Pore Fluids and the LGM Ocean Salinity—Reconsidered

Carl Wunsch*
Department of Earth and Planetary Sciences
Harvard University
Cambridge MA 02138
email: cwunsch@fas.harvard.edu

December 1, 2015

Abstract

Pore fluid chlorinity/salinity data from deep-sea cores related to the salinity maximum of the last glacial maximum (LGM) are analyzed using estimation methods deriving from linear control theory. With conventional diffusion coefficient values and no vertical advection, results show a very strong dependence upon initial conditions at -100 ky. Earlier inferences that the abyssal Southern Ocean was strongly salt-stratified in the LGM with a relatively fresh North Atlantic Ocean are found to be consistent within uncertainties of the salinity determination, which remain of order ± 1 g/kg. However, an LGM Southern Ocean abyss with an important relative excess of salt is an assumption, one not required by existing core data. None of the present results show statistically significant abyssal salinity values above the global average, and results remain consistent, apart from a general increase owing to diminished sea level, with a more conventional salinity distribution having deep values lower than the global mean. The Southern Ocean core does show a higher salinity than the North Atlantic one on the Bermuda Rise at different water depths. Although much more sophisticated models of the pore-fluid salinity can be used, they will only increase the resulting uncertainties, unless considerably more data can be obtained. Results are consistent with complex regional variations in abyssal salinity during deglaciation, but none are statistically significant.

1 Introduction

McDuff (1985) pointed out that pore-waters in deep-sea cores have a maximum chlorinity (salinity) at about 30m depth owing to the sea level reduction during the last glacial period. He

*Also, Department of Earth, Atmospheric and Planetary Sciences; Massachusetts Institute of Technology

emphasized, however, the basic million-year diffusive time-scale of change in cores of lengths of several hundred meters. Schrag and DePaolo (1993) pioneered the interpretation of the data, focussing on δ^{18}O in the pore water, and noted that in a diffusion-dominated system, the most useful signals would be confined to about the last 20,000 years. Subsequently, Schrag et al. (1996, 2002), Adkins et al. (2002), Adkins and Schrag (2003; hereafter denoted AS03) analyzed pore water data to infer the ocean abyssal water properties during the last glacial maximum (LGM) including chlorinity (interpreted as salinity) and $\delta^{18}O_w$. (The w subscript is used to distinguish the values from $\delta^{18}O_c$ in the calcite structures of marine organisms.)

The latter authors started with the uncontroversial inference that a reduction in sea level of about $\Delta h = -125$ m in an ocean of mean depth $h = 3800$ m would increase the oceanic average salinity, $\bar{S}$, by and amount ΔS as,

$$\frac{\Delta S}{\bar{S}} = -\frac{\Delta h}{h} = \frac{125}{3800} \approx 0.03. \tag{1}$$

With a modern average salinity of about 34.7 g/kg, $\Delta S \approx 1.04$ g/kg for a global average LGM salinity of about 35.7 g/kg. By calculating the salinity profile as a function of core depth, they drew the now widely accepted inference that the abyssal LGM ocean contained relatively more salt—with values above the LGM global mean—than it does today. A Southern Ocean core produced calculated values exceeding 37 g/kg. (AS03, used a somewhat higher value of 35.85 g/kg for the LGM mean. The difference is unimportant in what follows.)

Those inferences, coupled with analogous temperature estimates from $\delta^{18}O_w$ (Schrag et al., 2002) that the deep ocean was near freezing, has widespread consequences for the oceanic state, carbon storage, deglaciation mechanisms (e.g., Adkins et al., 2005), etc. A salty, very cold, Southern Ocean abyss has become a quasi-fact of the subject (e.g., Kobayashi et al., 2015).

In the interim, a few other analyses have been published. Insua et al. (2014), analyzed core pore-fluid data in the Pacific Ocean and came to roughly similar conclusions. Miller (2014) and Miller et al. (2015), using a Monte Carlo method, carried out a form of inversion of the available pore water profiles and drew the contradictory inference that the data were inadequate for any useful quantitative conclusion about the LGM salinity or $\delta^{18}O_w$.

Determining the stratification of the glacial ocean and its physical and dynamical consequences is where paleo-physical oceanography meets sedimentology and core chemistry; see Huybers and Wunsch (2010). The purpose of the present note is to carry out a more generic study of the problem of making inferences from one-dimensional time-dependent tracer profiles. For maximum simplicity, only chlorinity/salinity data are discussed, with an analysis of $\delta^{18}O_w$ postponed to a second paper. The question being addressed is *whether the chlorinity data alone determine the ocean salt stratification during the LGM?* The papers already cited can be in-

Core No.	Reference	Location	Water Depth (m)
ODP981	Jansen et al. (1996)	NE Atlantic, Feni Drift/Rockall	2200
ODP1063	Keigwin et al. (1998)	Bermuda Rise	4600
ODP1093	Gersonde et al. (1999)	Southern Ocean, SW Indian Ridge	3600
ODP1123	Carter et al. (1999)	E. of New Zealand, Chatham Rise	3300
ODP1239	Mix et al. (2002)	E. Tropical Pacific, Carnegie Ridge/Panama Basin	1400

Table 1: Cores from which chlorinity/salinity data were used, along with a reference to their initial description in the Ocean Drilling Program (ODP) and with a geographical label. A nominal water depth of the core-top is also listed.

terpreted as asking whether, given other knowledge of the LGM, the chlorinity data contradict their picture of that time?

Conventional inverse methods derived from control theory are used: these have a more intuitive methodology and interpretation relative to those of the more specifically Bayesian Markov Chain Monte Carlo (MCMC) method of Miller et al. (2015). Although the MCMC method produces full probability densities for the results, interpretations almost always begin with the mean and variance, quantities emerging from the more conventional methods. As in the formal Bayesian approaches, prior knowledge with statements of confidence is both needed and readily used. (For modern physical oceanographers, parallels exist with understanding the establishment through time of the "abyssal recipes" formulation of Munk (1966) although the parameter ranges are far different.) The approaches here are those used by Wunsch (1988) for oceanic passive tracers and by MacAyeal et al. (1991) to infer temperatures in ice boreholes (and see MacAyeal, 1995) .

Profiles

To set the stage and to provide some context, Fig. 1 shows the positions of the cores discussed by AS03 plotted on a contour map of modern salinity at 3600 m depth. A bimodal histogram of those salinity values is shown in Fig. 2. Figs. 3 and 4 display the data available from five cores whose positions are shown on the chart. Considerable variation is apparent in both space (latitude and longitude) and time (that is, with water depth and depth in the core).

Consider the Southern Ocean core ODP 1093 (see Gersonde et al., 1999, and Figs. 3, 5) analyzed by AS03 and by Miller et al. (2015). It was this core that displayed the highest apparent salinity during the LGM and which led to the inference of a strongly salinity-stratified ocean dominated by Antarctic Bottom Waters. (Its position on top of a major topographic feature, Fig. 5, raises questions about the one-dimensionality of the core physics, but that

problem is not pursued here.) The overall maximum of about 35.7 g/kg perceptible in the core is somewhere between 50 and 70 m depth below the core-top and is plausibly a residual of high salinity during the LGM (a very large value near 400 m depth is assumed to be an unphysical outlier). Initially, only the top 100 m of the measured cores, Fig. 4, will be dealt with here. The main questions pertain to the magnitude and timing of the maxima and their interpretations. Very great differences exist in the water depths of the cores (Table 1) and the physical regimes in which they are located are today very different. Differences amongst the core salinity profiles are unsupportive of simple global-scale change.

In what follows, only the Atlantic Cores 1063 (about 4500 m water depth) and 1093 (about 3600 m water depth) will be discussed. Notice (Fig. 4) that the maximum salinity observed in Core 1093 in the upper 100 m is at best at, but not above, the estimated oceanic LGM global mean salinity maximum of 37.08 g/kg of AS03.

In Core 1063, at the northeastern edge of the Bermuda Rise, the apparent maximum occurs somewhere in the vicinity of 40 m core depth, with a value of about 35.35 g/kg—above the modern water mean—but well-below the LGM mean. Core 1063 was used to infer that the deep North Atlantic Ocean had not become as saline as in the Southern Ocean, and hence with the results from Core 1093, that the Southern Ocean had an extreme value of abyssal salinity, relative to the rest of the LGM ocean. Generally, the salinity of Core 1063 is lower than that of 1093 except near the core-top where the variability in Core 1093 precludes any simple statement.

Inferences from pore water profiles correspond to what in the control literature is known as a "terminal constraint" problem (e.g., Luenberger, 1979; Brogan, 1991; Wunsch, 2006—hereafter W06): In a physical system, the externally prescribed disturbances are sought that will take the system from a given initial state to a known, within error-bars, final state.[1] Here the physical

[1] A more intuitive analogue of this problem may be helpful, one based upon the terminal control problem for a conventional robotic arm. An arm, with known electromechanical response to an externally imposed set of control signals, has to move from a three dimensional position, $\boldsymbol{\xi}_0 \pm \Delta\boldsymbol{\xi}_0$, at time $t = 0$, to a final position $\boldsymbol{\xi}_{t_f} \pm \Delta\boldsymbol{\xi}_f$ at time t_f. In three-dimensions, there exists an infinite number of pathways between the starting and ending position, excluding only those that are physically impossible (such as a movement over a time-interval physically too short for transit between the two positions). Even if the trajectory is restricted to a straight line, there will normally be an infinite indeterminacy involving speed and acceleration. The control designer "regularizes" the problem by using a figure-of-merit e.g., by demanding the fastest possible movement, or the least energy requiring one, or minimum induced accelerations etc. The designer might know e.g., that the arm must pass close to some known intermediate position $\boldsymbol{\xi}_i \pm \Delta\boldsymbol{\xi}_i$ and which can greatly reduce the order of the infinity of possible solutions. In the case of the pore fluid, the initial "position" (initial pore fluid value, $c\,(z, t_f)$, is at best a reasonable guess, and no intermediate values are known. The assumed prior control represents an initial guess at what controlling signals can be sent, e.g., that a voltage is unlikely to exceed some particular value. The "identification" problem would correspond to the situation in which the model or plant describing the reaction of the robotic arm to external signals was partially uncertain and had to be determined by experiment. And perhaps the response would also

system is an assumed advection-diffusion one, with the initial state being the salinity profile far in the past (perhaps -100,000 y), the terminal state is represented by data from the measured core pore-water. The disturbances sought are the abyssal water salinity—providing a time-varying boundary condition at the sediment-water interface. Readers familiar with advection-diffusion problems will recognize their dependence upon a long list of knowns, including the initial and boundary conditions, and the advective flows and diffusion coefficients governing the time-depth evolution. In the present case, flows and diffusion coefficients are expected to display structures varying in both space (depth in core) and time. In the best situation, only their terminal values can be measured in the core. The problem is further compounded by the dependence of advection and diffusion on the time history of the solid phase in the sediment containing the pore waters. Finally, and a question also ignored here, is whether a handful of core values can be used to infer global or regional mean properties (see Figs. 1, 2) with useful accuracy.

For the time being, the problem is reduced to a basic skeletal framework to understand its behavior under the most favorable conditions.

2 Models

General discussion of pore fluid behavior in sediments can be found in Berner (1980), Boudreau (1997), Fowler and Yang (1998), Einsele (2000), Bruna and Chapman (2015) and in the papers already cited. Simplifying a complex subject, LGM pore fluid studies reduce the vertical profiles to the one-space-dimension governing equation, the canonical model,

$$\frac{\partial c}{\partial t} + w\frac{\partial c}{\partial z} - \frac{\partial}{\partial z}\left(k\frac{\partial c}{\partial z}\right) = 0 \tag{2}$$

Here k is an "effective" diffusivity, and w is a non-divergent vertical velocity within the core fluid relative to the solid phase. Note that if $\partial k/\partial z \neq 0$, the diffusion term breaks up into two parts, one of which is indistinguishable from an apparent advective term, $w^* = -\partial k/\partial z$, so that Eq. (2) is,

$$\frac{\partial c}{\partial t} + \left(w - \frac{\partial k}{\partial z}\right)\frac{\partial c}{\partial z} - k\frac{\partial^2 c}{\partial z^2} = 0. \tag{3}$$

k depends upon the porosity, ϕ, and the tortuosity, θ, of the sediment through relations such as,

$$\theta^2 = \phi^{1-\alpha},\ \alpha \approx 1.8,\ k \propto \phi/\theta^2. \tag{4}$$

An upward increasing porosity (Fig. 6 and Eq. 4) produces, an apparent effective $w^* = -\partial k/\partial z$. The oceanographers' convention that z is positive upwards is being adopted, but here the origin is

depend upon time, involving the changing mechanical configuration, as occurs for example, in controlling the trajectory of aging spacecraft.

at 100m below the sediment-water interface, and where a boundary condition must be imposed. Experiments (not shown) with k changing linearly by a factor of two showed little change from the constant k values.

The core fluid is visualized as being contained in a vertical "pipe," extending from $z = 0$ at the base of the pipe to $z = h$ at the sediment-water interface. At that interface, it is subject to a time-dependent boundary condition $C_h(t) = c(z = h, t),\ 0 \le t \le t_f$ representing (here) the salinity of the abyssal water and whose values through time are sought. The only available data are the measured profile at $t = t_f$ over a depth range $0 \le z < h(t_f)$ where t_f is the date at which a core was drilled. h is here taken as 100m above the origin. Information about the initial condition, $C_0(z) = c(z, t = 0)$, and the boundary condition at $z = 0$, may, depending upon parameters, be essential.

Sediment continues to accumulate and erode over the time history recorded in the core. Thus the sediment-water interface, $z = h$, is time-dependent, and perhaps monotonically increasing. A somewhat typical sedimentation rate (they vary by more than an order of magnitude) might be about 5 cm/ky= 1.6×10^{-12} m/s. Following Berner (1980) and Boudreau (1997), h is fixed to the moving sediment-water interface, meaning that the solid material directly exposed to the abyssal salinity would be 5 m displaced from the initial surface at the end of 100,000 y. The assumption is thus made that while the particulate material is displaced, the *fluid* in contact with the overlying sea water remains the same. With $h(t)$ taken as a fixed point, a corresponding 5 m error at the core-base, $h = 0$, is incurred, and will be ignored.

The canonical model omits a complex set of boundary layers just below and just above the interface at the sea floor (e.g., Dade et al., 2001; Voermans, 2016), which in principle are observable at the core top, and which would affect the boundary condition there. These too, are being ignored.

Scale Analysis–Orders of Magnitude

Before doing any specific calculations, obtaining some rough orders of magnitude is helpful. Although every core is different, the time interval of most interest here is the LGM, taken to end nominally at $t = t_f - 20,000$ y= $t_f - 6.3 \times 10^{11}$ s (20 ky BP) following which deglaciation begins. For several cores, AS03 estimated $k \approx 3 \times 10^{-10} \mathrm{m}^2/\mathrm{s}$ and Miller (2014) a value of $k \approx 2 \times 10^{-10}$ m^2/s. In the purely diffusive limit with $w \approx 0$, the e-folding diffusion time to reach the whole core depth is $L^2/k \approx 1.6 \times 10^6$ y, with the latter value of k. The e-folding diffusion decay time at any depth is l^2/k, where l^2 is the vertical length scale of any disturbance in the profile. For l =10 m, $l^2/k \approx 16,000$ y not far from the time interval since the LGM. Depending directly upon the analytical sensitivities and the space/time scales of interest, a 100 m core can retain a signature of some disturbances dating back more than 1 million years, consistent with McDuff's (1985)

Notation	Variable	Definition
Initial Condition	$C_0(z)$	$c(t=0,z)$
Boundary Condition	$C_h(t)$	$c(t,z=0)$
Terminal Condition	$C_{\mathrm{term}}(z)$	$c(t=t_f,z)$

Table 2: Notation used for initial, final and boundary conditions and for algebraic expressions. In the discrete form, two time-steps of the concentration c make up the state vector, $x(t)$, and corresponding imposed conditions, Tildes over variables denote estimates. Matrices are bold upper case letters, column vectors are bold lower-case letters.

inference. In the shorter term, and as noted by Schrag et al. (2002), in a purely diffusive system the depth and attenuated amplitude of a local maximum represent competing dependencies on k, with smalleer scale signals not surviving beyond about 20,000 y.

Should the vertical velocity, w, of fluid within the core become significant, additional time and space scales emerge—depending upon the sign of w; the position and amplitude of maxima are then no longer simply related. For $w > 0$, a boundary layer familiar from Munk (1966) of vertical scale, k/w, appears with an establishment time of k/w^2. Should $w \approx 2 \times 10^{-10}$ m/s= 6 m/thousand years, the vertical scale is 1 m, with an establishment time of $5{\times}10^9$ s or about 150 y. Two Péclet numbers appear, one based upon L, the other upon l. If $w < 0$, the advection time L/w is relevant and the combined k, w scales are unimportant.

In a number of published results the initial conditions at some time, $t = 0$ in the core, $C_0(z)$, are simply assumed to be of little influence in the interpretation of the final profile $c(z, t = t_f) = C_{\mathrm{term}}(z)$, with most attention focussed on determining the temporal boundary condition control, $C_h(t)$. Whether the initial conditions are unimportant (the signal having decayed away) or dominant, given the long time scales within the core, will depend upon the magnitudes of w, k, the sign of w, as well as the core length.

In principle, w, k values can be calculated from the available data and various hypotheses, both physical and statistical the—"identification problem." Introducing further unknowns into what will be perceived as an already greatly underconstrained problem, leads necessarily to even greater uncertainty in the estimates $\tilde{C}_h(t), 0 \le t \le t_f$, or of $\tilde{C}_0(z)$ (see Table 2 for notation; tildes are used to denote estimates.) The simplest problem, with known k, w, produces a lower bound uncertainty on the results. Should that lower bound be too large for use, solving the nonlinear estimation problem involving k, w as additional unknowns to be extracted from the same limited data would not be justified.

Analytical Reference Solutions

In the simplest case with $w = 0$, and k constant, a variety of analytical solutions to Eq. (2) is available. These are again useful for understanding the solution structure. As a representative calculation, set $w = 0$ and $k = 2 \times 10^{-10} \mathrm{m}^2/\mathrm{s}$ at zero-Péclet number. Let $C_h(t) = H(t)$, $H(t)$ being the unit step (Heaviside) function, be the upper boundary value, and let $C_0(z)$ be the initial conditions. Fig. 7 displays the profiles from the analytical solution (Carslaw and Jaeger, 1986, p. 101) calculated as a summation here over 100 terms of a weighted cosine series, as a function of time,

$$c^{(1)}_{anal}(z,t) = 1 + \frac{2}{L}\sum_{n=0}^{\infty}\left\{\frac{2L(-1)^{n+1}}{(2n+1)\pi} + \int_0^L C_0(z')\cos\left(\frac{(2n+1)\pi z'}{2L}\right)dz'\right\}\times \qquad (5)$$
$$e^{-k(2n+1)^2\pi^2 t/4L^2}\cos\left(\frac{(2n+1)\pi z}{2L}\right),$$

for a unit amplitude surface boundary condition, zero flux at the bottom, and $C_0(z) = 0$. The terminal profile is also shown. Weighting, $\exp\left(-k(2n+1)^2\pi^2 t/4L^2\right)$, connects the dissipation rate to the vertical structure present in the solution,[2] and which rapidly removes even moderately high wavenumber, $(2n+1)\pi/2L$, structures whether present in the initial conditions, or emanating from the boundary condition (here, with zero initial conditions, only the boundary step-function gives rise to high wavenumbers). Fig. 8 shows the decay time to 1% of the initial value as a function of vertical scale in the initial conditions or those induced by the boundary conditions. Vertical scales shorter than about 12 m will have decayed by 99% after 20,000 years and need not be considered with this value of k.

The equivalent solution for zero initial conditions and a periodic surface boundary condition, $C_0(t) = \sin(\sigma t)$ is (Carslaw and Jaeger, 1986, p. 105),

$$c^{(2)}_{anal}(z,t) = \left\{\frac{\cosh 2(\sigma/2k)^{1/2} z - \cos 2(\sigma/2k)^{1/2} z}{\cosh 2(\sigma/2k)^{1/2} L - \cos 2(\sigma/2k)^{1/2} L}\right\}^{1/2}\sin(\sigma t + \varphi) + \qquad (6)$$
$$2\pi k\sum_{n=0}^{\infty}\frac{n(-1)^{n+1}\sigma L^2}{k^2 n^4\pi^4 + \sigma^2 L^4}e^{-kn^2\pi^2 t/L^2}\sin\left(\frac{n\pi z}{L}\right)$$
$$\varphi = \arctan\left\{\frac{\sinh\left[(\sigma/2k)^{1/2} z(1+i)\right]}{\sinh\left[(\sigma/2k)^{1/2} L(1+i)\right]}\right\}$$

where here the lower boundary condition is $c(0) = 0$ and placed at $z = -500$ m. The first term is the steady-state sinusoidal profile, whose amplitude is shown in Fig. 10a as a function of depth for varying σ and the above values of k, L. The second term is the starting transient with

[2] The scale used above, L^2/k, describes the lowest wavenumber response.

decay times shown in Fig. 10b. If the LGM were regarded as part of a quasi-periodic signal with the obliquity period of about 40,000 y, the signal would not penetrate much below 50 m. Even at 100 ky periods, no measurable signal reaches the base of a 100 m core.

Analytical solutions also exist for the case $w \neq 0$, but are not displayed here (see Wunsch, 2002 for references).

3 Representative Model Solutions

An axiom of inverse methods is that full understanding of the forward problem is a necessary preliminary. In a conventional forward calculation, solutions depend directly and jointly upon all of:

(1) The initial conditions, $C_0(z)$,

(2) The top boundary condition here, $c(z = h, t) = C_h(t)$

(3) The bottom boundary condition involving $c(z = 0, t)$ and/or its derivatives (here always a no-flux condition)

(4) Physical parameters, w, k.

Conventionally, these values are all perfectly known with the solution changing if any of them does.

In practice in the present case, only the final state of the solution, $C_{term}(z) = c(t_f, z)$, is approximately available. The inverse problem involves making inferences about the state, $c(z, t)$, $0 \leq t \leq t_f$, $0 \leq z \leq h$, initial and boundary conditions, and the w, k parameters from the limited supply of information. Use of prior information (assumptions) with statements of confidence becomes crucial. If incorrectly formulated, so-called inverse solutions to diffusive systems can become extremely unstable and demonstrably stable methods are required.

AS03 and subsequent authors have suggested that a good prior estimate of the boundary conditions on all cores consists of making $C_h^{\text{a priori}}(t)$ proportional to the best-estimated sea level curve. Fig. 11 shows the calculated global mean salinity over 120,000 years (Miller, 2014; Miller, et al., 2015) from a number of sources (their Table 1), and Eq. (1). Between about $-70,000$ y $< t < -25,000$ y, values higher than 35 g/kg are estimated owing to the reduction in sea level, reaching a maximum at a sea level minimum near $t = -20{,}000$ y. After that, the deglaciation leads to an estimated fall.[3]

[3] Adkins and Schrag (2003) used a considerably more structured estimated sea level curve. But because prior to -20 ky it was based upon measured δ^{18}O, which is one of the tracers under consideration in these cores Miller (2014), Miller et al. (2015) chose to avoid any possibility of circular reasoning. Much of the small scale structure present in the former curve would not survive the difffusive process in the core. Large scale structures are qualitatively the same in both approaches.

Thus, following the previous literature, the top boundary condition is, for now, assumed to be $C_h^{\text{a priori}}(t)$ and the bottom boundary condition to be one of zero diffusive flux. Initial conditions are problematic. The purely diffusive numerical calculation and the analytic solutions both show that disturbances at the surface will not penetrate significantly below about 40 m depth in 100,000 years. Structures in Figs. 3, 4 below that depth cannot have arisen from the core surface in the last 100 ky. At least four possibilities suggest themselves: (1) The structures are simply the noise in the core from measurements (see AS03 for of the technicalities and difficulties of shipboard measurements) or from processes not yet included in the model (time-space-dependent w, k, or clathrate formation, for example). (2) The structures arise from the memory of the initial conditions. (3) A purely diffusive model is inadequate. (4) The structures are the result of upward diffusion/advection across the base of the core, $z = 0$, noting in particular that $\partial C(0)/\partial z$ in general does not vanish. All of these possibilities could be at work.

The simplest interpretation of the solutions discussed by AS03 and others is based on assumption (1): that all structures other than the deep overall maximum represent errors in the data, and that *only the gross maximum feature must be reproduced*. In contrast, a more agnostic approach is taken here, in which an attempt is made to understand the extent to which some or all of the additional core-data features can be regarded as signals. For example, if structures in the initial conditions can persist in the core, they should be visible at the terminal time. Some of the published solutions have taken the sensible approach of maximum ignorance, and set $C_0(z) =$constant, where the constant might be the modern mean salinity. In that situation, either all of the terminal structure arises from $C_h(t)$, and/or non-uniform initial conditions are nonetheless also required by the terminal data. Another possibility is based upon the description of the glacial-interglacial cycles as being quasi-periodic, with glaciations recurring at intervals lying between 80,000 and 120,000 y, leading to a second plausible hypothesis that the initial condition at $t = -100$ ky is close to the observed terminal profile of the individual core (Fig. 3). Except where specifically stated otherwise, this quasi-periodic condition, but with different uncertainty estimates applied to the initial and terminal data, is used throughout this study. The initial condition uncertainty is always larger than the terminal one. A similar initial condition (set at -125 ky) was used by Miller et al. (2015).

4 Terminal Constraint Inversions

4.1 Numerical Model

Eq. (2) is now rendered into discrete form, simplifying the representation of noise processes. Discretization can be done in a number of ways. Here we follow Roache (1976; cf. Wunsch, 1987) in the use of what Roache calls the Dufort-Frankel leapfrog method. A uniform vertical grid, at spacing Δz indexed in $0 \leq i \leq N-1$, and at time intervals Δt produces,

$$c_i\left(t+\Delta t\right) = \alpha c_i\left(t-\Delta t\right) + \beta c_{i-1}\left(t\right) + \delta c_{i+1}\left(t\right), \tag{7}$$

$$\alpha = 1/\left(1+2d\right),\ \beta = 1/\left(1+2d\right)\left(2d+\bar{e}\right),\ \delta = 1/\left(1+2d\right)\left(2d-\bar{e}\right) \tag{8}$$

with $d = k\Delta t/\Delta z^2$, $\bar{e} = w\Delta t/\Delta z$ plus the boundary conditions, $c_N\left(t\right) = C_h\left(t\right)$, $c_1\left(t\right) - c_0\left(t\right) = 0$. The latter is an assumed no flux condition at the base of the core. Stability requires $d < 0.5$. Defining a state vector, $\mathbf{x}\left(t\right) = \left[\mathbf{c}\left(t-\Delta t\right), \mathbf{c}\left(t\right)\right]^T$ involving two time-steps, Eq. (7) has the canonical form (e.g., Stengel, 1986; Brogan, 1991; W06),

$$\mathbf{x}\left(t+\Delta t\right) = \mathbf{A}\mathbf{x}\left(t\right) + \mathbf{B}\mathbf{q}\left(t\right) + \mathbf{\Gamma}\mathbf{u}\left(t\right) \tag{9}$$

(Notation is that bold capitals are matrices, and bold lower case letters are column vectors.) For this particular discretization,

$$\mathbf{A} = \left\{\begin{array}{cccccc|cccccc} & & \mathbf{0}_{N-1} & & & & & & \mathbf{I}_{N-1} & & & \\ 0 & 0 & . & . & 0 & & 0 & 0 & . & . & 0 & \\ \alpha & 0 & 0 & 0 & . & 0 & 0 & \delta & 0 & 0 & . & 0 \\ 0 & \alpha & 0 & . & . & 0 & \beta & 0 & \delta & 0 & . & 0 \\ 0 & 0 & \alpha & 0 & . & 0 & 0 & \beta & 0 & \delta & . & 0 \\ . & . & . & . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & \alpha & 0 & 0 & 0 & 0 & 0 & . & \delta \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right\}$$

$$\mathbf{x}\left(t\right) = \left[c_1\left(t-\Delta t\right), c_2\left(t-\Delta t\right), ..., c_N\left(t-\Delta t\right), c_1\left(t\right), c_2\left(t\right), ..., c_N\left(t\right)\right]^T$$

The dimension of square matrix $\mathbf{A}$ is $2N \times 2N$ because of the need to carry two time-levels. Row $N+1$ forces an assumed no-flux condition at the bottom, and row $2N$ is all zeros, because the boundary condition is set at that grid point by putting $\mathbf{B}_{2N} = 1$, $\mathbf{B}_j = 0$, otherwise (here $\mathbf{B}$ is a $2N \times 1$ column vector) and $\mathbf{q}\left(t\right) = q\left(t\right)$ is the imposed scalar $C_h^{\text{a priori}}\left(t\right)$. In a conventional calculation of $\mathbf{x}\left(t\right)$, $\mathbf{\Gamma}$ is set to zero, but along with $\mathbf{u}\left(t\right)$ reappears in the inverse or control calculations, representing the controls when elements of $\mathbf{q}$ are regarded as unknown. $\mathbf{u}\left(t\right) = u\left(t\right)$, a scalar, with scalar variance $Q\left(t\right)$. t is now always a discrete value.

As a demonstration of the numerical model, let the time-step be $\Delta t = 4 \times 10^9$ s (127 years), $L = 100$ m, and Δz =1 m, $k = 2 \times 10^{-10}$ m^2/s, $w = 0$, for the Heaviside boundary condition, $q(t) = 1, 0 \leq t \leq t_f$ and zero initial condition, with result shown in Fig. 7 and the terminal state compared to the analytic solution. Consistent with the scale analysis and the analytic solution, the signature of the surface boundary condition has not reached beyond about 50 m after 100,000 years.

Now consider the quasi-periodic initial condition with the same w, k. Fig. 9 shows the results after 100 ky of forward model integration. The smallest scales present in the initial condition have vanished—as expected. However, much of the intermediate and largest structures at the terminal time originate with the initial conditions. In contrast, also shown is the state when $w = -|k|$. In that situation, the initial conditions are swept downward, out of the core, before the terminal time.

When $w \neq 0$, qualitative changes in the solutions occur. With $w > 0$, confinement of disturbances from $C_h^{\text{a priori}}(t)$ towards the surface is even more marked than for the purely diffusive case. When $w < 0$, structures in $C_h(t)$ can be carried much further down into the core than otherwise. The magnitude and sign of w thus become major issues.

4.2 Inversions/Control Solutions

Miller (2014) also discussed the linear time-dependent inverse problem of determining $u(t)$, the modification to $q(t) = C_h^{\text{a priori}}(t)$ and chose to solve it by "Tikhonov regularization." (Here both $\mathbf{q}$ and $\mathbf{u}$ are scalars—a special case). Although that method is a useful one for deterministic problems, it does not lend itself to a discussion of data and model error, nor of the uncertainties of the results owing to noise. Determining a best-solution involves not only the core physical properties and time-scales, but also the analytical accuracies, and systematic down-core errors.

Consider the problem of determining $C_h(t) = q(t) + u(t)$ (that is, $u(t)$) from the terminal values $\mathbf{x}(t_f)$, which involves assuming here that $\mathbf{c}(t_f - 1) = \mathbf{c}(t_f) = C_{term}(z)$. Now, $\mathbf{\Gamma} = \mathbf{B}$, and $u(t)$ is sought. Several standard methods exist. One approach is to explicitly write out the full set of simultaneous equations governing the system in space and time, recognizing that the only information about the state vector are its final values, $\mathbf{x}(t_f)$, and a guessed initial condition, $\mathbf{x}(0)$. In practice, neither will be known perfectly, and a covariance of the error in each is specified, here called $\mathbf{P}(0)$ and $\mathbf{R}(t_f)$ respectively. Writing out the full suite of governing

equations, setting $\Delta t = 1$ for notational simplicity but with no loss of generality,

$$\begin{aligned}
\mathbf{x}(0) + \mathbf{n}(0) &= \mathbf{C}_0(z) \\
\mathbf{x}(1) - \mathbf{A}\mathbf{x}(0) - \mathbf{\Gamma}u(0) + \mathbf{n}(1) &= \mathbf{B}q(0) \\
\mathbf{x}(2) - \mathbf{A}\mathbf{x}(1) - \mathbf{\Gamma}u(1) + \mathbf{n}(2) &= \mathbf{B}q(1) \\
\mathbf{x}(3) - \mathbf{A}\mathbf{x}(2) - \mathbf{\Gamma}u(2) + \mathbf{n}(3) &= \mathbf{B}q(2) \\
&\cdots \\
-\mathbf{A}\mathbf{x}(t_f - 1) - \mathbf{\Gamma}u(t_f - 1) + \mathbf{n}(t_f) &= -\mathbf{C}_{\text{term}}(z) + \mathbf{B}q(t_f - 1),
\end{aligned} \tag{10}$$

where all unknowns are on the left of the equals sign, and all known fields are on the right. Vectors $\mathbf{n}(t)$ represent the presence of errors in the starting and ending profiles and their propagation through the system. The $\mathbf{\Gamma}u(t)$ terms are the controls and which, more generally, include the model error, but here are specifically accounting only for the uncertainties in $\mathbf{B}q(t), q(t) = C_h^{\text{a priori}}(t)$. Equations (10) are a set of linear simultaneous equations which is, however, extremely sparse; unless t_f or N become very large, they can be solved by several methods for dealing with underdetermined systems. This route is not pursued here, but the existence of the set shows that any other method of solution is equivalent to solving it, and which can help greatly in the interpretation. The special structures present in the equations permit rapid and efficient solution algorithms not requiring explicitly inverting the resulting very large, albeit very sparse, matrix (a generalized-inverse would be involved in practice), and which is the subject of the next sections.

5 Lagrange Multipliers-Pontryagin Principle

5.1 Formulation

One approach uses ordinary least-squares and Lagrange multipliers to impose the model (Eq. 9) with an error represented by the controls, and minimizing the weighted quadratic misfit between the calculated value of $\mathbf{x}(0)$ and $\mathbf{x}_0$ and between the calculated $\mathbf{x}(t_f)$ and $\mathbf{C}_{term}$,

$$J = (\tilde{\mathbf{x}}(0) - \mathbf{C}_0(z))^T \mathbf{P}(0)^{-1} (\tilde{\mathbf{x}}(0) - \mathbf{C}_0(z)) + (\tilde{\mathbf{x}}(t_f) - \mathbf{C}_{\text{term}})^T \mathbf{R}(t_f)^{-1} (\tilde{\mathbf{x}}(t_f) - \mathbf{C}_{\text{term}}) + \tag{11}$$

$$\sum_{t=0}^{t_f-1} \tilde{\mathbf{u}}(t)^T \mathbf{Q}^{-1} \tilde{\mathbf{u}}(t).$$

respectively. Tildes denote estimates, but are sometimes omitted where the context makes clear what is being described. The third term renders the problem fully determined as a constrained

least-squares problem, by simultaneously minimizing the weighted mean square difference between $u(t)$ and its prior value (here written as zero), and with a result that is a form of the "Pontryagin Principle." The figure of merit in the L_2 norm attempts to minimize the mean square deviation of $u(t)$ from the prior, which as written here is zero, while simultaneously minimizing the squared difference from the assumed initial and final conditions in what is just a form of least-squares. (Other figures-of-merit such as maximum smoothness can be used. The problem can be reformulated too, using different norms such as L_1 or L_∞; see the references.) Because the system of equations (10) has a special block structure, a closed form solution can be obtained (W06, P. 218+, or the Appendix here) and which makes explicit the relationships between the initial and final states, and the control, all of which are subject to modification.

5.2 Using Lagrange Multipliers

With $w = 0$, pure diffusion, and the quasi-periodic initial condition taken from Core 1093, an integration is started at $t = -100,000$ y using the $C_h^{\text{a priori}}(t)$ in Fig. 11 with result shown in Fig. 9. Although a rough comparability to the core values occurs in the top 10-20m, they diverge qualitatively below that depth, both in the large-scale structures and in the high wavenumbers apparent in the core data. The first question to be answered is whether it is possible to modify $C_h^{\text{a priori}}(t)$ within acceptable limits, $\pm\sqrt{Q}$ so as to bring the two terminal profiles together within estimated error?

The second immediate question is whether the smaller scale structures in the core data are real structures or noise (issue (1) above)? Assume that they are uncorrelated white noise of RMS amplitude approximately 0.1 g/kg, and allowing the control $u(t)$ to have the possible large RMS fluctuation of 1 g/kg . The initial conditions are assumed to have a white noise (in depth) RMS error of 1.7 g/kg, the terminal data RMS uncertainty is 0.1 g/kg, and the result is shown in Fig. 12. Assuming that none of the structures visible in the core data, except the maximum in the vicinity of 40m from the core bottom, are just noise, this solution is a qualitatively acceptable one. AS03 noted, that their solutions *above* the maximum in depth did not produce a good fit. Much of the terminal state here is controlled by the initial conditions, not $C_h(t)$, except for the last few thousand years in the very upper parts of the core.

Introducing $w > 0$ exacerbates the confinement of the core-top disturbances to the upper core placing even more emphasis on the initial conditions.. On the other hand, permitting $w < 0$, here $w = -|k|$ m/s succeeds in producing a slightly better fit overall (see Fig. 14) and increases the sensitivity to $C_h(t)$ (using the same prior statistics, held fixed throughout this paper). The modification required to $C_h^{\text{a priori}}(t)$ is also shown along with the resulting total $q(t) + \tilde{u}(t)$ .This result decreases the maximum salinity estimated to 35.75 g/kg and delays its

timing to about -12,000 y, and is followed by a large variability. Notice that the estimated maximum salinity again lies *below* the LGM average—implying high salinities elsewhere. This solution is also a formally acceptable one, and if taken at face value, moves the salinity maximum several thousand years before that in the prior, and still below the LGM mean. The central question at the moment is whether any of the variations in $C_h(t) = q(t) + \tilde{u}(t)$ are significant? Further discussion of this result is postponed pending the calculation of its uncertainty.

The large negative value of w or w^*—required to carry information downward from the core top before diffusion erases the observed structures—is counter to the conventional wisdom that the appropriate model is nearly purely diffusive. No claim is made that the model here is "correct," only that if the magnitude of w is much smaller, or that it is positive upwards, then the canonical model cannot explain *any* of the pore-water salinity properties below about 20 m unless they originate in the initial conditions. On the other hand, the physics of fluid-solid interaction through hundreds of thousands of years is sufficiently unclear (see the references already cited) that ruling out large negative w is premature, particularly in partially saturated cores where the effects of sea level-induced pressure changes of hundreds of meters of water have not been accounted for. Violation of any of the other basic assumptions, including especially, that of a one-dimensional-space behavior, could render moot the entire discussion.

The Lagrange multiplier formalism does permit an affirmative answer to the question of whether a model can be fit to the top 100m of the core data within a reasonable error estimate? The stable flow of information, nominally "backwards" in time from the terminal state is particularly apparent (Eq. A1) via the transposed matrix $\mathbf{A}^T$ (the "adjoint matrix"). But it neither answers the question of whether this model is "correct" (or "valid" in modelling jargon), or if the model is nonetheless assumed correct, how uncertain is the estimate, $\tilde{C}_h(t) = q(t) + \tilde{u}(t)$? We next turn to this latter question.

6 Smoothers

A great advantage of the Lagrange multiplier approach is that it is computationally very efficient, not involving calculation of the uncertainty of $\tilde{\mathbf{u}}(t)$, (the adjustment to $C_h^{\text{a priori}}(t)$) nor of the intermediate time values of the profiles in $\tilde{\mathbf{x}}(t)$. On the other hand, the absence of those uncertainties is the greatest weakness of the estimated state and controls in problems such as this one. The need to find formal uncertainties leads to the alternative approach based upon the idea of "smoothers", which are recursive estimation methods for calculating the state and control vectors using data from a finite time-span. Several different smoothing algorithms exist depending upon the particular need. Perhaps the easiest to understand is the so-called RTS

(Rauch-Tung-Striebel) algorithm which involves two-passes through the system in time.

To start the RTS algorithm, a prediction algorithm known as the Kalman filter is used, beginning with the initial conditions and their uncertainty, employing the model (Eq. 9) to predict the state at the next time when more observations become available (perhaps many time-steps into the future). By weighting the prediction inversely to its uncertainty and the observations inversely to their errors, a new estimate is made combining the values appropriately, and determining the covariance matrix of the new combined estimate. With that new estimate, further predictions are made to times of new data. (Note that the state estimate jumps every time a new model-data combination is made, meaning that at those times the model evolution equation fails.) After arriving at the final data time, t_f, another algorithm is used to step backwards in time to $t = 0$, using the later-arriving data to correct the original predicted and combined values of $\mathbf{x}(t)$, and estimating the control vector $u(t)$ necessary to render the model exactly satisfied at *all* time-steps. Uncertainty estimates are required in the calculation for both state vector and control.

Because several covariance matrices are square of the dimension of $\mathbf{x}(t)$, for large systems the computational load can become enormous. Calculating the error covariance matrix of the state predicted by the Kalman filter is equivalent to running the model N^2 times at every time step, and which is why true Kalman filters and related smoothers are never used in real atmospheric or oceanic fluid systems. Nonetheless, in the present context, realistic calculations are feasible on modest computers. The state vector solution from the Lagrange multiplier method and from the RTS smoother can be shown to coincide (e.g., W06, P. 216) and the uncertainties may be of little interest as long as the controlling solution is physically acceptable.[4]

6.1 Using the Filter-Smoother

Consider again Core 1093. The Lagrange multiplier method shows that with $w = 0$ or $-|k|$, consistency can be found within varying estimated errors between the model and the measured terminal state. Those solutions, which minimize the square difference from $C_h^{\text{a priori}}(t)$, are not unique, and as in least-squares generally, an infinite number of solutions can exist, albeit with all others having a larger mean-square. The question to be answered is what the uncertainty of any particular solution is, given the existence of others? To do so, the filter-smoother algorithm is now invoked.

[4] For example, in operating a vehicle such as an aircraft, that a useful control exists may be the only concern, and with its non-uniqueness being of no interest.

6.2 The Filter Step

With the same initial condition and $C_h^{\text{a priori}}(t)$ as before, the model is run forward, one time step of 4×10^9s (127 y), from Eq. (9) as before, but with a slightly different notation,

$$\tilde{\mathbf{x}}(t+\Delta t,-) = \mathbf{A}\tilde{\mathbf{x}}(t,-) + \mathbf{B}q(t) + \mathbf{\Gamma}\tilde{u}(t), \tag{12}$$

with the minus sign showing that no data have been used in the model prediction one time-step into the future. This prediction based upon the state estimate at the previous time, and $\mathbf{B}q(t)$ set by $C_h^{\text{a priori}}(t)$. For now, $\mathbf{\Gamma}\tilde{u}=\mathbf{0}$. Simple algebra shows that the error covariance (uncertainty) of this one-step prediction is,

$$\mathbf{P}(t+\Delta t,-) = \mathbf{A}\mathbf{P}(t)\mathbf{A}^T + \mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^T, \tag{13}$$

where the first term arises from errors in the state estimate, $\tilde{\mathbf{x}}(t)$, and the term in $\mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^T$ represents the error from the unknown deviation, $u(t)$, from $q(t)$. The estimated prior covariance $\mathbf{Q}$ is here being treated as time-independent, and is also a scalar, Q. The progression is started with the given $\mathbf{P}(0)$. $\mathbf{P}(t+\Delta t,-)$ is the uncertainty at time t if no data at $t+\Delta t$ are used, and if no data are available then, $\mathbf{P}(t+\Delta t) = \mathbf{P}(t+\Delta t,-)$.

Let there be a time t' when measurement of the full profile is available, written for generality as,

$$\mathbf{y}(t') = \mathbf{E}\mathbf{x}(t') + \mathbf{n}(t'). \tag{14}$$

With a full profile observation, $\mathbf{E}=\mathbf{I}$, the identity matrix. $\mathbf{n}(t')$ is the zero-mean noise in each profile measurement, and with error covariance $\mathbf{R}(t')$. Evidently, at t', two estimates of the state vector, $\mathbf{x}(t')$ can be made: $\tilde{\mathbf{x}}(t',-)$ from the model prediction, and $\tilde{\mathbf{x}}_{\text{data}}(t') = \mathbf{E}^{+}\mathbf{y}(t')$, where $\mathbf{E}^{+}$ is a generalized-inverse of $\mathbf{E}$, but here is the identity, $\mathbf{I}$. Their corresponding uncertainties are $\mathbf{P}(t',-)$, and $\mathbf{R}(t')$. The gist of the Kalman filter is to make an improved estimate of $\tilde{\mathbf{x}}(t')$ by using the information available in these two (independent) estimates. With a bit of algebra (see any of the references), the best new estimate is the weighted average,

$$\tilde{\mathbf{x}}(t') = \tilde{\mathbf{x}}(t',-) + \mathbf{K}(t')\left[\mathbf{y}(t') - \mathbf{E}\tilde{\mathbf{x}}(t,-)\right], \tag{15a}$$

$$\mathbf{K}(t') = \mathbf{P}(t',-)\mathbf{E}^T\left[\mathbf{E}\mathbf{P}(t',-)\mathbf{E}^T + \mathbf{R}(t')\right]^{-1} \tag{15b}$$

and the new combined estimate has an uncertainty covariance matrix,

$$\mathbf{P}(t') = \mathbf{P}(t',-) - \mathbf{K}(t')\mathbf{E}(t')\mathbf{P}(t',-) \tag{16}$$

(variant algebraic forms exist). In the absence of data at t', $\tilde{\mathbf{x}}(t') = \tilde{\mathbf{x}}(t',-)$; $\mathbf{P}(t') = \mathbf{P}(t',-)$ because no new observational information is available. In this linear problem, Eqs. (13, 16)

are independent of the state $\mathbf{x}(t)$, and the uncertainties can be determined without calculating $\tilde{\mathbf{x}}(t)$ (and which is already available from the Lagrange multiplier solution).

In the present situation, only one time, the last one, exists where observations are available. Thus the model is run forward from the assumed initial conditions and two boundary conditions, making a prediction of $\tilde{\mathbf{x}}(t' = t_f, -)$, using the predicted profile from Eq. (12), along with an estimate of the error of that prediction (Eq. 13). Then from the weighted averaging in Eq. (15), a final profile is determined that uses both the information in the a priori model and in the data, paying due regard to their uncertainties.

6.3 The Smoothing Step

The Kalman filter is seen to be an optimal[5] *predictor* and, contrary to widespread misinterpretation, is not a general purpose estimator. The intermediate state $\tilde{\mathbf{x}}(t) = \tilde{\mathbf{x}}(t, -)$, $t < t_f$, is estimated without using any knowledge of the observed terminal profile and so cannot be the best estimate. $\tilde{\mathbf{x}}(t)$ does not satisfy the governing equation at the times when the predicted estimate and that from the data are combined and $\tilde{\mathbf{u}}(t)$ is not yet known. Thus in this particular algorithm (others exist, including direct inversion of the set, Eqs. (10)), the filter step is followed by the RTS algorithm, as written out in the Appendix and in the references. The calculation steps backward in time from the final, best estimate $\tilde{\mathbf{x}}(t_f)$ and its uncertainty, $\mathbf{P}(t_f)$, comparing the original $\tilde{\mathbf{x}}(t_f - \Delta t)$ (Eq. 15) and the prediction made from it, with the improved estimate now available at one time step in the future, which is both $\tilde{\mathbf{x}}(t_f)$ and its uncertainty. The RTS algorithm leads to a third, smoothed, estimate, $\tilde{\mathbf{x}}(t, +)$, (in addition to the existing $\tilde{\mathbf{x}}(t)$, $\tilde{\mathbf{x}}(t, -)$) and $\tilde{\mathbf{u}}(t)$ from the recursion given in the Appendix. The result includes the uncertainty, $\mathbf{P}(t, +)$, of the smoothed state, and $\mathbf{Q}(t, +)$ for $\tilde{\mathbf{u}}(t)$. Together, $\tilde{\mathbf{x}}(t)$ and $\tilde{\mathbf{u}}(t)$ satisfy the model at all times. At t_f, filtered and smoothed estimates are identical.

In the present special case, as in many control problems, the major changes in the scalar $\tilde{u}(t)$ occur near the end, as the terminal data are accounted for. Those data change $\tilde{\mathbf{x}}(t_f - \Delta t)$ and its uncertainty, leading to a change in its immediate predecessor, $\tilde{\mathbf{x}}(t_f - 2\Delta t)$, etc., commonly with a loss of amplitude the further the estimate recedes in time from the terminal data.

7 Results

7.1 Top 100 m

Core 1093

[5] The term "optimal" is only justified if the various statistics are correctly specified.

Fig. 12 displays the inferred modification, $\tilde{u}(t)$ to $C_h^{\text{a priori}}(t)$ and its standard deviation $\sqrt{\tilde{\mathbf{Q}}(t,+)}$. The terminal state itself is identical to that in Fig. 9 from the Lagrange multiplier method. The maximum value of $C_h(t)$ occurs at -12ky with a value 35.8±0.7. With an a priori uncertainty of $\mathbf{Q} = (1\text{g/kg})^2$, the information content of the terminal state alone is unable, except near the very end, to much reduce it. If the same calculation is done using $\mathbf{Q} = (0.1$ g/kg$)^2$ (not shown) the uncertainties are correspondingly reduced by producing a different $\tilde{u}(t)$, but the smaller permissible adjustments to $C_h^{\text{a priori}}(t)$ increase the terminal misfits. The a priori uncertainties are directly determining the accuracy with which $C_h(t)$ can be inferred from these data. A residual $\tilde{u}(t)$ uncertainty of $\pm(0.5\text{-}1$ g/kg$)^2$ precludes any interesting inference about LGM salinity changes.

That the general structure of the solution is nearly independent of the prior control is shown by Fig. 13 in which the prior was made a uniform value of the mean value of $C_h^{\text{a priori}}(t)$. Only in the last 20 ky does any structure appear, and it remains below the estimated LGM mean. A similar calculation with a very high prior of 37 g/kg (not shown), with $w = -|k|$, is *reduced* below the LGM mean in the last 20 ky. The inability here to obtain values as high as those found by AS03 and others lies in part with the requirement that the near-core-top data should also be fit, data that generally require a strong decrease in $C_h(t)$ in the last tens of thousands of years. Should the core-top data be regarded as noise, perhaps the result of unresolved boundary layers in the sediment, higher values of C_h could be obtained, particularly if the initial conditions are made uniform, and near-perfect, and so unchangeable by the estimation procedure.

Fig. 14 shows the result obtained by choosing the sea level prior, but allowing $w = -|k|$ m/s with the initial conditions nearly completely ineffective in the final state. The fit to the terminal state is somewhat improved, but the uncertainties for $\tilde{u}(t)$ remain O(1 g/kg) except at the very end.

Core 1063

For core 1063, with $w = 0$ and the same value of k, the results are shown in Fig. 15. The state estimate is generally within the estimated prior uncertainty. Control $\tilde{u}(t)$ produces the maximum at about -20ky, of 35.55± 0.85 g/kg, with a value for the total well-below the estimated LGM mean, but with an uncertainty encompassing it. The specific estimated maximum lies below that for the Southern Ocean core, consistent with the AS03 result, but here both nominally fall below the average. Following that maximum, a considerable variation again occurs, but it is without statistical significance.

The considerable structure in the estimated control (bottom water salinity) that emerges during the deglacial period is interesting, if only in its general variations (none of which are

statistically significant). During deglaciation, the injection of $\approx$ 125 m of freshwater and the shift in the entire ocean volume to the modern lower salinity, along with the major change in atmospheric winds and temperatures, must have generated a host of regional circulation and salinity shifts and with complicated spatial differences. Differences found here between the two cores do not support an hypothesis of any globally uniform shifts in abyssal salinity—although they cannot be ruled out.

Deeper Core Data

Using the values of k above, the purely diffusive system cannot explain disturbances down-core deeper than about 50m or from before about -20,000 y. If the possibility that the effective $w = -|k|$ is accepted, signals can penetrate from the surface far deeper into the sediments. Assume that the deeper structures are signal, and not measurement or geological noise. Then the smoother calculation was carried out for Cores 1093 and 1063 using data from -300 m to the surface with a start time of $t = -200$ ky, a 100 ky periodic $C_h^{\mathrm{a\ priori}}(t)$,and with results shown in Figs. 17, 18. The Southern Ocean core shows early excursions even exceeding the LGM mean at about -38 ky, while the Atlantic Ocean core is consistently below both the prior and the LGM mean. Although it is tempting to speculate about what these apparent excursions imply—attaching them to events such as the Bølling-Allerød, Heinrich events, etc.—none of them is statistically meaningful, and far more data would be needed to render them so.

Because the uncertainty, $\tilde{Q}$, of the control remains dominated by the prior assumption of independent increments in $u(t)$, the estimated values $\tilde{u}(t)$ remain largely uncorrelated. A plausible inference is that on the average over the LGM and the deglaciation that the near-Bermuda abyssal waters were considerably fresher than those near the Southwest Indian Ridge and to that extent supporting previously published inferences, but not the conclusion that the salinity in the latter region was above the global-volume mean.

8 Modifications and Extensions

Thus far, the models used have been purely nominal, one-dimensional with constant in space and time diffusion and fixed w, either zero, or $w = -|k|$. Neither of these models is likely very accurate; both parameters are subject to variations in time and space, including higher space dimensions which would permit non-zero values of $\partial w/\partial z$. The central difficulty is that using some of the information contained in either $C_0^{\mathrm{a\ priori}}(t)$, or in $C_{\mathrm{term}}(z)$ to find w or k *necessarily further increases* the calculated uncertainty of $\tilde{u}(t)$. More measurements with different tracers would help, as would a better understanding of the time-depth properties of pore fluids in abyssal

cores. More sophisticated use of the prior covariances (functions of depth and time) could also reduce the uncertainties—but only to the extent that they are accurate.

9 Summary and Conclusions

Reproducing pore-water chlorinity/salinity observations in a deep-sea core involves an intricate and sensitive tradeoff of assumptions concerning diffusion rates, k, magnitudes and signs of the fluid vertical velocity, w, prior estimates of lower and upper boundary conditions, and in some cores, the nature of the initial conditions, the one-dimensional behavior of an advection-diffusion equation and, crucially, strong assumptions about the nature of the recorded noise. Most previous inferences with $w = 0$,which have led to a picture of the abyssal ocean as particularly saline, have been based essentially on the assumption that only the salinity maximum appearing at tens of meters from the core-top is signal, and does not originate with the initial conditions. All remaining structures are supposed noise of unspecific origin.

In the more general, approach used here, initial condition structures in a purely diffusive 100 m core can persist for more than 100 ky, greatly complicating the inference that the terminal data are controlled by the sea level changes of the past 20 ky alone. When observed structures beyond the gross maximum in salinity are treated as signals related to abyssal water properties, a statistically acceptable fit can be obtained by permitting a substantial downward fluid flow, $w < 0$, and which removes the initial conditions from the system. The abyssal water property boundary condition (the system "control" in the present context) however, then displays a complex and rich structure, none of which is statistically distinguishable from the LGM mean salinity. Terminal time conditions, $C_{\mathrm{term}}(z)$, only weakly constrain the time history of the control, $C_h(t)$, insufficient in the two cores analyzed to reduce salinity uncertainties below about ± 0.5 g/kg at any time before a few hundred years ago. This inference is consistent with that of Miller et al. (2015), using the same pore-water data, but a different analysis methodology.

That the Southern Hemisphere ocean was heavily salt-stratified in the abyss, with values well above the LGM mean, remains a not-implausible *assumption* about the last glacial period ocean, one depending upon the claim that initial conditions have little or no effect at the core terminal state data or upon other data not used here. If that assumption is take at face value, it raises the question of what the initial conditions were in practice and why their effects are invisible at t_f? With this particular type of core data, reducing the resulting uncertainty requires among other elements, providing a prior estimate, $C_h^{\mathrm{a\ priori}}(t)$, with smaller levels of uncertainty (better than 0.1 g/kg), a requirement for which little prospect exists.

The uncertainties derived here are all lower bounds, and are based in part upon the assump-

tion of perfectly known, simple, core profiles of w, k. These parameters can, in a formal sense, be treated as further unknowns as a function of depth and time, but if the information contained in the terminal chlorinity data is used to estimate their values, the uncertainties of $\tilde{C}_h(t)$ will become even larger. No claim is made that the chosen parameters here, $k = 2 \times 10^{-10} \mathrm{m}^2/\mathrm{s}$, and $w = -|k|$ m/s are "correct", merely that they give a reasonable fit to the terminal data. If the equivalent of k is measured at the terminal time in the cores (via the porosity and tortuosity) the measurement errors are necessarily greater than the zero values used here in treating it as perfectly known. A further generalization estimates the uncertainty covariances as part of the calculation ("adaptive" filtering and smoothing; see e.g., Anderson and Moore, 1979), but it again necessarily further increases the estimated state and control vector uncertainties.

Of particular use would be pore water properties in regionally distributed cores. It would then become possible to better understand the background structures (are they regional covarying signals, or are they noise particular to one core?) and their geography. Such additional data would be a major step towards expanding the data base to the point where an accurate global average would become plausible.

Numerous interesting questions arise, at least within a theoretical framework. How the dynamical and kinematical response to an excess of evaporation, leading to formation of the continental ice sheets would have worked its way through the entire ocean volume, raising the mean by about 1 g/kg is far from obvious. Equally obscure is how the global volume salinity mean re-adjusted itself, much more rapidly, to its lower modern value through the excess runoff in the deglaciation. Complex transient behavior would be expected with time scales exceeding thousands of years. Amongst many such interesting issues, note that much of the salt in the modern upper North Atlantic Ocean arises from the highly saline Mediterranean Sea outflow. Paul et al. (2001) have discussed possibilities for LGM salinity changes there, also from pore fluid data. Whether any of the world-wide symptoms of these major re-adjustments can be detected in paleoceanographic data remains a challenging question.

To answer the two questions posed in the Introduction: An LGM ocean with greatly intensified salinity in the abyssal Southern Ocean is *not* required by the pore-water chlorinity data and, such an ocean is not contradicted by the pore water data within the large lower-bound residual error estimates.

Acknowledgments. Supported in part by National Science Foundation Grant OCE096713 to MIT. This work would not have been possible without long discussions with Dr. M. Miller and the data that were provided by her. I had essential suggestions and corrections from O. Marchal, R. Ferrari and P. Huybers. Special thanks to D. Schrag for a thoughtful review despite his thinking that the wrong questions were being posed.

Appendix-Control Algorithms

The algorithms for the Lagrange multiplier (or adjoint) solution, and for the filter-smoother are written out here for reference purposes; cf. W06.

Lagrange Multipliers

Assume a model (Eq. 9) with a state vector $\mathbf{x}(t)$, and a terminal data set, $\mathbf{x}_d(t_f)$, having error covariance $\mathbf{R}(t_f)$, $\tilde{\mathbf{x}}(0)$ is the initial condition with uncertainty $\mathbf{P}(0)$, and assuming for notational simplicity that none of $\mathbf{A}$, $\mathbf{B}$, or $\mathbf{\Gamma}$ is time-dependent. The covariance of the control, $\mathbf{u}(t)$, is $\mathbf{Q}(t)$. Let the objective or cost function be Eq. (11), the model is adjoined (appended) to J using a set of vector Lagrange multipliers, $\boldsymbol{\mu}(t)$. Generating the normal equations by differentiation in ordinary least-squares, $\boldsymbol{\mu}(t)$ satisfy a time-evolution equation

$$\boldsymbol{\mu}(t-1) = \mathbf{A}^T \boldsymbol{\mu}(t), \ t = 1, 2, .., t_f \tag{A1a}$$

$$\boldsymbol{\mu}(t_f) = \mathbf{R}^{-1}\left(\tilde{\mathbf{x}}(t_f) - \mathbf{x}_d(t_f)\right), \tag{A1b}$$

time appearing to "run backwards." The unknown controls are then,

$$\mathbf{u}(t) = -\mathbf{Q}\mathbf{\Gamma}^T \boldsymbol{\mu}(t+1) \tag{A2}$$

and

$$\begin{aligned} &\Big\{\mathbf{I} + \mathbf{A}^{(t_f-1)}\mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^T\mathbf{A}^{(t_f-1)T}\mathbf{R}^{-1} \\ &\quad + \mathbf{A}^{(t_f-2)}\mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^T\mathbf{A}^{(t_f-2)T}\mathbf{R}^{-1} + \cdots + \mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^T\mathbf{R}^{-1}\Big\}\,\tilde{\mathbf{x}}(t_f) \\ &= \mathbf{A}^{t_f}\tilde{\mathbf{x}}(0) + \Big\{\mathbf{A}^{(t_f-1)}\mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^T\mathbf{A}^{(t_f-1)T}\mathbf{R}^{-1} \\ &\qquad + \mathbf{A}^{(t_f-2)}\mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^T\mathbf{A}^{(t_f-2)T}\mathbf{R}^{-1} + \cdots + \mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^T\mathbf{R}^{-1}\Big\}\,\mathbf{x}_d\,. \end{aligned} \tag{A3}$$

explicitly relates the estimated terminal state, $\tilde{\mathbf{x}}(t_f)$, to the desired one, $\mathbf{x}_d$. Eq. (A1a) is then solved for $\boldsymbol{\mu}(t)$, and the entire state then follows from Eqs. (A2, 9). See W06, p.218+).

RTS Smoother

The RTS smoother uses a Kalman filter in the forwards-in-time direction, with the equations in the main text. In the time-reverse direction, the algorithm is more complicated in appearance because it takes account of the time-correlations in the error estimates that were built up in the

filter sweep. The resulting system, in the notation of W06, p. 208, is,

$$\tilde{\mathbf{x}}\left(t,+\right) = \tilde{\mathbf{x}}\left(t+\right)\mathbf{L}(t+1)\left[\tilde{\mathbf{x}}(t+1,+) - \tilde{\mathbf{x}}(t+1,-)\right], \tag{A4a}$$

$$\mathbf{L}(t+1) = \mathbf{P}(t)\mathbf{A}(t)^T\mathbf{P}(t+1,-)^{-1}, \tag{A4b}$$

$$\tilde{\mathbf{u}}(t+1) = \mathbf{M}(t+1)[\tilde{\mathbf{x}}(t+1,+) - \tilde{\mathbf{x}}(t+1,-)], \tag{A4c}$$

$$\mathbf{M}(t+1) = \mathbf{Q}(t)\mathbf{\Gamma}\left(t\right)^T\mathbf{P}(t+1,-)^{-1}, \tag{A4d}$$

$$\mathbf{P}(t,+) = \mathbf{P}(t) + \mathbf{L}(t+1)[\mathbf{P}(t+1,+) - \mathbf{P}(t+1,-)]\mathbf{L}(t+1)^T, \tag{A4e}$$

$$\tilde{\mathbf{Q}}(t,+) = \mathbf{Q}(t) + \mathbf{M}(t+1)[\mathbf{P}(t+1,+) - \mathbf{P}(t+1,-)]\mathbf{M}(t+1)^T, \tag{A4f}$$

$$t = 0, 1, ..., t_f - 1$$

Data do not appear, all information content having been used in the forward sweep.

References

Adkins, J.F., Ingersoll, A.P., Pasquero, C., 2005. Rapid climate change and conditional instability of the glacial deep ocean from the thermobaric effect and geothermal heating. Quat. Sci. Revs. 24, 581-594.

Adkins, J.F., McIntyre, K., Schrag, D.P., 2002. The salinity, temperature, and δ^{18}O of the glacial deep ocean. Science 298, 1769-1773.

Adkins, J.F., Schrag, D.P., 2003. Reconstructing Last Glacial Maximum bottom water salinities from deep-sea sediment pore fluid profiles. Earth Planet. Sc. Letts. 216, 109-123.

Anderson, B.D.O., Moore, J.B., 1979. Optimal Filtering. Prentice-Hall, Englewood Cliffs, N. J.

Berner, R.A., 1980. Early Diagenesis: A Theoretical Approach. Princeton University Press, Princeton, N.J.

Boudreau, B.P., 1997. Diagenetic Models And Their Implementation: Modelling Transport And Reactions In Aquatic Sediments. Springer, Berlin; New York.

Brogan, W.L., 1991. Modern Control Theory, 3rd Ed. Prentice-Hall/Quantum, Englewood Cliffs, N. J.

Bruna, M., Chapman, S.J., 2015. Diffusion in Spatially Varying Porous Media. SIAM J. Appl. Maths 75, 1648-1674.

Carslaw, H.S., Jaeger, J.C., 1986. Conduction of Heat in Solids. Oxford Un. Press.

Carter, R.M., McCave, I.N., Richter, C., Carter, L., et al., 1999. Proc. ODP, Init. Repts., 181.

Dade, W.B., Hogg, A.J., Boudreau, B.P., 2001. Physics of flow above the sediment-water interface, in: Boudreau, R.D., Jørgensen, B.B. (Eds.), The Benthic Boundary Layer. Transport Processes and Biogeochemistry. Oxford University Press, New York.

Einsele, G., 2000. Sedimentary Basins: Evolution, Facies, and Sediment Budget, 2nd, Ed. Springer, Berlin ; New York.

Forget, G., Campin, J.-M., Heimbach, P., Hill, C., Ponte, R., Wunsch, C., 2015. ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation. Geosci. Model Dev. 8, 3071–3104.

Fowler, A.C., Yang, X.S., 1998. Fast and slow compaction in sedimentary basins. SIAM J. App. Maths. 59, 365-385.

Gersonde, R., Hodell, D.A., Blum, P., et al., 1999. Proc. ODP, Init. Repts., 177

Huybers, P., Wunsch, C., 2010. Paleophysical oceanography with an emphasis on transport rates. Annu. Rev. Mar. Sci. 2, 1-34.

Insua, T.L., Spivack, A.J., Graham, D., D'Hondt, S., Moran, K., 2014. Reconstruction of Pacific Ocean bottom water salinity during the Last Glacial Maximum. Geophys. Res. Letts. 41, 2914-2920.

Jansen, E., Raymo, M.E., Blum, P., et al., 1996. Proc. ODP, Init. Repts., 162: College Station, TX (Ocean Drilling Program), 49–90. doi:10.2973/odp.proc.ir.162.103.1996.

Keigwin, L.D., Rio, D., Acton, G.D., et al., 1998. Proc. ODP, Init. Repts., 172: College Station, TX (Ocean Drilling Program), 251–308. doi:10.2973/odp.proc.ir.172.106.1998.

Kobayashi, H., Abe-Ouchi, A., Oka, A., 2015. Role of Southern Ocean stratification in glacial atmospheric CO2 reduction evaluated by a three-dimensional ocean general circulation model. Paleoceanography 30, 2015PA002786.

Luenberger, D.G., 1979. Introduction to Dynamic Systems. Theory, Models and Applications. John Wiley, New York.

Macayeal, D., 1995. Challenging an ice-core paleothermometer. Science 270, 444-445.

Macayeal, D.R., Firestone, J., Waddington, E., 1991. Paleothermometry by control methods. J Glaciol 37, 326-338.

McDuff, R.E., 1985. The chemistry of interstitial Waters, Deep-Sea Drilling Project Leg-86. Initial Reports of the Deep Sea Drilling Project 86, 675-687.

Miller, M.D., 2014. The Deep Ocean Density Structure at the Last Glacial Maximum. What Was It and Why? PhD Thesis Cal. Tech., p. 136.

Miller, M.D., Simons, M., Adkins, J.F., Minson, S.E., 2015 The information content of pore fluid δ^{18}O and [Cl^{-}]. J. Phys. Oc. 45, 2070-2094.

Mix, A., Tiedemann, R., Blum P., 2002. Initial Reports, Proc. Ocean Drilling Prog., 202.

Munk, W.H., 1966. Abyssal recipes. Deep-Sea Res. 13, 707-730.

Paul, H.A., Bernasconi, S.M., Schmid, D.W., McKenzie, J.A., 2001. Oxygen isotopic composition of the Mediterranean Sea since the Last Glacial Maximum: constraints from pore water analyses. Earth Planet. Sci. Letts. 192, 1-14.

Roache, P.J., 1976. Computational Fluid Dynamics. Hermosa, Albuqurque, N.M.

Schrag, D.P., Adkins, J.F., McIntyre, K., Alexander, J.L., Hodell, D.A., Charles, C.D., McManus, J.F., 2002. The oxygen isotopic composition of seawater during the Last Glacial Maximum. Quat. Sci. Revs. 21, 331-342.

Schrag, D.P., Depaolo, D.J., 1993. Determination of delta-o-18 of seawater in the deep ocean during the last glacial maximum. Paleoceanog. 8, 1-6.

Schrag, D.P., Hampt, G., Murray, D.W., 1996. Pore fluid constraints on the temperature and oxygen isotopic composition of the glacial ocean. Science 272, 1930-1932.

Stengel, R.F., 1986. Stochastic Optimal Control. Wiley-Interscience, N. Y.

Voermans, J., Ghisalberti, M., Ivey, G., 2016. Coherent vortex structures at the sediment-water-interface, 11th International Symposium on Ecohydraulics, Melbourne, Australia, pp. unpaged, in press.

Wunsch, C., 1987. Using transient tracers: The regularization problem. Tellus, 39B, 477-492.

Wunsch, C., 1988. Transient tracers as a problem in control theory. J. Geophys. Res., 93, 8099-8110.

Wunsch, C., 2002. Oceanic age and transient tracers: Analytical and numerical solutions. J. Geophys. Res., 107, 304810.1029/2001jc000797.

Wunsch, C., 2006. Discrete Inverse and State Estimation Problems: With Geophysical Fluid Applications. Cambridge University Press, Cambridge ; New York.

1. Core positions—white circles— used by Miller et al. (2015), Adkins and Schrag (2003). Shown on a chart of the modern 20-year average salinity at 3600 m from the ECCO 4 state estimate (e.g., Forget et al., 2015). The focus of attention here is on the North Atlantic core near Bermuda and the South Atlantic one southwest of the Cape of Good Hope. A modern average salinity calculated from these 5 positions might be useful but would not be very accurate. See Table 1 for descriptive references of each core, and the greatly varying water depths at each site. In the modern ocean, the North Atlantic at 3600 m is more saline than the Southern Ocean. The modern full volume average salinity is about 34.7 g/kg. The average value at this depth today is about 34.75 g/kg (not area weighted) and about 34.74 g/kg when weighted. A suite of charts for modern salinity and other properties in section and latitude-longitude form is available in the online WOCE Atlas. Variations are complex and defy a simple verbal description. In particular note that strong zonal structures in salinity exist in the abyssal Southern Ocean; it is not zonally homogeneous.

2. Histogram of the modern ocean salinity at 3600m as a time average over 20 years from the ECCO state estimate (Forget et al., 2015). Perhaps the glacial ocean was more homogeneous? The two modes roughly correspond to North Atlantic Deep Water and Antarctic Bottom Waters. The probability of an accurate global average from any handful of values is low and note that the core tops here lie at considerably different water depths (Table 1).

3. Salinity, g/kg over the full measured depth in each of the five cores. This paper focusses on Cores 1063, 1093 plotted as thicker lines. See Table 1 for a reference and geographical label for each core. Vertical dashed lines are the approximate modern global volume mean salinity, 34.7 g/kg and the approximate LGM value of 35.7 g/kg. and dotted line fragment shows the LGM maximum value of $C_h\,(t)$ estimated for this core by Adkins and Schrag (2003).

4. Same as Fig. 3 except expanded to show only the top 100m. The maximum measured value in core 1093 lies near the estimated LGM mean of 35.7 (vertical dashed line), but does not exceed it except slightly in short, possibly noise, events. Thick lines are the data from the two cores analyzed, 1063, 1093. Approximate modern mean salinity of 34.7 g/kg is also shown as a vertical dashed line. The salinity increase with depth in the much fresher Core 1239 is almost as large as that appearing in Core 1093. Near surface, the core is either undersampled, or the data are extremely noisy.

5.Location of ODP Core 1093 on the Southwest Indian Ridge. See Gersonde et al. (1999).

6. Measured porosity from all five cores. Tortuosity is assumed to follow Eq. (4). In the present calculations the corresponding diffusivity, k, is taken to be constant with depth. Experiments with linear k produced only slight changes from the solutions with a constant value.

7. Time-depth profile of a numerical solution using a Dufort-Frankel method (Roache, 1976) (a), and the analytic solution from Carslaw and Jaeger (1986, P. 101) for a zero-initial condition, (b) for $w = 0, k = 2 \times 10^{-10} \mathrm{m}^2/s$ in a 100 meter length "core" over a duration of 100,000 years. Panel (c) shows the terminal profile in the two solutions which are visually indistinguishable. Time scale zero is at -100 ky BP.

8. Time for a particular vertical scale to decay to 1% of its initial value (from Eq. 5). Horizontal dashed line is at 20 ky.

9. (a) Forward solution, with pure diffusion, using the quasi-periodic condition: starting with the observed Core 1093 at $t = -100$ ky and forced by $q(t) = C_h^{\mathrm{a\ priori}}(t)$. Much of the terminal structure originates with the initial conditions, although some small scale-structure is lost in the terminal-time profile (solid line) in (c). (b) Shows the forward calculation with the same initial and boundary conditions, except now $w = -|k|$ m/s. All of the initial condition data is swept out of the system before the end. (c) Observed terminal data (dashed) and the two terminal states for $w = 0$, and $w = -|k|$.

10. (a) Amplitude (Eq. 6) of the steady-periodic state component of $c(z)$ for different forcing periods using the same value of k. Short period ("high" frequency) responses rapidly diminish with depth. A 500 m depth was used because of the dependence on the zero boundary condition at the core bottom. (b) 99% decay times of the transient element of periodic surface forcing against vertical scale. Horizontal dashed line is at 20 ky.

11. Estimated global mean salinity from the sea level change curve (Miller, 2014, Miller et al., 2015, from a variety of sources). The direction of the time axis has been converted to the physics convention from the geological "age". This curve becomes the prior boundary condition $C_h^{\mathrm{a\ priori}}(t)$ .Dashed lines are the estimated volume average modern (lower) and LGM salinities (upper), the latter the value used by Adkins and Schrag (2003). Here the time scale represents time before the present.

12. (a) Kalman filter solution, Core 1093, pure diffusion ($w = 0$) and the sea level prior. The filter solution is identical to that shown from the pure forward calculation in Fig. 9 except,

nearly invisibly, near the terminal time. (b) Estimate of the state vector after application of the smoothing algorithm, and which changes the state as far back as its initial conditions. (c) Deviation of the terminal state estimate from the core data, along with one standard errors $\sqrt{P(t_f)} = \sqrt{P(t_f,+)}$ (d) $C_h^{\text{a priori}}(t)$ (solid), and the estimated $C_h(t) = q(t) + \tilde{u}(t)$ (dashed). Horizontal dashed lines are the modern and LGM global means, the latter the AS03 estimate. The estimated value of $C_h(t)$ remains below the global mean LGM salinity, as the deep maximum is controlled by the initial conditions in contrast to the solution of AS03 which reached 37.1 g/kg. (e) Last 5000 years of $\tilde{u}(t)$, and the one standard deviation uncertainty from $\pm\sqrt{Q(t)}$. Except at the very end, $\tilde{Q}(t,+)$ differs negligibly from $Q(t)$.(f) Terminal state from the Kalman filter just prior to the invocation of the terminal data (identical to the forward solution; solid line. Dashed line is the core data, dotted line the terminal state. Note that the Lagrange multiplier solution, the Kalman filter solution, and the smoothed solution are identical at $t = t_f$.

13. Core 1093 with a constant ("flat") prior of 34.93 g/kg, the mean of $C_h^{\text{a priori}}(t)$.(a) The solution as run forward in the Kalman filter sweep. (b) The same solution as modified by the smoothing sweep. (c) Deviation of the terminal state (either from Lagrange multipliers or the Kalman filter or the smoother) from the core data. (d) Flat, a priori control $q(t)$ (solid line) and the final estimated $C_h(t) = q(t) + \tilde{u}(t)$. (e) Last 5000 years of $\tilde{u}(t)$ and the one standard deviation uncertainty. (f) Comparison of the prediction, $\tilde{\mathbf{x}}(t_f,-), \tilde{\mathbf{x}}(t_f,+) = \tilde{\mathbf{x}}(t_f)$, and the Core 1093 data.

14. Same as Fig. 12 except for $w = -|k|$ m/s. Again $\tilde{C}_h(t)$ is always below the LGM mean.

15. Same as Figs. 12, except for Core 1063 on the Bermuda Rise. The terminal state does not well-match the core data in terms of depth and this solution probably should be rejected.

16. Core 1093 (Southwest Indian Ridge) results using 300m length, $w = -|k|$ and an initial time of -200 ky with the sea level curve treated as twice periodic over 100 ky. (a) The misfit to the core data and (b,c) the final control are shown. Now structure appears as far back as -60 ky, and includes an excursion above the LGM mean at around -38ky although not statistically significant.

17. Same as Fig. 17 except for Core 1063 on the Bermuda Rise.

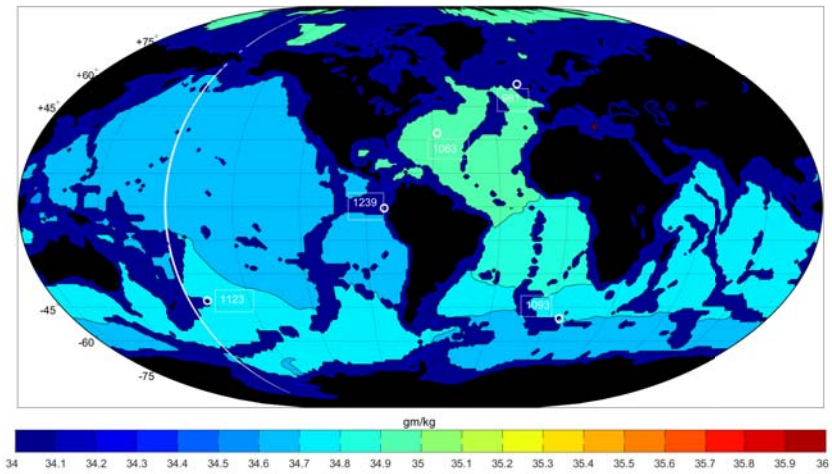

Figure 1: Core positions—white circles— used by Miller et al. (2015), Adkins and Schrag (2003). Shown on a chart of the modern 20-year average salinity at 3600 m from the ECCO 4 state estimate (e.g., Forget et al., 2015). The focus of attention here is on the North Atlantic core near Bermuda and the South Atlantic one southwest of the Cape of Good Hope. A modern average salinity calculated from these 5 positions might be useful but would not be very accurate. See Table 1 for descriptive references of each core, and the greatly varying water depths at each site. In the modern ocean, the North Atlantic at 3600 m is more saline than the Southern Ocean. The modern full volume average salinity is about 34.7 g/kg. The average value at this depth today is about 34.75 g/kg (not area weighted) and about 34.74 g/kg when weighted. A suite of charts for modern salinity and other properties in section and latitude-longitude form is available in the online WOCE Atlas. Variations are complex and defy a simple verbal description. In particular note that strong zonal structures in salinity exist in the abyssal Southern Ocean; it is not zonally homogeneous.

780 LaTeX

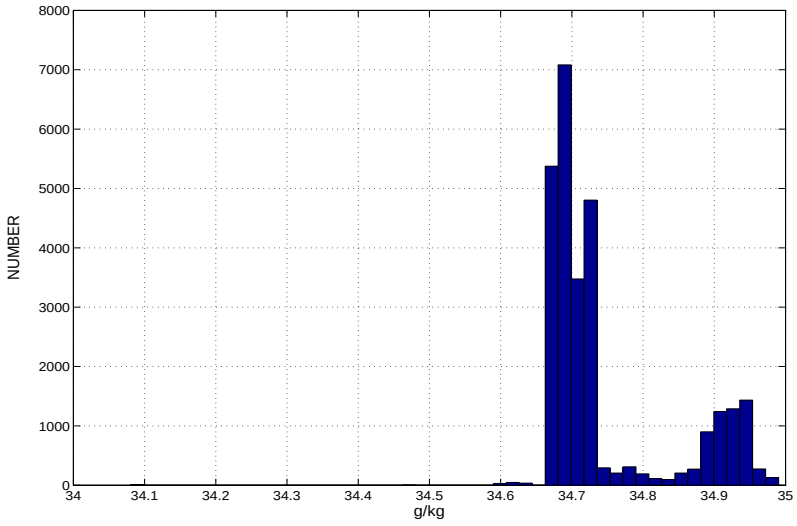

Figure 2: Histogram of the modern ocean salinity at 3600m as a time average over 20 years from the ECCO state estimate (Forget et al., 2015). Perhaps the glacial ocean was more homogeneous? The two modes roughly correspond to North Atlantic Deep Water and Antarctic Bottom Waters. The probability of an accurate global average from any handful of values is low and note that the core tops here lie at considerably different water depths (Table 1).

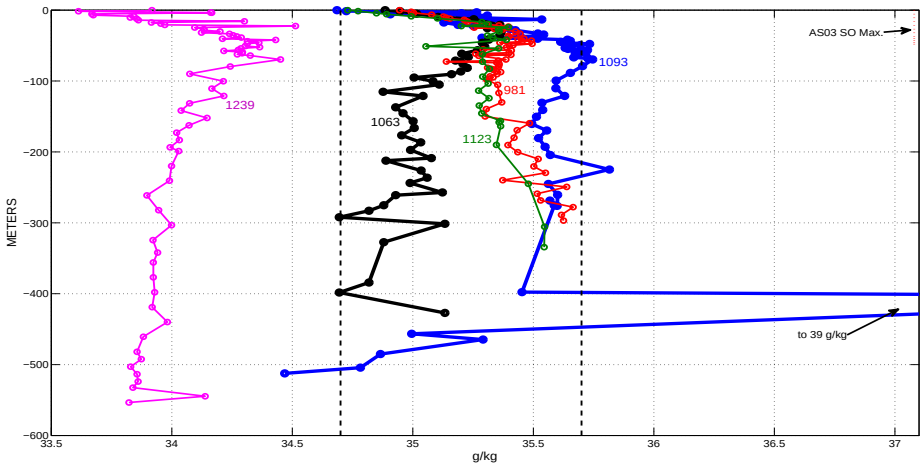

Figure 3: Salinity, g/kg over the full measured depth in each of the five cores. This paper focusses on Cores 1063, 1093 plotted as thicker lines. See Table 1 for a reference and geographical label for each core. Vertical dashed lines are the approximate modern global volume mean salinity, 34.7 g/kg and the approximate LGM value of 35.7 g/kg. and dotted line fragment shows the LGM maximum value of $C_h\,(t)$ estimated for this core by Adkins and Schrag (2003).

781 LaTeX

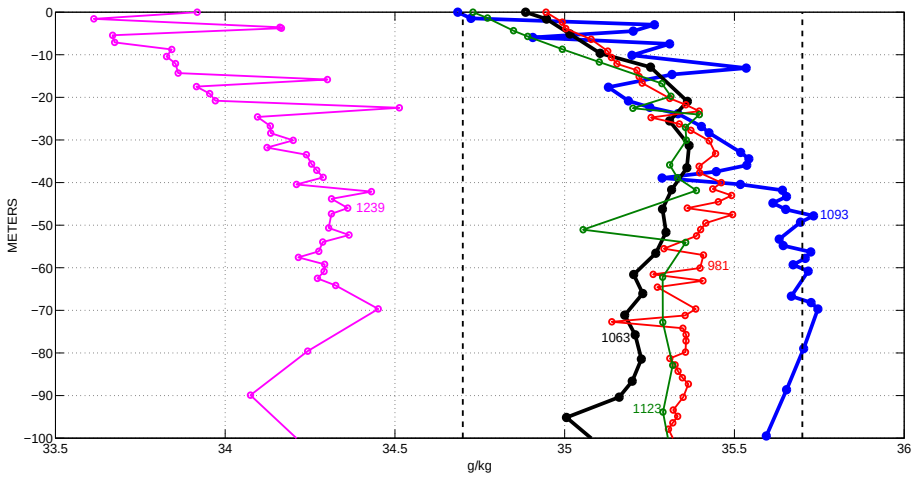

Figure 4: Same as Fig. 3 except expanded to show only the top 100m. The maximum measured value in core 1093 lies near the estimated LGM mean of 35.7 (vertical dashed line), but does not exceed it except slightly in short, possibly noise, events. Thick lines are the data from the two cores analyzed, 1063, 1093. Approximate modern mean salinity of 34.7 g/kg is also shown as a vertical dashed line. The salinity increase with depth in the much fresher Core 1239 is almost as large as that appearing in Core 1093. Near surface, the core is either undersampled, or the data are extremely noisy.

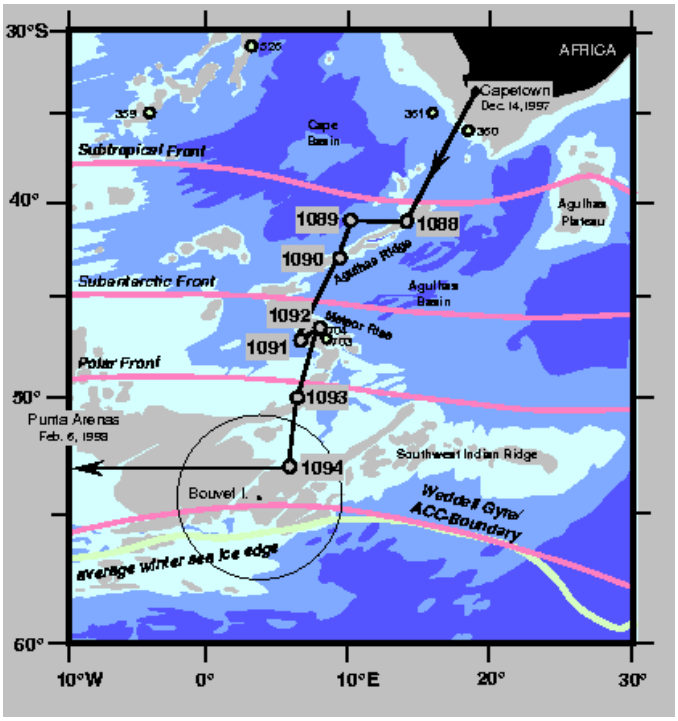

Figure 5: Location of ODP Core 1093 on the Southwest Indian Ridge. See Gersonde et al. (1999).

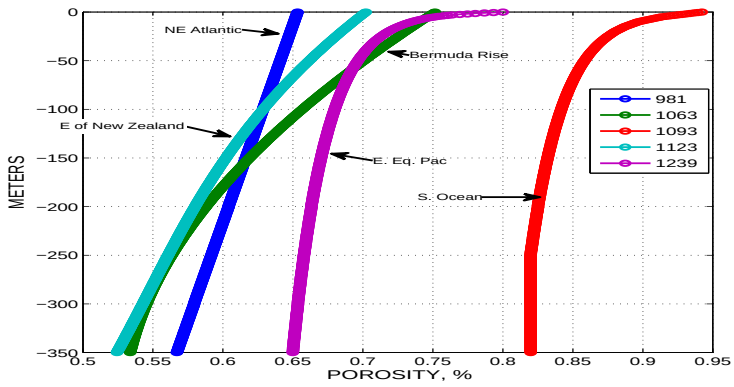

Figure 6: Measured porosity from all five cores. Tortuosity is assumed to follow Eq. (4). In the present calculations the corresponding diffusivity, k, is taken to be constant with depth. Experiments with linear k produced only slight changes from the solutions with a constant value.

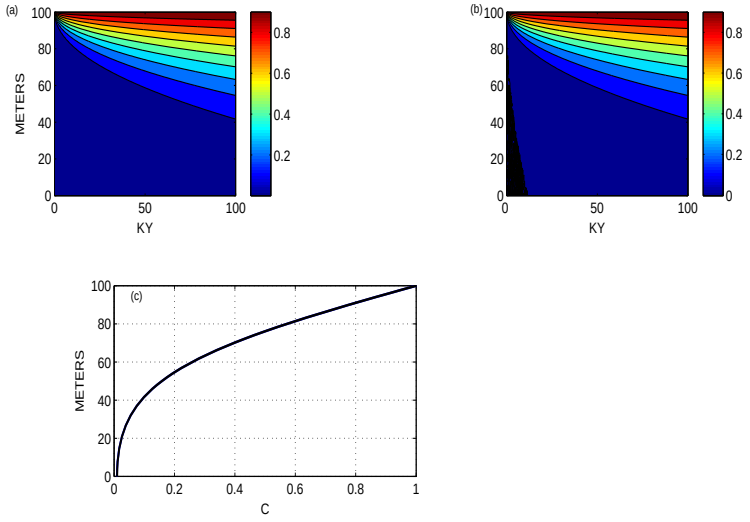

Figure 7: Time-depth profile of a numerical solution using a Dufort-Frankel method (Roache, 1976) (a), and the analytic solution from Carslaw and Jaeger (1986, P. 101) for a zero-initial condition, (b) for $w = 0, k = 2 \times 10^{-10} \mathrm{m}^2/s$ in a 100 meter length "core" over a duration of 100,000 years. Panel (c) shows the terminal profile in the two solutions which are visually indistinguishable. Time scale zero is at -100 ky BP.

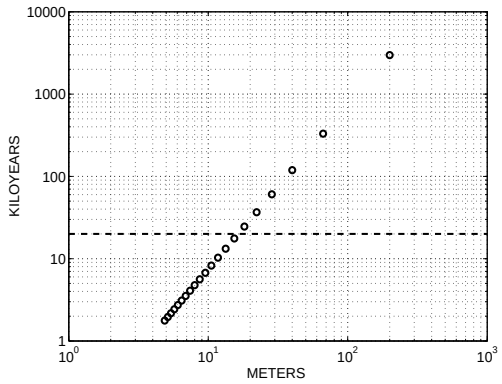

Figure 8: Time for a particular vertical scale to decay to 1% of its initial value (from Eq. 5). Horizontal dashed line is at 20 ky.

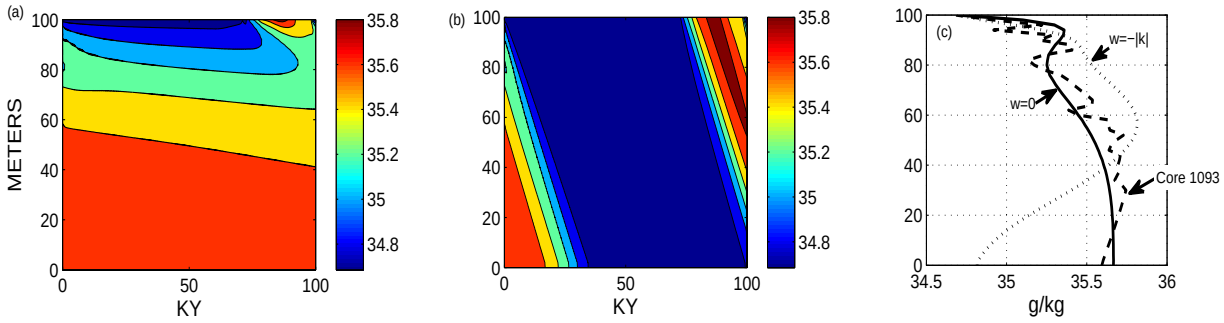

Figure 9: (a) Forward solution, with pure diffusion, using the quasi-periodic condition: starting with the observed Core 1093 at $t = -100$ ky and forced by $q(t) = C_h^{\text{a priori}}(t)$. Much of the terminal structure originates with the initial conditions, although some small scale-structure is lost in the terminal-time profile (solid line) in (c). (b) Shows the forward calculation with the same initial and boundary conditions, except now $w = -|k|$ m/s. All of the initial condition data is swept out of the system before the end. (c) Observed terminal data (dashed) and the two terminal states for $w = 0$, and $w = -|k|$.

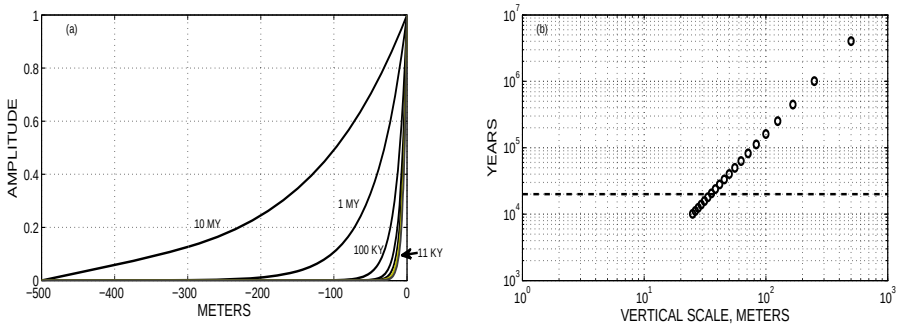

Figure 10: (a) Amplitude (Eq. 6) of the steady-periodic state component of $c\,(z)$ for different forcing periods using the same value of k. Short period ("high" frequency) responses rapidly diminish with depth. A 500 m depth was used because of the dependence on the zero boundary condition at the core bottom. (b) 99% decay times of the transient element of periodic surface forcing against vertical scale. Horizontal dashed line is at 20 ky.

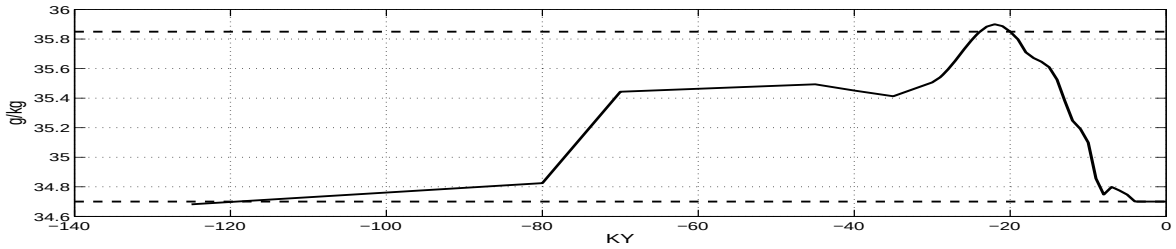

Figure 11: Estimated global mean salinity from the sea level change curve (Miller, 2014, Miller et al., 2015, from a variety of sources). The direction of the time axis has been converted to the physics convention from the geological "age". This curve becomes the prior boundary condition $C_h^{\mathrm{a\ priori}}(t)$ .Dashed lines are the estimated volume average modern (lower) and LGM salinities (upper), the latter the value used by Adkins and Schrag (2003). Here the time scale represents time before the present.

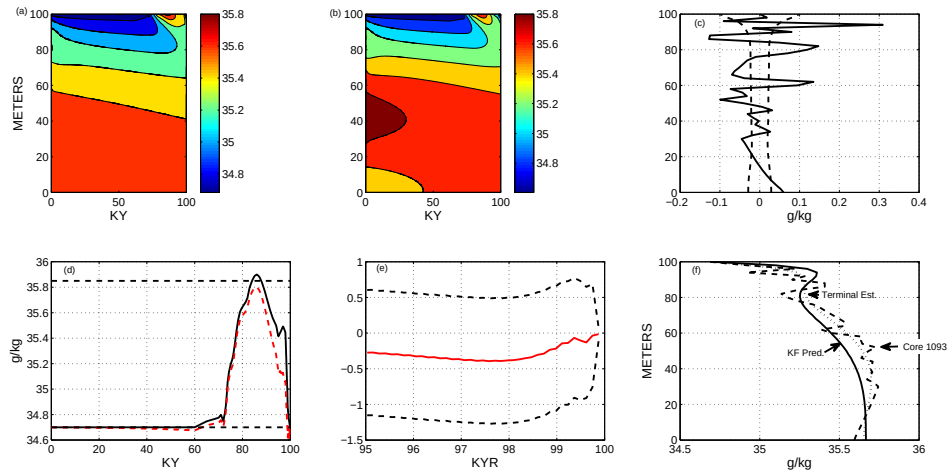

Figure 12: (a) Kalman filter solution, Core 1093, pure diffusion ($w = 0$) and the sea level prior. The filter solution is identical to that shown from the pure forward calculation in Fig. 9 except, nearly invisibly, near the terminal time. (b) Estimate of the state vector after application of the smoothing algorithm, and which changes the state as far back as its initial conditions. (c) Deviation of the terminal state estimate from the core data, along with one standard errors $\sqrt{P\left(t_f\right)} = \sqrt{P(t_f,+)}$ (d) $C_h^{\text{a priori}}\left(t\right)$ (solid), and the estimated $C_h\left(t\right) = q\left(t\right) + \tilde{u}\left(t\right)$ (dashed). Horizontal dashed lines are the modern and LGM global means, the latter the AS03 estimate. The estimated value of $C_h\left(t\right)$ remains below the global mean LGM salinity, as the deep maximum is controlled by the initial conditions in contrast to the solution of AS03 which reached 37.1 g/kg. (e) Last 5000 years of $\tilde{u}\left(t\right)$, and the one standard deviation uncertainty from $\pm\sqrt{Q\left(t\right)}$. Except at the very end, $\tilde{Q}\left(t,+\right)$ differs negligibly from $Q\left(t\right)$.(f) Terminal state from the Kalman filter just prior to the invocation of the terminal data (identical to the forward solution; solid line. Dashed line is the core data, dotted line the terminal state. Note that the Lagrange multiplier solution, the Kalman filter solution, and the smoothed solution are identical at $t = t_f$.

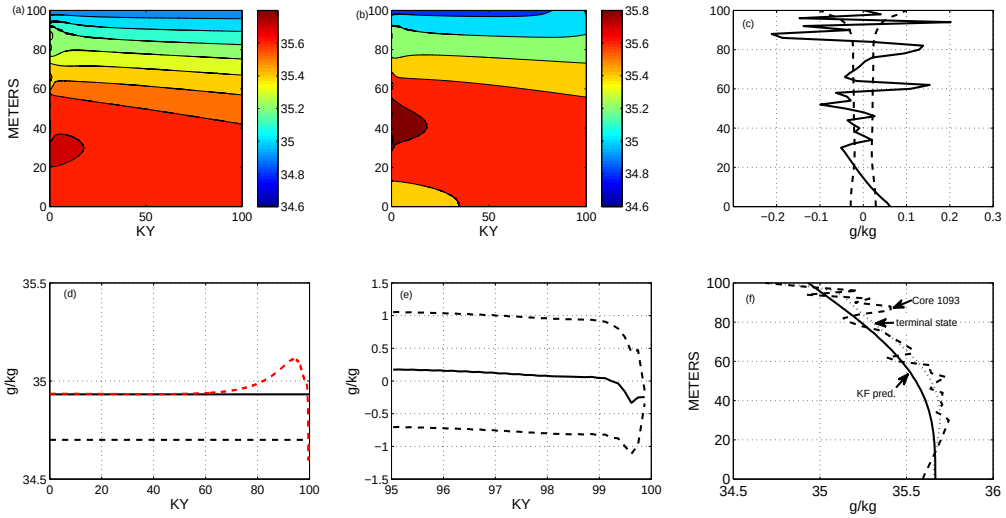

Figure 13: Core 1093 with a constant ("flat") prior of 34.93 g/kg, the mean of $C_h^{\text{a priori}}(t)$.(a) The solution as run forward in the Kalman filter sweep. (b) The same solution as modified by the smoothing sweep. (c) Deviation of the terminal state (either from Lagrange multipliers or the Kalman filter or the smoother) from the core data. (d) Flat, a priori control $q(t)$ (solid line) and the final estimated $C_h(t) = q(t) + \tilde{u}(t)$. (e) Last 5000 years of $\tilde{u}(t)$ and the one standard deviation uncertainty. (f) Comparison of the prediction, $\tilde{\mathbf{x}}(t_f, -), \tilde{\mathbf{x}}(t_f, +) = \tilde{\mathbf{x}}(t_f)$, and the Core 1093 data.

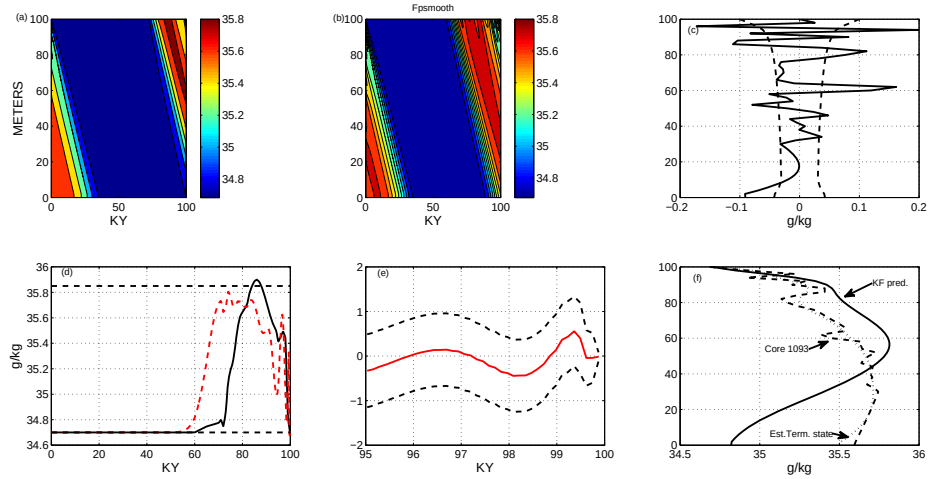

Figure 14: Same as Fig. 12 except for $w = -|k|$ m/s. Again $\tilde{C}_h(t)$ is always below the LGM mean.

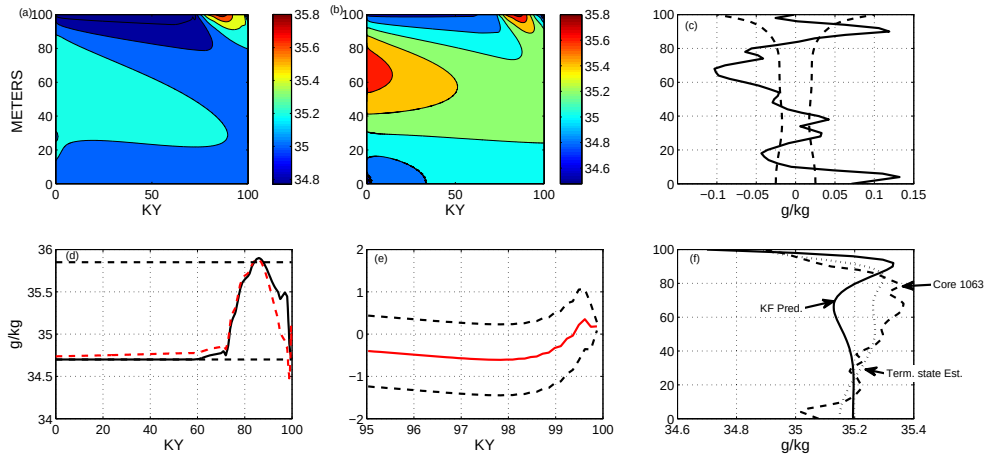

Figure 15: Same as Figs. 12, except for Core 1063 on the Bermuda Rise. The terminal state does not well-match the core data in terms of depth and this solution probably should be rejected.

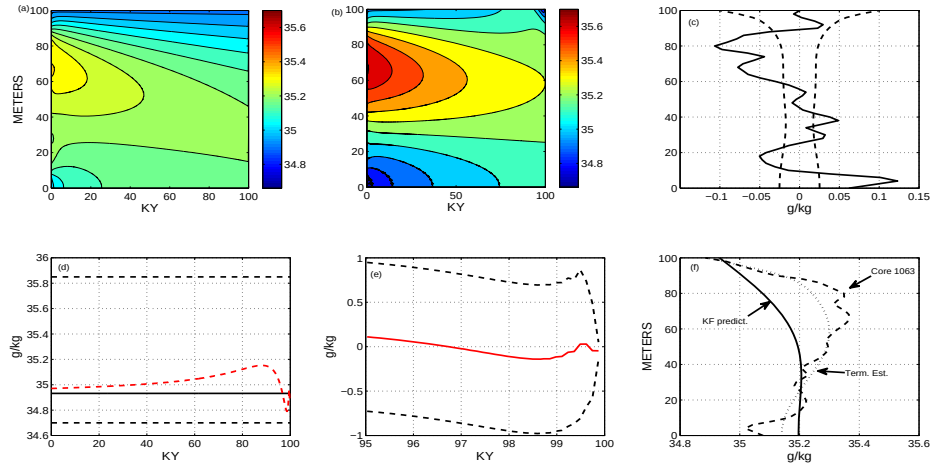

Figure 16: Same as Fig. 15 except for a constant prior $C_h^{\text{a priori}}(t)$.

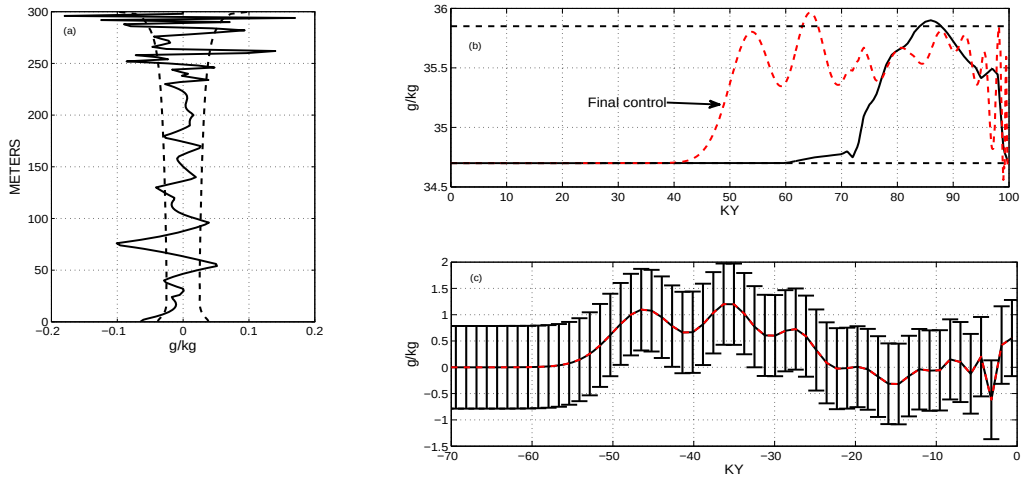

Figure 17: Core 1093 (Southwest Indian Ridge) results using 300m length, $w = -|k|$ and an initial time of -200 ky with the sea level curve treated as twice periodic over 100 ky. (a) The misfit to the core data and (b,c) the final control are shown. Now structure appears as far back as -60 ky, and includes an excursion above the LGM mean at around -38ky although not statistically significant.

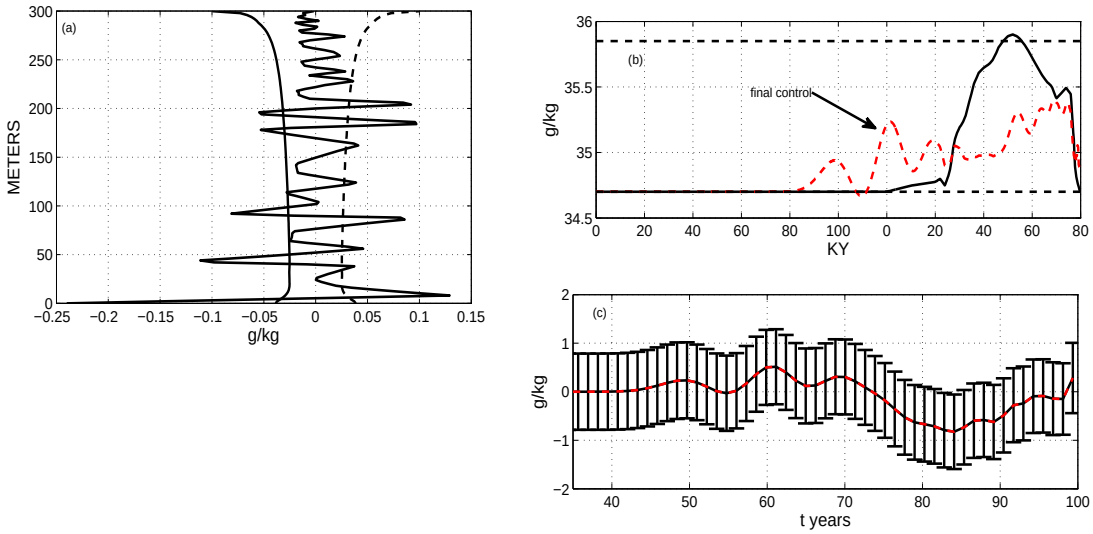

Figure 18: Same as Fig. 17 except for core 1063 on the Bermuda Rise.