Mantle pseudo-isochrons revisited

John F. Rudge [a,b,*]

[a] *Bullard Laboratories, Department of Earth Sciences, University of Cambridge, Madingley Road, Cambridge, CB3 0EZ, UK.*

[b] *Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK.*

Abstract

The 2.0 Ga pseudo-isochron age inferred from the mid-ocean ridge basalt $^{207}Pb/^{204}Pb$ against $^{206}Pb/^{204}Pb$ diagram is re-examined on the basis of a statistical box model of mantle processes. Simple equations are presented which relate the pseudo-isochron age to the decay constants and distribution of heterogeneity ages in the model mantle. In turn this age distribution is simply related to the history of melting. The equations are in good agreement with results from mantle convection simulations. The equations are different from but related to, and more general than, those found previously for mean box models. While the pseudo-isochron age does not signify a mean age in the usual sense, in the model presented it is related to a "generalised mean" over the distribution of heterogeneity ages. If a constant melt rate over the Earth's history is assumed, a mean remelting time of 0.5 Ga is required.

Key words: mantle, isochron, isotope ratios, heterogeneity, stirring, mixing

1 Introduction

When $^{207}Pb/^{204}Pb$ is plotted against $^{206}Pb/^{204}Pb$ for data from mid-ocean ridge basalt (MORB) or ocean island basalt (OIB) an approximate linear relationship is found (Figure 1). The slope of a regression line through these data points can be used to an infer an age by treating the regression line as if it were an isochron [1,2]. Formally an isochron age dates a single fractionation event, which is not the case for MORB and OIB; the isotopic systematics of these basalts result from multiple fractionations due to repeated melting and

* *Fax:* +44-1223-360779
Email address: rudge@esc.cam.ac.uk (John F. Rudge).

recycling over the course of the Earth's history. As such the ages calculated by the isochron method are often referred to as pseudo-isochron ages, and the aim of this work is to relate the pseudo-isochron ages to real physical parameters.

The isotopic systems we will study consist of a parent isotope p which decays to a daughter isotope d with decay constant λ. There is a reference isotope d' with respect to which these isotopes are measured. The reference isotope is of the same element as the daughter d, but neither decays nor is a decay product. There are two particular isochrons we will focus on, and we will refer to these as the parent-daughter isochron and the daughter-daughter isochron.

The parent-daughter isochron involves plotting d/d' against p/d'. The parent-daughter isochron age τ_{pdi} is related to the slope β of the regression line by

$$\mathrm{e}^{\lambda\tau_{\text{pdi}}} - 1 = \beta. \tag{1}$$

An example of this is the $^{147}\text{Sm}/^{144}\text{Nd}$-$^{143}\text{Nd}/^{144}\text{Nd}$ diagram. In practice parent-daughter isochrons are not very useful for analysing basalt as the melting that occurs shortly before sampling fractionates parent from daughter, destroying the underlying p/d' signal. As such our main focus will be on the daughter-daughter isochron which is not affected in this way by the melting that occurs before sampling. This is where $(d/d')_2$ is plotted against $(d/d')_1$ for two different isotopic systems which have the same parent and daughter elements. In practice there is just one case in which this applies: the $^{206}\text{Pb}/^{204}\text{Pb}$-$^{207}\text{Pb}/^{204}\text{Pb}$ diagram. The lead-lead (daughter-daughter) isochron age τ_{ddi} is related to the slope β of the regression line by

$$\frac{^{235}\text{U}}{^{238}\text{U}} \cdot \frac{\mathrm{e}^{\lambda_{235}\tau_{\text{ddi}}} - 1}{\mathrm{e}^{\lambda_{238}\tau_{\text{ddi}}} - 1} = \beta, \tag{2}$$

where λ_{235} and λ_{238} are the decay constants of ^{235}U and ^{238}U respectively, and the ratio $^{235}\text{U}/^{238}\text{U}$ has constant value of $1/137.88$ at the present day throughout the solar system. Measured MORB data leads to pseudo-isochron ages around 2.0 Ga (Figure 1) and it is this age that we are most keen on understanding.

There have been a number of attempts to model the MORB and OIB pseudo-isochrons, and they have all been based on the same physical principles of radioactive decay, stirring, melting and mixing. There have been three main approaches: mean box models [3,4,5], statistical box models [6], and numerical simulations of mantle convection [7,8]. This work aims to unite the different approaches and show that most of the modelling work that has been done so far can be best explained by some simple equations derived from a simplified statistical box model of mantle processes [9]. For easier reading, we have split the paper into two parts. The main text discusses the pseudo-isochron equations resulting from the model, and compares them to previous work. The

appendix presents the model derivations.

2 The model

The model used is a generalisation of the simple statistical box model described in [9], and we refer the reader to that paper for a more thorough discussion of the model. The model treats the mantle as a box containing a melting region which is visited by parcels of mantle material on an average time scale τ_{melt}. When parcels enter the melting region a melting event is performed which fractionates parent from daughter, producing a fraction F of melt and a fraction $1 - F$ of residue. G is the molar fraction of a chemical species that enters the melt, and depends on partition coefficients and the melt fraction F. For a given concentration C of species entering the melting region the concentration of the melt produced is CG/F and of the residue is $C(1 - G)/(1 - F)$ (Appendix A). The melt and residue produced are then recycled back into the box. Importantly, we assume that the stirring in the box is strong, and this allows us to treat melting as a Poisson process [10,11] with the statistics of material entering the melting region being the same as those over the whole box. Sampling is modelled by drawing N samples from the box and averaging [12], representative of the mixing that occurs after melting.

There are three crucial differences in the current model from that described in [9]. Firstly, the decay constant λ is no longer assumed small which allows $^{207}\text{Pb}/^{204}\text{Pb}$ to be modelled. Secondly, a steady state is not assumed, but instead the box is assumed to start with uniform concentrations of isotopes at an age τ_{s} before the present. Finally, the melt rate is no longer assumed constant over time but instead given by some function of age $\gamma_{\text{melt}}(\tau)$. For constant melt rate $\gamma_{\text{melt}}(\tau) = 1/\tau_{\text{melt}}$, and we now redefine τ_{melt} as just the melting time scale at the present day.

A number of key analytic results can be derived from the model in the asymptotic limit where $N \to \infty$ (heavy averaging), and all the results in the next section are based on this limit (Appendices B and C). Numerical simulation suggests that the dependence of the pseudo-isochron age on N is fairly weak (Figure 2), and so using the $N \to \infty$ asymptotics seems well justified. In this limit the distribution of isotopic ratios tends to a multivariate normal distribution, and expressions for the corresponding covariance matrix can be derived. In particular these expressions allow us to estimate the slopes of regression lines in plots of one isotopic ratio against another, and thus to derive expressions for the model pseudo-isochron ages.

There are many different ways of fitting a line to a cloud of data points. Following [3] we have focused on the geometric mean regression line (also

known as the reduced major axis regression line) whose slope is given by the ratio of the standard deviations of the two isotopic ratios in question (solid line in Figure 1). To give an estimate of the uncertainty in fitting a line we have also included results from using the two linear regression lines (dotted lines in Figure 1) in some of the figures. If the correlation is good all three lines should be similar.

3 The pseudo-isochron equations

The key to determining the pseudo-isochron ages in the model is the random variable $\hat{T}_{\mathrm{m}}$ which gives the distribution of parcel ages for those parcels that have passed through the melting region. The age of a parcel is defined as the time since last visit to the melting region. The parcels that have not visited the melting region are referred to as primordial parcels, and are not assigned an age. Let $\mathbb{E}$ be an expectation over a random variable, so that

$$\mathbb{E}f(\hat{T}_{\mathrm{m}}) = \int_0^{\tau_{\mathrm{s}}} f(\tau) q_{\mathrm{m}}(\tau)\,\mathrm{d}\tau, \tag{3}$$

where $q_{\mathrm{m}}(\tau)$ is the probability density function of parcel ages, and $f(\tau)$ is an arbitrary given function. The subscript 'm' is to emphasise that we consider only those parcels that have passed through the melting region. In this model the primordial parcels make no contribution to the pseudo-isochron ages, since primordial parcels have uniform isotopic concentrations equal to the mean over the whole box. In terms of $\hat{T}_{\mathrm{m}}$ the model pseudo-isochron equations are simply (Appendices D and E)

$$(\mathrm{e}^{\lambda\tau_{\mathrm{pdi}}} - 1)^2 = \mathbb{E}(\mathrm{e}^{\lambda\hat{T}_{\mathrm{m}}} - 1)^2, \tag{4}$$

$$\frac{(\mathrm{e}^{\lambda_{235}\tau_{\mathrm{ddi}}} - 1)^2}{(\mathrm{e}^{\lambda_{238}\tau_{\mathrm{ddi}}} - 1)^2} = \frac{\mathbb{E}(\mathrm{e}^{\lambda_{235}\hat{T}_{\mathrm{m}}} - 1)^2}{\mathbb{E}(\mathrm{e}^{\lambda_{238}\hat{T}_{\mathrm{m}}} - 1)^2}, \tag{5}$$

where τ_{pdi} and τ_{ddi} are the parent-daughter and lead-lead pseudo-isochron ages respectively. Note that these expressions depend only on the decay constants and the distribution of parcel ages. The expressions clearly show that the pseudo-isochron ages are just particular weighted averages over the parcel ages. In fact the pseudo-isochron ages are examples of "generalised means" [13,14](Appendix F). It also follows from these expressions that the pseudo-isochron ages are always greater than or equal to $\bar{\tau}_{\mathrm{m}} = \mathbb{E}\hat{T}_{\mathrm{m}}$ the mean age of parcels that have passed through the melting region. This is an important result as it means that models with $\bar{\tau}_{\mathrm{m}}$ greater than the observed lead-lead pseudo-isochron age of around 2.0 Ga cannot be compatible with the isotopic observations (such as the Davies model [15]). If the parcel ages $\ll 1/\lambda$ then (4) reduces to $\tau_{\mathrm{pdi}} = \sqrt{\mathbb{E}\hat{T}_{\mathrm{m}}^2}$, a result which is independent of λ. Hence

parent-daughter pseudo-isochron ages will be the same for all slowly decaying isotopes, namely for the ^{147}Sm/^{144}Nd-^{143}Nd/^{144}Nd, ^{87}Rb/^{86}Sr-^{87}Sr/^{86}Sr, ^{176}Lu/^{177}Hf-^{176}Hf/^{177}Hf, and ^{232}Th/^{204}Pb-^{208}Pb/^{204}Pb diagrams.

The history of the rate of melting can be directly related to the distribution of parcel ages in the model. If $\gamma_{\rm melt}(\tau)$ is the melt rate as a function of age then $q_{\rm m}(\tau) = q(\tau)/\int_0^{\tau_{\rm s}} q(\tau)\,{\rm d}\tau$ where (Appendix G)

$$q(\tau) = \gamma_{\rm melt}(\tau) \exp\left(-\int_0^{\tau} \gamma_{\rm melt}(\tau)\,{\rm d}\tau\right). \tag{6}$$

An important special case is where melt rate is constant $\gamma_{\rm melt}(\tau) = 1/\tau_{\rm melt}$, where $\tau_{\rm melt}$ is a constant melting time scale. In this case $q_{\rm m}(\tau) = {\rm e}^{-\tau/\tau_{\rm melt}}/(\tau_{\rm melt}(1-{\rm e}^{-\tau_{\rm s}/\tau_{\rm melt}}))$ and the model parent-daughter pseudo-isochron equation is (Appendix H)

$$({\rm e}^{\lambda\tau_{\rm pdi}}-1)^2 = \frac{\int_0^{\tau_{\rm s}}({\rm e}^{\lambda\tau}-1)^2\,{\rm e}^{-\tau/\tau_{\rm melt}}\,{\rm d}\tau}{\tau_{\rm melt}\,(1-{\rm e}^{-\tau_{\rm s}/\tau_{\rm melt}})}. \tag{7}$$

The most important feature of this equation is that it depends only on the three time scale parameters in the problem: the melting time scale $\tau_{\rm melt}$, the start age $\tau_{\rm s}$, and the decay constant λ; and not on any of the other parameters. Figure 3 plots solutions to (7) for ^{143}Nd/^{144}Nd-^{147}Sm/^{144}Nd ($\lambda = 0.00654$ Ga^{-1}) in two different ways. Figure 4 shows similar graphs for ^{235}U/^{204}Pb-^{207}Pb/^{204}Pb ($\lambda = 0.985$ Ga^{-1}), a case for which the decay is not linearisable. Note that there is a reasonable uncertainty in the parent-daughter pseudo-isochron ages indicated by the wide grey region in Figures 3b and 4b. This expected (and indeed observed) lack of very good correlation is another reason why parent-daughter pseudo-isochron ages are not particularly useful.

The more useful pseudo-isochron equation is the lead-lead pseudo-isochron equation given for constant melt rate by (Appendix I)

$$\frac{({\rm e}^{\lambda_{235}\tau_{\rm ddi}}-1)^2}{({\rm e}^{\lambda_{238}\tau_{\rm ddi}}-1)^2} = \frac{\int_0^{\tau_{\rm s}}({\rm e}^{\lambda_{235}\tau}-1)^2\,{\rm e}^{-\tau/\tau_{\rm melt}}\,{\rm d}\tau}{\int_0^{\tau_{\rm s}}({\rm e}^{\lambda_{238}\tau}-1)^2\,{\rm e}^{-\tau/\tau_{\rm melt}}\,{\rm d}\tau}. \tag{8}$$

Figure 5 plots solutions to (8). Note that the lead-lead pseudo-isochron ages are fairly well constrained as the model correlation is good (as indicated by the narrow grey regions of Figure 5b).

The pseudo-isochron equations only encode age information, and do not involve the parameters G and F. However, if we are concerned with the variance of isotopic ratios, or the slopes in plots of one isotopic ratio against another which do not have common parent and daughter elements, then these parameters are involved. This was the main focus of [9], and the corresponding generalisation of (1) of [9] for the standard deviation σ of d/d' ratios after

sampling is (Appendix C)

$$\sigma = \frac{\bar{p}}{\bar{d}'} \cdot \frac{|G_p - G_d|}{\sqrt{NF\,(1-F)}} \cdot \sqrt{\int_0^{\tau_s} \left(\mathrm{e}^{\lambda\tau} - 1\right)^2 q(\tau)\,\mathrm{d}\tau}, \tag{9}$$

where $\bar{p}/\bar{d}'$ is the ratio of mean parent isotope concentration to mean reference isotope concentration over the whole box at the present day.

4 Linear pseudo-isochron equations

There are three models based on linear evolution equations that have recently been proposed for the pseudo-isochrons [3,4,5]. In fact, all three models produce identical pseudo-isochron equations and are closely related. Albarède [4] and Donnelly et al. [5] both consider a two reservoir model, and derive the pseudo-isochron equations in precisely the same way. Allègre and Lewin [3] consider a different set of linear evolution equations for isotopic dispersion, which with some rearranging are almost equivalent to the two interacting reservoir equations.

The governing equations for two interacting reservoirs derived by [4] and [5] are

$$\frac{\mathrm{d}n_1}{\mathrm{d}t} = -\frac{n_1}{\tau_1} + \frac{n_2}{\tau_2} - \lambda n_1, \tag{10}$$

$$\frac{\mathrm{d}n_2}{\mathrm{d}t} = \frac{n_1}{\tau_1} - \frac{n_2}{\tau_2} - \lambda n_2, \tag{11}$$

$$\frac{\mathrm{d}m_1}{\mathrm{d}t} = -\frac{m_1}{\theta_1} + \frac{m_2}{\theta_2} + \lambda n_1, \tag{12}$$

$$\frac{\mathrm{d}m_2}{\mathrm{d}t} = \frac{m_1}{\theta_1} - \frac{m_2}{\theta_2} + \lambda n_2, \tag{13}$$

$$\frac{\mathrm{d}s_1}{\mathrm{d}t} = -\frac{s_1}{\theta_1} + \frac{s_2}{\theta_2}, \tag{14}$$

$$\frac{\mathrm{d}s_2}{\mathrm{d}t} = \frac{s_1}{\theta_1} - \frac{s_2}{\theta_2}, \tag{15}$$

where the notation of [4] has been followed. n_i, m_i and s_i are the *total number of moles* of parent, daughter and reference isotopes respectively in reservoir i (these should not be confused with p, d and d' used throughout the appendix to represent *concentrations*). τ_i is the residence time of the parent element in reservoir i, θ_i the corresponding residence time for the daughter element. Let Σ_n, Σ_m and Σ_s be the total number of moles of parent, daughter and reference isotopes in both reservoirs: $\Sigma_n = n_1 + n_2$, $\Sigma_m = m_1 + m_2$, and $\Sigma_s = s_1 + s_2$.

Then

$$\frac{\mathrm{d}\Sigma_n}{\mathrm{d}t} = -\lambda\Sigma_n, \quad \frac{\mathrm{d}\Sigma_m}{\mathrm{d}t} = \lambda\Sigma_n, \quad \frac{\mathrm{d}\Sigma_s}{\mathrm{d}t} = 0, \tag{16}$$

and the governing equations for reservoir 1 can be rewritten as

$$\frac{\mathrm{d}n_1}{\mathrm{d}t} = -\frac{n_1}{\tau} + \frac{\Sigma_n}{\tau_2} - \lambda n_1, \tag{17}$$

$$\frac{\mathrm{d}m_1}{\mathrm{d}t} = -\frac{m_1}{\theta} + \frac{\Sigma_m}{\theta_2} + \lambda n_1, \tag{18}$$

$$\frac{\mathrm{d}s_1}{\mathrm{d}t} = -\frac{s_1}{\theta} + \frac{\Sigma_s}{\theta_2}, \tag{19}$$

where τ and θ are the relaxation times of the two elements: the harmonic means of the residence times in the two reservoirs

$$\frac{1}{\tau} = \frac{1}{\tau_1} + \frac{1}{\tau_2}, \quad \frac{1}{\theta} = \frac{1}{\theta_1} + \frac{1}{\theta_2}. \tag{20}$$

We now rewrite these equations for closer comparison with [3]. Introduce new variables $n_1^\star$ and $m_1^\star$ defined by

$$n_1^\star = \frac{1}{\Sigma_s}\left(n_1 - s_1\left(\frac{n}{s}\right)_{1+2}\right), \tag{21}$$

$$m_1^\star = \frac{1}{\Sigma_s}\left(m_1 - s_1\left(\frac{m}{s}\right)_{1+2}\right), \tag{22}$$

where the subscript 1+2 refers to the total system: $(n/s)_{1+2} = \Sigma_n/\Sigma_s$ and $(m/s)_{1+2} = \Sigma_m/\Sigma_s$. Note that $n_2^\star = -n_1^\star$ and $m_2^\star = -m_1^\star$. The governing equations in the starred variables can then be written as

$$\frac{\mathrm{d}n_1^\star}{\mathrm{d}t} = \left(\frac{1}{\tau_2} - \frac{1}{\theta_2}\right)\left(\frac{n}{s}\right)_{1+2} - \lambda n_1^\star - \frac{1}{\Sigma_s}\left(\frac{n_1}{\tau} - \frac{s_1}{\theta}\left(\frac{n}{s}\right)_{1+2}\right), \tag{23}$$

$$\frac{\mathrm{d}m_1^\star}{\mathrm{d}t} = \lambda n_1^\star - \frac{m_1^\star}{\theta}. \tag{24}$$

These governing equations take on a particularly simple form if the relaxation times for parent and daughter elements are the same, $\tau = \theta = 1/\gamma$ say. For later convenience, define $g_n = \tau/\tau_1$ and $g_m = \theta/\theta_1$, and note that $0 \le g_n, g_m \le 1$. Then

$$\frac{\mathrm{d}n_1^\star}{\mathrm{d}t} = -\left(g_n - g_m\right)\gamma\left(\frac{n}{s}\right)_{1+2} - \left(\lambda + \gamma\right)n_1^\star, \tag{25}$$

$$\frac{\mathrm{d}m_1^\star}{\mathrm{d}t} = \lambda n_1^\star - \gamma m_1^\star. \tag{26}$$

These can be compared to the governing equations (1) and (2) of Allègre and

Lewin [3]

$$\frac{\mathrm{d}\langle\mu\rangle(t)}{\mathrm{d}t} = A(t) - (\lambda + M(t))\,\langle\mu\rangle(t), \tag{27}$$

$$\frac{\mathrm{d}\langle\alpha\rangle(t)}{\mathrm{d}t} = \lambda\langle\mu\rangle(t) + B(t) - M(t)\langle\alpha\rangle(t). \tag{28}$$

These governing equations are identical to (25) and (26) if the chemical dispersion $\langle\mu\rangle(t) = n_1^\star$, the isotopic dispersion $\langle\alpha\rangle(t) = m_1^\star$, the rate of injection of chemical heterogeneity $A(t) = -(g_n - g_m)\,\gamma\,(n/s)_{1+2}$, the rate of injection of isotopic heterogeneity $B(t) = 0$, and the stirring parameter $M(t) = \gamma$. The appendix of [3] discusses solutions to these equations in the simplified form ((A1-1) and (A1-2) of [3])

$$\frac{\mathrm{d}\langle\mu\rangle(t)}{\mathrm{d}t} = \frac{\Delta\langle\mu\rangle\,\mathrm{e}^{-\lambda t}}{R} - \left(\lambda + \tau_{\mathrm{stir}}^{-1}\right)\langle\mu\rangle(t), \tag{29}$$

$$\frac{\mathrm{d}\langle\alpha\rangle(t)}{\mathrm{d}t} = \lambda\langle\mu\rangle(t) + \frac{\Delta\langle\alpha\rangle}{R} - \frac{\langle\alpha\rangle(t)}{\tau_{\mathrm{stir}}}, \tag{30}$$

where by comparison $\Delta\langle\mu\rangle/R = -(g_n - g_m)\,\gamma(n/s)_{1+2}(0)$, $\Delta\langle\alpha\rangle = 0$, and $\tau_{\mathrm{stir}} = \gamma^{-1}$. These are solved subject to initial conditions $\langle\mu\rangle(0) = 0$, $\langle\alpha\rangle(0) = 0$ which corresponds to the two reservoirs having initially identical isotopic ratios $(n/s)_1(0) = (n/s)_2(0) = (n/s)_{1+2}(0)$ and $(m/s)_1(0) = (m/s)_2(0) = (m/s)_{1+2}(0)$. Allègre and Lewin define the slope of their parent-daughter pseudo-isochron by the ratio $\langle\alpha\rangle(t)/\langle\mu\rangle(t)$. However,

$$\frac{\langle\alpha\rangle(t)}{\langle\mu\rangle(t)} = \frac{m_1^\star}{n_1^\star} = \frac{m_1 - s_1\,(m/s)_{1+2}}{n_1 - s_1\,(n/s)_{1+2}} = \frac{(m/s)_1 - (m/s)_{1+2}}{(n/s)_1 - (n/s)_{1+2}}. \tag{31}$$

Hence the definition of the slope by the ratio $\langle\alpha\rangle(t)/\langle\mu\rangle(t)$ is the same as the slope of the line through the reservoirs 1, 2 and 1+2 on the parent-daughter isochron diagram. Thus provided the relaxation times are equal in the two reservoir model, and there is no excess isotopic heterogeneity $\Delta\langle\alpha\rangle$ in the Allègre and Lewin model, the pseudo-isochron equations are exactly the same. Importantly, the stirring time τ_{stir} of Allègre and Lewin can be reinterpreted as the common relaxation time in the two reservoir model.

In one sense the Allègre and Lewin model is more general than the two reservoir model because of the excess isotopic heterogeneity term $\Delta\langle\alpha\rangle$, but in practice this was always set to zero for their pseudo-isochron calculations. On the other hand, the two reservoir model can be thought of as more general since it allows the parent and daughter elements to have different relaxation times. If parent and daughter are fractionated by the same melting process we might expect relaxation times to be the same. Also, note that in secular equilibrium only the relaxation time θ of the daughter element determines the pseudo-isochron [4].

Standard deviations are not conservative and should not be modelled by linear evolution equations such as (27) and (28). The justification of these equations by Allègre and Lewin is rather ad hoc. The two reservoir model and the Allègre and Lewin model are both linear models and thus essentially concerned with mean values, whereas standard deviations are fundamentally nonlinear. In the statistical box model it possible to discuss standard deviations, variances and covariances, as well as means, because the underlying probability distributions are being modelled. Note that the starred variables which relate the Allègre and Lewin model to the two reservoir model arise naturally from the linearisation

$$\frac{x}{y} - \frac{\bar{x}}{\bar{y}} \approx \frac{1}{\bar{y}} \left(x - y \frac{\bar{x}}{\bar{y}} \right), \tag{32}$$

valid for $|x - \bar{x}| \ll \bar{x}$ and $|y - \bar{y}| \ll \bar{y}$. The starred variables turn out to be particularly useful when considering asymptotics for large N in the statistical box model (Appendix B).

We can make an important connection between the statistical box model and the two reservoir model by dividing the parcels in the box into two groupings. Labelling the residue parcels and a fraction $1 - F$ of the primordial parcels as one reservoir, and the melt parcels and a fraction F of the primordial parcels as another reservoir, gives a two reservoir system with a common relaxation time. The connection is made if $\gamma = \gamma_{\text{melt}}$, $g_n = G_p$ and $g_m = G_d$ (Appendix J). Furthermore, this suggests an alternative way of writing the linear pseudo-isochron equations of Allègre and Lewin in terms of an age distribution. These are the same as (4) and (5) except with the squareds removed

$$\mathrm{e}^{\lambda \tau_{\text{pdil}}} - 1 = \mathbb{E}(\mathrm{e}^{\lambda \hat{T}_{\text{m}}} - 1), \tag{33}$$

$$\frac{\mathrm{e}^{\lambda_{235} \tau_{\text{ddil}}} - 1}{\mathrm{e}^{\lambda_{238} \tau_{\text{ddil}}} - 1} = \frac{\mathbb{E}(\mathrm{e}^{\lambda_{235} \hat{T}_{\text{m}}} - 1)}{\mathbb{E}(\mathrm{e}^{\lambda_{238} \hat{T}_{\text{m}}} - 1)}, \tag{34}$$

where τ_{pdil} and τ_{ddil} are the linear parent-daughter and lead-lead pseudo-isochron ages respectively. It should be noted that the linear pseudo-isochron ages will always be less than the corresponding ages obtained from (4) and (5). The pseudo-isochron equations (4) and (5) for our problem are different because they involve the variance of a mixture of melt, residue and primordial parcels, whereas the above equations result from mean values of reservoirs. The squareds reflect the difference between looking at a mean value and looking at a variance. To emphasise the similarities and differences between the linear pseudo-isochron equations and our pseudo-isochron equations, Figures 3, 4, and 5 mimic Figures 3, 4 and 9 of Allègre and Lewin [3].

A key result used in [3] to estimate a stirring time for the mantle is the lead-lead pseudo-isochron age relationship $\tau_{\text{ddil}} \sim 2\tau_{\text{stir}}$ for vigorous stirring. In the context of this paper the corresponding asymptotic result is $\tau_{\text{ddi}} \sim 3\tau_{\text{melt}}$ for rapid remelting. In practice this result is only accurate for very rapid

remelting, as it requires that $\tau_{\rm melt} \ll \tau_{\rm s}$ and $\tau_{\rm melt} \ll 1/\lambda_{235} = 1.0$ Ga. Hence comparison between the linear pseudo-isochron equations and our pseudo-isochron equations is best done using the full equations rather than any rapid remelting asymptotics. Using (A2-9) of [3], if $\tau_{\rm s} = 4.5$ Ga then $\tau_{\rm stir} = 0.82$ Ga is needed to produce a lead-lead pseudo-isochron age of 2.0 Ga. Using our (8), if $\tau_{\rm s} = 4.5$ Ga then $\tau_{\rm melt} = 0.45$ Ga is needed, which is almost a factor of 2 less. For parent-daughter isochrons the vigorous stirring relationship $\tau_{\rm pdil} \sim \tau_{\rm stir}$ in [3] becomes $\tau_{\rm pdi} \sim \sqrt{2}\tau_{\rm melt}$ for rapid remelting ($\tau_{\rm melt} \ll \tau_{\rm s}$ and $1/\lambda$).

The statistical model of Kellogg et al. [16] also has a $\tau_{\rm stir}$ parameter, but note that this takes on a different meaning to the $\tau_{\rm stir}$ of the Allègre and Lewin model [3]. The Kellogg et al. $\tau_{\rm stir}$ reduces the length scale of heterogeneities before sampling: essentially it relates to the parameter N in our model. The $\tau_{\rm stir}$ in the Allègre and Lewin model describes the destruction of heterogeneity by repeated melting, and thus is $\tau_{\rm melt}$ in our model.

5 Numerical simulations of mantle convection

Christensen and Hofmann [7] put isotopic tracers into a numerical simulation of mantle convection. 252,000 tracers were used, and sampling was performed by dividing the domain into 40 $\times$ 20 sampling cells, and averaging over those cells. Thus N for their model is around 300, although note that their model is slightly different to that presented here as different cells may have different numbers of tracers. They have a constant rate of melting in their model, so (8) applies. Their standard model has $\tau_{\rm s} = 3.6$ Ga and $\tau_{\rm melt} = -1.36/\log(1-0.9) = 0.59$ Ga (inferred from the statement "in 1.36 Ga, statistically 90% of the total basalt content has been cycled through a melting zone in the model"). They found a lead-lead pseudo-isochron age of 2.10 Ga which compares very favourably to the figure of 2.15 Ga that is predicted by (8) (and not as favourably with the figure of 1.39 Ga predicted by the linear pseudo-isochron equation (A2-9) of [3]).

Xie and Tackley [8] used a similar approach to Christensen and Hofmann [7] for a different numerical simulation. Their model differs in having a melt rate which changes over time; melting being more vigorous in the past. 400,000 tracers were used, and the domain divided into 256 $\times$ 64 sampling cells, and thus N for their model is around 25. This smaller amount of averaging probably explains why their arrays show less correlation than the Christensen and Hofmann arrays. Their model has less frequent remelting than the Christensen and Hofmann model, which is why larger isotopic ages are observed. A rough rule of thumb for less frequent melting ($\tau_{\rm melt}$ = 1-2 Ga) is $\tau_{\rm ddi} \approx 0.75\tau_{\rm s}$ (Figure 5a). The rule works reasonably well at estimating the pseudo-isochron ages they found, but the simple constant melt rate formula of (8) does not actually

apply in this case. To do a more accurate comparison we need to examine carefully the distribution of ages the Xie and Tackley model produces.

In Figures 5c and 5e of [8] the integrated crustal production and crustal production rate are plotted against time for a $\tau_{\mathrm{s}} = 4.5$ Ga run. We can use this information in our statistical box model to predict the distribution of ages (Appendix G). Figure 6 shows there is good agreement between these calculated distributions and the age distribution observed (Figure 10a of [8]). Using (4) and (5) we can use these age distributions to calculate pseudo-isochron ages. The integrated crustal production gives a lead-lead pseudo-isochron age of 3.58 Ga, the crustal production rate an age of 3.76 Ga, and the observed age distribution an age of 3.67 Ga. These estimates are all slightly greater than the observed pseudo-isochron age of 3.39 Ga (Figure 8a of [8]). The small differences may be due to various factors, including differences in line fitting, the effects of binning and windowing, and the small value of N used in their simulations.

Xie and Tackley discussed the effect of greater rates of melting in the past on pseudo-isochron age. To examine this they introduced a box model and calculated a mean parcel age for three different scenarios. Using our statistical box model we can go further and calculate not only the mean age but also the age distributions and pseudo-isochron ages. Their three different scenarios were (1) melt rate constant, (2) melt rate proportional to the square of mantle heating rate (referred to as H^2), and (3) as H^2 but also proportional to depth of melting (referred to as $H^2 d$). Mantle heating rate was assumed to decay with a half-life of 2.247 Ga. This corresponds to an exponentially growing history of melt rate $\gamma_{\mathrm{melt}}(\tau) = \mathrm{e}^{\nu\tau}/\tau_{\mathrm{melt}}$ where $\nu = (2\log 2)/2.247 = 0.62$ Ga^{-1} and τ_{melt} is the melting time scale at the present day. We have approximated the $H^2 d$ case by a similar exponential growth with $\nu = 0.74$ Ga^{-1}; only a small discrepancy in the mean ages results from this approximation. The age distributions are plotted in Figure 7 and the mean and pseudo-isochron ages are given in Table 1.

An important distinction should be drawn between the two mean ages given in Table 1. Xie and Tackley [8] calculated the mean mantle age $\bar{\tau}_{\mathrm{total}}$ which includes all parcels, and in particular it includes the primordial parcels which are assigned an age of τ_{s}. This is a misleading mean age to relate to the pseudo-isochron ages, since for the pseudo-isochron ages it is only the parcels that have passed through the melting region which matter. Hence of greater interest is $\bar{\tau}_{\mathrm{m}} = \mathbb{E}\hat{T}_{\mathrm{m}}$ which is a mean age that only includes the parcels that have passed through the melting region. It can be formally shown that the pseudo-isochron ages must be greater than or equal to $\bar{\tau}_{\mathrm{m}}$, but there is no such constraint on $\bar{\tau}_{\mathrm{total}}$. Indeed Table 1 shows pseudo-isochron ages both greater than and less than $\bar{\tau}_{\mathrm{total}}$.

Note that the smallest Pb-Pb pseudo-isochron age in Table 1 is 2.97 Ga, so none of the scenarios are compatible with the measured MORB data. Furthermore note that more frequent remelting in the past can have the effect of both increasing and decreasing the pseudo-isochron age (e.g. $\tau_{\rm melt} = 13.8$ Ga in Table 1). This is due to two competing effects: if melting is more rapid in the past it can mean a lot of the early heterogeneity is destroyed, but it can also mean that a lot of heterogeneity is created early on. For $\nu = 0.62$ Ga^{-1} and $\nu = 0.74$ Ga^{-1} to get lead-lead pseudo-isochron ages of 2.0 Ga the present melting time scale $\tau_{\rm melt}$ must be 1.45 Ga and 1.83 Ga respectively.

6 Conclusions

The key results of this work are the simple pseudo-isochron equations (4) and (5) which relate in a straightforward way melting history and pseudo-isochron age. The equations are similar to those examined by previous authors but differ in certain important details. In the case of constant melt rate these relationships can expressed in terms of just two unknown parameters: the age $\tau_{\rm s}$ at which you begin the model, and $\tau_{\rm melt}$ the melting time scale. The natural choice for $\tau_{\rm s}$ is the age of the Earth, but the unmodeled process of continent formation motivates younger choices for $\tau_{\rm s}$ such as 3.6 Ga [7]. To produce the observed lead-lead pseudo-isochron ages of 2.0 Ga values of $\tau_{\rm melt} = 0.45$ Ga (for $\tau_{\rm s} = 4.5$ Ga) or $\tau_{\rm melt} = 0.52$ Ga (for $\tau_{\rm s} = 3.6$ Ga) are needed. If the melt rate is not constant but is instead greater in the past, then slower present day melting time scales may be needed, for example $\tau_{\rm melt} = 1.45$ Ga found earlier when the rate of melting was proportional to the square of mantle heating rate. Figure 8 plots the corresponding probability density functions for a constant melt rate case and a variable melt rate case.

There is quantitative agreement between the predictions of this simple statistical box model and the much more complicated mantle convection calculations that have been done by previous authors. This helps justify some of the key assumptions that are made in the statistical box model, and in particular the neglect of the particular details of the underlying flow. Statistical box models are still in their infancy, but they seem to provide a powerful way of approaching the problem of isotopic heterogeneity in the mantle. A great advantage is that general analytical results can derived which can lead to better understanding of the problem without the need to run lots of numerical simulations. However, mantle convection simulations still have an important role to play; of most interest for the pseudo-isochron problem is knowing the kinds of melting history that they can produce. Future simulations should investigate further the validity of these simple analytic relationships between melt rate and age distribution, and between age distribution and pseudo-isochron age.

An important caveat to bear in mind when using this model, and indeed in the convection simulations as well, is that no attempt is made to model continent formation. This is an important process to consider as early heterogeneity may become locked up in the continents and not be recycled into the mantle. It is this idea which prompts the use of younger values of τ_{s}. The multi-reservoir statistical box models of [6] have made some attempt at addressing this issue, although as is remarked it is somewhat difficult to see through the myriad of parameters used in that model. A key problem in modelling continent formation is that it requires modelling properly the mean lead isotopic evolution of mantle, which in itself is still not clearly understood. There is still yet to be any consensus on the resolution of the various lead paradoxes.

For isotopic ratios d/d' for which the decay is linearisable the outcome of the model we present here is little changed from that described in [9]. In particular the slopes in plots of one isotopic system against another are unchanged ((5) of [9]). There thus remains the fundamental inconsistency between the Pb isotopes and the Nd, Sr and Hf isotopes. Furthermore, the value of $\tau_{\mathrm{melt}} \sim 0.5$ Ga inferred here from the lead-lead pseudo-isochron is a factor of 3 different from the value of $\tau_{\mathrm{melt}} \sim 1.7$ Ga inferred from the standard deviations of the Nd, Sr and Hf isotopic ratios in [9]. In fact, $\tau_{\mathrm{melt}} \sim 0.5$ Ga provides a much better match to the observed standard deviations of the Pb isotopic ratios than $\tau_{\mathrm{melt}} \sim 1.7$ Ga (in [9], model lead isotopic standard deviations were around 3 times larger than observed). Note that the pseudo-isochron is a much more robust way of estimating τ_{melt} as it depends only on the decay constants and start age, whereas the standard deviation of isotopic ratios depends on all the parameters in the model. It is still not possible in the model as it stands to match both the Pb isotopes and the Nd, Sr and Hf isotopes with a single set of parameters.

As outlined in the conclusion of [9] there are still many possible generalisations of this model that might prove useful. One particular generalisation that we are interested in developing is to look at removal of species from the box during a melting event. This could model continent formation or the degassing of isotopes of noble gases (such as ^{3}He and ^{4}He). Again, further investigation is needed.

Acknowledgements

I thank Dan McKenzie and Peter Haynes for all their support and advice. I particularly thank Francis Albarède for his thoughtful review and very helpful discussions. This project was supported under the Environmental Mathematics and Statistics initiative of NERC and EPSRC.

A Statistical box model

The model used is the simple statistical box model described in [9] with three restrictions lifted: the decay is no longer linearised, a steady state is not assumed, and the melt rate may vary with time. There are three types of parcel in our model mantle: melt parcels (parcels last formed as melt in the melting region), residue parcels (parcels last formed as residue in the melting region), and primordial parcels. Melt and residue parcels have an associated age τ of when they last visited the melting region. The distribution of parcel ages is given by a random variable $\hat{T}$, which is defined so that those parcels of age $\geq \tau_\mathrm{s}$ are primordial. Let $q(\tau)$ be the probability density function of $\hat{T}$. Thus a fraction $1 - \int_0^{\tau_\mathrm{s}} q(\tau)\,\mathrm{d}\tau$ of parcels are primordial.

We track the concentrations of a parent isotope p, a daughter d, and a reference d'. G is the molar fraction of a chemical species that enters the melt, and depends on the melt fraction F and the partition coefficients. The corresponding concentrations of the three parcel types at the present day are:

$$p_\mathrm{prim} = \bar{p}, \tag{A.1}$$
$$d_\mathrm{prim} = \bar{d}, \tag{A.2}$$
$$d'_\mathrm{prim} = \bar{d}'. \tag{A.3}$$

$$p_\mathrm{melt}(\tau) = \frac{G_p}{F}\bar{p}, \tag{A.4}$$
$$d_\mathrm{melt}(\tau) = \frac{G_d}{F}\bar{d} + \frac{G_p - G_d}{F}\left(\mathrm{e}^{\lambda\tau} - 1\right)\bar{p}, \tag{A.5}$$
$$d'_\mathrm{melt}(\tau) = \frac{G_d}{F}\bar{d}'. \tag{A.6}$$

$$p_\mathrm{res}(\tau) = \frac{1 - G_p}{1 - F}\bar{p}, \tag{A.7}$$
$$d_\mathrm{res}(\tau) = \frac{1 - G_d}{1 - F}\bar{d} - \frac{G_p - G_d}{1 - F}\left(\mathrm{e}^{\lambda\tau} - 1\right)\bar{p}, \tag{A.8}$$
$$d'_\mathrm{res}(\tau) = \frac{1 - G_d}{1 - F}\bar{d}', \tag{A.9}$$

where $\bar{p}$, $\bar{d}$, and $\bar{d}'$ are the mean concentrations of the isotopes in the box at the present day, and λ is the decay constant.

B The asymptotics of averaging ratio quantities

As before, we model sampling by taking a number N of independent identically distributed (i.i.d.) samples from our model mantle and averaging. Again, it is important to note that we are interested in ratio quantities. This section describes some important general results on the asymptotics of averaging ratio quantities.

Consider i.i.d. pairs of random variables $\{\hat{x}_i, \hat{y}_i\}$, $i = 1, 2, ...N$. Suppose $\hat{y}_i > 0$. We are interested in the asymptotic behaviour of the ratio of sums

$$\hat{Z} = \sum_i \hat{x}_i / \sum_i \hat{y}_i \tag{B.1}$$

for N large. Let

$$\hat{x}^\star = \frac{1}{\bar{y}} \left(\hat{x} - \frac{\bar{x}}{\bar{y}} \hat{y} \right), \tag{B.2}$$

$$\hat{y}^\star = \frac{1}{\bar{y}} \left(\hat{y} - \bar{y} \right), \tag{B.3}$$

where $\bar{x} = \mathbb{E}(\hat{x})$, $\bar{y} = \mathbb{E}(\hat{y})$. Note that $\mathbb{E}(\hat{x}^\star) = \mathbb{E}(\hat{y}^\star) = 0$. Note also that

$$\hat{Z} = \frac{\sum \hat{x}_i}{\sum \hat{y}_i} = \frac{\bar{x}}{\bar{y}} + \frac{\sum \hat{x}_i^\star / N}{1 + \sum \hat{y}_i^\star / N}. \tag{B.4}$$

It can be shown [17] that $\hat{Z}$ is asymptotically normal for large N under appropriate assumptions (assumptions which give the central limit theorem for $\sum \hat{x}_i^\star$ and give the law of large numbers for $\sum \hat{y}_i$). The condition $\hat{y}_i > 0$ ensures that moments of $\hat{Z}$ are well defined, and Cauchy distribution problems do not arise. By Taylor expanding (B.4) and taking expectations the following expressions for the asymptotic moments can be derived:

$$\mu = \bar{Z} = \mathbb{E}(\hat{Z}) = \frac{\bar{x}}{\bar{y}} - \frac{1}{N} \mathbb{E}\left(\hat{x}^\star \hat{y}^\star \right) + O\left(\frac{1}{N^2} \right), \tag{B.5}$$

$$\mu_2 = \sigma^2 = \mathbb{E}\left(\hat{Z} - \bar{Z} \right)^2 = \frac{1}{N} \mathbb{E}\left(\hat{x}^{\star 2} \right) + O\left(\frac{1}{N^2} \right), \tag{B.6}$$

$$\mu_3 = \mathbb{E}\left(\hat{Z} - \bar{Z} \right)^3 = \frac{1}{N^2} \left(\mathbb{E}\left(\hat{x}^{\star 3} \right) - 6\mathbb{E}\left(\hat{x}^\star \hat{y}^\star \right) \mathbb{E}\left(\hat{x}^{\star 2} \right) \right) + O\left(\frac{1}{N^3} \right). \tag{B.7}$$

Hence the skew parameter γ_1 is

$$\gamma_1 = \frac{\mu_3}{(\mu_2)^{3/2}} = \frac{\mathbb{E}\left(\hat{x}^{\star 3} \right) - 6\mathbb{E}\left(\hat{x}^\star \hat{y}^\star \right) \mathbb{E}\left(\hat{x}^{\star 2} \right)}{N^{1/2} \left(\mathbb{E}\left(x^{\star 2} \right) \right)^{3/2}} + O\left(\frac{1}{N^{3/2}} \right). \tag{B.8}$$

The kurtosis and higher order moments can be derived similarly by expanding to higher orders.

In this work we are most concerned with plots of one ratio against another. So consider two sets of i.i.d. pairs of random variables $\{\hat{x}_i, \hat{y}_i\}_1$ and $\{\hat{x}_i, \hat{y}_i\}_2$, $i = 1, 2, ...N$. Then the covariance of $\hat{Z}_1$ and $\hat{Z}_2$ is given by

$$\begin{aligned} \operatorname{cov}\left(\hat{Z}_1, \hat{Z}_2\right) &= \mathbb{E}\left(\left(\hat{Z}_1 - \bar{Z}_1\right)\left(\hat{Z}_2 - \bar{Z}_2\right)\right) = \frac{1}{N}\mathbb{E}\left(\hat{x}_1^\star \hat{x}_2^\star\right) + O\left(\frac{1}{N^2}\right) \\ &= \frac{1}{N}\operatorname{cov}\left(\hat{x}_1^\star, \hat{x}_2^\star\right) + O\left(\frac{1}{N^2}\right), \end{aligned} \tag{B.9}$$

with corresponding correlation

$$\begin{aligned} r_{12} = \operatorname{cor}\left(\hat{Z}_1, \hat{Z}_2\right) &= \frac{\operatorname{cov}\left(\hat{Z}_1, \hat{Z}_2\right)}{\sigma_1 \sigma_2} = \frac{\mathbb{E}\left(\hat{x}_1^\star \hat{x}_2^\star\right)}{\sqrt{\mathbb{E}\left(\hat{x}_1^{\star 2}\right)\mathbb{E}\left(\hat{x}_2^{\star 2}\right)}} + O\left(\frac{1}{N}\right) \\ &= \operatorname{cor}\left(\hat{x}_1^\star, \hat{x}_2^\star\right) + O\left(\frac{1}{N}\right). \end{aligned} \tag{B.10}$$

We are most interested in calculating the slope of a regression line of one system against another. There are many different methods for fitting regression lines to a cloud of data points which make various assumptions about the underlying data. Three commonly used estimates are

$$\beta_{12} = r_{12}\frac{\sigma_2}{\sigma_1} = \frac{\mathbb{E}(\hat{x}_1^\star \hat{x}_2^\star)}{\mathbb{E}(\hat{x}_1^{\star 2})} + O\left(\frac{1}{N}\right), \tag{B.11}$$

$$\bar{\beta} = \frac{\sigma_2}{\sigma_1} = \sqrt{\frac{\mathbb{E}(\hat{x}_2^{\star 2})}{\mathbb{E}(\hat{x}_1^{\star 2})}} + O\left(\frac{1}{N}\right), \tag{B.12}$$

$$\beta_{21} = \frac{1}{r_{12}}\frac{\sigma_2}{\sigma_1} = \frac{\mathbb{E}(\hat{x}_2^{\star 2})}{\mathbb{E}(\hat{x}_1^\star \hat{x}_2^\star)} + O\left(\frac{1}{N}\right), \tag{B.13}$$

where β_{12} is the slope of the linear least squares regression line of system 2 on system 1, and β_{21} is the same line but for system 1 on system 2. $\bar{\beta}$ is the slope of the geometric mean regression line, and following [3] it is this estimate of the slope we will focus on here. Note that the sign of $\bar{\beta}$ is chosen to be the same as r_{12} and that $\beta_{12} \leq \bar{\beta} \leq \beta_{21}$. If the correlation is good all three estimates will be similar.

C Calculating the asymptotic moments

Based on the general analysis of the previous section, introduce new variables $p^\star$, $d^\star$ and $d'^\star$ defined by:

$$p^\star = \frac{1}{\bar{d}'}\left(p - \frac{\bar{p}}{\bar{d}'}d'\right), \tag{C.1}$$

$$d^\star = \frac{1}{\bar{d}'}\left(d - \frac{\bar{d}}{\bar{d}'}d'\right), \tag{C.2}$$

$$d'^\star = \frac{1}{\bar{d}'}\left(d' - \bar{d}'\right). \tag{C.3}$$

In these new variables the different parcel types are given by

$$p^\star_{\text{prim}} = 0, \tag{C.4}$$

$$d^\star_{\text{prim}} = 0, \tag{C.5}$$

$$d'^\star_{\text{prim}} = 0. \tag{C.6}$$

$$p^\star_{\text{melt}}(\tau) = \frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{F}, \tag{C.7}$$

$$d^\star_{\text{melt}}(\tau) = \frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{F}\left(\mathrm{e}^{\lambda\tau} - 1\right), \tag{C.8}$$

$$d'^\star_{\text{melt}}(\tau) = \frac{G_d - F}{F}. \tag{C.9}$$

$$p^\star_{\text{res}}(\tau) = -\frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{1 - F}, \tag{C.10}$$

$$d^\star_{\text{res}}(\tau) = -\frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{1 - F}\left(\mathrm{e}^{\lambda\tau} - 1\right), \tag{C.11}$$

$$d'^\star_{\text{res}}(\tau) = -\frac{G_d - F}{1 - F}. \tag{C.12}$$

Key moments are then

$$\mathbb{E}(\hat{p}^{\star 2}) = \left(\frac{\bar{p}}{\bar{d}'}\right)^2 \frac{(G_p - G_d)^2}{F(1-F)} \int_0^{\tau_{\text{s}}} q(\tau)\,\mathrm{d}\tau, \tag{C.13}$$

$$\mathbb{E}(\hat{d}^\star \hat{p}^\star) = \left(\frac{\bar{p}}{\bar{d}'}\right)^2 \frac{(G_p - G_d)^2}{F(1-F)} \int_0^{\tau_{\text{s}}} \left(\mathrm{e}^{\lambda\tau} - 1\right) q(\tau)\,\mathrm{d}\tau, \tag{C.14}$$

$$\mathbb{E}(\hat{d}^{\star 2}) = \left(\frac{\bar{p}}{\bar{d}'}\right)^2 \frac{(G_p - G_d)^2}{F(1-F)} \int_0^{\tau_{\text{s}}} \left(\mathrm{e}^{\lambda\tau} - 1\right)^2 q(\tau)\,\mathrm{d}\tau. \tag{C.15}$$

Let $\hat{Z}_d = \sum_i \hat{d}_i / \sum_i \hat{d}'_i$ be the random variable giving the distribution of isotopic ratios d/d' after averaging, and $\hat{Z}_p = \sum_i \hat{p}_i / \sum_i \hat{d}'_i$ be the random variable giving

the distribution of isotopic ratios p/d' after averaging. From (B.6) we have that

$$\operatorname{var} \hat{Z}_d \sim \frac{1}{N}\mathbb{E}\left(\hat{d}^{\star 2}\right) = \left(\frac{\bar{p}}{\bar{d}'}\right)^2 \frac{(G_p - G_d)^2}{NF\,(1-F)} \int_0^{\tau_\mathrm{s}} \left(\mathrm{e}^{\lambda\tau} - 1\right)^2 q(\tau)\,\mathrm{d}\tau, \tag{C.16}$$

$$\operatorname{var} \hat{Z}_p \sim \frac{1}{N}\mathbb{E}\left(\hat{p}^{\star 2}\right) = \left(\frac{\bar{p}}{\bar{d}'}\right)^2 \frac{(G_p - G_d)^2}{NF\,(1-F)} \int_0^{\tau_\mathrm{s}} q(\tau)\,\mathrm{d}\tau. \tag{C.17}$$

(C.16) is the generalisation of (1) of [9]. A similar generalisation for the skew (4) of [9] can obtained using (B.8).

D Parent-daughter isochrons

It is convenient to introduce $\hat{T}_\mathrm{m}$: this is the random variable giving the distribution of parcel ages restricted to those parcels that have passed through the melting region. $q_\mathrm{m}(\tau) = q(\tau)/\int_0^{\tau_\mathrm{s}} q(\tau)\,\mathrm{d}\tau$ is then the probability density function of $\hat{T}_\mathrm{m}$. Suppose we plot d/d' against p/d'. Then the correlation is estimated by (B.10) as

$$r_{pd} \sim \frac{\mathbb{E}(\hat{d}^\star \hat{p}^\star)}{\sqrt{\mathbb{E}(\hat{d}^{\star 2})\mathbb{E}(\hat{p}^{\star 2})}} = \frac{\mathbb{E}(\mathrm{e}^{\lambda \hat{T}_\mathrm{m}} - 1)}{\sqrt{\mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_\mathrm{m}} - 1\right)^2}}. \tag{D.1}$$

The slopes of the regression lines are estimated by (B.11), (B.12), and (B.13) as

$$\beta_{pd} \sim \frac{\mathbb{E}(\hat{d}^\star \hat{p}^\star)}{\mathbb{E}(\hat{p}^{\star 2})} = \mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_\mathrm{m}} - 1\right), \tag{D.2}$$

$$\bar{\beta} \sim \sqrt{\frac{\mathbb{E}(\hat{d}^{\star 2})}{\mathbb{E}(\hat{p}^{\star 2})}} = \sqrt{\mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_\mathrm{m}} - 1\right)^2}, \tag{D.3}$$

$$\beta_{dp} \sim \frac{\mathbb{E}(\hat{d}^{\star 2})}{\mathbb{E}(\hat{d}^\star \hat{p}^\star)} = \frac{\mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_\mathrm{m}} - 1\right)^2}{\mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_\mathrm{m}} - 1\right)}. \tag{D.4}$$

The parent-daughter pseudo-isochron age is related to the slope of the regression line by

$$\mathrm{e}^{\lambda \tau_\mathrm{pdi}} - 1 = \beta, \tag{D.5}$$

and thus using the geometric mean regression line (D.3) the model parent-daughter pseudo-isochron equation is

$$\left(\mathrm{e}^{\lambda \tau_\mathrm{pdi}} - 1\right)^2 = \mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_\mathrm{m}} - 1\right)^2. \tag{D.6}$$

Note that when the decay is linearisable ($\hat{T}_{\rm m} \ll 1/\lambda$) (D.1) and (D.6) reduce to $r_{pd} = \mathbb{E}\hat{T}_{\rm m}/\sqrt{\mathbb{E}\hat{T}_{\rm m}^2}$ and $\tau_{\rm pdi} = \sqrt{\mathbb{E}\hat{T}_{\rm m}^2}$.

E Daughter-daughter isochrons

Now suppose we plot $(d/d')_2$ against $(d/d')_1$ for two different isotopic systems 1 and 2. Key moments are then

$$\mathbb{E}(\hat{d}_1^{\star 2}) = \left(\frac{\bar{p}_1}{\bar{d}_1'}\right)^2 \frac{(G_{p1}-G_{d1})^2}{F\,(1-F)} \int_0^{\tau_{\rm s}} \left(\mathrm{e}^{\lambda_1 \tau} - 1\right)^2 q(\tau)\,\mathrm{d}\tau, \tag{E.1}$$

$$\mathbb{E}(\hat{d}_1^{\star}\hat{d}_2^{\star}) = \left(\frac{\bar{p}_1}{\bar{d}_1'}\right)\left(\frac{\bar{p}_2}{\bar{d}_2'}\right) \frac{(G_{p1}-G_{d1})(G_{p2}-G_{d2})}{F\,(1-F)} \int_0^{\tau_{\rm s}} \left(\mathrm{e}^{\lambda_1 \tau} - 1\right)\left(\mathrm{e}^{\lambda_2 \tau} - 1\right) q(\tau)\,\mathrm{d}\tau, \tag{E.2}$$

$$\mathbb{E}(\hat{d}_2^{\star 2}) = \left(\frac{\bar{p}_2}{\bar{d}_2'}\right)^2 \frac{(G_{p2}-G_{d2})^2}{F\,(1-F)} \int_0^{\tau_{\rm s}} \left(\mathrm{e}^{\lambda_2 \tau} - 1\right)^2 q(\tau)\,\mathrm{d}\tau. \tag{E.3}$$

The correlation is then given by

$$r_{12} \sim \frac{\mathbb{E}(\mathrm{e}^{\lambda_1 \hat{T}_{\rm m}} - 1)(\mathrm{e}^{\lambda_2 \hat{T}_{\rm m}} - 1)}{\sqrt{\mathbb{E}\left(\mathrm{e}^{\lambda_1 \hat{T}_{\rm m}} - 1\right)^2 \mathbb{E}\left(\mathrm{e}^{\lambda_2 \hat{T}_{\rm m}} - 1\right)^2}} \operatorname{sgn}\left((G_{p1}-G_{d1})(G_{p2}-G_{d2})\right). \tag{E.4}$$

Note that $|r_{12}| \sim 1$ if the decay is linearisable ($\hat{T}_{\rm m} \ll 1/\lambda_1$ and $1/\lambda_2$), which is why in Figure 4 of [9] the model data form almost perfect straight lines. However, when the decay is not linearisable we will not get perfect correlation. The slope of the geometric mean regression line is given by

$$\bar{\beta} \sim \frac{\left(\bar{p}_2/\bar{d}_2'\right)(G_{p2}-G_{d2})}{\left(\bar{p}_1/\bar{d}_1'\right)(G_{p1}-G_{d1})} \sqrt{\frac{\mathbb{E}\left(\mathrm{e}^{\lambda_2 \hat{T}_{\rm m}} - 1\right)^2}{\mathbb{E}\left(\mathrm{e}^{\lambda_1 \hat{T}_{\rm m}} - 1\right)^2}}. \tag{E.5}$$

Note that if the decay is linearisable this reduces to

$$\bar{\beta} \sim \frac{\left(\bar{p}_2/\bar{d}_2'\right)(G_{p2}-G_{d2})\,\lambda_2}{\left(\bar{p}_1/\bar{d}_1'\right)(G_{p1}-G_{d1})\,\lambda_1}, \tag{E.6}$$

which is precisely (5) of [9]. (E.5) is the generalisation of (5) of [9]. Hence for those isotopic systems for which a linear decay approximation is valid the slopes are unchanged from [9].

We are particularly interested in a special case of (E.5). If the parent and daughter elements are the same, and the reference isotope is also the same,

then (E.5) reduces to

$$\bar{\beta} \sim \frac{\bar{p}_2}{\bar{p}_1} \sqrt{\frac{\mathbb{E}\left(\mathrm{e}^{\lambda_2 \hat{T}_\mathrm{m}} - 1\right)^2}{\mathbb{E}\left(\mathrm{e}^{\lambda_1 \hat{T}_\mathrm{m}} - 1\right)^2}}. \tag{E.7}$$

This is the generalisation of (6) of [9]. The other estimates of the slope of the regression line in this special case are given by

$$\beta_{12} \sim \frac{\bar{p}_2}{\bar{p}_1} \cdot \frac{\mathbb{E}(\mathrm{e}^{\lambda_1 \hat{T}_\mathrm{m}} - 1)(\mathrm{e}^{\lambda_2 \hat{T}_\mathrm{m}} - 1)}{\mathbb{E}\left(\mathrm{e}^{\lambda_1 \hat{T}_\mathrm{m}} - 1\right)^2}, \quad \beta_{21} \sim \frac{\bar{p}_2}{\bar{p}_1} \cdot \frac{\mathbb{E}\left(\mathrm{e}^{\lambda_2 \hat{T}_\mathrm{m}} - 1\right)^2}{\mathbb{E}(\mathrm{e}^{\lambda_1 \hat{T}_\mathrm{m}} - 1)(\mathrm{e}^{\lambda_2 \hat{T}_\mathrm{m}} - 1)}. \tag{E.8}$$

The daughter-daughter pseudo-isochron age is related to the slope of the regression lines by

$$\frac{\bar{p}_2}{\bar{p}_1} \cdot \frac{\mathrm{e}^{\lambda_2 \tau_\mathrm{ddi}} - 1}{\mathrm{e}^{\lambda_1 \tau_\mathrm{ddi}} - 1} = \beta. \tag{E.9}$$

Hence, combining (E.7) and (E.9) the model pseudo-isochron age τ_ddi satisfies the simple relationship

$$\frac{(\mathrm{e}^{\lambda_2 \tau_\mathrm{ddi}} - 1)^2}{(\mathrm{e}^{\lambda_1 \tau_\mathrm{ddi}} - 1)^2} = \frac{\mathbb{E}(\mathrm{e}^{\lambda_2 \hat{T}_\mathrm{m}} - 1)^2}{\mathbb{E}(\mathrm{e}^{\lambda_1 \hat{T}_\mathrm{m}} - 1)^2}. \tag{E.10}$$

Note that when the decay is linearisable ($\hat{T}_\mathrm{m} \ll 1/\lambda_1$ and $1/\lambda_2$) (E.10) reduces to $\tau_\mathrm{ddi} = \mathbb{E}\hat{T}_\mathrm{m}^3 / \mathbb{E}\hat{T}_\mathrm{m}^2$ (by Taylor series expansion).

F Means

It is important to distinguish between different definitions of the mean age of parcels. Since the heterogeneity we are interested in is generated by fractionation on melting, an important mean age is $\bar{\tau}_\mathrm{m} = \mathbb{E}\hat{T}_\mathrm{m}$, the mean age of the parcels that have passed through the melting region. Primordial parcels do not contribute to the pseudo-isochron ages. The mean mantle age $\bar{\tau}_\mathrm{total}$ is often defined by including the primordial parcels and assigning them an age of τ_s, and it is this definition that is used in [8]. Hence $\bar{\tau}_\mathrm{total} \geq \bar{\tau}_\mathrm{m}$. We have

$$\bar{\tau}_\mathrm{total} = \int_0^{\tau_\mathrm{s}} \tau q(\tau)\, \mathrm{d}\tau + \tau_\mathrm{s} \left(1 - \int_0^{\tau_\mathrm{s}} q(\tau)\, \mathrm{d}\tau\right), \tag{F.1}$$

$$\bar{\tau}_\mathrm{m} = \mathbb{E}\hat{T}_\mathrm{m} = \int_0^{\tau_\mathrm{s}} \tau q_\mathrm{m}(\tau)\, \mathrm{d}\tau. \tag{F.2}$$

The parent-daughter pseudo-isochron age τ_pdi is an example of a generalised mean [13]. A generalised mean M_ϕ is defined by $M_\phi(\hat{X}) = \phi^{-1}\left(\mathbb{E}\phi(\hat{X})\right)$, where ϕ is a strictly monotonic function, and $\hat{X}$ is a random variable. Commonly

encountered examples include $\phi(x) = x$ (arithmetic mean), $\phi(x) = x^r$ (power mean), and $\phi(x) = \log x$ (geometric mean). In the case of the parent-daughter pseudo-isochron (D.6), $\phi(x) = \left(\mathrm{e}^{\lambda x} - 1\right)^2$. Generalised means have a number of important properties, but of most interest to us is the notion of 'comparability': whether there is always an inequality between different means regardless of the distribution of the random variable $\hat{X}$. A common example of comparability is the arithmetic mean - geometric mean inequality. There is an important theorem which states that if ψ and χ are monotonically increasing functions, $\phi = \chi\psi^{-1}$, and $\phi'' > 0$, then $M_\psi \leq M_\chi$ (Theorem 96 of [13]). By use of this theorem we find the following inequalities are satisfied by the parent-daughter pseudo-isochron age:

$$\min \hat{T}_\mathrm{m} \leq \bar{\tau}_\mathrm{m} \leq \sqrt{\mathbb{E}\hat{T}_\mathrm{m}^2} \leq \tau_\mathrm{pdi}(\lambda_1) \leq \tau_\mathrm{pdi}(\lambda_2) \leq \max \hat{T}_\mathrm{m}, \tag{F.3}$$

where $0 < \lambda_1 < \lambda_2$, and $\min \hat{T}_\mathrm{m}$ and $\max \hat{T}_\mathrm{m}$ are the smallest and largest ages respectively with any probability mass.

The daughter-daughter pseudo-isochron age τ_ddi is not a generalised mean as described by [13], but it is an example of a generalised abstracted mean (Definition 2.4 of [14]). A generalised abstracted mean is defined by $M_{\phi_1,\phi_2}(\hat{X}) = (\phi_1/\phi_2)^{-1}\left(\mathbb{E}\phi_1(\hat{X})/\mathbb{E}\phi_2(\hat{X})\right)$ where ϕ_1/ϕ_2 is a strictly monotonic function. In the case of the daughter-daughter pseudo-isochron equation (E.10), $\phi_1(x) = \left(\mathrm{e}^{\lambda_1 x} - 1\right)^2$ and $\phi_2(x) = \left(\mathrm{e}^{\lambda_2 x} - 1\right)^2$. The generalised abstracted mean shares many of the properties of the generalised mean of [13], and under a suitable transformation of the probability density function can be written in the same form. However, for our purposes what is of most interest is the inequality

$$\tau_\mathrm{pdi}(\lambda_2) \leq \tau_\mathrm{ddi}(\lambda_1, \lambda_2) \leq \max \hat{T}_\mathrm{m}. \tag{F.4}$$

The above inequalities (F.3) and (F.4) are strict inequalities unless all the probability mass is concentrated at a single age. In this case all the inequalities are equalities, and all the means yield this single age.

G Relating melt rate and parcel ages

Suppose melt rate as a function of age τ is $\gamma_\mathrm{melt}(\tau)$. Define τ_melt to be the melting time scale at the present day, so that $\gamma_\mathrm{melt}(0) = 1/\tau_\mathrm{melt}$. Let $Q(\tau) = \mathbb{P}(\hat{T} \leq \tau)$ be the proportion of material in the box with age less than τ (the cumulative distribution function), with $1 - Q(\tau_\mathrm{s})$ being the proportion of primordial material. $Q(\tau)$ satisfies

$$\frac{\mathrm{d}Q(\tau)}{\mathrm{d}\tau} = \gamma_\mathrm{melt}(\tau)\left(1 - Q(\tau)\right), \quad Q(0) = 0, \tag{G.1}$$

and thus

$$Q(\tau) = 1 - \exp\left(-\int_0^\tau \gamma_{\text{melt}}(\tau)\,\mathrm{d}\tau\right). \tag{G.2}$$

Hence the probability density functions are

$$q(\tau) = \frac{\mathrm{d}Q(\tau)}{\mathrm{d}\tau} = \gamma_{\text{melt}}(\tau)\exp\left(-\int_0^\tau \gamma_{\text{melt}}(\tau)\,\mathrm{d}\tau\right), \tag{G.3}$$

$$q_{\text{m}}(\tau) = \frac{q(\tau)}{Q(\tau_{\text{s}})} = \frac{1}{Q(\tau_{\text{s}})}\frac{\mathrm{d}Q(\tau)}{\mathrm{d}\tau}. \tag{G.4}$$

In some cases it is more convenient to work with the cumulative distribution function (cdf) rather than the pdf. Note that if $f(0) = 0$ (as it is for all functions we consider) then

$$\mathbb{E}f(\hat{T}_{\text{m}}) = \int_0^{\tau_{\text{s}}} f'(\tau)(1 - Q_m(\tau))\,\mathrm{d}\tau, \tag{G.5}$$

where the cdf $Q_m(\tau)$ is given by

$$Q_m(\tau) = \int_0^\tau q_m(\tau)\,\mathrm{d}\tau = \frac{Q(\tau)}{Q(\tau_{\text{s}})}. \tag{G.6}$$

Xie and Tackley [8] quote results from their numerical simulations in terms of the crustal production rate given by $b\gamma_{\text{melt}}(\tau)$, where b is their basalt fraction ($b = 0.3$ in their standard runs). Their integrated crustal production (integrated forward in time) is given as a function of age by

$$c(\tau) = 1 - \exp\left(-\int_\tau^{\tau_{\text{s}}} b\gamma_{\text{melt}}(\tau)\,\mathrm{d}\tau\right), \tag{G.7}$$

which implies

$$\exp\left(-\int_0^\tau \gamma_{\text{melt}}(\tau)\,\mathrm{d}\tau\right) = \left(\frac{1 - c(0)}{1 - c(\tau)}\right)^{1/b}. \tag{G.8}$$

Hence the cdf $Q_m(\tau)$ (and thus also pseudo-isochron ages) can be calculated directly from their integrated crustal production or their crustal production rate.

For a constant melt rate $\gamma_{\text{melt}}(\tau) = 1/\tau_{\text{melt}}$ and $\hat{T}$ is an exponential random variable with parameter $1/\tau_{\text{melt}}$. The corresponding probability density functions are

$$q(\tau) = \mathrm{e}^{-\tau/\tau_{\text{melt}}}/\tau_{\text{melt}}, \tag{G.9}$$

$$q_{\text{m}}(\tau) = \mathrm{e}^{-\tau/\tau_{\text{melt}}}/\left(\tau_{\text{melt}}\left(1 - \mathrm{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}}\right)\right), \tag{G.10}$$

and corresponding means are

$$\bar{\tau}_{\text{total}} = \tau_{\text{melt}}(1 - \mathrm{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}}), \tag{G.11}$$

$$\bar{\tau}_{\text{m}} = \mathbb{E}\hat{T}_{\text{m}} = \tau_{\text{melt}} \frac{1 - \mathrm{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}}(1 + \tau_{\text{s}}/\tau_{\text{melt}})}{1 - \mathrm{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}}}, \tag{G.12}$$

$$\sqrt{\mathbb{E}\hat{T}_{\text{m}}^2} = \sqrt{2}\tau_{\text{melt}} \sqrt{\frac{1 - \mathrm{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}}\left(1 + \tau_{\text{s}}/\tau_{\text{melt}} + \frac{1}{2}\left(\tau_{\text{s}}/\tau_{\text{melt}}\right)^2\right)}{1 - \mathrm{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}}}}. \tag{G.13}$$

H Constant melt rate: Parent-daughter isochrons

For constant melt rate the parent-daughter pseudo-isochron equation is

$$(\mathrm{e}^{\lambda\tau_{\text{pdi}}} - 1)^2 = \frac{\int_0^{\tau_{\text{s}}} (\mathrm{e}^{\lambda\tau} - 1)^2 \, \mathrm{e}^{-\tau/\tau_{\text{melt}}} \, \mathrm{d}\tau}{\tau_{\text{melt}} \left(1 - \mathrm{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}}\right)}. \tag{H.1}$$

To gain some insight into the behaviour of this equation we now consider some simple asymptotics.

H.1 Asymptotics when τ_{melt} *or* $\tau_{\text{s}} \ll 1/\lambda$

Since we often study slowly decaying isotopes, the most important asymptotics are when τ_{melt} or $\tau_{\text{s}} \ll 1/\lambda$. In this limit the correlation becomes

$$r_{pd} = \frac{1 - \mathrm{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}}(1 + \tau_{\text{s}}/\tau_{\text{melt}})}{\sqrt{2\left(1 - \mathrm{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}}\right)\left(1 - \mathrm{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}}\left(1 + \tau_{\text{s}}/\tau_{\text{melt}} + \frac{1}{2}\left(\tau_{\text{s}}/\tau_{\text{melt}}\right)^2\right)\right)}}. \tag{H.2}$$

Furthermore, if $\tau_{\text{s}}/\tau_{\text{melt}} \gg 1$ then $r_{pd} \sim 1/\sqrt{2} \approx 0.71$. Alternatively if $\tau_{\text{s}}/\tau_{\text{melt}} \ll 1$ then $r_{pd} \sim \sqrt{3}/2 \approx 0.87$. For linearisable decay r_{pd} always lies between these two values. It is important to note that there is thus never a perfect correlation between d/d' and p/d'.

When τ_{melt} or $\tau_{\text{s}} \ll 1/\lambda$ the parent-daughter pseudo-isochron age becomes simply $\tau_{\text{pdi}} = \sqrt{\mathbb{E}\hat{T}_{\text{m}}^2}$, which is given for constant melt rate by (G.13). If $\tau_{\text{s}}/\tau_{\text{melt}} \gg 1$ (G.13) simplifies to $\tau_{\text{pdi}} \sim \sqrt{2}\tau_{\text{melt}}$, which determines the slope of curves near the origin in Figures 3b and 4b, and the asymptotes for large τ_{s} in Figure 3a. On the other hand, if $\tau_{\text{s}}/\tau_{\text{melt}} \ll 1$ (G.13) simplifies to $\tau_{\text{pdi}} \sim \tau_{\text{s}}/\sqrt{3}$, which determines the slope of curves near the origin in Figures 3a and 4a, and the asymptotes for large τ_{melt} in Figure 3b.

H.2 Asymptotics when $\tau_s \ll \tau_{\text{melt}}$

When the decay is not linearisable, asymptotics based solely on $\tau_s \ll \tau_{\text{melt}}$ can be found. The pseudo-isochron equation (H.1) becomes

$$(\mathrm{e}^{\lambda\tau_{\text{pdi}}}-1)^2 = \frac{2\lambda\tau_{\text{s}} + 3 - 4\,\mathrm{e}^{\lambda\tau_{\text{s}}} + \mathrm{e}^{2\lambda\tau_{\text{s}}}}{2\lambda\tau_{\text{s}}}, \tag{H.3}$$

which is independent of τ_{melt}. This equation determines the $\tau_{\text{melt}} = \infty$ Ga curve in Figures 3a and 4a, and asymptotes for large τ_{melt} in Figures 3b and 4b.

H.3 Asymptotics when $\tau_s \gg \tau_{\text{melt}}$

(H.1) has two regimes of asymptotic behaviour when $\tau_{\text{s}}/\tau_{\text{melt}} \gg 1$. If $\lambda\tau_{\text{melt}} > 1/2$ then

$$\mathrm{e}^{2\lambda\tau_{\text{pdi}}} = \frac{\mathrm{e}^{(2\lambda\tau_{\text{melt}}-1)\tau_{\text{s}}/\tau_{\text{melt}}}}{2\lambda\tau_{\text{melt}} - 1}, \tag{H.4}$$

whereas if $\lambda\tau_{\text{melt}} < 1/2$ then

$$(\mathrm{e}^{\lambda\tau_{\text{pdi}}}-1)^2 = \frac{2(\lambda\tau_{\text{melt}})^2}{(1-\lambda\tau_{\text{melt}})(1-2\lambda\tau_{\text{melt}})}. \tag{H.5}$$

Note that (H.4) depends on τ_{s} whereas (H.5) is independent of τ_{s}. This is why in Figure 4a curves with $\tau_{\text{melt}} < 1/2\lambda$ flatten out for large τ_{s}, while curves with $\tau_{\text{melt}} > 1/2\lambda$ grow linearly for large τ_{s} with slope $1 - 1/2\lambda\tau_{\text{melt}}$.

I Constant melt rate: Daughter-daughter isochrons

The constant melt rate daughter-daughter pseudo-isochron equation is

$$\frac{(\mathrm{e}^{\lambda_2\tau_{\text{ddi}}}-1)^2}{(\mathrm{e}^{\lambda_1\tau_{\text{ddi}}}-1)^2} = \frac{\int_0^{\tau_{\text{s}}}(\mathrm{e}^{\lambda_2\tau}-1)^2\,\mathrm{e}^{-\tau/\tau_{\text{melt}}}\,\mathrm{d}\tau}{\int_0^{\tau_{\text{s}}}(\mathrm{e}^{\lambda_1\tau}-1)^2\,\mathrm{e}^{-\tau/\tau_{\text{melt}}}\,\mathrm{d}\tau}. \tag{I.1}$$

Again, insights into the behaviour of (I.1) can be gained by some simple asymptotics, although linearising the decay is not as relevant here. We will assume without loss of generality that $\lambda_2 > \lambda_1$ for the subsequent asymptotics.

I.1 Asymptotics when $\tau_s \ll \tau_{\rm melt}$

When $\tau_s \ll \tau_{\rm melt}$ the pseudo-isochron equation (I.1) becomes

$$\frac{(\mathrm{e}^{\lambda_2\tau_{\rm ddi}}-1)^2}{(\mathrm{e}^{\lambda_1\tau_{\rm ddi}}-1)^2} = \frac{\lambda_1}{\lambda_2}\cdot\frac{2\lambda_2\tau_{\rm s}+3-4\,\mathrm{e}^{\lambda_2\tau_{\rm s}}+\mathrm{e}^{2\lambda_2\tau_{\rm s}}}{2\lambda_1\tau_{\rm s}+3-4\,\mathrm{e}^{\lambda_1\tau_{\rm s}}+\mathrm{e}^{2\lambda_1\tau_{\rm s}}}. \tag{I.2}$$

Note that this equation is independent of $\tau_{\rm melt}$. This equation determines the asymptotes for large $\tau_{\rm melt}$ in Figure 5b, and the $\tau_{\rm melt} = \infty$ Ga curve in Figure 5a. Furthermore, if it is also the case that $\lambda\tau_s \ll 1$, we can Taylor expand both sides to find

$$\frac{\lambda_2^2}{\lambda_1^2}\left(1+(\lambda_2-\lambda_1)\tau_{\rm ddi}+...\right) = \frac{\lambda_2^2}{\lambda_1^2}\left(1+\frac{3(\lambda_2-\lambda_1)}{4}\tau_s+...\right), \tag{I.3}$$

and thus get the simple result that $\tau_{\rm ddi} \sim \frac{3}{4}\tau_s$, which determines the slope of curves near the origin in Figure 5a. If instead $\lambda\tau_s \gg 1$ then we can approximate both sides as

$$\mathrm{e}^{2(\lambda_2-\lambda_1)\tau_{\rm ddi}} = \frac{\lambda_1}{\lambda_2}\,\mathrm{e}^{2(\lambda_2-\lambda_1)\tau_{\rm s}}, \tag{I.4}$$

which demonstrates that the slope of the $\tau_{\rm melt} = \infty$ Ga curve in Figure 5a will approach 1 for large τ_s.

I.2 Asymptotics when $\tau_s \gg \tau_{\rm melt}$

We have three different regimes of asymptotic behaviour for (I.1) when $\tau_s \gg \tau_{\rm melt}$. If $\tau_{\rm melt} > 1/2\lambda_1$ the pseudo-isochron equation is

$$\mathrm{e}^{2(\lambda_2-\lambda_1)\tau_{\rm ddi}} = \mathrm{e}^{2(\lambda_2-\lambda_1)\tau_{\rm s}}\frac{2\lambda_1\tau_{\rm melt}-1}{2\lambda_2\tau_{\rm melt}-1}, \tag{I.5}$$

if $1/2\lambda_2 < \tau_{\rm melt} < 1/2\lambda_1$ then

$$\mathrm{e}^{2(\lambda_2-\lambda_1)\tau_{\rm ddi}} = \frac{\mathrm{e}^{(2\lambda_2\tau_{\rm melt}-1)\tau_{\rm s}/\tau_{\rm melt}}(1-\lambda_1\tau_{\rm melt})(1-2\lambda_1\tau_{\rm melt})}{(\lambda_1\tau_{\rm melt})^2\,(2\lambda_2\tau_{\rm melt}-1)}, \tag{I.6}$$

and if $\tau_{\rm melt} < 1/2\lambda_2$ then

$$\frac{(\mathrm{e}^{\lambda_2\tau_{\rm ddi}}-1)^2}{(\mathrm{e}^{\lambda_1\tau_{\rm ddi}}-1)^2} = \frac{\lambda_2^2(1-\lambda_1\tau_{\rm melt})(1-2\lambda_1\tau_{\rm melt})}{\lambda_1^2(1-\lambda_2\tau_{\rm melt})(1-2\lambda_2\tau_{\rm melt})}. \tag{I.7}$$

The most important feature of (I.7) is that it is independent of τ_s, whereas in the other two asymptotic regimes there is a dependence on τ_s. This is why for values of $\tau_{\rm melt} < 1/2\lambda_2$ in Figure 5a the curve flattens out for large $\tau_{\rm s}$ whereas in the other regimes they grow linearly for large $\tau_{\rm s}$ with slopes

$(\lambda_2 - 1/2\tau_{\text{melt}})/(\lambda_2 - \lambda_1)$ and 1 respectively. (I.7) can be further simplified if $\lambda\tau_{\text{melt}} \ll 1$. Both sides can be Taylor expanded to yield

$$\frac{\lambda_2^2}{\lambda_1^2}\left(1 + (\lambda_2 - \lambda_1)\tau_{\text{ddi}} + ...\right) = \frac{\lambda_2^2}{\lambda_1^2}\left(1 + 3(\lambda_2 - \lambda_1)\tau_{\text{melt}} + ...\right), \tag{I.8}$$

leading to the simple result that $\tau_{\text{ddi}} \sim 3\tau_{\text{melt}}$, which determines the slope of curves near the origin in Figure 5b.

J Relationship to linear evolution models

There is an important connection between our statistical box model and the pseudo-isochron equations derived from linear evolution models by previous authors [3,4,5]. Suppose we divide the box into two based on parcel type. Let all the residue parcels and a fraction $1 - F$ of the primordial parcels be called reservoir 1, and all the melt parcels and a fraction F of the primordial parcels be reservoir 2. These two reservoirs do not change in size over time. Reservoir 1 is a fraction $1 - F$ of the box, and reservoir 2 a fraction F. At an age τ_{s} before the present all parcels are primordial and thus both reservoirs have the same uniform isotopic concentrations. The mean concentrations of p, d and d' at the present in each reservoir can be calculated by integrating (A.1-A.9) over the age distribution of parcels. For the isochron calculation it is simplest to consider instead the $p^\star$ and $d^\star$ values. Let $\hat{p}_1^\star$ be the random variable giving the distribution of $p^\star$ in reservoir 1, and $\hat{d}_1^\star$, $\hat{p}_2^\star$ and $\hat{d}_2^\star$ be defined similarly. Note that the subscripts 1 and 2 now refer to the different reservoirs, rather than the different isotopic systems as in earlier sections. Integration of (C.4-C.12) yields the mean values as

$$\mathbb{E}\hat{p}_1^\star = -\frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{1 - F}\int_0^{\tau_{\text{s}}} q(\tau)\,\mathrm{d}\tau, \tag{J.1}$$

$$\mathbb{E}\hat{d}_1^\star = -\frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{1 - F}\int_0^{\tau_{\text{s}}} \left(\mathrm{e}^{\lambda\tau} - 1\right) q(\tau)\,\mathrm{d}\tau, \tag{J.2}$$

$$\mathbb{E}\hat{p}_2^\star = \frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{F}\int_0^{\tau_{\text{s}}} q(\tau)\,\mathrm{d}\tau, \tag{J.3}$$

$$\mathbb{E}\hat{d}_2^\star = \frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{F}\int_0^{\tau_{\text{s}}} \left(\mathrm{e}^{\lambda\tau} - 1\right) q(\tau)\,\mathrm{d}\tau. \tag{J.4}$$

By multiplying these expressions by the fraction of the box each reservoir occupies these can be converted into molar values $n^\star$ and $m^\star$ ((21) and (22)) as

$$n_1^\star = (1 - F)\,\mathbb{E}\hat{p}_1^\star = -\frac{\bar{p}}{\bar{d}'}\left(G_p - G_d\right)\int_0^{\tau_{\text{s}}} q(\tau)\,\mathrm{d}\tau, \tag{J.5}$$

$$m_1^\star = (1 - F)\,\mathbb{E}\hat{d}_1^\star = -\frac{\bar{p}}{\bar{d}'}\left(G_p - G_d\right)\int_0^{\tau_{\text{s}}} \left(\mathrm{e}^{\lambda\tau} - 1\right) q(\tau)\,\mathrm{d}\tau. \tag{J.6}$$

where $n_2^\star = -n_1^\star$ and $m_2^\star = -m_1^\star$. (J.5) and (J.6) are solutions of (25) and (26), and are the backward in time versions of (A1-4) and (A1-5) of [3] with $\Delta\langle\alpha\rangle = 0$. The number of moles of each isotope species in the two reservoirs must be modelled by the linear evolution equations (10-15); the only question is finding expressions for the residence times. Here the relaxation times for parent and daughter are both given by $1/\gamma_{\text{melt}}$, since they are both fractionated by the same melting process on the same timescale. The individual residence times for each reservoir are determined by the fraction of moles G that enter the melt for each chemical species as

$$\tau_1 = \frac{1}{G_p \gamma_{\text{melt}}}, \quad \tau_2 = \frac{1}{(1-G_p)\,\gamma_{\text{melt}}}, \tag{J.7}$$

$$\theta_1 = \frac{1}{G_d \gamma_{\text{melt}}}, \quad \theta_2 = \frac{1}{(1-G_d)\,\gamma_{\text{melt}}}. \tag{J.8}$$

Consider a pseudo-isochron defined by the where the two reservoirs and the whole box lie on the isochron diagram. The slope in the parent-daughter isochron diagram is given by

$$\frac{(m/s)_1 - (m/s)_{1+2}}{(n/s)_1 - (n/s)_{1+2}} = \frac{m_1^\star}{n_1^\star} = \frac{\int_0^{\tau_{\text{s}}} (\mathrm{e}^{\lambda\tau} - 1) q(\tau)\,\mathrm{d}\tau}{\int_0^{\tau_{\text{s}}} q(\tau)\,\mathrm{d}\tau} = \mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_{\text{m}}} - 1\right). \tag{J.9}$$

Hence the corresponding pseudo-isochron equations are

$$\mathrm{e}^{\lambda\tau_{\text{pdil}}} - 1 = \mathbb{E}(\mathrm{e}^{\lambda \hat{T}_{\text{m}}} - 1), \tag{J.10}$$

$$\frac{\mathrm{e}^{\lambda_2\tau_{\text{ddil}}} - 1}{\mathrm{e}^{\lambda_1\tau_{\text{ddil}}} - 1} = \frac{\mathbb{E}(\mathrm{e}^{\lambda_2 \hat{T}_{\text{m}}} - 1)}{\mathbb{E}(\mathrm{e}^{\lambda_1 \hat{T}_{\text{m}}} - 1)}, \tag{J.11}$$

which are just (D.6) and (E.10) with squareds removed, and hence we will refer to these as the linear pseudo-isochron equations.

An alternative reservoir representation is to consider the box split into three reservoirs, one with all the melt parcels, one with all the residue parcels, and one with all the primordial parcels. These three reservoirs change in size over time, with the melt and residue reservoirs growing at the expense of the primordial reservoir. This is analogous to model I of Jacobsen and Wasserburg [18] where their depleted mantle reservoir 2 and crust reservoir 3 grow from an homogenous undepleted mantle reservoir 1. Compare the statistical box model result, obtained by integrating (A.4-A.9),

$$\frac{\mathbb{E}\hat{d}_{\text{res}}}{\mathbb{E}\hat{d}'_{\text{res}}} - \frac{\bar{d}}{\bar{d}'} = -\frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{1 - G_d}\mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_{\text{m}}} - 1\right), \tag{J.12}$$

$$\frac{\mathbb{E}\hat{d}_{\text{melt}}}{\mathbb{E}\hat{d}'_{\text{melt}}} - \frac{\bar{d}}{\bar{d}'} = \frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{G_d}\mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_{\text{m}}} - 1\right), \tag{J.13}$$

with (19) of [18], rewritten using an integration by parts,

$$\frac{N_{d,j}(\tau)}{N_{s,j}(\tau)} - \frac{N_{d,1}(\tau)}{N_{s,1}(\tau)} = \frac{N_{r,1}(\tau)}{N_{s,1}(\tau)} f_j^{r/s} \int_0^\tau \left(\mathrm{e}^{\lambda(\tau-\xi)} - 1\right) \frac{1}{M_j(\tau)} \frac{\mathrm{d}M_j(\xi)}{\mathrm{d}\xi}\,\mathrm{d}\xi. \quad \text{(J.14)}$$

(J.12) and (J.13) are equivalent to (J.14), although note that (J.14) is written with τ running forward in time. The fractionation factors relate through $1 + f_2 = (1 - G_p)/(1 - G_d)$ and $1 + f_3 = G_p/G_d$. The corresponding probability density function of parcels ages is $(1/M_j(\tau))\mathrm{d}M_j(\xi)/\mathrm{d}\xi$ (compare with (G.4)). Note that the relationship between melt rate and age distribution will be slightly different between the two models, since there is recycling in the statistical box model (a melt parcel may become a residue parcel and vice-versa), but not in model I. The pseudo-isochron equations that result from where the three reservoirs plot on the isochron diagram are also (J.10) and (J.11).

The linear parent-daughter pseudo-isochron age τ_{pdil} is also an example of a generalised mean, with $\phi(x) = \mathrm{e}^{\lambda x} - 1$. The linear pseudo-isochron ages satisfy

$$\min \hat{T}_{\mathrm{m}} \leq \bar{\tau}_{\mathrm{m}} \leq \tau_{\mathrm{pdil}}(\lambda_1) \leq \tau_{\mathrm{pdil}}(\lambda_2) \leq \tau_{\mathrm{ddil}}(\lambda_1, \lambda_2) \leq \max \hat{T}_{\mathrm{m}}, \quad \text{(J.15)}$$

where $0 < \lambda_1 < \lambda_2$. For linearisable decay ($\hat{T}_{\mathrm{m}} \ll 1/\lambda$) the linear parent-daughter pseudo-isochron equation (J.10) reduces to $\tau_{\mathrm{pdil}} = \mathbb{E}\hat{T}_{\mathrm{m}} = \bar{\tau}_{\mathrm{m}}$, and the linear daughter-daughter pseudo-isochron equation (J.11) to $\tau_{\mathrm{ddil}} = \mathbb{E}\hat{T}_{\mathrm{m}}^2/\mathbb{E}\hat{T}_{\mathrm{m}}$. It can be shown that the linear ages are always less than the corresponding quadratic ages, namely that $\tau_{\mathrm{pdil}}(\lambda) \leq \tau_{\mathrm{pdi}}(\lambda)$, and $\tau_{\mathrm{ddil}}(\lambda_1, \lambda_2) \leq \tau_{\mathrm{ddi}}(\lambda_1, \lambda_2)$.

For the case of constant melt rate the linear parent-daughter pseudo-isochron equation is

$$\begin{aligned} \mathrm{e}^{\lambda\tau_{\mathrm{pdil}}} - 1 &= \frac{\int_0^{\tau_{\mathrm{s}}} (\mathrm{e}^{\lambda\tau} - 1)\,\mathrm{e}^{-\tau/\tau_{\mathrm{melt}}}\,\mathrm{d}\tau}{\tau_{\mathrm{melt}}(1 - \mathrm{e}^{-\tau_{\mathrm{s}}/\tau_{\mathrm{melt}}})} \\ &= \frac{\lambda\tau_{\mathrm{melt}} + (1 - \lambda\tau_{\mathrm{melt}})\,\mathrm{e}^{-\tau_{\mathrm{s}}/\tau_{\mathrm{melt}}} - \mathrm{e}^{(\lambda\tau_{\mathrm{melt}}-1)\tau_{\mathrm{s}}/\tau_{\mathrm{melt}}}}{(1 - \lambda\tau_{\mathrm{melt}})\,(1 - \mathrm{e}^{-\tau_{\mathrm{s}}/\tau_{\mathrm{melt}}})}, \end{aligned} \quad \text{(J.16)}$$

which is the parent-daughter pseudo-isochron equation (A2-4) of [3]. This should be compared with our (H.1). The corresponding linear daughter-daughter pseudo-isochron equation is

$$\begin{aligned} \frac{\mathrm{e}^{\lambda_2\tau_{\mathrm{ddil}}} - 1}{\mathrm{e}^{\lambda_1\tau_{\mathrm{ddil}}} - 1} &= \frac{\int_0^{\tau_{\mathrm{s}}} (\mathrm{e}^{\lambda_2\tau} - 1)\,\mathrm{e}^{-\tau/\tau_{\mathrm{melt}}}\,\mathrm{d}\tau}{\int_0^{\tau_{\mathrm{s}}} (\mathrm{e}^{\lambda_1\tau} - 1)\,\mathrm{e}^{-\tau/\tau_{\mathrm{melt}}}\,\mathrm{d}\tau} \\ &= \frac{1 - \lambda_1\tau_{\mathrm{melt}}}{1 - \lambda_2\tau_{\mathrm{melt}}} \cdot \frac{(1 - \lambda_2\tau_{\mathrm{melt}})\,\mathrm{e}^{-\tau_{\mathrm{s}}/\tau_{\mathrm{melt}}} + \lambda_2\tau_{\mathrm{melt}} - \mathrm{e}^{(\lambda_2\tau_{\mathrm{melt}}-1)\tau_{\mathrm{s}}/\tau_{\mathrm{melt}}}}{(1 - \lambda_1\tau_{\mathrm{melt}})\,\mathrm{e}^{-\tau_{\mathrm{s}}/\tau_{\mathrm{melt}}} + \lambda_1\tau_{\mathrm{melt}} - \mathrm{e}^{(\lambda_1\tau_{\mathrm{melt}}-1)\tau_{\mathrm{s}}/\tau_{\mathrm{melt}}}}, \end{aligned} \quad \text{(J.17)}$$

which is precisely the daughter-daughter pseudo-isochron equation (A2-9) of [3]. This should be compared with our (I.1). Furthermore, when $\tau_{\mathrm{melt}} \ll \tau_{\mathrm{s}}$

and $\lambda\tau_{\text{melt}} < 1$ (J.16) and (J.17) become

$$e^{\lambda\tau_{\text{pdil}}} - 1 = \frac{\lambda\tau_{\text{melt}}}{1 - \lambda\tau_{\text{melt}}} \implies \tau_{\text{pdil}} = -\frac{1}{\lambda}\log\left(1 - \lambda\tau_{\text{melt}}\right), \tag{J.18}$$

$$\frac{e^{\lambda_2\tau_{\text{ddil}}} - 1}{e^{\lambda_1\tau_{\text{ddil}}} - 1} = \frac{\lambda_2\left(1 - \lambda_1\tau_{\text{melt}}\right)}{\lambda_1\left(1 - \lambda_2\tau_{\text{melt}}\right)}, \tag{J.19}$$

which are the pseudo-isochron equations (A2-6) and (A2-10) of [3], (31) and (32) of [4], and (7) and (9) of [5]. The corresponding equations in our model are (H.5) and (I.7). The $\tau_{\text{melt}} \ll \tau_{\text{s}}$ and $\lambda\tau_{\text{melt}} < 1$ limit is precisely the same as the conditions for secular equilibrium in [4] and [5]. [3] explores various other asymptotic regimes for (J.16) and (J.17), as we have for our constant melt rate equations in Appendices H and I.

References

[1] C. Brooks, S. R. Hart, A. Hofmann, D. E. James, Rb-Sr mantle isochrons from oceanic regions, Earth Planet. Sci. Lett. 32 (1976) 51–61.

[2] M. Tatsumoto, Isotopic composition of lead in oceanic basalt and its implication to mantle evolution, Earth Planet. Sci. Lett. 38 (1978) 63–87.

[3] C. J. Allègre, E. Lewin, Isotopic systems and stirring times of the earth's mantle, Earth Planet. Sci. Lett. 136 (1995) 629–646.

[4] F. Albarède, Radiogenic ingrowth in systems with multiple reservoirs: applications to the differentiation of the mantle-crust system, Earth Planet. Sci. Lett. 189 (2001) 59–73, doi:10.1016/S0012-821X(01)00350-8.

[5] K. E. Donnelly, S. L. Goldstein, C. H. Langmuir, M. Spiegelman, Origin of enriched ocean ridge basalts and implications for mantle dynamics, Earth Planet. Sci. Lett. 226 (2004) 347–366, doi:10.1016/j.epsl.2004.07.019.

[6] J. B. Kellogg, Towards an understanding of chemical and isotopic heterogeneity in the Earth's mantle, Ph.D. thesis, Harvard University (2004).

[7] U. R. Christensen, A. W. Hofmann, Segregation of subducted oceanic crust in the convecting mantle, J. Geophys. Res. 99 (1994) 19867–19884.

[8] S. Xie, P. J. Tackley, Evolution of U-Pb and Sm-Nd systems in numerical models of mantle convection and plate tectonics, J. Geophys. Res. 109 (2004) B11204, doi:10.1029/2004JB003176.

[9] J. F. Rudge, D. McKenzie, P. H. Haynes, A theoretical approach to understanding the isotopic heterogeneity of mid-ocean ridge basalt, Geochim. Cosmochim. Acta 69 (2005) 3873–3887, doi:10.1016/j.gca.2005.03.004.

[10] F. Albarède, The survival of mantle geochemical heterogeneities, in: R. D. van der Hilst, J. D. Bass, J. Matas, J. Trampert (Eds.), Earth's Deep Mantle: Structure, Composition, and Evolution, Vol. 160, AGU Geophysical Monograph Series, 2005, pp. 27–46.

[11] F. Albarède, Geochemistry: An Introduction, Cambridge University Press, 2003, ISBN 0-521-81468-5.

[12] A. Meibom, D. L. Anderson, The statistical upper mantle assemblage, Earth Planet. Sci. Lett. 217 (2004) 123–139, doi:10.1016/S0012-821X(03)00573-9.

[13] G. H. Hardy, J. E. Littlewood, G. Pólya, Inequalities, Cambridge University Press, 1934.

[14] F. Qi, Generalized abstracted mean values, J. Inequal. Pure and Appl. Math. 1 (2000) Art. 4.

[15] G. F. Davies, Stirring geochemistry in mantle convection models with stiff plates and slabs, Geochim. Cosmochim. Acta 66 (2002) 3125–3142, doi:10.1016/S0016-7037(02)00915-8.

[16] J. B. Kellogg, S. B. Jacobsen, R. J. O'Connell, Modelling the distribution of isotopic ratios in geochemical reservoirs, Earth Planet. Sci. Lett. 204 (2002) 183–202, doi:10.1016/S0012-821X(02)00981-0.

[17] S. Y. Novak, On the distribution of the ratio of sums of random variables, Theory Probab. Appl. 41 (1996) 479–503.

[18] S. B. Jacobsen, G. J. Wasserburg, The mean age of mantle and crustal reservoirs, J. Geophys. Res. 84 (1979) 7411–7427.

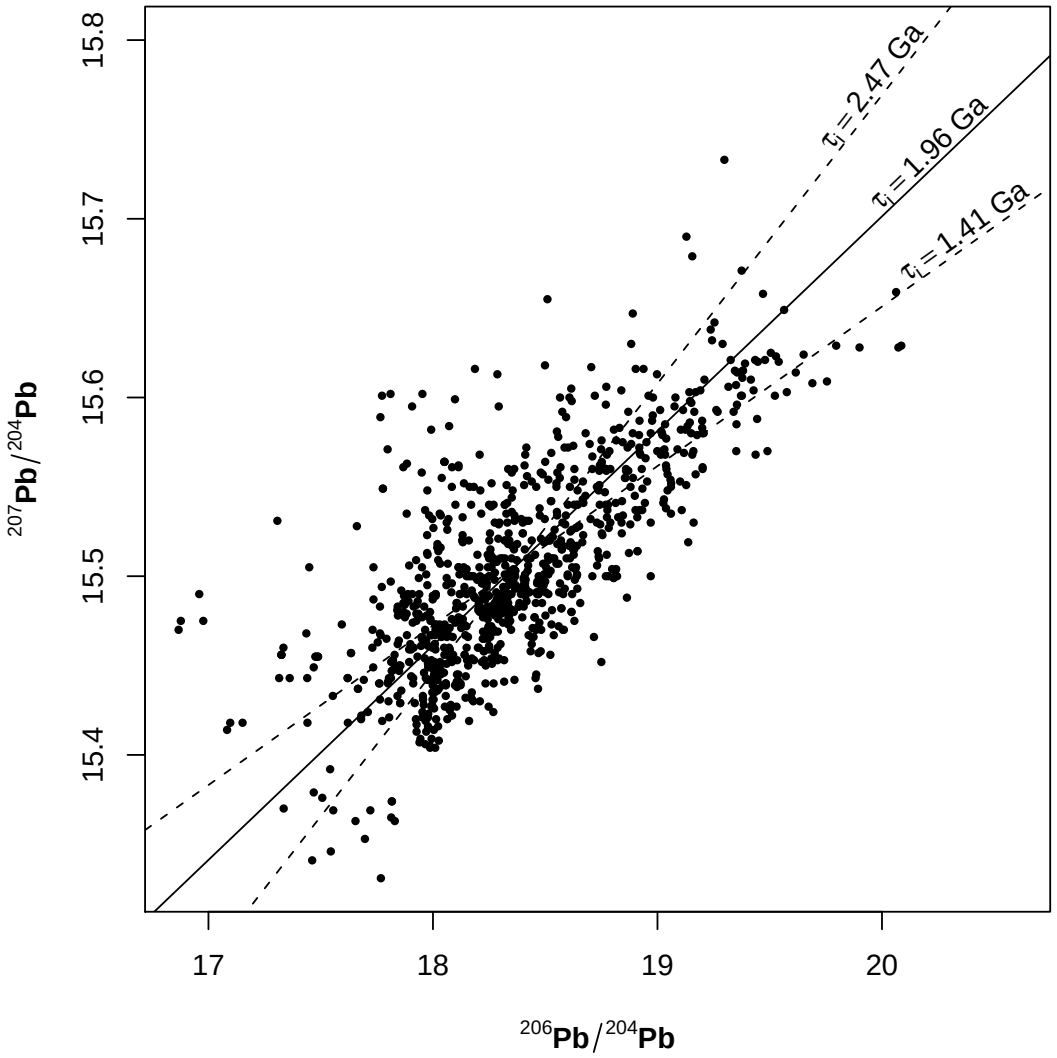

Fig. 1. Scatterplot of measured MORB data (see Appendix F of [9] for a list of data sources). Three regression lines are plotted, and the corresponding pseudo-isochron ages shown. Linear regression of ^{207}Pb/^{204}Pb against ^{206}Pb/^{204}Pb gives a pseudo-isochron age of 1.41 Ga. Linear regression of ^{206}Pb/^{204}Pb against ^{207}Pb/^{204}Pb gives a pseudo-isochron age of 2.47 Ga. The geometric mean regression line has a pseudo-isochron age of 1.96 Ga.

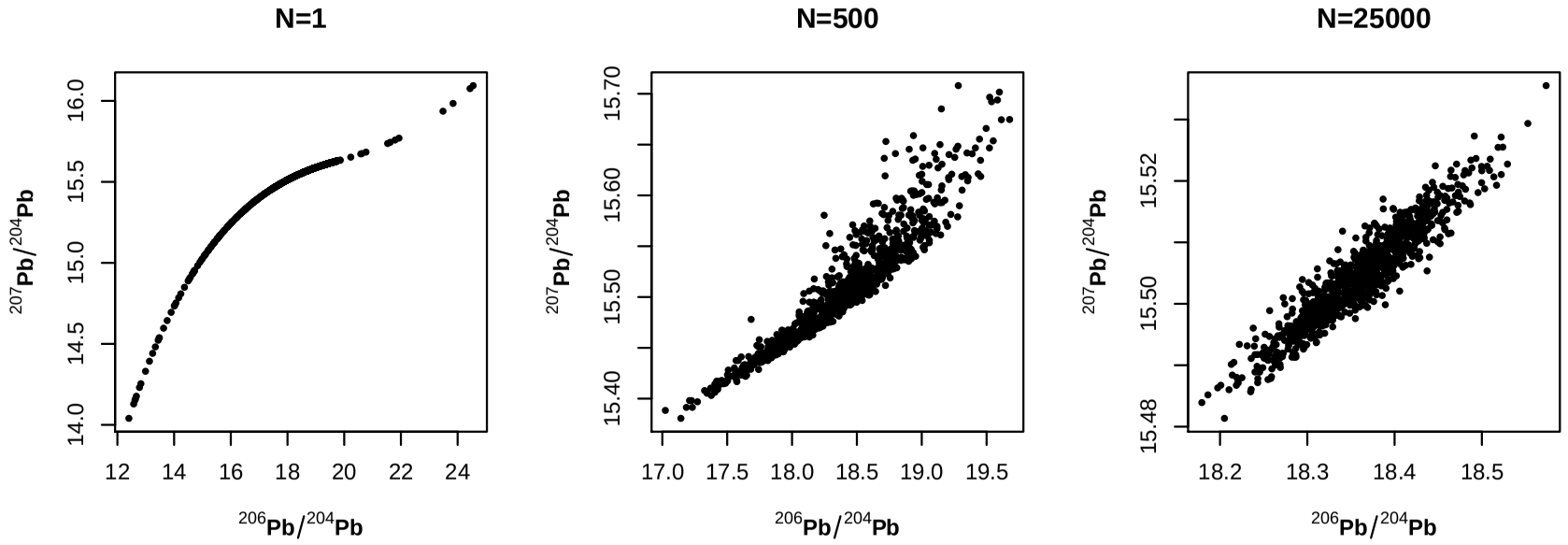

Fig. 2. Constant melt rate model numerical simulations with $N = 1$, 500, and 25000, $\tau_{\text{melt}} = 0.75$ Ga, $\tau_{\text{s}} = 3.0$ Ga, $F = 1.5\%$, and a sample size of 1000. Partition coefficients are those used in [9], giving $G_{\text{Pb}} = 0.42$ and $G_{\text{U}} = 1.00$ ((2) of [9]). The middle picture has similar variance and slope as the measured MORB data (Figure 1). The left picture shows unaveraged compositions, the right heavily averaged compositions. Note that the data forms an ellipse under heavy averaging, indicative of a bivariate normal distribution as expected by the central limit theorem [17]. The geometric mean regression line for the three cases leads to pseudo-isochron ages of 2.30 Ga, 1.98 Ga, and 2.06 Ga respectively. The pseudo-isochron equation (8) predicts an age of 2.03 Ga as $N \to \infty$.

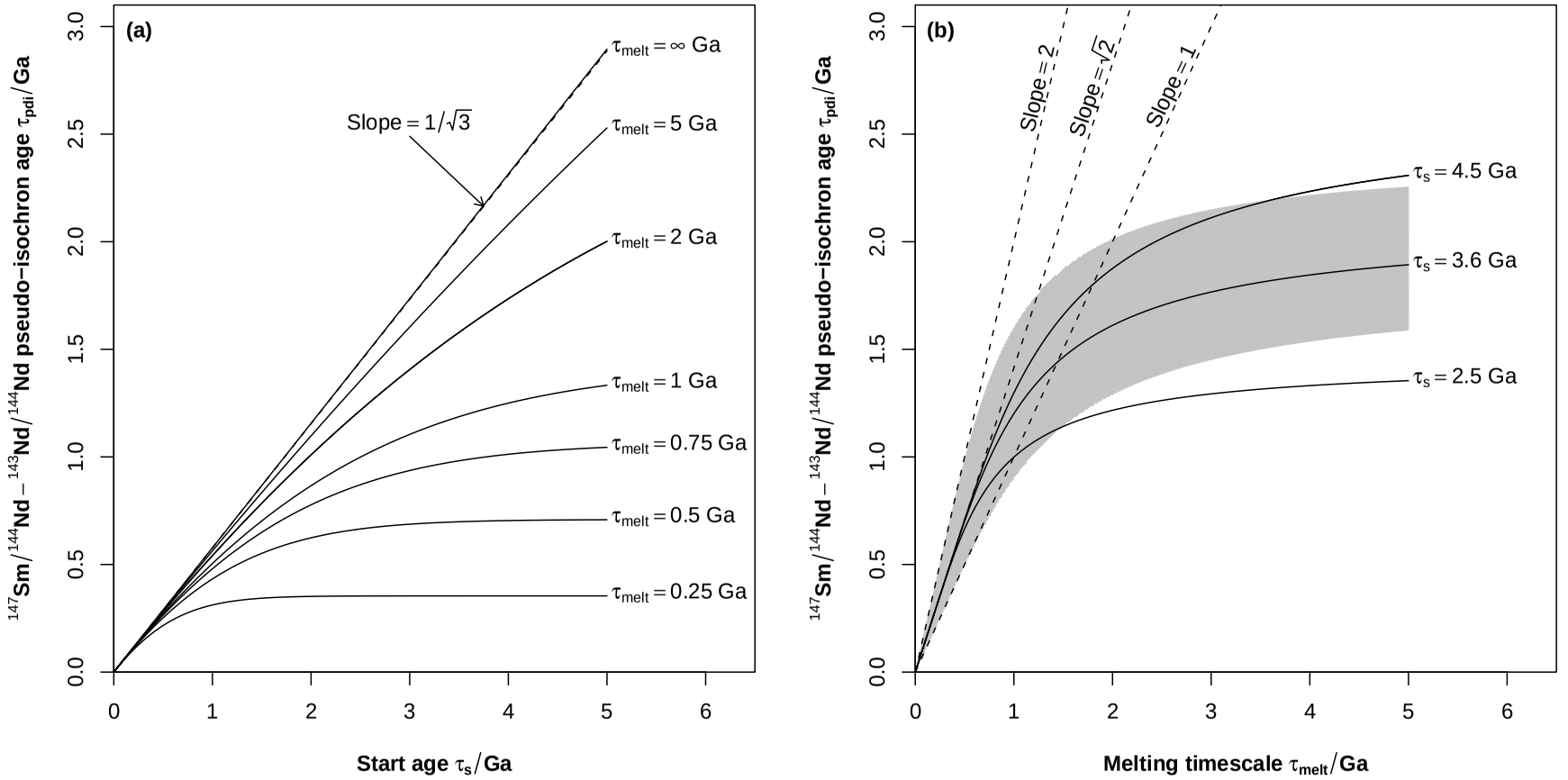

Fig. 3. Plot showing the variation of ^{147}Sm/^{144}Nd-^{143}Nd/^{144}Nd ($\lambda = 0.00654$ Ga^{-1}) model pseudo-isochron age $\tau_{\rm pdi}$ (7). Near identical curves will be produced for other isotopic systems where the decay is linearisable ($\tau_{\rm melt}$ or $\tau_{\rm s} \ll 1/\lambda$), such as ^{87}Rb/^{86}Sr-^{87}Sr/^{86}Sr, ^{176}Lu/^{177}Hf-^{176}Hf/^{177}Hf, and ^{232}Th/^{204}Pb-^{208}Pb/^{204}Pb. (a) plots $\tau_{\rm pdi}$ against start age $\tau_{\rm s}$ for different values of the melting time scale $\tau_{\rm melt}$. For small values of τ_s the curves have slope $= 1/\sqrt{3}$. The curves shown asymptote to $\sqrt{2}\tau_{\rm melt}$ for large values of τ_s with the exception of the $\tau_{\rm melt} = \infty$ Ga curve. (b) plots $\tau_{\rm pdi}$ against $\tau_{\rm melt}$ for different values of $\tau_{\rm s}$. Values of $\tau_s = 2.5$ Ga, 3.6 Ga and 4.5 Ga have been chosen for comparison with numerical simulations of mantle convection [7,8]. For small values of $\tau_{\rm melt}$ the curves have a slope of $\sqrt{2}$. For large values of $\tau_{\rm melt}$ the curves asymptote to pseudo-isochron ages of 1.45 Ga, 2.08 Ga and 2.60 Ga respectively ($\sim \tau_{\rm s}/\sqrt{3}$). Shown in grey is the uncertainty in fitting a line for $\tau_{\rm s} = 3.6$ Ga: the geometric mean regression age (solid line) lies between the two linear regression ages which bound the grey region. The corresponding regions for $\tau_{\rm s} = 2.5$ Ga and 4.5 Ga have not been plotted as they substantially overlap the 3.6 Ga region.

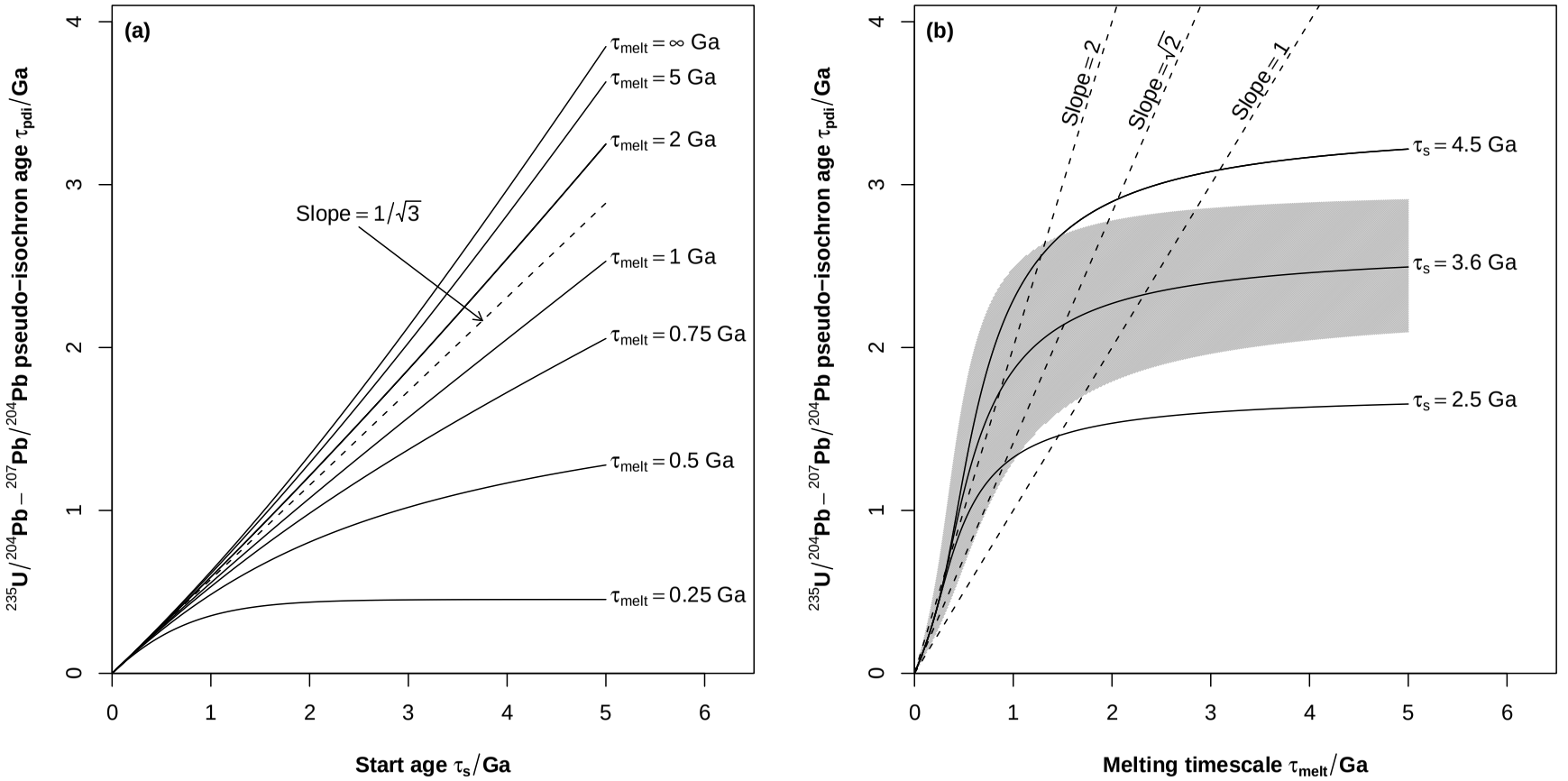

Fig. 4. As Figure 3 but for $^{235}U/^{204}Pb$-$^{207}Pb/^{204}Pb$ ($\lambda = 0.985$ Ga^{-1}), where the decay is not linearisable. (a) plots τ_{pdi} against τ_s for different values of τ_{melt}. For $\tau_{melt} < 1/2\lambda = 0.51$ Ga the curves flatten out for large τ_s. For $\tau_{melt} > 1/2\lambda$ the curves grow linearly with τ_s for large τ_s. The $\tau_{melt} = \infty$ Ga curve approaches a slope of 1 for large τ_s. (b) plots τ_{pdi} against τ_{melt} for different values of τ_s. For large values of τ_{melt} the curves asymptote to pseudo-isochron ages of 1.73 Ga, 2.63 Ga and 3.40 Ga respectively.

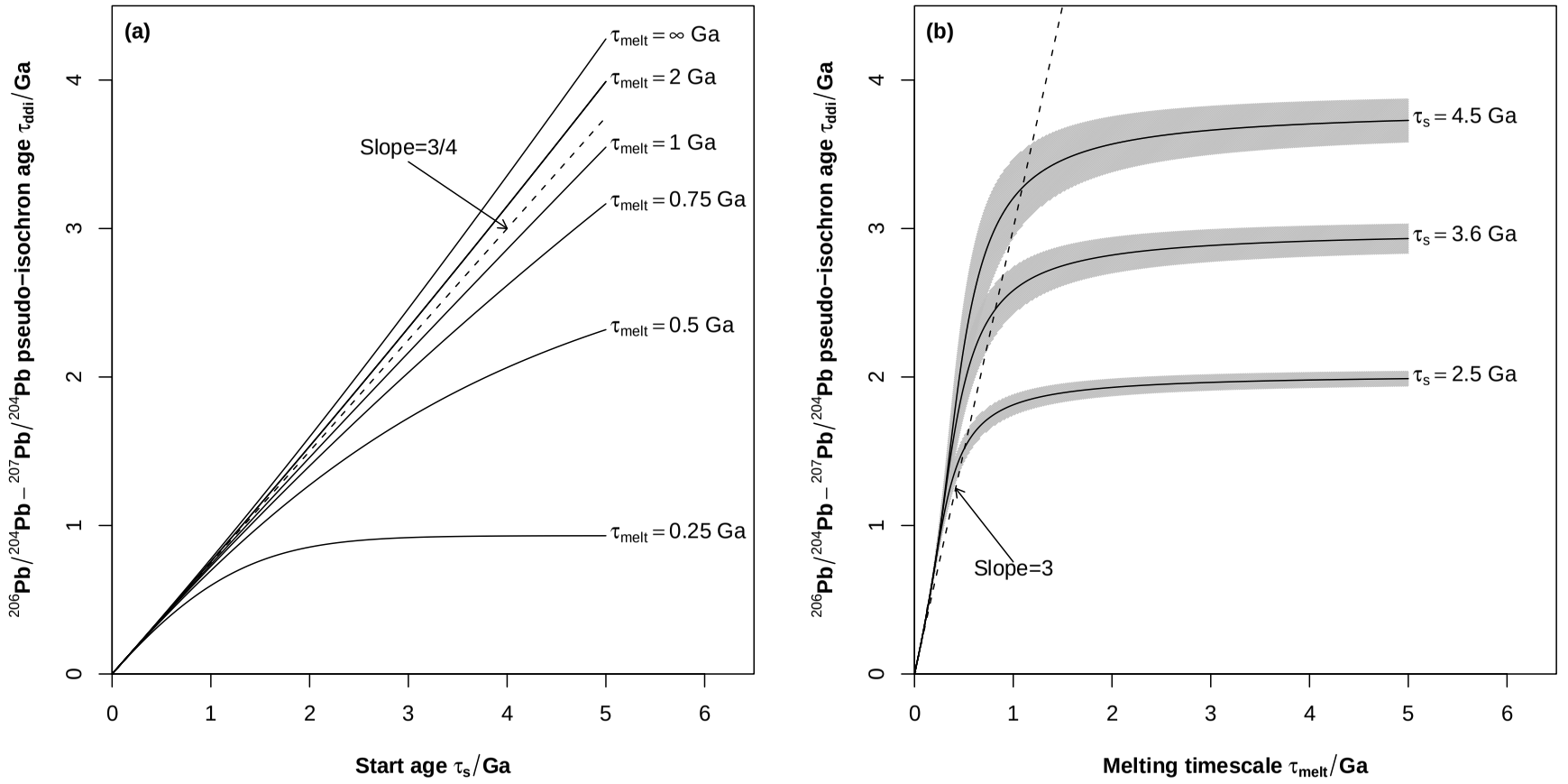

Fig. 5. As Figure 3 but for the ^{206}Pb/^{204}Pb-^{207}Pb/^{204}Pb ($\lambda_{238} = 0.155$ Ga^{-1}, $\lambda_{235} = 0.985$ Ga^{-1}) model pseudo-isochron age $\tau_{\rm ddi}$ (8). (a) plots $\tau_{\rm ddi}$ against $\tau_{\rm s}$ for different values of $\tau_{\rm melt}$. For small values of τ_s the curves have slope = 0.75. For values of $\tau_{\rm melt}$ around 1-2 Ga the approximate relationship $\tau_{\rm ddi} \approx 0.75\tau_s$ holds for the range of τ_s plotted. For $\tau_{\rm melt} < 1/2\lambda_{235} = 0.51$ Ga the curves flatten out for large $\tau_{\rm s}$. For $\tau_{\rm melt} > 1/2\lambda_{235}$ the curves grow linearly with $\tau_{\rm s}$ for large $\tau_{\rm s}$. The $\tau_{\rm melt} = \infty$ Ga curve approaches a slope of 1 for large $\tau_{\rm s}$. (b) plots $\tau_{\rm ddi}$ against $\tau_{\rm melt}$ for different values of $\tau_{\rm s}$. For small values of $\tau_{\rm melt}$ the curves have a slope of 3. For large values of $\tau_{\rm melt}$ the curves asymptote to pseudo-isochron ages of 2.02 Ga, 2.99 Ga and 3.81 Ga respectively. Shown in grey is the uncertainty in fitting a line for each of the $\tau_{\rm s}$ values.

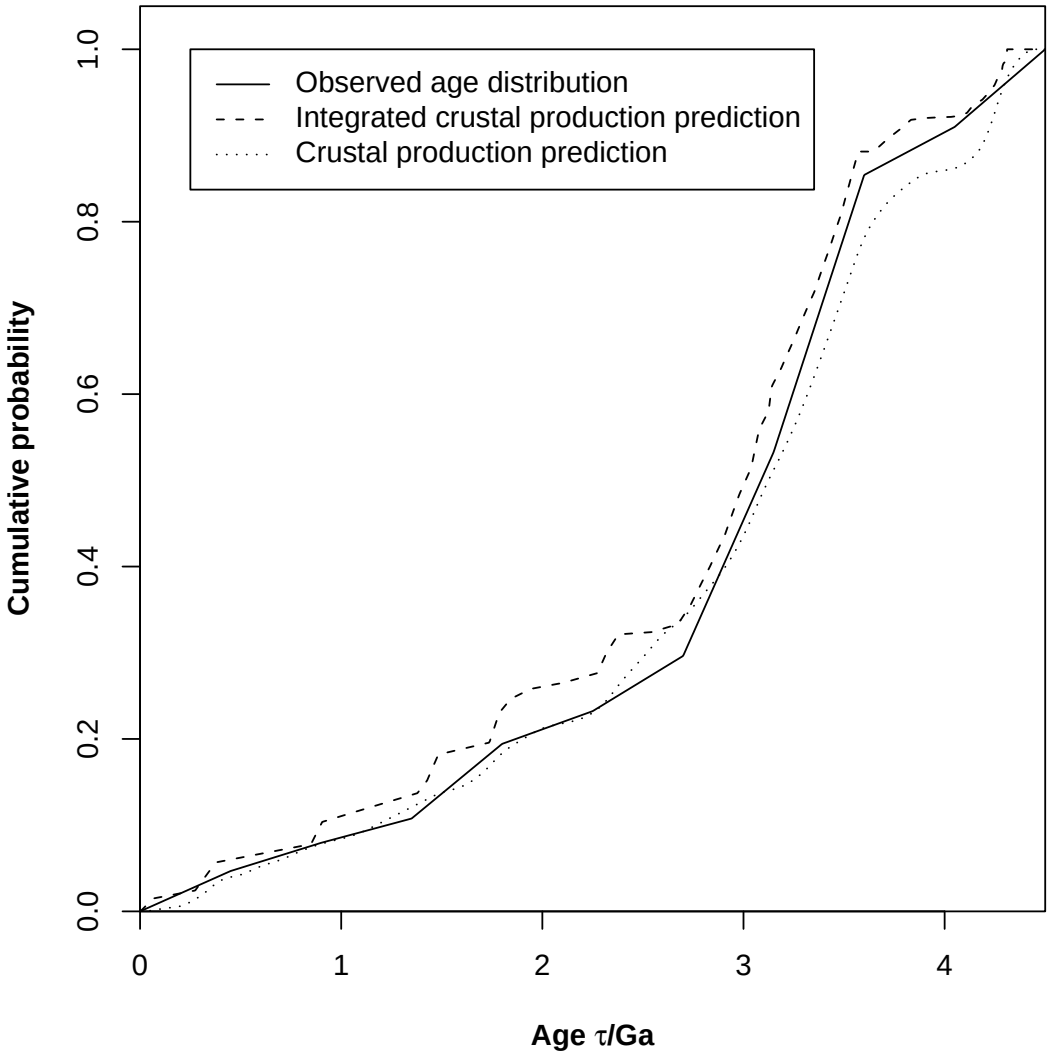

Fig. 6. Plot of cumulative distribution function $Q_m(\tau) = \int_0^\tau q_m(\tau)\,\mathrm{d}\tau$ for a Xie and Tackley [8] numerical simulation. The solid line shows the observed age distribution (their Figure 10a), where residue and basalt tracers have been lumped together. The dashed line is a calculated age distribution (Appendix G) based on their observed integrated crustal production (solid line of their Figure 5c). The dotted line is a calculated age distribution based on their observed crustal production rate (solid line of their Figure 5e), although note that this rate is actually an average over three slightly different runs. The three age distributions are very similar. Using (4) and (5), the age distributions correspond to lead-lead pseudo-isochron ages τ_{ddi} of 3.67 Ga, 3.58 Ga, and 3.76 Ga, Sm-Nd pseudo-isochron ages τ_{pdi} of 3.00 Ga, 2.88 Ga, and 3.06 Ga, and mean ages $\bar{\tau}_{\mathrm{m}}$ of 2.81 Ga, 2.67 Ga, and 2.86 Ga respectively.

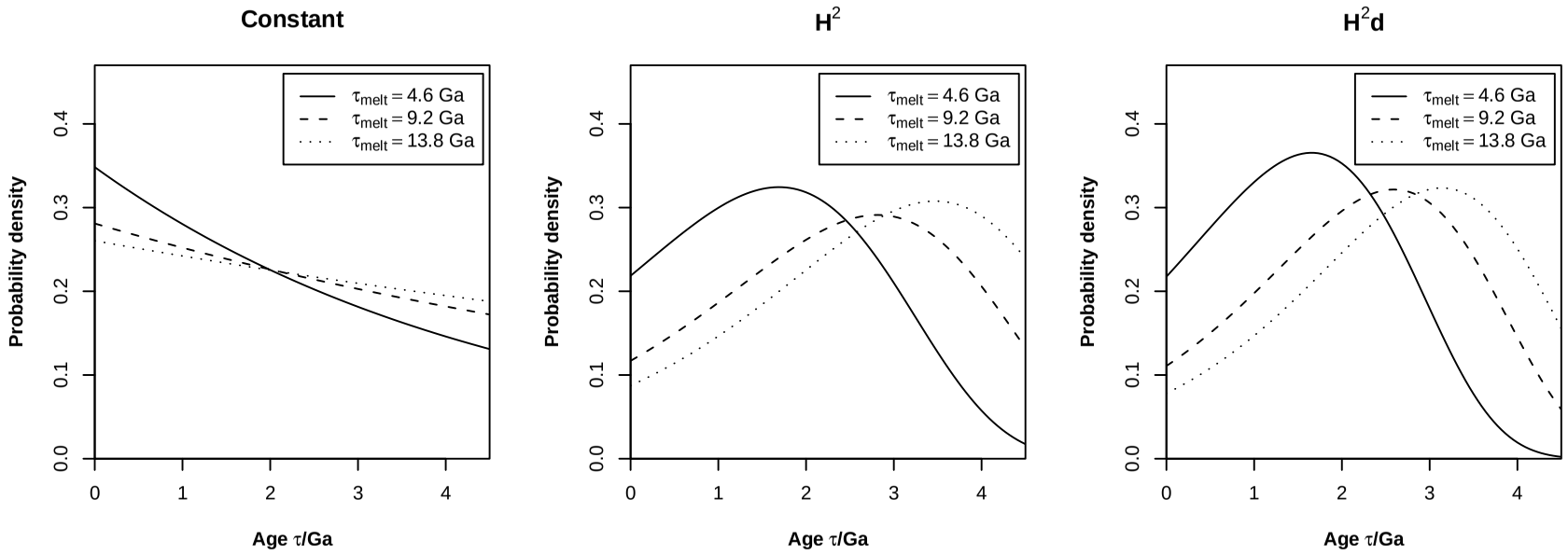

Fig. 7. Plot of probability density function $q_{\rm m}(\tau)$. The melt rate is given by $\gamma_{\rm melt}(\tau) = {\rm e}^{\nu\tau}/\tau_{\rm melt}$, and the three cases are constant ($\nu = 0$ Ga^{-1}), H^2 ($\nu = 0.62$ Ga^{-1}), and $H^2 d$ ($\nu = 0.74$ Ga^{-1}).

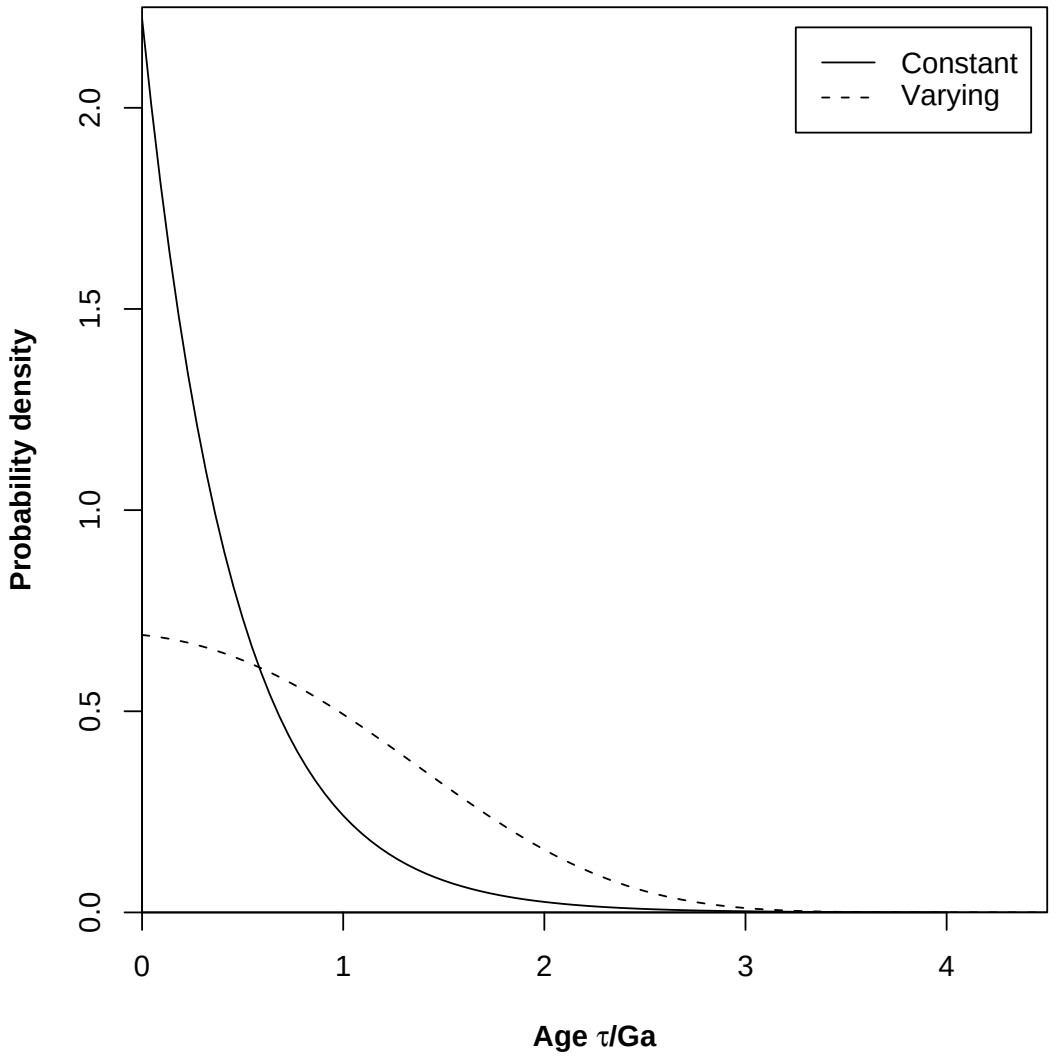

Fig. 8. Plot of probability density function $q_{\mathrm{m}}(\tau)$. Two cases are shown which both yield a lead-lead pseudo-isochron age of 2.0 Ga. The solid line has constant melt rate $\gamma_{\mathrm{melt}}(\tau) = 1/\tau_{\mathrm{melt}}$ where $\tau_{\mathrm{melt}} = 0.45$ Ga. The dashed line has varying melt rate $\gamma_{\mathrm{melt}}(\tau) = \mathrm{e}^{\nu\tau}/\tau_{\mathrm{melt}}$ where $\nu = 0.62$ Ga^{-1} and $\tau_{\mathrm{melt}} = 1.45$ Ga. While the lead-lead ages are identical for the two cases, the parent-daughter and mean ages are not. Sm-Nd $\tau_{\mathrm{pdi}} = 0.64$ Ga and 1.10 Ga, and $\bar{\tau}_{\mathrm{m}} = 0.45$ Ga and 0.90 Ga respectively. Note that in both cases the very old (> 3.0 Ga) heterogeneity has been essentially eliminated.

Table 1
Age calculations for varying melt rate $\gamma_{\rm melt}(\tau) = {\rm e}^{\nu\tau}/\tau_{\rm melt}$ (compare with Table 4 of [8]). The three cases are constant ($\nu = 0$ Ga^{-1}), H^2 ($\nu = 0.62$ Ga^{-1}), and $H^2 d$ ($\nu = 0.74$ Ga^{-1}). The present day crustal production rate $b/\tau_{\rm melt} = 1/46$ Ga^{-1} in all cases. $b = 0.1$, 0.2, and 0.3, and hence $\tau_{\rm melt} = 4.6$ Ga, 9.2 Ga and 13.8 Ga. $\tau_{\rm s} = 4.5$ Ga and all ages are quoted in Ga.

	Constant			H^2			$H^2 d$		
$\tau_{\rm melt}$	4.6	9.2	13.8	4.6	9.2	13.8	4.6	9.2	13.8
$\bar{\tau}_{\rm total}$	2.87	3.56	3.84	1.82	2.54	2.96	1.67	2.31	2.72
$\bar{\tau}_{\rm m}$	1.89	2.07	2.13	1.81	2.39	2.64	1.67	2.27	2.58
$\tau_{\rm pdi}$[Sm-Nd]	2.28	2.44	2.50	2.09	2.66	2.89	1.92	2.53	2.82
$\tau_{\rm ddi}$[Pb-Pb]	3.72	3.77	3.78	3.29	3.69	3.81	2.97	3.54	3.73