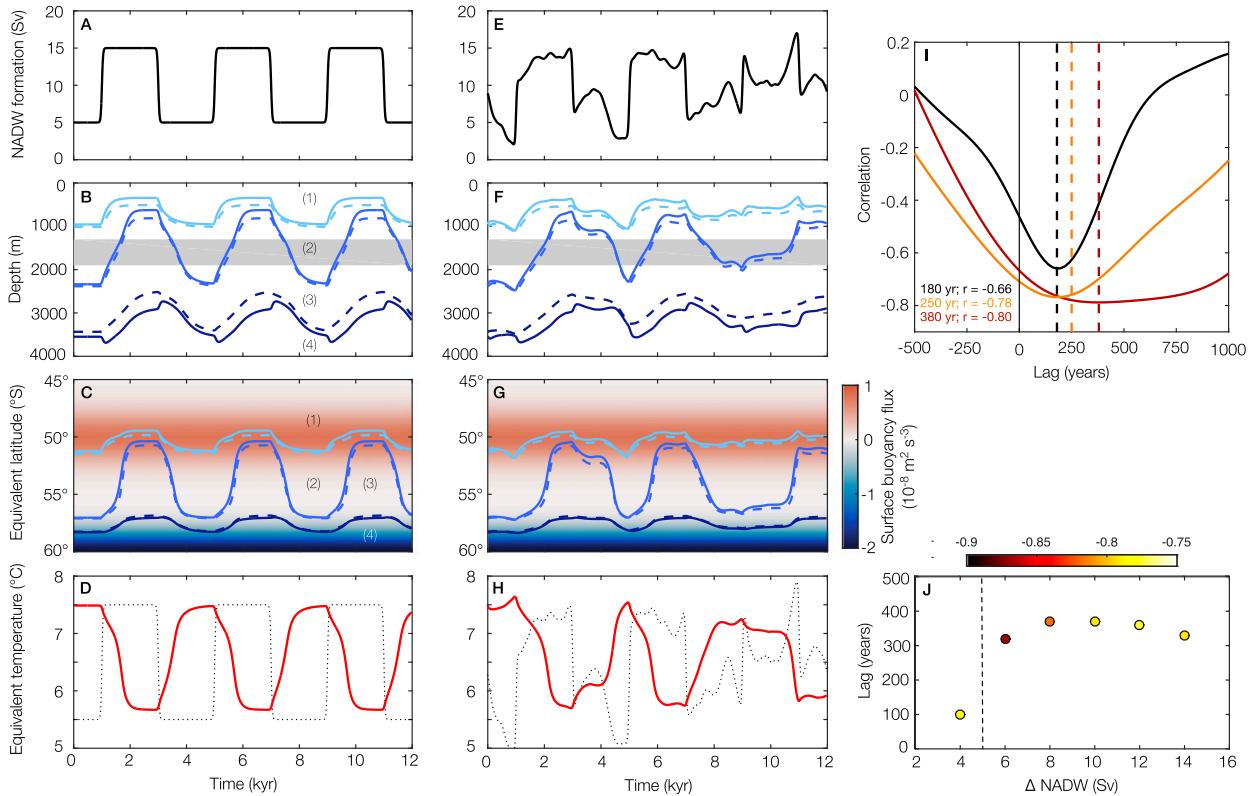


FIG. 4. Modeled DO cycle for a four-layer, two-basin simulation. (a) Prescribed NADW formation  $T_{\text{NADW}}$ , which varies by 10 Sv; the change occurs over a period of  $\sim 20$  yr and persists for a period of 2000 yr. (b) Low-latitude stratification with curves indicating the interfaces between density classes; the gray bar corresponds to the  $\kappa$  transition region (Fig. 2d). (c) Southern Ocean surface outcrop positions with the meridional distribution of the surface buoyancy flux shown in color as in Fig. 2a. (d) Parameterized mean surface temperature in the Southern Ocean (red) showing the antiphasing between the NH and SH (see text for details); the dotted line is reproduced from (a). (e)–(h) As in (a)–(d), but for a modified NADW formation that includes noise (section 2b); the dotted curve in (h) is reproduced from (e). (i) Lagged correlation between NGRIP  $\delta^{18}\text{O}$  and WDC  $d(\delta^{18}\text{O})/dt$  time series (black), and lagged correlation between NADW formation strength and the time derivative of the Southern Ocean equivalent surface temperature without noise (red) and with noise (orange). (j) The SH lag as a function of the amplitude of NADW top-hat variations [as in (a)]; the color indicates the maximum (negative) correlation. For  $\Delta T_{\text{NADW}} < 5$  Sv (the dashed line), a transition in outcrop position does not occur. In (b), (c), (f), and (g), the solid and dashed curves indicate interface positions in the Atlantic and Pacific basins, respectively. In (b) and (c), the layer numbers are indicated.

$T_{\text{SO}}$  only changes sign for the interface separating layers 2 and 3, which we discuss next.

Numerical models have shown that the Southern Ocean eddy overturning nearly balances the wind-driven overturning across a range of wind stress amplitudes—a behavior validated in both idealized and realistic eddy-resolving ocean simulations (Meredith and Hogg 2006; Munday et al. 2013; Bishop et al. 2016). The implication is that near steady state  $\dot{y}_{\text{wind}} + \dot{y}_{\text{eddy}} \approx 0$ , such that that variations in the isopycnal slope  $s_p$  should be small. Indeed, recent work proposes  $s_p$  tends toward a constant,  $s_0 = \tau/(\rho_0 K)$ , even across very different climates (Ferrari et al. 2014). Here, the evolution of  $s_p$  (Fig. 6 depicts the slope of the interface between layers 2 and 3) is shown to vary by  $\pm 30\%$  about the reference value  $s_p/s_0 = 1$ . As required by Eq. (2),  $s_p$  is steeper than  $s_0$  when the

outcrop position lies in a negative buoyancy forcing region. Here, the eddy overturning “outcompetes” the wind-driven overturning to balance the negative water mass transformation. When the interface outcrops in the positive buoyancy forcing region, the wind-driven overturning is stronger, such that  $s_p$  is shallower than  $s_0$  (Fig. 6b). This behavior is somewhat dependent on the assumption of a uniform eddy diffusivity  $K$  in the channel, but the sense of the outcrop position migration is always to keep  $s_p/s_0 \approx 1$ . Thus, as the depth of the interface in the basin  $z$  deepens (e.g., the low-latitude stratification increases at depth from modifications in  $T_{\text{NADW}}$ ) the outcrop position  $y$  migrates poleward, and vice versa (Figs. 2, 4, and 6).

Coupling to the atmosphere and the cryosphere is required to determine the climate’s full response to DO

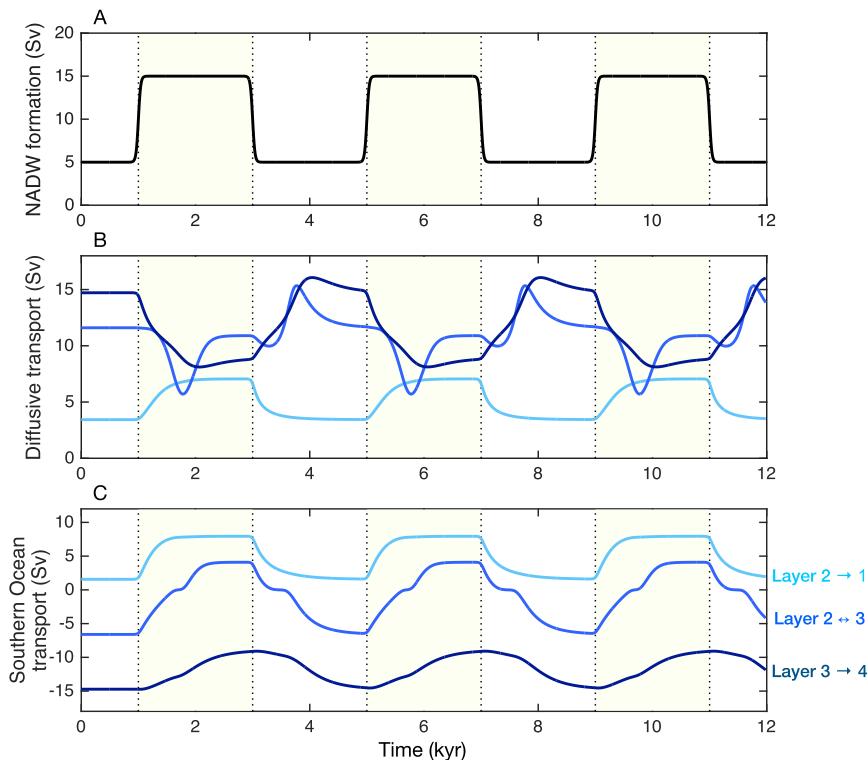


FIG. 5. Additional diagnostics from the model simulation shown in Fig. 4. (a) Model forcing: NADW formation rate. (b) The total (upward) diffusive transport (Sv) summed over both Atlantic and Pacific basins. Dark, medium, and light blue indicate transport from layer 4 to layer 3, between layer 3 and layer 2, and from layer 2 to layer 1, respectively; the colors correspond to the density-class interfaces in Fig. 4. (c) The total water mass transformation in the Southern Ocean due to direct surface buoyancy forcing (Sv), combining both the Atlantic and Pacific basins. Positive and negative values indicate transformation to lighter and heavier density classes, respectively.

events, but here we provide a relatively crude estimate of the Southern Ocean surface temperature by assigning temperatures to each density class and calculating a weighted average based on their meridional extent in the ACC (Fig. 4d; Table 1). As argued earlier, poleward outcropping of lighter density classes leads to a warmer SST and an increase in SH temperatures. Acknowledging the simplicity of this model, we prefer to emphasize the mechanism that explains why Southern Ocean SSTs should change in response to  $T_{\text{NADW}}$  and why this produces a bipolar seesaw without attempting to be predictive about the magnitude of the temperature changes in the climate record. The primary result is that a change in NADW production modifies the basin stratification, and must be accompanied by a rearrangement of the surface density distribution in the Southern Ocean. This rearrangement is dynamically constrained to lead to SH warming during reduced NADW regimes and SH cooling during intensified NADW regimes. This relationship holds even when

noise is added to the NADW forcing time series (Figs. 4e–h). Furthermore, using the same lagged correlation analysis as in the ice core records (Buizert et al. 2015a), the model output shows that SH surface temperature lags the NH by roughly 300 years (Fig. 4j), providing consistency between the ice core observations and our dynamically constrained theory.

#### b. DO phasing

A key feature of these simulations is that the SH lag is weakly dependent on the amplitude of  $T_{\text{NADW}}$  fluctuations  $\Delta T_{\text{NADW}}$ , provided that  $\Delta T_{\text{NADW}}$  is sufficiently large (Fig. 4j). This behavior, at least in this model, occurs because the lag is set by a fast adjustment of the outcrop position (separating layers 2 and 3) as it transitions between regions of positive and negative surface buoyancy forcing (Figs. 4c and 6). Immediately after  $T_{\text{NADW}}$  changes, the outcrop position  $y$  begins to adjust slowly. A later abrupt shift in  $y$ , associated with a faster speed of  $\sim 1 \text{ km yr}^{-1}$  (Fig. 6c), dominates the temporal

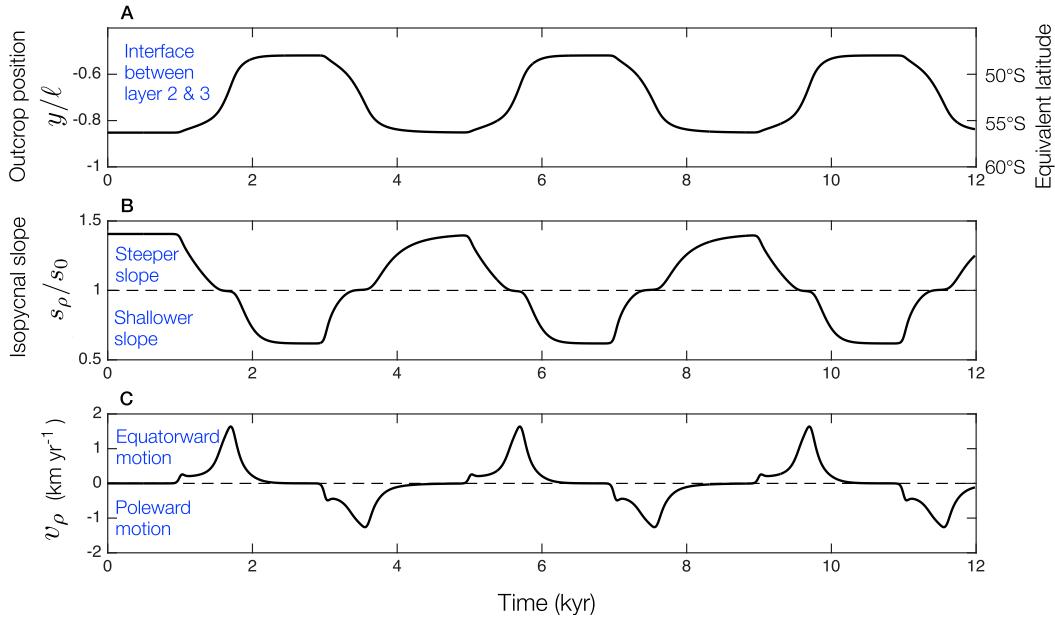


FIG. 6. (a) Time series of the outcrop location of the interface between layers 3 and 2 in the Atlantic basin corresponding to the simulation in Fig. 4. (b) Nondimensional isopycnal slope across the ACC for the same interface;  $s_0 = \tau/(\rho_0 f K)$ . The slope  $s_\rho$  is always negative (note  $s_0 < 0$ ); values that are greater (less) than 1 correspond to steeper (shallower) slopes. (c) Speed of outcrop migration  $\dot{y}$  ( $\text{km yr}^{-1}$ ) in Eq. (2), for the same interface.

evolution of  $y$  and isolates the multicentennial lag. This transition is linked to both the Southern Ocean surface buoyancy flux  $F_b$ , and the vertical distribution of low-latitude diffusive mixing  $\kappa$  (Figs. 2a,d). The rapid displacement occurs as  $y$  transits the region between sites of sea ice formation and sea ice melt where  $F_b \approx 0$ . Furthermore, during this period the isopycnal slope is close to the critical slope,  $s_\rho \approx s_0 = \tau/(\rho_0 f K)$  (Fig. 6b). From Eq. (2), this implies that wind and eddy overturning contributions compensate and, when combined with  $F_b \approx 0$ , suggests that  $\dot{y} \approx 0$ , or that  $y$  should be stationary. However, the surface buoyancy distribution is coupled to the low-latitude stratification through  $s_\rho$ , and the low-latitude stratification is still adjusting, such that  $\dot{z} \neq 0$ . Thus  $\dot{y}_{\text{eddy}}$ , and therefore  $\dot{y}$ , continue to change in Eq. (2). During this period, the interface between upper and lower layers in the basins lies between the depths of high and low  $\kappa$ , and  $T_{\text{diff}}$  is the dominant term in Eq. (1). Given the relationship  $T_{\text{diff}} = \kappa A/h$  from Eq. (3), where  $h$  is the thickness of layer 2, the vertical velocity of the interface in the basin is approximately given by  $w \approx \kappa/h$ . Since  $s_\rho \approx s_0$  throughout this period, the meridional speed of the outcrop position is  $v_\rho \approx w/s_0 \approx \kappa/(hs_0)$  and the time scale for adjustment is  $\tau_{\text{trans}} \approx \ell_i/v_\rho \approx hs_0 \ell_i/\kappa$ , where  $\ell_i$  is the distance between sea ice melt and formation regions. This relationship for  $\tau_{\text{trans}}$  is notable because it does not depend on the amplitude of  $K$  or  $\tau$  alone, which may have varied

significantly throughout Earth's history, but rather on the assumption that  $s_\rho \approx s_0$  (Munday et al. 2013; Ferrari et al. 2014). Plugging in representative values from the model, assuming  $z$  resides in the transition region,

$$\begin{aligned} \tau_{\text{trans}} &= \frac{hs_0 \ell_i}{\kappa} \approx \frac{1000 \text{ m} \times 10^{-3} \times 500 \text{ km}}{5 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}} \\ &= 10^{10} \text{ s} \approx 300 \text{ yr}. \end{aligned} \quad (12)$$

Again, the emphasis here is on the mechanism: a gap between extremes in Southern Ocean buoyancy forcing leads to an abrupt shift in outcrop positions in the Southern Ocean that gives rise to the phasing. While the parameters in Eq. (12) may vary during different climate states, especially  $\ell_i$ , they are unlikely to produce a time scale that is significantly greater than or less than a few hundred years; observed lags for individual DO events also vary about the stacked value of 200 yr. The smaller  $\tau_{\text{trans}}$  in Fig. 4j at small  $\Delta T_{\text{NADW}}$  occurs when  $\Delta T_{\text{NADW}}$  is not large enough to trigger a shift in outcrop positions in the Southern Ocean between regions of negative and positive surface buoyancy forcing. A new steady state can be achieved without a change in a sense of the transformation in the Southern Ocean for the different density classes. The Pa/Th ratio shown in Fig. 1c indicates that substantial changes in  $\Delta T_{\text{NADW}}$  can occur across individual DO cycles; however, a direct empirical relationship between this ratio and  $T_{\text{NADW}}$  is not agreed

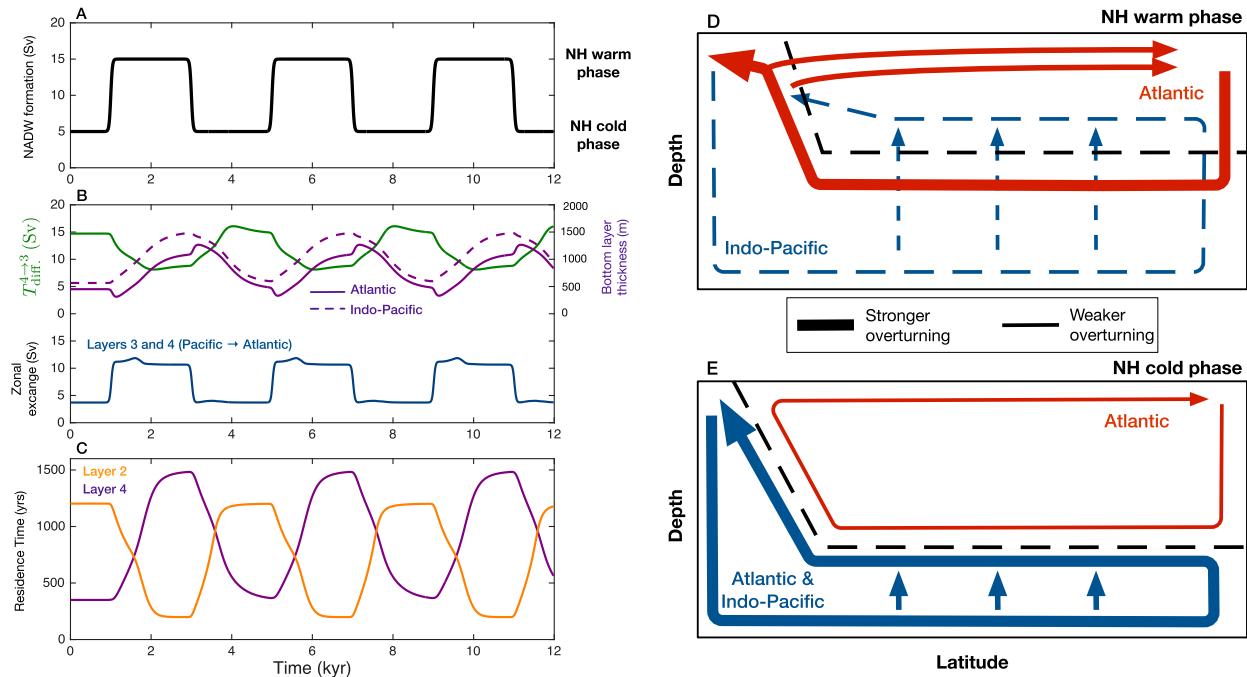


FIG. 7. Residence times and overturning pathways for the four-layer, two-basin model. (a) The  $T_{\text{NADW}}$  forcing as in Figs. 4a–4d. (b) Diffusive upwelling  $T_{\text{diff}}$  in the abyssal ocean (green curve), thickness of the bottom-most layer in the model (purple curves), and zonal exchange from the Pacific to the Atlantic in the abyssal ocean (layers 3 and 4; blue curve). (c) Residence times for layers 2 and 4 for the simulation in Figs. 4a–4d. Residence time is defined as the total volume of the layer (Atlantic and Pacific basins) divided by the mass transport out of the layer. Also shown are schematics of the overturning pathways for NH (d) warm and (e) cold phases.

upon, which makes it difficult to quantitatively identify a threshold from the paleorecord.

Additional feedbacks—for example, changes to sea ice extent as a result of changing Southern Ocean temperatures—will affect longer-term climate variability and the exact timing of the SH lag. These aspects warrant exploration in more-comprehensive ocean models or coupled models. However, they should not qualitatively impact this explanation, which is based on observationally constrained features of the climate system, such as the ocean’s vertical mixing profile (stronger mixing at depth) and the freshwater cycle associated with sea ice growth and melt (spatially separated regions of ice formation and melt).

### c. Residence times

Beyond capturing the interhemispheric asymmetry and temporal lag of DO events, our proposed mechanism also implies an inverse relationship between NADW production and the strength of both abyssal stratification and abyssal overturning (Figs. 7a,b). This inverse relationship has a large impact on deep-ocean residence times. A strong stratification implies both a smaller volume (thickness) for a given layer and a stronger diffusive transport, which together, ventilates

the layer more rapidly. This relationship between stratification and abyssal overturning strength has been noted previously in the context of changes to SH forcing (De Boer and Hogg 2014; Jansen 2017). The emergence of similar behavior in response to changes in NADW formation indicates that this inverse relationship is robust to different forcing scenarios. For this simulation, the residence time of bottom-water layers varies between 400 and 1500 yr (Fig. 7c).

An important difference between the two-layer and four-layer simulations is that, in the latter, dense waters formed at NH and SH high latitudes do not fill the same density class: NADW overlies dense bottom waters originating from the SH. For strong NADW formation regimes, a significant fraction of the deep water produced in the Atlantic (layer 3) is transformed into bottom water in the Southern Ocean (layer 4). This behavior can only be sustained if there is a compensating transformation of NADW into lighter density classes (layer 2), which principally occurs in the Indo-Pacific where there is no northern source of deep water (Talley 2013; Jones and Cessi 2016; T16). To accommodate this, a larger zonal exchange between basins occurs during an NADW-on state (Ferrari et al. 2014; T16) (Fig. 7b). This model goes beyond earlier studies

(Ferrari et al. 2014) by predicting not only a change in overturning pathways, but also linking this to a change in the intensity of upper and abyssal circulations. Thus over a DO cycle, this model suggests that there is an adjustment of the full ocean overturning and stratification that involves a transition between a two-cell structure (NH cold phase) and a single interleaved circulation that transits through both basins (NH warm phase) (Figs. 7d,e).

Direct application of these results to the paleorecord needs to be taken with caution. For instance, in T16 (their Fig. 9), it was shown that the overturning transitions to a state with two distinct overturning cells as  $T_{\text{NADW}}$  increases. This apparent discrepancy with the discussion above is due to the fact that in T16, NADW was exchanged between layers 1 and 2, rather than layers 1 and 3. This would represent a lighter form of NADW. Thus, strong formation of a lighter version of NADW leads to two separated cells, which may be a good representation of the Last Glacial Maximum (LGM; Curry and Oppo 2005). Hines et al. (2019) provides further insight into the sensitivity of the global overturning and stratification to the partitioning of NADW between different layers, but a full exploration of changes to both the strength and the density of NADW is needed.

#### 4. Discussion

The idealized nature of the model used in this study was chosen to highlight the transient adjustment between different overturning states (Hines et al. 2019), which is linked to Southern Ocean surface buoyancy forcing that is imposed by Antarctic sea ice dynamics. However, there are a number of limitations to this approach. First, because of the small number of density classes considered, the parameterizations used to diagnose water mass transformation are basic. For example, in a more realistic model, the diffusive transformation will depend on layer thicknesses both above and below a given interface. However, as the stratification weakens with depth, the layer above tends to dominate the transformation. Similarly, transformation due to surface buoyancy forcing in the Southern Ocean is chosen to be evaluated at the outcrop position, rather than integrating the surface buoyancy flux over the outcrop area of each density class. This approach becomes more justifiable as more layers are added. Furthermore, the lightest and heaviest density classes, which feel net positive and negative buoyancy forcing, respectively, cannot become lighter or denser in this isopycnal model. Overturning circulations due to variations in the surface buoyancy flux that occur within a

single layer are simply not expressed in this low-order model. In addition to the parameterizations, the geometry of the model is highly idealized. While the leading-order dynamics described in Eq. (2), valid for the upper part of the ACC, that is, above the sill depth, have been largely realized in high-resolution ocean models (Hallberg and Gnanadesikan 2006; Bishop et al. 2016), the formation of dense waters in the SH occurs over the Antarctic continental shelf and depends on complicated processes involving the polar gyres, the circulation over the Antarctic continental slope and interactions with sea ice and floating ice shelves over the continental shelf (Orsi et al. 1999; Whitworth et al. 1998; Miller et al. 2012). Because these processes are confined to the shelf and slope, they are likely to have a smaller impact on the outcrop position of deep water in the Southern Ocean, although they do influence the magnitude of  $F_b$ .

Another major limitation of this ocean-only model is that feedbacks involving the atmosphere and the cryosphere are not included. Perhaps most importantly, changes to the Southern Ocean surface buoyancy (or temperature) distribution will influence sea ice extent. This coupling will influence our proposed mechanism as the time scale for the abrupt transition depends on the distance between regions of positive and negative buoyancy forcing, or  $\ell_i$  in Eq. (12) (Buizert and Schmittner 2015). While DO events are characteristic of the last glacial period, they are not apparent in the ice core records during the Holocene. A simple explanation is that fluctuations in NADW formation during the current interglacial period are not sufficiently large to trigger a DO event. However, during the last glacial period, an expanded sea ice extent would have led to a larger separation between regions of positive and negative surface buoyancy forcing. This behavior implies that large changes in both the Southern Ocean surface density distribution and the low-latitude vertical stratification are required to produce the bipolar seesaw end members shown in Figs. 2a,b. In the modern ocean, regions of sea ice formation and melt are less separated than they would have been during glacial periods (Ferrari et al. 2014). This reduced separation suggests that density classes may be able to transition between outcropping in different water mass transformation regions (e.g., sea ice growth or melt regions) with smaller changes in the surface density distribution and therefore weaker swings in Antarctic climate.

The lack of coupling with sea ice and the atmosphere means that these simulations are limited in their ability to reproduce the full response of the climate system to a perturbation in NADW formation. Nevertheless, the dynamics here are deemed to be sufficiently robust that

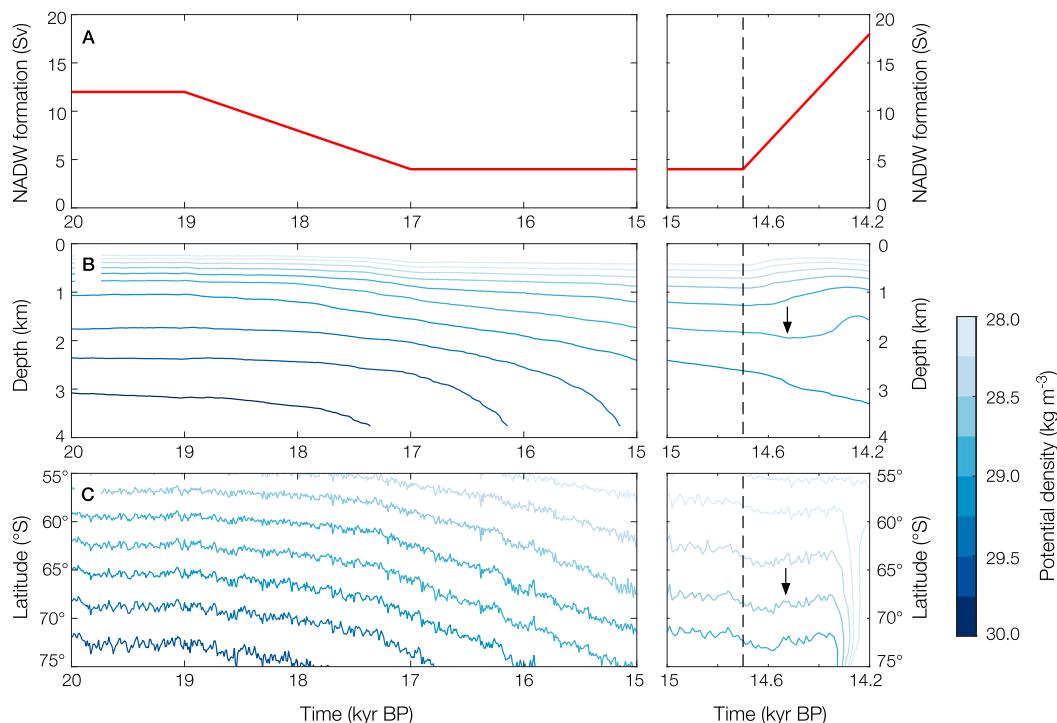


FIG. 8. Stratification and Southern Ocean outcrop evolution from the TraCE simulation (Liu et al. 2009; He et al. 2013) transitioning out of the LGM. (a) NADW formation rate (Sv) as reported in Fig. 1d of Liu et al. (2009), which is a response to a prescribed freshwater flux applied variably to the North Atlantic and to the Gulf of Mexico. (b) Temporal evolution of the low-latitude vertical stratification averaged between 20°S and 20°N in the Atlantic basin; the contours show potential density referenced to the surface. (c) Zonally averaged outcropping position of potential density surfaces in the Southern Ocean. A large freshwater flux, occurring at  $-14.2$  kyr BP, is prescribed in the Southern Ocean of this simulation. The dashed line indicates the abrupt increase in NADW formation, and the arrows show evidence of a lagged response after a multicentennial period.

key predictions of this model should be evident in more realistic general circulation models. Indeed, simulations of the overturning at the LGM, forced by reducing atmospheric temperatures, show an intensification of the low-latitude deep-ocean stratification and an intensification of the abyssal overturning cell (Jansen 2017; Marzocchi and Jansen 2017). While these experiments do not perturb NADW formation directly, a reduction in  $T_{\text{NADW}}$  is linked to a thickening of upper-ocean density classes and a weakening of the upper-ocean overturning cell. However, there is not a consensus between different models—for instance, PMIP3 models showed a deepening of the Atlantic meridional overturning circulation (AMOC) at the LGM (Muglia and Schmittner 2015). There are few simulations that have addressed the transient response of the global overturning, but output from the Simulation of the Transient Climate Evolution (TraCE) (Liu et al. 2009; He et al. 2013) does provide some support for the transient mechanisms and adjustment time scale discussed above (Fig. 8). While the forcing of the TraCE simulation

(Fig. 8a) is more complicated than an abrupt perturbation to NADW, the steady reduction in NADW formation between 19 and 17 kyr before present (BP) is associated with a deepening of the low-latitude stratification and is coupled to a poleward retreat of the outcropping positions in the Southern Ocean (Figs. 8b,c), as predicted by our proposed bipolar seesaw mechanism. A relatively abrupt increase in NADW at the Bølling–Allerød does not have an immediate response in either the low-latitude stratification or the Southern Ocean outcrop positions; in fact, for the latter, isopycnals continue to move poleward for over a century. After a period of roughly a couple hundred years, there is an increase in stratification in the northern basins and an equatorward shift of Southern Ocean outcrop positions associated with deep density classes (Figs. 8b,c, arrows). While these changes are not confirmation of our mechanism for the phasing of DO events, they are presented here to show that changes to the basin-scale stratification may occur over a period of a few hundred years. A longer response cannot be

tracked because of a large freshwater forcing in the south induced at 14.2 kyr BP.

Galbraith and de Lavergne (2019) carried out a comprehensive suite of simulations using an Earth system model that considers the climate's response to changes in atmospheric CO<sub>2</sub>, orbital forcing, and ice sheet size. A number of results from these simulations resonate with our proposed mechanism. Cold climate regimes and expanded Antarctic sea ice extent are linked to an increase in abyssal waters sourced from the Southern Ocean. The change in volume fraction is also strongly tied to NH–SH differences in surface temperature, salinity, and density. Galbraith and de Lavergne (2019) find that an enhanced NADW volume fraction (our NH warm phase) is associated with colder and fresher surface waters in the Southern Ocean, consistent with greater sea ice extent. However, these simulations consider steady configurations and not the transient adjustments across DO events that are the focus of this study. Calculations of water mass transformation, especially in the Southern Ocean, were not presented in this study but should be explored in the future.

## 5. Summary

An idealized ocean model is used to provide a mechanistic explanation for both the interhemispheric asymmetry and the multicentennial SH lag associated with DO events. We show that closure of the global overturning circulation requires a modification of the surface density distribution in the Southern Ocean in response to NADW perturbations. This constraint implies, in this model, that SH cooling (warming) occurs during periods of increased (reduced) NADW production. The transient adjustment of density surfaces spanning the Southern Ocean is critical to the mechanism and is represented in this study by a time-dependent version of the residual mean overturning in the Southern Ocean (Marshall and Radko 2003, T16). This model captures the coupling between upper and lower overturning cells, and while our results are consistent with changes to the circulation in the Atlantic basin during DO events (e.g., Gottschalk et al. 2015), we emphasize that a full-ocean response is needed to understand observed SH temperature variations. Residence times are shown to be sensitive to transient changes in NADW formation, and the degree to which this helps to explain the temporal evolution of  $\Delta^{14}\text{C}$  at various depths at the end of the last glacial (Skinner et al. 2017) will be explored in future work.

This mechanism explicitly accounts for the dynamics of outcropping density surfaces in an eddy-rich Southern Ocean and their link to surface buoyancy forcing imposed

by Antarctic sea ice. This mechanism marks a departure from earlier models (Stocker and Johnsen 2003; Pedro et al. 2018) that require a “reservoir” of heat to build up in the system. We argue that appealing to the buoyancy budget, rather than just the heat budget, is necessary to understand the interhemispheric propagation of climate signals by the ocean (Galbraith and de Lavergne 2019). While this model surely misses key dynamics involving the coupling between an evolving atmosphere and cryosphere, this study provides a dynamical description of the SH response to abrupt transitions in NADW formation rates, a response that has been coined the bipolar seesaw. This mechanism can and should be interrogated in more sophisticated ocean and coupled numerical simulations.

*Acknowledgments.* We thank Raffaele Ferrari and David Marshall for constructive comments to an earlier draft of this study and three anonymous reviewers for their insightful comments. We acknowledge helpful conversations with Paola Cessi, Malte Jansen, Brad Markle, Emily Newsom, Andrew Stewart, and Laure Zanna. We also thank Feng He for providing the TraCE simulation output. AFT received support from the David and Lucille Packard Foundation and from National Science Foundation (NSF) Grant OCE-1235488; SKH received support from NSF Grant P2C2-1503129 and the Lamont-Doherty Earth Observatory Postdoctoral Fellowship.

## REFERENCES

- Abernathy, R. P., I. Cerovečki, P. R. Holland, E. Newsom, M. Mazloff, and L. D. Talley, 2016: Water-mass formation by sea ice in the upper branch of the Southern Ocean overturning. *Nat. Geosci.*, **9**, 596–601, <https://doi.org/10.1038/ngeo2749>.
- Andersen, K. K., and Coauthors, 2004: High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature*, **431**, 147–151, <https://doi.org/10.1038/nature02805>.
- Armour, K., J. Marshall, J. R. Scott, A. Donohoe, and E. R. Newsom, 2016: Southern Ocean warming delayed by circumpolar upwelling and equatorward transport. *Nat. Geosci.*, **9**, 549–554, <https://doi.org/10.1038/ngeo2731>.
- Barker, S., P. Diz, M. J. Vautravers, J. Pike, G. Knorr, I. R. Hall, and W. S. Broecker, 2009: Interhemispheric Atlantic seesaw response during the last deglaciation. *Nature*, **457**, 1097–1102, <https://doi.org/10.1038/nature07770>.
- Bishop, S. P., P. R. Gent, F. O. Bryan, A. F. Thompson, M. C. Long, and R. Abernathy, 2016: Southern Ocean overturning compensation in an eddy-resolving climate simulation. *J. Phys. Oceanogr.*, **46**, 1575–1592, <https://doi.org/10.1175/JPO-D-15-0177.1>.
- Blunier, T., and E. J. Brook, 2001: Timing of millennial-scale climate change in Antarctica and Greenland during the last glacial period. *Science*, **291**, 109–112, <https://doi.org/10.1126/science.291.5501.109>.

- Broecker, W. S., 1991: The great ocean conveyor. *Oceanography*, **4**, 79–89, <https://doi.org/10.5670/oceanog.1991.07>.
- Buizert, C., and A. Schmittner, 2015: Southern Ocean control of glacial AMOC stability and Dansgaard–Oeschger interstadial duration. *Paleoceanogr.*, **30**, 1595–1612, <https://doi.org/10.1002/2015PA002795>.
- , and Coauthors, 2015a: Precise inter polar phasing of abrupt climate change during the last ice age. *Nature*, **520**, 661–664, <https://doi.org/10.1038/nature14401>.
- , and Coauthors, 2015b: The WAIS Divide deep ice core WD2014 chronology—Part 1: Methane synchronization (68–31 ka BP) and the gas age–ice age difference. *Climate Past*, **11**, 153–173, <https://doi.org/10.5194/cp-11-153-2015>.
- Cimadoribus, A. A., S. S. Drijfhout, and H. A. Dijkstra, 2014: Meridional overturning circulation: Stability and ocean feedbacks in a box model. *Climate Dyn.*, **42**, 311–328, <https://doi.org/10.1007/s00382-012-1576-9>.
- Crowley, T. J., 1992: North Atlantic deep water cools the Southern Hemisphere. *Paleoceanogr.*, **7**, 489–497, <https://doi.org/10.1029/92PA01058>.
- Curry, W. B., and D. W. Oppo, 2005: Glacial water mass geometry and the distribution of  $\delta^{13}\text{C}$  of  $\text{ECO}_2$  in the western Atlantic Ocean. *Paleoceanogr.*, **20**, PA1017, <https://doi.org/10.1029/2004PA001021>.
- Dansgaard, W., and Coauthors, 1993: Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, **364**, 218–220, <https://doi.org/10.1038/364218a0>.
- De Boer, A. M. D., and A. McC. Hogg, 2014: Control of the glacial carbon budget by topographically induced mixing. *Geophys. Res. Lett.*, **41**, 4277–4284, <https://doi.org/10.1002/2014GL059963>.
- de Lavergne, C., G. Madec, F. Roquet, R. M. Holmes, and T. J. McDougall, 2017: Abyssal ocean overturning shaped by sea-floor distribution. *Nature*, **551**, 181–186, <https://doi.org/10.1038/nature24472>.
- Ferrari, R., M. Jansen, J. F. Adkins, A. Burke, A. L. Stewart, and A. F. Thompson, 2014: An ocean tale of two climates: Modern and Last Glacial Maximum. *Proc. Natl. Acad. Sci. USA*, **111**, 8753–8758, <https://doi.org/10.1073/pnas.1323922111>.
- , A. Mashayek, T. J. McDougall, M. Nikurashin, and J.-M. Campin, 2016: Turning ocean mixing upside down. *J. Phys. Oceanogr.*, **46**, 2239–2261, <https://doi.org/10.1175/JPO-D-15-0244.1>.
- Galbraith, E., and C. de Lavergne, 2019: Response of a comprehensive climate model to a broad range of external forcings: Relevance for deep ocean ventilation and the development of late Cenozoic ice ages. *Climate Dyn.*, **52**, 653–679, <https://doi.org/10.1007/s00382-018-4157-8>.
- Gent, P. R., and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.*, **20**, 150–155, [https://doi.org/10.1175/1520-0485\(1990\)020<0150:IMIOCM>2.0.CO;2](https://doi.org/10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2).
- Gnanadesikan, A., 1999: A simple predictive model of the structure for the oceanic pycnocline. *Science*, **283**, 2077–2079, <https://doi.org/10.1126/science.283.5410.2077>.
- Gordon, A. L., 1986: Interocean exchange of thermocline water. *J. Geophys. Res.*, **91**, 5037–5046, <https://doi.org/10.1029/JC091iC04p05037>.
- Gottschalk, J., L. C. Skinner, S. Misra, C. Waelbroeck, L. Menviel, and A. Timmermann, 2015: Abrupt changes in the southern extent of North Atlantic Deep Water during Dansgaard–Oeschger events. *Nat. Geosci.*, **8**, 950–954, <https://doi.org/10.1038/ngeo2558>.
- Hallberg, R., and A. Gnanadesikan, 2006: The role of eddies in determining the structure and response of the wind-driven Southern Hemisphere overturning: Results from the Modeling Eddies in the Southern Ocean (MESO) project. *J. Phys. Oceanogr.*, **36**, 2232–2252, <https://doi.org/10.1175/JPO2980.1>.
- Haumann, F. A., N. Gruber, M. Münnich, I. Frenger, and S. Kern, 2016: Sea-ice transport driving Southern Ocean salinity and its recent trends. *Nature*, **537**, 89–92, <https://doi.org/10.1038/nature19101>.
- He, F., J. D. Shakun, P. U. Clark, A. E. Carlson, Z. Liu, B. L. Otto-Bliessner, and J. E. Kutzbach, 2013: Northern Hemisphere forcing of Southern Hemisphere climate during the last deglaciation. *Nature*, **494**, 81–85, <https://doi.org/10.1038/nature11822>.
- Henry, L. G., J. F. McManus, W. B. Curry, N. L. Roberts, A. M. Piotrowski, and L. D. Keigwin, 2016: North Atlantic Ocean circulation and abrupt climate change during the last glaciation. *Science*, **353**, 470–474, <https://doi.org/10.1126/science.aaf5529>.
- Hines, S. K. V., A. F. Thompson, and J. F. Adkins, 2019: The role of the Southern Ocean in abrupt transitions and hysteresis in glacial ocean circulation. *Paleoceanogr. Paleoclimatol.*, **34**, 490–510, <https://doi.org/10.1029/2018PA003415>.
- Jansen, M. F., 2017: Glacial ocean circulation and stratification explained by reduced atmospheric temperature. *Proc. Natl. Acad. Sci. USA*, **114**, 45–50, <https://doi.org/10.1073/pnas.1610438113>.
- , and L.-P. Nadeau, 2019: A toy model for the response of the residual overturning circulation to surface warming. *J. Phys. Oceanogr.*, **49**, 1249–1268, <https://doi.org/10.1175/JPO-D-18-0187.1>.
- Johnson, G. C., and H. L. Bryden, 1989: On the size of the Antarctic Circumpolar Current. *Deep-Sea Res.*, **36**, 39–53, [https://doi.org/10.1016/0198-0149\(89\)90017-4](https://doi.org/10.1016/0198-0149(89)90017-4).
- Johnson, H. L., and D. P. Marshall, 2002: A theory for the surface Atlantic response to thermohaline variability. *J. Phys. Oceanogr.*, **32**, 1121–1132, [https://doi.org/10.1175/1520-0485\(2002\)032<1121:ATFTSA>2.0.CO;2](https://doi.org/10.1175/1520-0485(2002)032<1121:ATFTSA>2.0.CO;2).
- , —, and D. A. J. Sproson, 2007: Reconciling theories of a mechanically driven meridional overturning circulation with thermohaline forcing and multiple equilibria. *Climate Dyn.*, **29**, 821–836, <https://doi.org/10.1007/s00382-007-0262-9>.
- Jones, C. S., and P. Cessi, 2016: Interbasin transport of the meridional overturning circulation. *J. Phys. Oceanogr.*, **46**, 1157–1169, <https://doi.org/10.1175/JPO-D-15-0197.1>.
- Kawase, M., 1987: Establishment of deep ocean circulation driven by deep-water production. *J. Phys. Oceanogr.*, **17**, 2294–2317, [https://doi.org/10.1175/1520-0485\(1987\)017<2294:EODOCD>2.0.CO;2](https://doi.org/10.1175/1520-0485(1987)017<2294:EODOCD>2.0.CO;2).
- Landais, A., and Coauthors, 2015: A review of the bipolar see-saw from synchronized and high resolution ice core water stable isotope records from Greenland and East Antarctica. *Quat. Sci. Rev.*, **114**, 18–32, <https://doi.org/10.1016/j.quascirev.2015.01.031>.
- Lippold, J., J. Grützner, D. Winter, Y. Lahaye, A. Mangini, and M. Christl, 2009: Does sedimentary  $^{231}\text{Pa}/^{230}\text{Th}$  from the Bermuda Rise monitor past Atlantic meridional overturning circulation. *Geophys. Res. Lett.*, **36**, L12601, <https://doi.org/10.1029/2009GL038068>.
- Liu, Z., and Coauthors, 2009: Transient simulation of last deglaciation with a new mechanism for Bølling–Allerød warming. *Science*, **325**, 310–314, <https://doi.org/10.1126/science.1171041>.
- Lumpkin, R., and K. Speer, 2007: Global ocean meridional overturning. *J. Phys. Oceanogr.*, **37**, 2550–2562, <https://doi.org/10.1175/JPO3130.1>.

- Lynch-Stieglitz, J., 2017: The Atlantic Meridional Overturning Circulation and abrupt climate change. *Annu. Rev. Mar. Sci.*, **9**, 83–104, <https://doi.org/10.1146/annurev-marine-010816-060415>.
- Markle, B. R., and Coauthors, 2017: Global atmospheric teleconnections during Dansgaard–Oeschger events. *Nat. Geosci.*, **10**, 36–42, <https://doi.org/10.1038/ngeo2848>.
- Marshall, D., 1997: Subduction of water masses in an eddying ocean. *J. Mar. Res.*, **55**, 201–222, <https://doi.org/10.1357/0022240973224373>.
- , and L. Zanna, 2014: A conceptual model of ocean heat uptake under climate change. *J. Climate*, **27**, 8444–8465, <https://doi.org/10.1175/JCLI-D-13-00344.1>.
- Marshall, J., and T. Radko, 2003: Residual-mean solutions for the Antarctic Circumpolar Current and its associated overturning circulation. *J. Phys. Oceanogr.*, **33**, 2341–2354, [https://doi.org/10.1175/1520-0485\(2003\)033<2341:RSFTAC>2.0.CO;2](https://doi.org/10.1175/1520-0485(2003)033<2341:RSFTAC>2.0.CO;2).
- , and K. Speer, 2012: Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nat. Geosci.*, **5**, 171–180, <https://doi.org/10.1038/ngeo1391>.
- Marzocchi, A., and M. F. Jansen, 2017: Connecting Antarctic sea ice to deep-ocean circulation in modern and glacial climate simulations. *Geophys. Res. Lett.*, **44**, 6286–6295, <https://doi.org/10.1002/2017GL073936>.
- McManus, J. F., R. Francois, J.-M. Gherardi, L. D. Keigwin, and S. Brown-Leger, 2004: Collapse and rapid resumption of Atlantic meridional circulation linked to deglacial climate changes. *Nature*, **428**, 834–837, <https://doi.org/10.1038/nature02494>.
- Meredith, M. P., and A. McC. Hogg, 2006: Circumpolar response of Southern Ocean eddy activity to a change in the southern annular mode. *Geophys. Res. Lett.*, **33**, L16608, <https://doi.org/10.1029/2006GL026499>.
- Miller, M. D., J. F. Adkins, D. Menemenlis, and M. P. Schodlok, 2012: The role of ocean cooling in setting glacial southern source bottom water salinity. *Paleoceanogr.*, **27**, PA3207, <https://doi.org/10.1029/2012PA002297>.
- Muglia, J., and A. Schmittner, 2015: Glacial Atlantic overturning increased by wind stress in climate models. *Geophys. Res. Lett.*, **42**, 9862–9868, <https://doi.org/10.1002/2015GL064583>.
- Munday, D. R., H. L. Johnson, and D. P. Marshall, 2013: Eddy saturation of equilibrated circumpolar currents. *J. Phys. Oceanogr.*, **43**, 507–532, <https://doi.org/10.1175/JPO-D-12-095.1>.
- Nikurashin, M., and G. Vallis, 2011: A theory of deep stratification and overturning circulation in the ocean. *J. Phys. Oceanogr.*, **41**, 485–502, <https://doi.org/10.1175/2010JPO4529.1>.
- , and —, 2012: A theory of the interhemispheric meridional overturning circulation and associated stratification. *J. Phys. Oceanogr.*, **42**, 1652–1667, <https://doi.org/10.1175/JPO-D-11-0189.1>.
- Orsi, A. H., G. C. Johnson, and J. L. Bullister, 1999: Circulation, mixing, and production of Antarctic Bottom Water. *Prog. Oceanogr.*, **43**, 55–109, [https://doi.org/10.1016/S0079-6611\(99\)00004-X](https://doi.org/10.1016/S0079-6611(99)00004-X).
- Pedro, J. B., M. Jochum, C. Buizert, F. He, S. Barker, and S. O. Rasmussen, 2018: Beyond the bipolar seesaw: Toward a process understanding of interhemispheric coupling. *Quat. Sci. Rev.*, **192**, 27–46, <https://doi.org/10.1016/j.quascirev.2018.05.005>.
- Pellichero, V., J.-B. Sallée, C. C. Chapman, and S. Downes, 2018: The Southern Ocean meridional overturning in the sea-ice sector is driven by freshwater fluxes. *Nat. Commun.*, **9**, 1789, <https://doi.org/10.1038/s41467-018-04101-2>.
- Polzin, K. L., J. M. Toole, J. R. Ledwell, and R. W. Schmitt, 1997: Spatial variability of turbulent mixing in the abyssal ocean. *Science*, **276**, 93–96, <https://doi.org/10.1126/science.276.5309.93>.
- Shakespeare, C. J., and A. McC. Hogg, 2012: An analytical model of the response of the meridional overturning circulation to changes in wind and buoyancy forcing. *J. Phys. Oceanogr.*, **42**, 1270–1287, <https://doi.org/10.1175/JPO-D-11-0198.1>.
- Skinner, L. C., and Coauthors, 2017: Radiocarbon constraints on the glacial ocean circulation and its impact on atmospheric CO<sub>2</sub>. *Nat. Commun.*, **8**, 16010, <https://doi.org/10.1038/ncomms16010>.
- Stocker, T. F., and S. J. Johnsen, 2003: A minimum thermodynamic model for the bipolar seesaw. *Paleoceanogr.*, **18**, 1087, <https://doi.org/10.1029/2003PA000920>.
- Sun, S., I. Eisenman, and A. L. Stewart, 2016: The influence of Southern Ocean surface buoyancy forcing on glacial–interglacial changes in the global deep ocean stratification. *Geophys. Res. Lett.*, **43**, 8124–8132, <https://doi.org/10.1002/2016GL070058>.
- Talley, L. D., 2013: Closure of the global overturning circulation through the Indian, Pacific, and Southern Oceans: Schematics and transports. *Oceanography*, **26**, 80–97, <https://doi.org/10.5670/oceanog.2013.07>.
- Thompson, A. F., A. L. Stewart, and T. Bischoff, 2016: A multi-basin residual-mean model for the global overturning circulation. *J. Phys. Oceanogr.*, **46**, 2583–2604, <https://doi.org/10.1175/JPO-D-15-0204.1>.
- Toggweiler, J. R., and B. Samuels, 1995: Effect of Drake Passage on the global thermohaline circulation. *Deep-Sea Res. II*, **42**, 477–500, [https://doi.org/10.1016/0967-0637\(95\)00012-U](https://doi.org/10.1016/0967-0637(95)00012-U).
- Walín, G., 1982: On the relation between sea-surface heat flow and thermal circulation in the ocean. *Tellus*, **34**, 187–195, <https://doi.org/10.3402/tellusa.v34i2.10801>.
- Waterhouse, A. F., and Coauthors, 2014: Global patterns of diapycnal mixing from measurements of the turbulent dissipation rate. *J. Phys. Oceanogr.*, **44**, 1854–1872, <https://doi.org/10.1175/JPO-D-13-0104.1>.
- Whitworth, T., III, A. H. Orsi, S.-J. Kim, W. D. Nowlin, and R. A. Locarnini, 1998: Water masses and mixing near the Antarctic Slope Front. *Ocean Ice and Atmosphere: Interactions at the Antarctic Continental Margin*, S. S. Jacobs and R. F. Weiss, Eds., Antarctic Research Series, Vol. 75, Amer. Geophys. Union, 1–27, <https://doi.org/10.1029/AR075p0001>.
- Wolfe, C. L., and P. Cessi, 2011: The adiabatic pole-to-pole overturning circulation. *J. Phys. Oceanogr.*, **41**, 1795–1810, <https://doi.org/10.1175/2011JPO4570.1>.
- , and —, 2015: Multiple regimes and low-frequency variability in the quasi-adiabatic overturning circulation. *J. Phys. Oceanogr.*, **45**, 1690–1708, <https://doi.org/10.1175/JPO-D-14-0095.1>.
- Wunsch, C., 2006: Abrupt climate change: An alternative view. *Quat. Res.*, **65**, 191–203, <https://doi.org/10.1016/j.yqres.2005.10.006>.