Geophys. J. R. astr. Soc. (1982) **70**, 755–787

Optimal estimation of scalar seismic moment

Paul G. Silver and Thomas H. Jordan *Geological Research Division, Scripps Institution of Oceanography, La Jolla, California 92093, USA*

Received 1982 January 9; in original form 1981 September 9

Summary. For any seismic source specified by a frequency-dependent moment-rate tensor $\mathbf{M}(\omega)$, we define the total moment $M_T(\omega) = \|\mathbf{M}\|/\sqrt{2}$ and the isotropic moment $M_I(\omega) = |\operatorname{tr}\mathbf{M}|/\sqrt{6}$. A method is presented for estimating these scalar seismic moments from noisy seismic data when the source mechanism is uncertain or completely unknown. Our formulation exploits the linear relation between squared moment M^2 (total or isotropic) and the product of two seismic spectra; in a particular frequency band, an estimate $\tilde{M}^2$ is constructed as a linear combination of power and cross-spectra integrated across the band. The coefficients yielding an exact estimate from perfect data are the solution to a linear system of equations involving spectral integrals of the transfer functions that relate $\mathbf{M}$ to the seismograms. The failure to solve this system exactly induces an error in $\tilde{M}^2$ whose statistics can be calculated from a likelihood function for the source mechanism, which we model using the hyperspherical normal distribution and its Gaussian approximation and generalizations. We also develop expressions for the bias and variance induced in $\tilde{M}^2$ by ambient seismic noise and by transfer-function errors due to aspherical heterogeneity. To optimize the estimate, the coefficients specifying $\tilde{M}^2$ are computed by minimizing a non-negative-definite quadratic form constructed from these statistics. We have applied the method to IDA records of the deep-focus Honshu earthquake of 1978 March 7 and the shallow-focus Oaxaca earthquake of 1978 November 29. For each event, estimates of M_T have been obtained with good precision over disjunct 1-mHz bands spanning the frequency interval 1–11 mHz; their relative standard deviations range from 10 to 22 per cent. Our best estimate of M_T averaged over the entire 1–11 mHz interval is 0.43×10^{27} dyne cm for Honshu and 2.8×10^{27} dyne cm for Oaxaca. The isotropic component of the Oaxaca event, as measured by the ratio $\tilde{M}_I^2/\tilde{M}_T^2$, is negligibly small ($< 0.1$). In the case of Honshu, however, this ratio averages about 0.34; all six estimates at frequencies greater than 5 mHz are significantly greater than zero at the 90 per cent confidence level, and four of the six are significant at the 95 per cent level. This observation lends credence to the conjecture made previously by seismologists that isotropic compression accompanies some deep-focus earthquakes.

1 Introduction

The concept of seismic moment as an appropriate measure of earthquake size arises from the dynamical equivalence of a dislocation – the idealization of slip on a fault – and a double couple (Burridge & Knopoff 1964). Aki (1966) was the first to estimate seismic moment by a systematic analysis of seismic data, and he exploited the equivalence of these two source descriptions to infer the average fault displacement for the 1964 Niigata Earthquake. The success of Aki's procedure, in which the data are fit by assuming exact prior knowledge of the source mechanism, has led to its wide application in other earthquake studies (e.g. Kanamori 1970a, b).

This paper presents a much different algorithm for estimating seismic moment. The procedure is optimized for the inversion of noisy seismic data in the realistic situation when some information about the source mechanism may (or may not) be available, but when this prior information does not specify the mechanism exactly.

The problem of moment estimation will be formulated in the general context of a moment-tensor description of the seismic source, of which the double couple is an important special case. Source mechanisms more general than a double couple but with radiation patterns that can be expanded in harmonics of zeroth and second degree were discussed by Knopoff & Randall (1970) and applied to deep-focus events by Randall & Knopoff (1970). Randall (1971) and Gilbert (1971a, 1973a) recognized that any such mechanism can be specified by a symmetric, second-order moment tensor $\mathcal{M}$ representing a torque-free linear combination of couples at a point. The general theory of moment-tensor densities for extended seismic sources was developed by Backus & Mulcahy (1976a, b). These authors demonstrated that all indigenous sources can be described by an excess stress, or stress glut, and that the ordinary time-dependent moment tensor $\mathcal{M}(t)$ is the zeroth-degree spatial moment in a polynomial expansion of the stress glut; that is, $\mathcal{M}(t)$ is the volume integral of the stress glut over the physical source region.

Attention in this paper is confined to localized seismic sources. For our purposes a source is 'localized' if the stress glut is sufficiently concentrated around its centroid $(\mathbf{r}_0, t_0)$ so that: (1) the truncation of higher-degree spatial moments is justified for the wavelengths of interest, and (2) the Fourier spectrum of the moment-rate tensor $\mathbf{M}(t) \equiv d\mathcal{M}/dt$ is slowly varying across the frequency band of interest. A discussion of these conditions, including the precise definitions of spatial and temporal centroids, is provided by Backus (1977).

The displacement field produced by a localized seismic source is linearly related to the seismic moment tensor, a fact first recognized by Gilbert (1971a, 1973a) and exploited by Dziewonski & Gilbert (1974) and Gilbert & Dziewonski (1975). The complex Fourier spectrum of the seismogram u at an observation point $\mathbf{r}$ admits the normal-mode expansion

$$u(\mathbf{r}, \omega) = \sum_k C_k(\omega)\, \mathscr{A}_k(\mathbf{r}, \mathbf{r}_0) : \mathbf{M}(\omega), \tag{1.1}$$

where $C_k(\omega)$ is the resonance spectrum of the kth mode and $\mathscr{A}_k$ is its tensor-valued excitation function (Gilbert & Dziewonski 1975). Definition of the second-order tensor

$$\mathbf{G}(\mathbf{r}, \mathbf{r}_0, \omega) = \sum_k C_k(\omega)\, \mathscr{A}_k(\mathbf{r}, \mathbf{r}_0) \tag{1.2}$$

allows equation (1.1) to be succinctly expressed as

$$u(\mathbf{r}, \omega) = \mathbf{G}(\mathbf{r}, \mathbf{r}_0, \omega) : \mathbf{M}(\omega). \tag{1.3}$$

If convenient, the standing-wave sum (1.2) can be recast as a superposition of travelling modes or generalized rays, or directly in terms of the impulse response (Gilbert 1976; Backus & Mulcahy 1976a; Stump & Johnson 1977; Aki & Richards 1980).

The problem of estimating the moment-rate tensor by the direct inversion of equation (1.3) has been formulated for various types of seismic signals, including free oscillations (Gilbert & Dziewonski 1975; Buland & Gilbert 1976), surface waves (McCowan 1976; Mendiguren 1977), and body waves (Strelitz 1980; Ward 1980a, b). The application of these linear inversion techniques to digitally recorded data is yielding important new information about the nature of the seismic source (Dziewonski, Chou & Woodhouse 1981; Masters & Gilbert 1981, private communication).

Additional work on the methodology of source recovery is justified, however. Recent studies have demonstrated that significant bias can be introduced into moment-tensor estimates by the use of incorrect transfer functions (Patton & Aki 1979). Variations in the transfer functions caused by the Earth's lateral heterogeneity are a particularly serious source of bias, especially at low frequencies where the moment-rate tensor must be estimated from long time series and sparse networks of stations. In lieu of transfer functions that explicitly account for aspherical heterogeneity, more complicated schemes than the direct inversion of (1.3) may be required to reduce this bias.

This paper does not address the general inverse problem for the moment-rate tensor; rather, it attacks the restricted problem of estimating scalar seismic moment from low-frequency seismic data. In constructing an optimized inversion procedure for scalar moment explicit consideration is given to the effects of seismic noise and errors in the transfer function **G**. In particular, we consider the case where estimates of **G** are biased by ignoring lateral heterogeneity. The normal mode representation (1.2) is used throughout our analysis.

The total (scalar) seismic moment is here defined to be the positive, real function of frequency whose square is one-half the squared Euclidian length of **M**,

$$M_T^2(\omega) = \tfrac{1}{2}\,\mathbf{M}^*(\omega) : \mathbf{M}(\omega). \tag{1.4}$$

The asterisk indicates complex-conjugate (Hermitian) transpose. A point dislocation, or double couple, is specified by a unit vector $\hat{\mathbf{f}}$ normal to the fault plane, a unit slip vector $\hat{\mathbf{s}}$, and a step-function time history $H(t)$. In this special case the moment-rate tensor is independent of frequency,

$$\mathbf{M}(\omega) = M_0\,[\hat{\mathbf{f}}\hat{\mathbf{s}} + \hat{\mathbf{s}}\hat{\mathbf{f}}] \qquad \text{(double couple)}, \tag{1.5}$$

and M_T reduces to the ordinary scalar moment M_0. Requiring this correspondence imposes the factor of $\tfrac{1}{2}$ in equation (1.4). This definition of total moment allows any moment-rate tensor to be written

$$\mathbf{M}(\omega) = \sqrt{2} M_T(\omega)\,\hat{\mathbf{M}}(\omega), \tag{1.6}$$

where $\hat{\mathbf{M}}$ is a symmetric tensor of unit length:

$$\hat{\mathbf{M}}^*(\omega) : \hat{\mathbf{M}}(\omega) = 1. \tag{1.7}$$

Generalizing the terminology commonly applied to the bracketed expression in equation (1.5), we shall refer to the normalized moment-rate tensor $\hat{\mathbf{M}}$ as the source mechanism. Let $\mathbf{M}_D$ be the deviatoric part of **M**, so that tr $\mathbf{M}_D = 0$ and

$$\mathbf{M}(\omega) = \mathbf{M}_D + \tfrac{1}{3}(\mathrm{tr}\,\mathbf{M})\,\mathbf{I}. \tag{1.8}$$

Then, a deviatoric moment M_D and an isotropic moment M_I are defined by the equations

$$M_D^2(\omega) = \tfrac{1}{2}\,\mathbf{M}_D^*(\omega) : \mathbf{M}_D(\omega) \tag{1.9}$$

$$M_I^2(\omega) = \tfrac{1}{6}\,|\,\mathrm{tr}\,\mathbf{M}(\omega)\,|^2 = \tfrac{1}{6}\,|\,\mathbf{I} : \mathbf{M}(\omega)\,|^2. \tag{1.10}$$

The three types of scalar moment satisfy the simple relation

$$M_T^2(\omega) = M_D^2(\omega) + M_I^2(\omega). \tag{1.11}$$

M_T is a rotational invariant and can be written in terms of the three eigenvalues of $\mathbf{M}$; i.e.

$$M_T = \left(\frac{1}{2}\sum_{i=1}^{3} |\lambda_i|^2\right)^{1/2}.$$

Of course, other satisfactory generalizations of scalar seismic moment exist. Total moment could instead be defined as the maximum of $|\lambda_i|$, the so-called spectral norm of $\mathbf{M}$, which also reduces to M_0 in the special case of a double couple. Such a definition does not lend itself to the computational procedures proposed in this paper, however.

At zero frequency the values of total moment given by equation (1.4) is simply related to what Backus (1977, equation 5.4) calls the total 'charge' of the moment(glut)-rate density, denoted $c_M^{(0)}$ and defined to be the squared Euclidian norm of $\mathbf{M}(0)$: $c_M^{(0)} = \|\mathbf{M}(0)\|^2 = 2M_T^2(0)$. Both $M_T(0)$ and $c_M^{(0)}$ measure the total 'size' of a seismic event. Randall's (1971) 'shear invariant' L is linearly related to the deviatoric moment by $L = \mu^{-1} M_D(0)$, where μ is the shear modulus. Both $M_D(0)$ and L measure the size of an event's deviatoric part.

Much can be learned about earthquakes by studying the scalar moments M_T, M_D and M_I. The accurate determination of M_T at low frequencies provides the most useful information about earthquake size, radiated energy and fault slippage, especially for the very large events that dominate global strain release and saturate standard magnitude scales (Kanamori 1977; Hanks & Kanamori 1979). The variation of M_T with frequency places constraints on the temporal behaviour of the seismic source and should be diagnostic of the anomalous time functions associated with so-called 'slow earthquakes' (Kanamori & Cipar 1974; Kanamori & Stewart 1979; Sacks *et al.* 1978, 1981). The existence of a non-zero M_I is indicative of an apparent isotropic component of strain release, a phenomenon observed for some deep-focus events (Dziewonski & Gilbert 1974; Gilbert & Dziewonski 1975) and perhaps diagnostic of the volume contraction due to a phase change (Evison 1963; Benioff 1963; Randall 1972).

To resolve anomalous temporal behaviour or prove the existence of an isotropic component using these methods requires that we be able to test various hypotheses regarding the numerical values of the scalar moments. Careful attention must therefore be given to the statistical properties of the estimates and to schemes for reducing the estimation errors.

The problems encountered in dealing with the errors in scalar moment are illustrated by a simple example: suppose an estimate $\tilde{\mathbf{M}}$ of the moment-rate tensor is available at some particular frequency, so that the error in this estimate is represented by the tensor $\mathbf{\Delta M} \equiv \tilde{\mathbf{M}} - \mathbf{M}$, considered to be a sample of a stochastic process. Then, the expected value of the estimated squared moment $\tilde{M}_T^2 = \frac{1}{2}\|\tilde{\mathbf{M}}\|^2$ is

$$\langle \tilde{M}_T^2 \rangle = M_T^2 + \mathrm{Re}[\mathbf{M}^* : \langle \mathbf{\Delta M} \rangle] + \tfrac{1}{2}\langle \|\mathbf{\Delta M}\|^2 \rangle. \tag{1.12}$$

Thus, even if $\tilde{\mathbf{M}}$ is unbiased ($\langle \mathbf{\Delta M} \rangle = 0$), the estimate $\tilde{M}_T^2$, and hence its square root $\tilde{M}_T$, is not: random errors in $\tilde{\mathbf{M}}$ introduce bias into $\tilde{M}_T^2$ in proportion to their variance (second statistical moment). Furthermore, the expression for the variance of $\tilde{M}_T^2$ generally involves the first four statistical moments of $\mathbf{\Delta M}$. It is clear, therefore, that procedures designed to yield unbiased, minimum-variance estimates of the moment-rate tensor do not necessarily provide similarly optimal estimates of scalar moment.

2 A linear inverse problem for squared moments

Our method of moment retrieval rests on the demonstration that the squared scalar moments M_T^2, M_I^2 and M_D^2 can be specified by certain linear combinations of the power and

cross-spectra from a sparse network of seismic stations. We show in this section that the coefficients of these linear combinations are themselves the solutions to linear systems of equations. Hence, the retrieval of the squared moments can be formulated as a linear problem, fully amenable to the rigorous methods of linear estimation theory.

The basic data function used in our formulation is the cross-spectrum for the record pair (p, q) selected from a set of P seismograms and integrated over the positive frequency interval (ω_a, ω_b):

$$U_{pq} = \int_{\omega_a}^{\omega_b} u_p^*(\omega)\, u_q(\omega)\, d\omega. \tag{2.1}$$

The frequency interval (ω_a, ω_b) is assumed to be small enough and the spectrum of the moment-rate tensor smooth enough so that any variation of $\mathbf{M}(\omega)$ across this band can be ignored. Equation (2.1) can thus be written

$$U_{pq} = \mathbf{M}^* : \mathscr{H}_{pq} : \mathbf{M}, \tag{2.2}$$

where $\mathscr{H}_{pq}$ is the fourth-order outer-product tensor

$$\mathscr{H}_{pq} = \int_{\omega_a}^{\omega_b} \mathbf{G}_p^*(\omega)\, \mathbf{G}_q(\omega)\, d\omega. \tag{2.3}$$

To relate these data functionals to the scalar seismic moments we consider the symmetrized fourth-order isotropic tensor

$$\mathscr{C}_{ijkl}(\lambda, \mu) = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{2.4}$$

and its associated quadratic form

$$M^2(\lambda, \mu) = \mathbf{M}^* : \mathscr{C}(\lambda, \mu) : \mathbf{M}. \tag{2.5}$$

It is easily verified that the squared moments defined in the previous section correspond to values of (2.5) for particular choices of the scalar coefficients λ and μ; namely

$$M_{\mathrm{T}}^2 = M^2(0, \tfrac{1}{4}) \tag{2.6a}$$

$$M_{\mathrm{I}}^2 = M^2(\tfrac{1}{6}, 0) \tag{2.6b}$$

$$M_{\mathrm{D}}^2 = M^2(-\tfrac{1}{6}, \tfrac{1}{4}). \tag{2.6c}$$

The similarity between equations (2.2) and (2.5) suggests seeking estimates of the squared moments that are linear combinations of the U_{pq}s. For some set of complex-valued coefficients $\{a_{pq}\}$ we define the scalar quantity

$$\tilde{M}^2 = \sum_{pq} a_{pq}\, U_{pq}. \tag{2.7}$$

The index pairs (p, q) included in the summation are permitted to range over the entire set of P^2 record pairs or can be limited to some subset; for example, the summation can be restricted to include only power spectra $(p = q)$. We require, however, that $\tilde{M}^2$ be strictly real. Since the U_{pq}s have Hermitian symmetry $(U_{qp} = U_{pq}^*)$, this is assured if the a_{pq}s are Hermitian $(a_{pq} = a_{pq}^*)$ and if the summation is always taken over both permutations of any record pair.

Substituting (2.2) into (2.7) yields

$$\tilde{M}^2 = \mathbf{M}^* : \tilde{\mathscr{C}} : \mathbf{M} \tag{2.8}$$

where

$$\tilde{\mathscr{C}} \equiv \sum_{pq} a_{pq} \mathscr{H}_{pq}. \tag{2.9}$$

If the coefficients $\{a_{pq}\}$ can be constructed such that $\tilde{\mathscr{C}} = \mathscr{C}(\lambda, \mu)$ for any particular pair (λ, μ), then $\tilde{M}^2 = M^2(\lambda, \mu)$. Estimating squared moment is thus reduced to solving the linear system

$$\sum_{pq} a_{pq} \mathscr{H}_{pq} = \mathscr{C}(\lambda, \mu). \tag{2.10}$$

The components of the complex-valued tensor $\mathscr{H}_{pq}$ satisfy

$$\mathscr{H}^{pq}_{ijkl} = \mathscr{H}^{pq}_{jikl} = \mathscr{H}^{pq}_{ijlk} = \mathscr{H}^{qp*}_{klij}. \tag{2.11}$$

Because $a_{pq} = a^*_{pq}$, $\tilde{\mathscr{C}}$ has the symmetries

$$\tilde{\mathscr{C}}_{ijkl} = \tilde{\mathscr{C}}_{jikl} = \tilde{\mathscr{C}}_{ijkl} = \tilde{\mathscr{C}}^*_{klij}. \tag{2.12}$$

The tensor defined by (2.4) is real-valued and shares the same symmetries as the real part of $\tilde{\mathscr{C}}$. Hence, the 81 complex equations in the linear system (2.10) can be reduced to 36 real equations, 21 requiring $\mathrm{Re}[\tilde{\mathscr{C}}] = \mathscr{C}$ and 15 requiring $\mathrm{Im}[\tilde{\mathscr{C}}] = 0$.

Suppose the summation in (2.7) comprises all P^2 integrals of power and cross-spectra available from a network of P seismographs, so that the set of coefficients $\{a_{pq}\}$ has P^2 degrees of freedom. Then, for buried sources, $P \geqslant 6$ is sufficient to ensure that the system (2.10) has at least one solution, barring accidental degeneracies. Sparse global arrays such as the International Deployment of Accelerometers (IDA) and Seismic Research Observatory (SRO) networks should therefore be adequate for the inversion of (2.10) in individual frequency bands.

Of course, even fewer records are required if the moment-rate tensor is a smooth enough function of frequency to permit the mixing of U_{pq}s from more than one frequency band. Under the assumption of frequency independence, as few as two vertical-component seismograms or one horizontal-component seismogram are sufficient to constrain the entire moment-rate tensor (Gilbert & Buland 1976).

The number of records needed to solve (2.10) exactly can also be reduced by assuming a special form for the moment-rate tensor. In general, the six independent components of $\mathbf{M}$ at any particular frequency are complex with differing amplitudes and phases. Suppose, however, that the phases of these components are in fact equal, so that the moment-rate tensor can be written

$$\mathbf{M}(\omega) = |\mathbf{M}(\omega)| \exp[i\phi(\omega)], \tag{2.13}$$

where $|\mathbf{M}|$ is the modulus of $\mathbf{M}$ and ϕ is a real-valued function of frequency. Then, any two components of $\mathbf{M}$ are related by a zero-phase filter, and their time-domain cross-correlation functions are symmetric about zero lag. Such a moment-rate tensor describes what we shall call a *synchronous source*. This terminology conforms to the fact that the six time functions of a general synchronous source all have the same temporal centroid, as defined by Backus (1977, equation 5.5e). Clearly, any moment-rate tensor whose components share the same time history is a special case of a synchronous source.

For a synchronous source the symmetries expressed in (2.12) imply

$$\mathbf{M}^* : \mathrm{Im}[\tilde{\mathscr{C}}] : \mathbf{M} \equiv 0 \qquad \text{(synchronous source)}. \tag{2.14}$$

Hence, the imaginary part of $\tilde{\mathscr{C}}$ need not be constrained to vanish, and only the 21 real, scalar equations that guarantee $\mathrm{Re}[\tilde{\mathscr{C}}] = \mathscr{C}$ are sufficient to yield $\tilde{M}^2 = M^2$. If no degeneracies occur, then five or more records are adequate to construct the squared moments of a synchronous source in a single frequency band.

Assuming there exists at least one solution to (2.10), or, in the case of a synchronous source, the real part of (2.10), it can be found by applying an appropriate generalized inverse to this linear system. Scalar moments can then be derived from equation (2.7) without the need for any prior information about the source mechanism.

To translate this theoretical procedure into a viable algorithm some accounting must be made for the presence of noise in the observed values of the integrated power and cross-spectra. Furthermore, if any information about the source mechanism does exist, as is often the case, it should be used in constructing the moment estimates. The development of an algorithm optimized for noisy data and capable of incorporating prior information about the source mechanism is the subject of the remainder of this paper.

3 Convenient isomorphisms

The analysis of the moment-retrieval problem is simplified considerably by introducing an isomorphism between symmetric, second-order tensors of dimension three and vectors of dimension six. By fixing an appropriate Cartesian reference frame in three-space (usually a local geographic frame at the source) we define for every symmetric tensor $\mathbf{M}$ a six-vector $\mathbf{m}$:

$$\left.\begin{aligned} &m_1 = M_{11}, \qquad m_2 = M_{22}, \qquad m_3 = M_{33} \\ &m_4 = \sqrt{2}\,M_{12}, \qquad m_5 = \sqrt{2}\,M_{13}, \qquad m_6 = \sqrt{2}\,M_{23}. \end{aligned}\right\} \tag{3.1}$$

This isomorphism, written $\mathbf{M} \Leftrightarrow \mathbf{m}$, has the advantage of preserving the Euclidian norm:

$$\| \mathbf{m} \|^2 \equiv \mathbf{m}^* \cdot \mathbf{m} = \mathbf{M}^* : \mathbf{M} = \| \mathbf{M} \|^2 \tag{3.2}$$

A similar isomorphism was employed by Gilbert & Dziewonski (1975, equation 17), but they omit the factors of $\sqrt{2}$, and their isomorphism is not norm-preserving.

Analogous to (3.1) is an isomorphism between fourth-order tensors of dimension three satisfying the symmetry conditions (2.12) and Hermitian matrices of dimension six, constructed to preserve the quadratic form (2.8). The correspondence $\tilde{\mathscr{C}} \Leftrightarrow \tilde{\mathbf{C}}$ is expressed by the equations

$$\tilde{C}_{mn} = \gamma_{mn}\,\tilde{\mathscr{C}}_{ijkl}, \tag{3.3}$$

where m and n are specified by the index pairs ij and kl according to (3.1), respectively, and the coefficients $\{\gamma_{mn}\}$ form the symmetric array

$$\begin{bmatrix} 1 & 1 & 1 & \sqrt{2} & \sqrt{2} & \sqrt{2} \\ & 1 & 1 & \sqrt{2} & \sqrt{2} & \sqrt{2} \\ \vdots & & 1 & \sqrt{2} & \sqrt{2} & \sqrt{2} \\ & & & 2 & 2 & 2 \\ & & & & 2 & 2 \\ & \cdots & & & & 2 \end{bmatrix} \tag{3.4}$$

In particular, the isotropic tensor $\mathscr{C}(\lambda, \mu)$ is isomorphic to the matrix

$$\mathbf{C}(\lambda,\mu) = \begin{bmatrix} \lambda+2\mu & \lambda & \lambda & 0 & 0 & 0 \\ & \lambda+2\mu & \lambda & 0 & 0 & 0 \\ \vdots & & \lambda+2\mu & 0 & 0 & 0 \\ & & & 2\mu & 0 & 0 \\ & & & & 2\mu & 0 \\ & \cdots & & & & 2\mu \end{bmatrix} \tag{3.5}$$

Similarly, the fourth-order tensor $\mathscr{H}_{pq}$ is isomorphic to the (generally non-Hermitian) matrix

$$\mathbf{H}_{pq} = \int_{\omega_a}^{\omega_b} \mathbf{g}_p^*(\omega)\, \mathbf{g}_q(\omega)\, d\omega. \tag{3.6}$$

where $\mathbf{G}_p \Leftrightarrow \mathbf{g}_p$, so that equation (2.2) can be expressed

$$U_{pq} = \mathbf{m}^* \cdot \mathbf{H}_{pq} \cdot \mathbf{m}. \tag{3.7}$$

These isomorphisms will be employed throughout the remaining analysis.

4 Statistics of squared-moment estimates

For any seismic source described by a moment-rate vector $\mathbf{m}$, the observed cross-spectra integrals $\hat{U}_{pq}$ differ from their theoretical values U_{pq} given by (3.7) because the former are contaminated by ambient seismic noise and the latter are computed using incorrect transfer functions. We seek an estimate of the squared moment $M^2 = M^2(\lambda, \mu)$ of the form

$$\tilde{M}^2 = \sum_{pq} a_{pq} \hat{U}_{pq}. \tag{4.1}$$

To optimize a scheme for choosing the coefficients $\{a_{pq}\}$ and to assess the uncertainty of the estimate requires that we formulate the estimation error induced by inexact data and faulty seismological assumptions.

Let $\Delta U_{pq}^{(n)}$ be the error in $\hat{U}_{pq}$ due to ambient noise and $\Delta U_{pq}^{(g)}$ be the error due to incorrect transfer functions, so that to first order

$$\tilde{M}^2 = \sum_{pq} a_{pq} (U_{pq} + \Delta U_{pq}^{(n)} + \Delta U_{pq}^{(g)}). \tag{4.2}$$

The estimate $\tilde{M}^2$ differs from its true value M^2 by an amount e that is a sum over these errors plus a term proportional to the difference matrix

$$\mathbf{E} \equiv \tilde{\mathbf{C}} - \mathbf{C} = \sum_{pq} a_{pq} \mathbf{H}_{pq} - \mathbf{C},$$

$$\begin{aligned} e &\equiv \tilde{M}^2 - M^2 \\ &= \sum_{pq} a_{pq} \Delta U_{pq}^{(n)} + \sum_{pq} a_{pq} \Delta U_{pq}^{(g)} + \mathbf{m}^* \cdot \mathbf{E} \cdot \mathbf{m} \\ &\equiv e_{\mathrm{n}} + e_{\mathrm{g}} + e_{\mathrm{m}}. \end{aligned} \tag{4.3}$$

The quantities e_{n} and e_{g} are referred to as the 'ambient-noise error' and 'transfer-function error,' respectively, whereas e_{m} represents the error caused by the failure to solve equation (2.10) exactly and is termed the 'modelling error.'

We shall treat e_n, e_g and e_m as stochastic variables and derive their statistics from simple models of the error processes. The standard statistical treatments of ambient seismic noise are, of course, easily adapted to this problem. Errors in the transfer functions can also be handled by statistical methods if the perturbations due to, say, aspherical heterogeneity are sufficiently randomized over the station network.

The modelling error e_m derives its statistical properties from an incomplete knowledge of the source mechanism. Equation (1.6) allows us to write

$$e_m = 2M_T^2 \hat{\mathbf{m}}^* \cdot \mathbf{E} \cdot \hat{\mathbf{m}} \equiv 2M_T^2 \hat{e}_m, \tag{4.4}$$

where $\hat{\mathbf{m}}$ is the unit vector describing the mechanism. If $\hat{\mathbf{m}}$ were exactly known, the normalized modelling error $\hat{e}_m$ could be calculated directly from the a_{pq}s, and the value of $\hat{M}^2$ could be corrected accordingly. In practice, however, either $\hat{\mathbf{m}}$ is unknown or only an estimate is available, so that $\hat{e}_m$ must also be treated as a stochastic variable.

Our error analysis is simplified by assuming that the total error e is sufficiently well characterized by its first two statistical moments; namely a bias $\beta \equiv \langle e \rangle$ and a variance $\sigma^2 \equiv \mathrm{Var}[e] \equiv \langle (e-\beta)^2 \rangle$. The bias obtained by taking the expected value of equation (4.3) is

$$\beta = \sum_{pq} a_{pq} (B_{pq}^{(n)} + B_{pq}^{(g)}) + 2M_T^2 \langle \hat{e}_m \rangle, \tag{4.5}$$

where $B_{pq}^{(n,g)} \equiv \langle \Delta U_{pq}^{(n,g)} \rangle$. The variance is calculated by assuming the error processes in (4.3) are uncorrelated; its expression in terms of the covariance matrices $V_{pqst}^{(n,g)} \equiv \mathrm{Cov}[\Delta U_{pq}^{(n,g)*} \Delta U_{st}^{(n,g)}]$ is

$$\sigma^2 = \sum_{pqst} a_{pq}^* a_{st} (V_{pqst}^{(n)} + V_{pqst}^{(g)}) + 4M_T^4 \mathrm{Var}[\hat{e}_m]. \tag{4.6}$$

4.1 AMBIENT-NOISE ERROR

We assume the ambient noise $n_p(t)$ contaminating the pth seismogram is Gaussian and white with a zero mean and a variance (power-spectral density) of ν_p^2 and is uncorrelated with the noise at other stations:

$$\langle n_p(t) \rangle = 0, \tag{4.7}$$

$$\langle n_p(t) n_q(t') \rangle = \nu_p^2 \delta_{pq} \delta(t-t'). \tag{4.8}$$

The cross-spectral integrals U_{pq} are estimated from transient signals on a finite time interval, taken to be $[0, T]$. The Fourier transform of $n_p(t)$ over this window,

$$\bar{n}_p(\omega) = \int_0^T n_p(t) \exp(-i\omega t)\, dt, \tag{4.9}$$

defines a random variable for which

$$\langle \bar{n}_p(\omega) \rangle = 0, \tag{4.10}$$

$$\langle \bar{n}_p^*(\omega) \bar{n}_q(\omega') \rangle = \frac{[\![\exp[i(\omega-\omega')T] - 1]\!]}{i(\omega-\omega')} \nu_p^2 \delta_{pq}. \tag{4.11}$$

In terms of this random variable the error in the cross-spectral integral is

$$\Delta U_{pq}^{(n)} = \int_{\omega_a}^{\omega_b} [u_p^*(\omega) \bar{n}_q(\omega) + \bar{n}_p^*(\omega) u_q(\omega) + \bar{n}_p^*(\omega) \bar{n}_q(\omega)]\, d\omega. \tag{4.12}$$

Defining the bandwidth $\Delta\omega = \omega_b - \omega_a$ and taking the expected value of (4.12) we obtain for the bias

$$B_{pq}^{(n)} \equiv \langle \Delta U_{pq}^{(n)} \rangle = \Delta\omega T \nu_p^2 \delta_{pq}. \tag{4.13}$$

It is convenient to introduce the signal-to-noise ratio (snr) for the pth power-spectral integral, defined by

$$R_p = (\Delta\omega T \nu_p^2)^{-1} U_{pp}, \tag{4.14}$$

so that the bias becomes

$$B_{pq}^{(n)} = R_p^{-1} U_{pp} \delta_{pq}. \tag{4.15}$$

The power-spectral integrals thus have a relative bias $R\bar{p}^1$ owing to ambient noise, whereas the cross-spectral integrals are unbiased by this error process.

The formulation of the covariance matrix $V_{pqst}^{(n)}$ is outlined in Appendix A. Rewriting equation (A.6) in terms of the snr yields

$$V_{pqst}^{(n)} = \Delta\omega^{-1}(R_p^{-1} U_{pp} U_{qt} \delta_{ps} + R_q^{-1} U_{qq} U_{ps}^* \delta_{qt} + R_p^{-1} R_q^{-1} U_{pp} U_{qq} \delta_{ps} \delta_{qt}). \tag{4.16}$$

The covariance between $\Delta U_{pq}^{(n)}$ and $\Delta U_{st}^{(n)}$ is identically zero unless they have at least one record in a common position ($p = s$ or $q = t$). The third term in (4.16) is second order in R_p^{-1} and, for snrs typical of IDA accelerograms of large earthquakes (> 15 dB), can be neglected. The relative power-spectral variance is thus $U_{pp}^{-2} V_{pppp} \simeq 2\Delta\omega^{-1} R_p^{-1}$, which will decrease as the integration bandwidth increases. In contrast, if the signals are from geographically separated stations, u_p and u_q will be essentially uncorrelated; therefore, the cross-spectral integral U_{pq} will increase with $\Delta\omega$ at approximately the rate of a random walk process ($|U_{pq}|^2 \sim \Delta\omega^{-1} U_{pp} U_{qq}$), and the relative cross-spectral variance

$$|U_{pq}|^{-2} V_{pqpq}, \qquad p \neq q,$$

will remain roughly constant.

4.2 TRANSFER-FUNCTION ERROR

The transfer functions used in low-frequency source-recovery procedures are calculated from spherically symmetric, non-rotating, isotropic earth models which only approximate the less ideal, real Earth. Errors introduced by these approximations can severely bias moment-tensor inversions and lead to large errors in scalar moment (e.g. Patton & Aki 1980).

Our analysis of this problem is based on the normal-mode representation (1.1). The seismogram for the pth station is

$$u_p(\omega) = \sum_k A_{kp} C_k(\omega), \tag{4.17}$$

where C_k is the resonance spectrum for the kth multiplet and A_{kp} is its complex-valued initial amplitude,

$$A_{kp} \equiv \mathbf{A}_{kp} \cdot \mathbf{m} \equiv \mathscr{A}_k(r_p, r_0) : \mathbf{M}. \tag{4.18}$$

The spherically symmetric earth model used in the calculations has degenerate eigenfrequencies $\pm\omega_k (\omega_k > 0)$ and mode quality factors $Q_k \gg 1$; hence, the kth multiplet is collapsed to a singlet with an attenuation half-width $\alpha_k = \omega_k / 2Q_k$. It is convenient to intro-

duce the complex frequency $\nu_k = \omega_k + i\alpha_k$. In the vicinity of ω_k the resonance spectrum for a record of length T beginning at $t = 0$ is approximately

$$C_k(\omega) = [\![1 - \exp[-i(\omega - \nu_k)T]]\!]/2i(\omega - \nu_k). \tag{4.19}$$

Errors in the transfer functions include those due to mislocation of the source and those due to an incorrect earth model. For large, propagating ruptures the standard location procedures based on high-frequency body-wave arrival times do not necessarily yield the best estimates of the source centroid (Dziewonski *et al.* 1981). The data functionals U_{pq} are independent of shifts in an event's temporal centroid, but they are affected by spatial mislocation. Location errors, particularly in hypocentral depths of shallow earthquakes, alter the relative excitations of the normal modes. The partial derivatives of the transfer functions with respect to the location parameters (Dziewonski *et al.* 1981) could be used to map the statistics of the mislocation error into the bias matrix $B_{pq}^{(g)}$ and the covariance matrix $V_{pqst}^{(g)}$; the formalism for developing such statistics directly from the mode data is implicit in the treatments of Backus (1977) and Dziewonski *et al.* (1981), but too few earthquakes have been studied at low frequencies to establish an empirical basis adequate for this purpose. Hence, in our formal error analysis we shall ignore source mislocation and interpret our error estimates as uncertainties conditional on a particular choice of spatial centroid. A synthetic experiment illustrating the effects of depth mislocation on moment estimation for shallow-focus and deep-focus earthquakes is presented in Section 6.2.

Our analysis of transfer-function error is concentrated on the use of a radial earth model which is incorrect because it does not adequately represent the spherically averaged earth and does not account for the existence of aspherical heterogeneity. Our treatment is approximate. Although the kth multiplet recorded at the pth station may be split by asphericities, we assume it can still be characterized as a singlet with an apparent initial amplitude $\tilde{A}_{kp}$ and an apparent complex frequency $\tilde{\nu}_{kp}$. These mode parameters will generally differ from those predicted by the radial reference model:

$$\tilde{A}_{kp} = A_{kp} + \Delta A_{kp} \tag{4.20}$$

$$\begin{aligned}\tilde{\nu}_{kp} &= \nu_k + \Delta\nu_{kp} \\ &= (\omega_k + \Delta\omega_{kp}) + i(\alpha_k + \Delta\alpha_{kp}).\end{aligned} \tag{4.21}$$

Substitution into equations (4.17) and (4.19) yields

$$\begin{aligned}U_{pq} + \Delta U_{pq}^{(g)} = \sum_{kk'} &(A_{kp} + \Delta A_{kp})^* (A_{k'q} + \Delta A_{k'q}) \\ &\times \int_{\omega_a}^{\omega_b} C_k^*(\omega - \Delta\nu_{kp}^*)\, C_{k'}(\omega - \Delta\nu_{k'q})\, d\omega.\end{aligned} \tag{4.22}$$

The analysis is simplified by several further approximations. In calculating the error $\Delta U_{pq}^{(g)}$ from (4.22) we assume the modes are sufficiently isolated for cross-terms ($k \neq k'$) to be neglected and the mode widths are sufficiently small for the limits of integration to be extended to $\pm\infty$. The necessary integrals can then be evaluated explicitly:

$$\int_{-\infty}^{\infty} C_k^*(\omega - \Delta\nu_{kp})\, C_k(\omega - \Delta\nu_{kq}^*)\, d\omega = \frac{\pi}{4\alpha_k}\left[\frac{1 - \exp[-2\alpha_k T(1 + \epsilon_{kpq})]}{1 + \epsilon_{kpq}}\right]. \tag{4.23}$$

We have here introduced the dimensionless parameter

$$\begin{aligned}\epsilon_{kpq} &\equiv i(\Delta\nu_{kp}^* - \Delta\nu_{kq})/2\alpha_k \\ &= [\Delta\alpha_{kp} + \Delta\alpha_{kq} + i(\Delta\omega_{kp} - \Delta\omega_{kq})]/2\alpha_k.\end{aligned} \tag{4.24}$$

In the short-record limit ($\alpha_k T \ll 1$) the integral (4.23) is approximately $\pi T/2$ and is independent of ϵ_{kpq}. In practice the record length is usually large enough ($\alpha_k T \gtrsim 1$) and ϵ_{kpq} is small enough ($\langle \epsilon_{kpq}^2 \rangle^{1/2} \lesssim 0.5$) so that the dependence of the numerator of (4.23) on ϵ_{kpq} can be ignored. Under these conditions (4.22) reduces to

$$\Delta U_{pq}^{(g)} = \frac{\pi}{4} \sum_{k \in \mathscr{S}} A_{kp}^* A_{kq} \frac{[1 - \exp(-2\alpha_k T)]}{\alpha_k} h_{kpq}, \tag{4.25}$$

where $\mathscr{S}$ is the set of modes contained in the interval (ω_a, ω_b), and

$$h_{kpq} = \frac{(1 + \Delta A_{kp}/A_{kp})^* (1 + \Delta A_{kq}/A_{kq})}{(1 + \epsilon_{kpq})} - 1. \tag{4.26}$$

The statistics of $\Delta U_{pq}^{(g)}$ are therefore derivable from the statistics of the dimensionless quantity h_{kpq}.

The shift in the apparent centre frequency of the multiplet can be expressed as the sum of two terms,

$$\Delta\omega_{kp} = \Delta\bar{\omega}_k + \delta\omega_{kp}. \tag{4.27}$$

The first is the difference between the degenerate eigenfrequency of the spherically averaged earth and that of the radial reference model used to compute the transfer functions: $\Delta\bar{\omega}_k \equiv \bar{\omega}_k - \omega_k$ (quantities pertaining to the spherically averaged earth are indicated by overbars). $\Delta\bar{\omega}_k$ is the same for all stations, so its contributions to ϵ_{kpq} cancel; hence, to this order of approximation errors in the degenerate eigenfrequencies do not contribute to the transfer-function error $\Delta U_{pq}^{(g)}$.

The second term represents the shift in the apparent centre frequency of the multiplet by aspherical heterogeneity. Such shifts have been studied theoretically by Jordan (1978) and Dahlen (1979) and empirically by Silver & Jordan (1981). To first order in the heterogeneity $\delta\omega_{kp}$ equals the multiplet location parameter λ_{kp} defined by Jordan (1978); λ_{kp} is a linear functional of the heterogeneity which tends asymptotically ($l \gg n$) to the local eigenfrequency averaged over the great circle connecting the source and receiver. The diagonal-sum rule of the degenerate perturbation theory (Gilbert 1971b) can be used to show that λ_{kp} has zero mean when averaged over the sphere (Jordan 1980). We therefore take

$$\langle \delta\omega_{kp} \rangle = 0. \tag{4.28}$$

We specify the cross-correlation between any two modes in $\mathscr{S}$ by assuming the frequency shifts are uncorrelated at different stations but strongly correlated at the same station, having Gaussian distributions with variances proportional to α_k^2:

$$\langle \delta\omega_{kp} \, \delta\omega_{k'q} \rangle = \alpha_k \alpha_{k'} \xi_\omega^2 \, \delta_{pq}. \tag{4.29}$$

To assume the locations of nearby multiplets are strongly correlated on the same record is consistent with the predictions of the asymptotic theory (Jordan 1978; Dahlen 1979) and is reasonably well justified by observations (Silver & Jordan 1981). Rms values of $\delta\omega_{kp}/\alpha_k$ derived from fundamental-mode IDA data are shown in Fig. 1; a typical estimate is

$$\xi_\omega = \langle \delta\omega_{kp}^2 \rangle^{1/2}/\alpha_k \simeq 0.5. \tag{4.30}$$

The error in the apparent attenuation parameter of the multiplet can be similarly partitioned,

$$\Delta\alpha_{kp} = \Delta\bar{\alpha}_{kp} + \delta\alpha_{kp}. \tag{4.31}$$

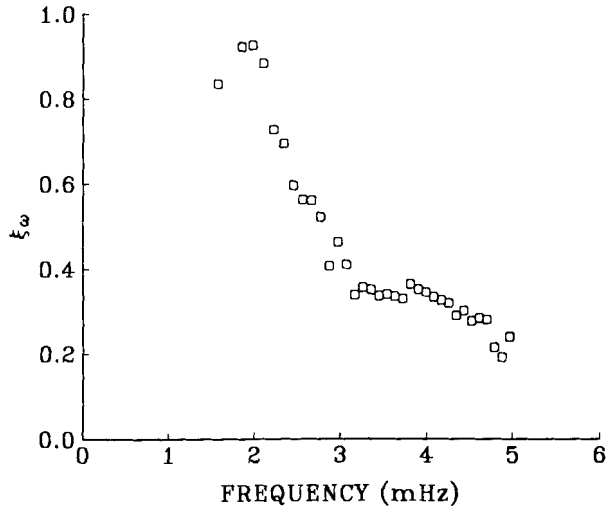

Figure 1. The frequency-shift parameter ξ_ω of equation (4.29) for the fundamental modes ${}_0S_5 - {}_0S_{43}$ derived from the data of Silver & Jordan (1981).

We assume a Gaussian statistical model analogous to that for the frequency shifts:

$$\langle \delta\alpha_{kp} \rangle = 0, \tag{4.32}$$

$$\langle \delta\alpha_{kp} \delta\alpha_{k'p} \rangle = \alpha_k \alpha_{k'} \xi_\alpha^2 \delta_{pq}. \tag{4.33}$$

For the fundamental spheroidal modes Riedesel, Jordan & Masters (1981) estimate

$$\xi_\alpha = \langle \delta\alpha_{kp}^2 \rangle^{1/2} / \alpha_k \simeq 0.2. \tag{4.34}$$

They find that the fluctuations in apparent half-width are not strongly correlated with the multiplet locations, so we assume $\delta\alpha_{kp}$ and $\delta\omega_{kp}$ have independent Gaussian distributions.

Unlike $\Delta\bar{\omega}_k$, $\Delta\bar{\alpha}_k$ does not cancel out of the expression for ϵ_{kpq}; errors in the radial part of the anelastic structure thus contribute to $\Delta U_{pq}^{(g)}$. The expected value of (4.24) is

$$\langle \epsilon_{kpq} \rangle = \Delta\bar{\alpha}_k / \alpha_k. \tag{4.35}$$

To simplify the expressions for $B_{pq}^{(g)}$ and $V_{pqst}^{(g)}$ we henceforth assume the radial reference model is good enough to allow us to ignore this bias; the validity of this assumption for the particular radial attenuation structure used in our calculations will be discussed in Section 6.2. Taking $\Delta\bar{\alpha}_k = 0$ yields

$$\langle \epsilon_{kpq}^* \epsilon_{k'st} \rangle = \tfrac{1}{4}(\xi_\alpha^2 + \xi_\omega^2)(\delta_{ps} + \delta_{qt}) + \tfrac{1}{4}(\xi_\alpha^2 - \xi_\omega^2)(\delta_{pt} + \delta_{qs}). \tag{4.36}$$

To first order the error in the initial amplitude can also be written as the sum of the radial and aspherical parts,

$$\Delta A_{kp} = \Delta\bar{A}_{kp} + \delta A_{kp}. \tag{4.37}$$

Calculations with published earth models suggest that the size of the relative error $\Delta\bar{A}_{kp}/A_{kp}$ is not large; for example, the amplitudes of the fundamental spheroidal modes computed for

sub-crustal earthquakes using models 1066A and 1066B (Gilbert & Dziewonski 1975) generally vary by less than 5 per cent, even though the details of the crustal and upper mantle structures are quite different. We therefore take

$$\Delta\bar{A}_{kp} = 0. \tag{4.38}$$

An expression for the ratio $\delta A_{kp}/A_{kp}$ has been derived by Jordan & Silver (1981) from perturbation theory; to first order it is a sum over mode-mode interactions and, like the multiplet location, is a linear functional of the aspherical heterogeneity. As yet, little theoretical work has been done to quantify the magnitude of the amplitude fluctuations, but some empirical results have been obtained by Riedesel *et al.* (1981), who compared observed initial amplitudes with those predicted from good source models. Well-excited fundamental spheroidal modes show an rms scatter in relative amplitude (observed/synthetic) of ~ 40 per cent. Although larger than perhaps anticipated, these fluctuations are not strongly correlated among nearby modes and are thus reduced by integrating over frequency. A Gaussian model consistent with these results is specified by

$$\langle \delta A_{kp} \rangle = 0 \tag{4.39}$$

$$\langle \delta A^*_{kp} \delta A_{k'q} \rangle = |A_{kp}|^2 \xi_A^2 \, \delta_{pq} \, \delta_{kk'} \tag{4.40}$$

$$\xi_A = \langle |\delta A_{kp}/A_{kp}|^2 \rangle^{1/2} \simeq 0.4. \tag{4.41}$$

The distributions of δA_{kp} and ϵ_{kpq} are taken to be independent.

We expand the denominator of (4.26) in powers of ϵ_{kpq} and apply our statistical assumptions. Truncating terms of higher order than ξ^2 yields the expected value

$$\langle h_{kpq} \rangle \simeq \tfrac{1}{2}(\xi_\alpha^2 - \xi_\omega^2) + [\xi_A^2 + (\xi_\alpha^2 + \xi_\omega^2)] \, \delta_{pq}. \tag{4.42}$$

A consistent expression for the covariance requires keeping terms to $\mathcal{O}(\xi^4)$. The result is

$$\begin{aligned}
\mathrm{Cov}\,[h^*_{kpq} \, h_{k'st}] = {} & \tfrac{1}{8} \xi_\alpha^2 \Delta^{(2+)}_{pqst} [2 + \xi_\alpha^2 (\Delta^{(2+)}_{pqst} + 3\Delta^{(1)}_{pqst} + 6)] \\
& + \tfrac{1}{8} \xi_\omega^2 \Delta^{(2-)}_{pqst} [2 + \xi_\omega^2 (\Delta^{(2-)}_{pqst} + 3\Delta^{(1)}_{pqst} - 6)] \\
& + \tfrac{3}{8} \xi_\alpha^2 \xi_\omega^2 [\Delta^{(2+)}_{pqst} (\Delta^{(1)}_{pqst} - 2) + \Delta^{(2-)}_{pqst} (\Delta^{(1)}_{pqst} + 2)] \\
& + \xi_A^2 \, \delta_{kk'} [\Delta^{(2+)}_{pqst} + \xi_A^2 (\delta_{ps}\delta_{qt} + \delta_{pt}\delta_{qs})] \\
& + \tfrac{1}{4} \xi_\alpha^2 \xi_A^2 \, \delta_{kk'} [2\Delta^{(2+)}_{pqst} (\Delta^{(1)}_{pqst} + 2) + \Delta^{(2+)}_{pqst} (\Delta^{(2+)}_{pqst} + \Delta^{(1)}_{pqst})] \\
& + \tfrac{1}{4} \xi_\omega^2 \xi_A^2 \, \delta_{kk'} [2\Delta^{(2+)}_{pqst} (\Delta^{(1)}_{pqst} - 2) + \Delta^{(2-)}_{pqst} (\Delta^{(2+)}_{pqst} + \Delta^{(1)}_{pqst})],
\end{aligned} \tag{4.43}$$

where $\Delta^{(1)}_{pqst}$ and $\Delta^{(2\pm)}_{pqst}$ are sums of Kronecker deltas:

$$\Delta^{(1)}_{pqst} \equiv \delta_{pq} + \delta_{st} \tag{4.44}$$

$$\Delta^{(2\pm)}_{pqst} \equiv \delta_{ps} \pm \delta_{pt} \pm \delta_{qs} + \delta_{qt} \tag{4.45}$$

Because (4.42) is independent of the mode index k, it factors out of the expected value of equation (4.25), and the bias matrix reduces to

$$B^{(g)}_{pq} = \{\tfrac{1}{2}(\xi_\alpha^2 - \xi_\omega^2) + [\xi_A^2 + \tfrac{1}{2}(\xi_\alpha^2 + \xi_\omega^2)] \, \delta_{pq}\} \, U_{pq}. \tag{4.46}$$

Similarly, the terms in (4.43) independent of k contribute to the covariance matrix in proportion to $U^*_{pq} U_{st}$, whereas those involving $\delta_{kk'}$ are proportional to the spectral integral

$$\mathscr{U}_{pqst} \equiv \int_{\omega_a}^{\omega_b} u_p(\omega) u^*_q(\omega) u^*_s(\omega) u_t(\omega)\, d\omega. \tag{4.47}$$

Substitution gives

$$\begin{aligned} V^{(g)}_{pqst} = {}& \tfrac{1}{8} U^*_{pq} U_{st} \{2\xi^2_\alpha \Delta^{(2+)}_{pqst} + 2\xi^2_\omega \Delta^{(2-)}_{pqst} + \xi^4_\alpha \Delta^{(2+)}_{pqst}(\Delta^{(2+)}_{pqst} + 3\Delta^{(1)}_{pqst} + 6) \\ & + \xi^4_\omega \Delta^{(2-)}_{pqst}(\Delta^{(2-)}_{pqst} + 3\Delta^{(1)}_{pqst} - 6) + 3\xi^2_\alpha \xi^2_\omega [\Delta^{(2+)}_{pqst}(\Delta^{(1)}_{pqst} - 2) \\ & + \Delta^{(2-)}_{pqst}(\Delta^{(1)}_{pqst} + 2)]\} + \xi^2_A \mathscr{U}_{pqst} \{\Delta^{(2+)}_{pqst} + \xi^2_A (\delta_{ps}\delta_{qt} + \delta_{pt}\delta_{qs}) \\ & + \tfrac{1}{4}\xi^2_\alpha [2\Delta^{(2+)}_{pqst}(\Delta^{(1)}_{pqst} + 2) + \Delta^{(2+)}_{pqst}(\Delta^{(2+)}_{pqst} + \Delta^{(1)}_{pqst})] \\ & + \tfrac{1}{4}\xi^2_\omega [2\Delta^{(2+)}_{pqst}(\Delta^{(1)}_{pqst} - 2) + \Delta^{(2-)}_{pqst}(\Delta^{(2+)}_{pqst} + \Delta^{(1)}_{pqst})]\} . \end{aligned} \tag{4.48}$$

In the special case of power spectra ($p = q$, $s = t$), all terms proportional to ξ^2_ω drop out of equations (4.46) and (4.48), as expected, since the power-spectral integrals are insensitive to errors in the multiplet centre frequencies. Such variations introduce a negative bias to the cross-spectral integrals (the coherence of the observed time series is decreased by aspherical heterogeneity), which competes against a positive bias caused by variations in apparent attenuation. According to this model, amplitude fluctuations contribute a positive bias to the power spectra but none to the cross spectra, a behaviour identical to ambient noise (cf. equation 4.15).

To $\mathcal{O}(\xi^2)$ the 'diagonal' elements of the covariance matrix ($p = s, q = t$) are

$$V^{(g)}_{pqpq} = \tfrac{1}{2} |U_{pq}|^2 [\xi^2_\alpha (1 + \delta_{pq}) + \xi^2_\omega (1 - \delta_{pq})] + 2\mathscr{U}_{pqpq} \xi^2_A (1 + \delta_{pq}). \tag{4.49}$$

As the bandwidth of integration increases $|U_{pq}|^2$ grows more rapidly than $\mathscr{U}_{pqpq}$, and the amplitude fluctuations contribute less to the relative variance of $\tilde{U}_{pq}$. This is a consequence of presuming the variations in $\tilde{\alpha}_{kp}$ and $\tilde{\omega}_{kp}$ are correlated for all modes in this band at a given station but the amplitude fluctuations are not.

4.3 MODELLING ERROR

The normalized modelling error $\hat{e}_m$ in equation (4.4) is defined in terms of the source mechanism $\hat{\mathbf{m}}$ and the error matrix $\mathbf{E} = \tilde{\mathbf{C}} - \mathbf{C}$ by the quadratic form

$$\hat{e}_m = \hat{\mathbf{m}}^* \cdot \mathbf{E} \cdot \hat{\mathbf{m}} = (\hat{\mathbf{m}}^* \hat{\mathbf{m}}) : \mathbf{E}. \tag{4.50}$$

To derive the statistics of $\hat{e}_m$ we impose as a probability distribution on $\hat{\mathbf{m}}$ an appropriate likelihood function constructed from prior information about the source; i.e. we consider $\hat{\mathbf{m}}$ to be a stochastic process from which the actual (unknown) mechanism is a sample. The bias and variance of $\hat{e}_m$ are completely specified by the second- and fourth-order tensors

$$\boldsymbol{\mu}^{(2)} = \langle \hat{\mathbf{m}}^* \hat{\mathbf{m}} \rangle, \tag{4.51}$$

$$\boldsymbol{\mu}^{(4)} = \langle \hat{\mathbf{m}} \hat{\mathbf{m}}^* \hat{\mathbf{m}}^* \hat{\mathbf{m}} \rangle. \tag{4.52}$$

The expressions are

$$\hat{\beta}_m \equiv \langle \hat{e}_m \rangle = \boldsymbol{\mu}^{(2)} : \mathbf{E} \tag{4.53}$$

$$\begin{aligned} \hat{\sigma}^2_m &\equiv \mathrm{Var}[\hat{e}_m] \\ &= \mathbf{E}^* : (\boldsymbol{\mu}^{(4)} - \boldsymbol{\mu}^{(2)*} \boldsymbol{\mu}^{(2)}) : \mathbf{E}. \end{aligned} \tag{4.54}$$

Let us first consider the important special case of a synchronous source (equation 2.13), where, with no loss of generality, $\hat{\mathbf{m}}$ can be taken to be purely real. A useful and simple probability density function for the orientation of a real unit vector of dimension n is the hyperspherical normal distribution

$$S_n(\hat{\mathbf{m}}) = C_n^{-1}(\kappa) \exp(\kappa \hat{\mathbf{m}}_0 \cdot \hat{\mathbf{m}}). \tag{4.55}$$

This distribution is parametrized by a most probable unit vector $\hat{\mathbf{m}}_0$ and a dimensionless scale factor $\kappa \geqslant 0$, whose inverse is a measure of the dispersion. Large κ corresponds to a distribution of unit vectors tightly clustered about $\hat{\mathbf{m}}_0$, whereas $\kappa = 0$ corresponds to a uniform distribution on the n-dimensional unit hypersphere Ω_n. $C_n(\kappa)$ is the normalization integral taken over this domain,

$$C_n(\kappa) = \int_{\Omega_n} \exp(\kappa \hat{\mathbf{m}}_0 \cdot \hat{\mathbf{m}})\, d\Omega_n(\hat{\mathbf{m}}). \tag{4.56}$$

$S_2(\hat{\mathbf{m}})$ is the circular normal, or von Mises, distribution (von Mises 1918), and $S_3(\hat{\mathbf{m}})$ is the spherical normal, or Fisher, distribution (Fisher 1953); the latter is commonly employed in the statistical manipulation of palaeomagnetic pole positions (e.g. Watson 1966). The general hyperspherical case ($n > 3$) needed for our purposes has evidently received less mathematical attention than either of these special cases (Johnson & Kotz 1970, for example, give an incorrect form for the probability density function); the low-order moments of the hyperspherical normal distribution are therefore derived in Appendix B.

The results of Appendix B lead to the following expressions for the bias and variance of $\hat{e}_m$ for synchronous sources:

$$\hat{\beta}_m = f_1(\kappa)\, \mathrm{tr}\, \mathbf{E}_R + f_2(\kappa)\, \hat{\mathbf{m}}_0 \cdot \mathbf{E}_R \cdot \hat{\mathbf{m}}_0, \tag{4.57}$$

$$\begin{aligned}\hat{\sigma}_m^2 = {} & f_3(\kappa)\, (\mathrm{tr}\, \mathbf{E}_R)^2 + f_4(\kappa)\, \mathbf{E}_R : \mathbf{E}_R + f_5(\kappa)\, (\mathrm{tr}\, \mathbf{E}_R)\, (\hat{\mathbf{m}}_0 \cdot \mathbf{E}_R \cdot \hat{\mathbf{m}}_0) \\ & + f_6(\kappa)\, \| \mathbf{E}_R \cdot \hat{\mathbf{m}}_0 \|^2 + f_7(\kappa)\, (\hat{\mathbf{m}}_0 \cdot \mathbf{E}_R \cdot \hat{\mathbf{m}}_0)^2,\end{aligned} \tag{4.58}$$

where $\mathbf{E}_R \equiv \mathrm{Re}\, \mathbf{E}$. The functions $\{f_i(\kappa) : i = 1, 2, \ldots, 7\}$ are given in Table 1 in terms of the coefficients $a_\nu^{(q)}(\kappa)$ of equation (B8), together with explicit expressions for the special cases of a uniform distribution ($\kappa = 0$) and the Gaussian limit ($\kappa \to \infty$).

To generalize equations (4.57) and (4.58) to asynchronous sources we use an isomorphism between 6-dimensional complex unit vectors $\hat{\mathbf{m}} \equiv \mathbf{m}_R + i\, \mathbf{m}_I$ and 12-dimensional real unit vectors $\hat{\mathbf{m}}' \equiv [\mathbf{m}_R \mid \mathbf{m}_I]$. Since $\mathbf{E} \equiv \mathbf{E}_R + i\, \mathbf{E}_I$ is Hermitian ($\mathbf{E}_R^T = \mathbf{E}_R$, $\mathbf{E}_I^T = -\mathbf{E}_I$), the normalized modelling error can be written $\hat{e}_m = \hat{\mathbf{m}}' \cdot \mathbf{E}' \cdot \hat{\mathbf{m}}'$, where

$$\mathbf{E}' = \left[\begin{array}{c|c} \mathbf{E}_R & -\mathbf{E}_I \\ \hline \mathbf{E}_I & \mathbf{E}_R \end{array}\right]. \tag{4.59}$$

Hence, the appropriate probability density function is $S_{12}(\hat{\mathbf{m}}')$, and the moments of $\hat{e}_m$ can be computed by applying the formulae of Appendix B to the case $n = 12$. Recast in terms of the complex-valued quantities $\hat{\mathbf{m}}_0$ and $\mathbf{E}$, the results for general asynchronous sources are

$$\hat{\beta}_m = f_1(\kappa)\, \mathrm{tr}\, \mathbf{E} + f_2(\kappa)\, \hat{\mathbf{m}}_0^* \cdot \mathbf{E} \cdot \hat{\mathbf{m}}_0, \tag{4.60}$$

$$\begin{aligned}\hat{\sigma}_m^2 = {} & f_3(\kappa)\, (\mathrm{tr}\, \mathbf{E})^2 + f_4(\kappa)\, \mathbf{E}^* : \mathbf{E} \\ & + f_5(\kappa)\, (\mathrm{tr}\, \mathbf{E})\, (\hat{\mathbf{m}}_0^* \cdot \mathbf{E} \cdot \hat{\mathbf{m}}_0) \\ & + f_6(\kappa)\, \| \mathbf{E} \cdot \hat{\mathbf{m}}_0 \|^2 + f_7(\kappa)\, (\hat{\mathbf{m}}_0^* \cdot \mathbf{E} \cdot \hat{\mathbf{m}}_0)^2,\end{aligned} \tag{4.61}$$

where the functions $\{f_i(\kappa) : i = 1, 2, \ldots, 7\}$ are again given in Table 1.

Table 1. Expressions for $f_i(\kappa)$.

	Synchronous sources*			Asynchronous sources†		
i	General $0 \leqslant \kappa \leqslant \infty$	Uniform $\kappa = 0$	Gaussian limit $\kappa \gg 1$	General $0 \leqslant \kappa \leqslant \infty$	Uniform $\kappa = 0$	Gaussian limit $\kappa \gg 1$
1	$a_2^{(1)}$	1/6	κ^{-1}	$2a_5^{(1)}$	1/6	$2\kappa^{-1}$
2	$\kappa^2 a_2^{(3)}$	0	$1-\kappa^{-1}$	$\kappa^2 a_5^{(2)}$	0	$1-\kappa^{-1}$
3	$a_2^{(2)}-(a_2^{(1)})^2$	$-1/144$	0	$4[a_5^{(2)}-(a_5^{(1)})^2]$	$-1/252$	0
4	$2a_2^{(2)}$	1/24	$2\kappa^{-2}$	$4a_5^{(2)}$	1/42	$4\kappa^{-2}$
5	$2\kappa^2(a_2^{(3)}-a_2^{(1)}a_2^{(2)})$	0	0	$4\kappa^2(a_5^{(3)}-a_5^{(1)}a_5^{(2)})$	0	0
6	$4\kappa^2 a_2^{(3)}$	0	$4\kappa^{-1}-2\kappa^{-2}$	$4\kappa^2 a_5^{(3)}$	0	$4\kappa^{-1}-2\kappa^{-2}$
7	$\kappa^4[a_2^{(4)}-(a_2^{(2)})^2]$	0	$-4\kappa^{-1}$	$\kappa^4[a_5^{(4)}-(a_5^{(2)})^2]$	0	$-4\kappa^{-1}$

*Pertain to equations (4.57) and (4.58).
†Pertain to equations (4.61) and (4.62).

The choice $\kappa = 0$ corresponds to the interesting special case of a uniform distribution on the hypersphere of possible source mechanisms. All mechanisms have equal probability density and no *a priori* information is imposed on the nature of the source. Equations (4.60) and (4.61) reduce to the intuitively satisfying expressions

$$\hat{\beta}_m = \tfrac{1}{6}\,\mathrm{tr}\,\mathbf{E} \qquad (\kappa = 0), \tag{4.62}$$

$$\hat{\sigma}_m^2 = \tfrac{1}{42}\,[\mathbf{E}^* : \mathbf{E} - \tfrac{1}{6}(\mathrm{tr}\,\mathbf{E})^2] = \tfrac{1}{42}\,\mathbf{E}_D^* : \mathbf{E}_D \qquad (\kappa = 0), \tag{4.63}$$

where $\mathbf{E}_D \equiv \mathbf{E} - \frac{1}{6}(\mathrm{tr}\,\mathbf{E})\mathbf{I}$ is the deviatoric part of $\mathbf{E}$. If we want to require the source to be synchronous but impose no other constraints, we can substitute $\kappa = 0$ into equations (4.57) and (4.58); the results are similar to those above, except $\mathbf{E}$ and $\mathbf{E}^*$ are replaced by $\mathbf{E}_R$ and the geometrical factor 1/42 is replaced by 1/24.

The distributions of $\hat{\mathbf{m}}$ discussed thus far are isotropic about a most probable mechanism $\hat{\mathbf{m}}_0$; they provide no means of asserting that some components of $\hat{\mathbf{m}}_0$, or combinations of components, are better known than others, as is often the case in seismology. The generalization of the complete hyperspherical normal distribution to include anisotropic dispersion involves certain mathematical difficulties we wish to avoid here, but it is straightforward in the Gaussian limit of small dispersion (large κ). As discussed in Appendix B, $S_n(\hat{\mathbf{m}})$ is asymptotically Gaussian with a mean $\hat{\mathbf{m}}_0$ and a variance matrix

$$\hat{\mathbf{V}}_m = \kappa^{-1}(\mathbf{I} - \hat{\mathbf{m}}_0\hat{\mathbf{m}}_0). \tag{4.64}$$

Table 1 shows that, for synchronous sources, equations (4.57) and (4.58) reduce in this limit to

$$\hat{\beta}_m = \hat{\mathbf{V}}_m : \mathbf{E}_R + \hat{\mathbf{m}}_0 \cdot \mathbf{E}_R \cdot \hat{\mathbf{m}}_0, \tag{4.65}$$

$$\hat{\sigma}_m^2 = 4\hat{\mathbf{m}}_0 \cdot \mathbf{E}_R \cdot \hat{\mathbf{V}}_m \cdot \mathbf{E}_R \cdot \hat{\mathbf{m}}_0 + 2\,\|\,\hat{\mathbf{V}}_m \cdot \mathbf{E}_R\,\|^2. \tag{4.66}$$

To include anisotropic dispersion all we need do is generalize the variance matrix $\hat{\mathbf{V}}_m$ appropriately. Suppose, for example, a symmetrical, real, but possibly non-diagonal, variance

matrix $\mathbf{V}_m$ was obtained from some estimation procedure yielding the synchronous source $\mathbf{m}_0$, such as the direct inversion of equation (1.3). Then, the appropriate variance matrix for use in (4.65) and (4.66) would be computed by projecting and normalizing $\mathbf{V}_m$ thus:

$$\hat{\mathbf{V}}_m = \|\mathbf{m}_0\|^{-2}(\mathbf{I} - \hat{\mathbf{m}}_0\hat{\mathbf{m}}_0)\cdot\mathbf{V}_m\cdot(\mathbf{I} - \hat{\mathbf{m}}_0\hat{\mathbf{m}}_0). \tag{4.67}$$

The extension of this Gaussian generalization to asynchronous sources is most easily accomplished by applying the isomorphism used to derive (4.60) and (4.61); equations (4.64)–(4.67) retain the same form except all vectors and matrices are 12-dimensional rather than 6.

As a final topic of this section we consider appropriate forms for the modelling errors in the special case when the source mechanism is presumed to be purely deviatoric, i.e. when both the best estimate $\hat{\mathbf{m}}_0$ and its associated errors are perpendicular to any purely isotropic mechanism of the form $\hat{\mathbf{m}}_I \exp(i\phi_I)$, where

$$\hat{\mathbf{m}}_I = (3)^{-1/2}[1\ \ 1\ \ 1\ \ 0\ \ 0\ \ 0], \tag{4.68}$$

and ϕ_I is an arbitrary phase. The domain over which a likelihood function must be specified is the intersection of the unit hypersphere with the hyperplane defined by $\hat{\mathbf{m}}_I$. All of the equations given in this section for $\hat{\beta}_m$ and $\hat{\sigma}_m$ retain the same form except that the error matrix $\mathbf{E}$ is replaced by its projection

$$(\mathbf{I} - \hat{\mathbf{m}}_I\hat{\mathbf{m}}_I)\cdot\mathbf{E}\cdot(\mathbf{I} - \hat{\mathbf{m}}_I\hat{\mathbf{m}}_I). \tag{4.69}$$

The dimension of the manifold, n, is reduced from 6 to 5 for synchronous sources and from 12 to 10 for asynchronous sources; hence, in the specification of the functions $f_i(\kappa)$, $i = 1, 2, \ldots, 7$, the coefficients $a_2^{(q)}$ of Table 1 are replaced by $a_{3/2}^{(q)}$, and the coefficients $a_5^{(q)}$ are replaced by $a_4^{(q)}$.

We shall make use of the likelihood functions for purely deviatoric mechanisms in our algorithm for detecting the existence of an isotropic part (Section 5.2).

5 Optimization algorithms

The statistical models developed in Section 4 permit, for any particular set of coefficients $\{a_{pq}\}$, the estimation of the bias and variance induced in $\tilde{M}^2$ by ambient seismic noise, transfer-function error, and the failure to solve equation (2.10) exactly ('modelling error'). We shall optimize the calculation of the coefficients $\{a_{pq}\}$ by minimizing certain linear combinations of non-negative-definite quadratic forms involving these statistics. Our optimization algorithms thus belong to the general class of 'least-squares' estimation procedures. Characteristic of these procedures is the existence of trade-off curves between two competing measures of estimation error (Backus & Gilbert 1970).

Consider, for example, the trade-off curve between the squared bias β^2 (equation 4.5) and the variance σ^2 (equation 4.6) generated on the locus $\theta \in (0, \pi/2)$ by minimizing the positive-definite quadratic form $\beta^2\cos\theta + \sigma^2\sin\theta$ with respect to a variation of the a_{pq}s. As $\theta \to 0$, a solution (or solutions) is approached which minimizes σ^2 subject to the constraint that β^2 (and therefore β) be zero. As $\theta \to \pi/2$, σ^2 decreases to zero and β approaches $-M^2$, corresponding to the solution $a_{pq} = 0$. The intermediate point on this trade-off curve ($\theta = \pi/4$) yields the solution(s) minimizing

$$\epsilon^2 = \beta^2 + \sigma^2. \tag{5.1}$$

By definition ϵ^2 equals $\langle e^2\rangle$, the so-called 'mean squared error', whose minimization is a common criterion for optimality in power spectral estimation (Jenkins & Watts 1968, p. 247). We adopt this criterion in our optimization scheme.

Notation can be simplified by defining the 'total observational error' $e_O = e_n + e_g$. The mean and variance of this random variable are

$$\beta_O \equiv \sum_{pq} a_{pq} B_{pq}^{(O)} = \sum_{pq} a_{pq} (B_{pq}^{(n)} + B_{pq}^{(g)}), \tag{5.2}$$

$$\sigma_O^2 \equiv \sum_{pqst} a_{pq}^* a_{st} V_{pqst}^{(O)} = \sum_{pqst} a_{pq}^* a_{st} (V_{pqst}^{(n)} + V_{pqst}^{(g)}). \tag{5.3}$$

Of course, for any specified set of coefficients $\{a_{pq}\}$ we have available only estimates of β_O and σ_O^2, because quantities such as U_{pq} appearing in the expressions for the bias and variance matrices (equations 4.15, 4.16, 4.46 and 4.48) must be approximated by their observed values $\tilde{U}_{pq}$. Statistics obtained by these substitutions are indicated by superposed tildes, e.g. $\tilde{\beta}_O$ and $\tilde{\sigma}_O^2$.

5.1 ESTIMATION OF M_T^2

An appropriate measure of the error induced in an estimate of total squared moment

$$\tilde{M}_T^2 = \sum_{pq} a_{pq} \tilde{U}_{pq}$$

by ambient seismic noise and aspherical heterogeneity is the mean squared observational error $\tilde{\epsilon}_O^2 = \tilde{\beta}_O^2 + \tilde{\sigma}_O^2$. A similarly appropriate measure of the failure to solve exactly the constraint equations

$$\sum_{pq} a_{pq} \mathbf{H}_{pq} = \mathbf{C}(0, \tfrac{1}{4}) \equiv \mathbf{C}_T$$

is the normalized mean squared modelling error $\hat{\epsilon}_m^2 = \hat{\beta}_m^2 + \hat{\sigma}_m^2$. The former is an absolute error with dimensions of M_T^4, whereas the latter is a dimensionless relative error.

To optimize the estimation of M_T^2 we consider the solutions lying on the trade-off curve between $\tilde{\epsilon}_O^2$ and $\hat{\epsilon}_m^2$, i.e. the sets of coefficients $\{a_{pq}\}$ minimizing the function

$$f(\theta) = \epsilon_O^2 + w \hat{\epsilon}_m^2 \tan\theta, \qquad \theta \in (0, \pi/2). \tag{5.4}$$

w is a constant with dimensions of M_T^4 which scales the two types of errors and renders (5.4) dimensionally homogeneous; a convenient choice is discussed below.

Minimizing (5.4) with respect to the a_{pq}s leads to a system of normal equations

$$\sum_{st} \Gamma_{pqst} a_{st} = d_{pq}, \tag{5.5}$$

where

$$\Gamma_{pqst} = \mathbf{H}_{pq}^* : \boldsymbol{\mu}^{(4)} : \mathbf{H}_{st} + w^{-1} \cot\theta \, (\tilde{V}_{pqst}^{(O)} + \tilde{B}_{pq}^{(O)*} \tilde{B}_{st}^{(O)}), \tag{5.6}$$

$$d_{pq} = \mathbf{H}_{pq}^* : \boldsymbol{\mu}^{(4)} : \mathbf{C}_T, \tag{5.7}$$

and $\boldsymbol{\mu}^{(4)}$ is the fourth-order tensor defined by (4.52). If the indices in (5.5) are permitted to range over an entire set of P records, then (5.5) represents a system of P^2 complex-valued equations for the P^2 complex-valued unknowns $\{a_{pq}\}$. However, because the matrices a_{pq} and d_{pq} are Hermitian and $\mathbf{H}_{pq}$ obeys the symmetries embodied in equation (2.11), the use of appropriate isomorphisms allows this system to be transformed into a square system of P^2 real-valued equations, considerably reducing the computational labour required for its solution.

The choice of the scaling constant w is formally arbitrary, since it only governs the parametrization of the trade-off curve by θ (Backus & Gilbert 1970, p. 145), but inspection of equations (4.3) and (4.4) shows that the two types of error are commensurate for $w = 4M_T^4$. Selecting w to be near this value insures the 'knee' of the trade-off curve will be in the vicinity of $\theta = \pi/4$. This intermediate point then represents a nearly optimal compromise between the extremal solutions at $\theta = 0$ ($a_{pq} = 0$, $\tilde{\epsilon}_O^2 = 0$) and $\theta = \pi/2$ (minimum of $\tilde{\epsilon}_m^2$).

To implement the opimization algorithm we make a rough estimate of M_T to fix w and calculate the coefficients $\{a_{pq}\}$ for $\theta = \pi/4$ by solving (5.5). If the system is singular or ill-conditioned, which is commonly the case if the number of record pairs is large, the inversion is accomplished by a regularization procedure which adds a small, positive constant to the 'diagonal' matrix components Γ_{pqpq}. The estimate $\tilde{M}_T^2$ obtained from equation (4.1) is corrected for the relative modelling bias $2\hat{\beta}_m$ and the absolute observational bias β_O and is assigned a relative modelling variance $4\hat{\sigma}_m^2$ and an absolute observational variance $\hat{\sigma}_O^2$. The results of the optimization algorithm can thus be written

$$\tilde{M}_T^2 = (1 - 2\hat{\beta}_m) \sum_{pq} a_{pq} \tilde{U}_{pq} - \tilde{\beta}_O, \tag{5.8}$$

$$\mathrm{Var}[\tilde{M}_T^2] = 4\tilde{M}_T^4 \hat{\sigma}_m^2 + \tilde{\sigma}_O^2. \tag{5.9}$$

Seismologists prefer to speak in terms of moment, rather than its square. Since the square-root of (5.9) is typically small compared to $\tilde{M}_T^2$, we can characterize the error in $\tilde{M}_T$ by a relative standard deviation whose square is

$$\rho_T^2 \equiv \tilde{M}_T^{-2} \mathrm{Var}[\tilde{M}_T] \simeq \tfrac{1}{4} \tilde{M}_T^{-4} \mathrm{Var}[\tilde{M}_T^2]. \tag{5.10}$$

Defining the normalized observation errors by

$$\hat{\beta}_O = \tilde{\beta}_O / 2\tilde{M}_T^2, \tag{5.11}$$

$$\hat{\sigma}_O = \tilde{\sigma}_O / 2\tilde{M}_T^2, \tag{5.12}$$

allows us to write the relative bias correction in $\tilde{M}_T$ as $\hat{\beta}_m + \hat{\beta}_O$ and the square of its relative standard deviation as

$$\rho_T^2 = \hat{\sigma}_m^2 + \hat{\sigma}_O^2. \tag{5.13}$$

5.2 DETECTION OF AN ISOTROPIC PART

One of the primary motivations for this work is the need for better algorithms to search for an isotropic component of the seismic moment tensor. This detection problem is best set up as a test of the null hypothesis that the moment tensor is purely deviatoric, i.e. $M_T = M_D$ and $M_I = 0$. A detection is deemed successful when the null hypothesis has been rejected at some pre-assigned confidence level.

We begin by adopting the null hypothesis in our specification of the likelihood function for the source mechanism; $\hat{\mathbf{m}}_0$ is chosen as the best deviatoric estimate of the mechanism, and the tensors $\boldsymbol{\mu}^{(2)}$ and $\boldsymbol{\mu}^{(4)}$ are constructed from the assumption that errors in $\hat{\mathbf{m}}_0$ are perpendicular to the isotropic mechanism $\hat{\mathbf{m}}_I$. This can be accomplished by the methods outlined at the end of Section 4.3.

Such a specification has the property that $\boldsymbol{\mu}^{(4)} : \mathbf{C}_I = 0$, where $\mathbf{C}_I \equiv \mathbf{C}(\tfrac{1}{6}, 0)$, so that minimizing the function $f(\theta)$ of (5.4) yields a system of normal equations with $d_{pq} = 0$. The

trade-off curve obtained by the regularized inversion of (5.5) thus collapses to the trivial solution $a_{pq} = 0$.

This behaviour is a consequence of imposing the deviatoric hypothesis we seek to test. As formulated, $\hat{\epsilon}_m^2$, and hence $f(\theta)$, is independent of the quantity

$$\hat{\beta}_I = \hat{\mathbf{m}}_I \cdot \mathbf{E} \cdot \hat{\mathbf{m}}_I$$
$$= 2\tilde{\mathbf{C}} : \mathbf{C}_I - 1/2. \tag{5.14}$$

$2\hat{\beta}_I$ is just the bias in M_I^2 due to the isotropic component of the modelling error.

In our detection algorithm the estimate $\tilde{M}_I^2$ is obtained by minimizing $f(\theta)$ subject to the constraint $\hat{\beta}_I = 0$. As an optimal point on the trade-off curve we select $w \tan\theta = 4\tilde{M}_T^4$, where $\tilde{M}_T$ is the estimate of total moment given by equation (5.8). The resulting normal equations have the form of (5.5), where

$$\Gamma_{pqst} = \mathbf{H}_{pq}^* : \boldsymbol{\mu}^{(4)} : \mathbf{H}_{st} + \tfrac{1}{4}\tilde{M}_T^{-4}(\tilde{V}_{pqst}^{(O)} + \tilde{B}_{pq}^{(O)*}\tilde{B}_{st}^{(O)}) \tag{5.15}$$

$$d_{pq} = \lambda\, \mathbf{H}_{pq}^* : \mathbf{C}_I, \tag{5.16}$$

and λ is a Lagrange multiplier. λ is chosen such that $\hat{\beta}_I = 0$ or, equivalently,

$$\sum_{pq} a_{pq}\, \mathbf{H}_{pq} : \mathbf{C}_I = 1/4. \tag{5.17}$$

Again, this system of complex-valued equations can be reduced to an equivalent system of real-valued equations by employing appropriate isomorphisms.

The estimate of the squared isotropic moment obtained by this procedure is

$$\tilde{M}_I^2 = \sum_{pq} a_{pq}\, \tilde{U}_{pq} - 2\tilde{M}_T^2\, \hat{\beta}_m - \tilde{\beta}_O, \tag{5.18}$$

$$\mathrm{Var}[\tilde{M}_I^2] = 4\tilde{M}_T^4\, \hat{\sigma}_m^2 + \tilde{\sigma}_O^2. \tag{5.19}$$

By assuming the distribution of errors in $\tilde{M}_I^2$ is adequately modelled by a Gaussian density function with zero mean and a variance given by (5.19), the null hypothesis $M_I^2 = 0$ can be tested against the alternative $M_I^2 > 0$ at the $(1-\alpha)$ per cent confidence level according to standard procedures. The null hypothesis is rejected if

$$\tilde{M}_I^2 > \kappa_\alpha (4\tilde{M}_T^4\, \hat{\sigma}_m^2 + \tilde{\sigma}_O^2)^{1/2}, \tag{5.20}$$

where κ_α is the appropriate critical value for a one-sided tests, e.g. at the 95 per cent confidence level, $\alpha = 0.05$ and $\kappa_\alpha = 1.64$.

6 Applications to real and synthetic data

The techniques of scalar-moment estimation formulated in this paper have been applied to two well-studied events, the deep-focus Honshu earthquake of 1978 March 7 (Dziewonski *et al.* 1981; Masters & Gilbert 1981, private communication) and the shallow-focus Oaxaca

Table 2. Earthquakes studied.

Location	Date	Origin time (UT)	Latitude (deg)	Longitude (deg)	Depth (km)	m_b	M_s
Honshu, Japan	1978 Mar. 7	2:48:47.6	32.005 N	137.609 E	439	6.9	–
Oaxaca, Mexico	1978 Nov. 29	19:52:47.6	16.010 N	96.591 W	18	6.4	7.7

Table 3. Source mechanisms and variance matrices for the Honshu and Oaxaca events.

Honshu (after Masters & Gilbert 1981, private communication)

	$\hat{\mathbf{m}}_0$	$\hat{\mathbf{V}}_\mathbf{m} \times 10$					
		m_{rr}	$m_{\theta\theta}$	$m_{\phi\phi}$	$m_{r\theta}$	$m_{r\phi}$	$m_{\theta\phi}$
m_{rr}	0.488	0.115	−0.122	0.007	−0.056	0.000	0.206
$m_{\theta\theta}$	0.053		0.782	−0.660	0.116	0.677	0.088
$m_{\phi\phi}$	−0.541	⋮		0.654	−0.060	−0.676	−0.294
$m_{r\theta}$	0.384				0.061	0.073	−0.042
$m_{r\phi}$	−0.551					0.718	0.281
$m_{\theta\phi}$	−0.121		...				0.770

Oaxaca (after Stewart *et al.* 1981)

	$\hat{\mathbf{m}}_0$	$\hat{\mathbf{V}}_\mathbf{m} \times 10$					
		m_{rr}	$m_{\theta\theta}$	$m_{\phi\phi}$	$m_{r\theta}$	$m_{r\phi}$	$m_{\theta\phi}$
m_{rr}	0.262	0.565	−0.308	−0.256	−0.269	0.150	0.356
$m_{\theta\theta}$	−0.260		0.825	−0.517	0.337	−0.075	0.066
$m_{\phi\phi}$	−0.002	⋮		0.773	−0.068	−0.074	−0.442
$m_{r\theta}$	0.717				0.347	0.118	−0.164
$m_{r\phi}$	−0.574					0.262	0.049
$m_{\theta\phi}$	0.141		...				0.492

earthquake of 1978 November 29 (Ward 1980b; Stewart, Chael & McNally 1981; Reichle, Orcutt & Priestly 1982). The Preliminary Determination of Epicentre (PDE) bulletins of the United States Geological Survey (USGS) assign a body-wave magnitude of 6.9 to the former and a surface-wave magnitude of 7.7 to the latter. Estimates of the location parameters and source mechanisms are listed in Tables 2 and 3.

Both events were recorded on the vertical-component accelerometers at nine stations of the IDA network. The data $\{\tilde{U}_{pq}; p, q = 1, 2, \ldots, 9\}$ were obtained for each record pair by fast-Fourier transforming Hanning windowed seismograms of 10-hr lengths and integrating the cross-spectra over 1 millihertz (mHz) bands. For each station estimates of the snr were made in each band by integrating the spectra of noise samples taken prior to the events.

The bias matrices $B^{(n)}_{pq}$ and $B^{(g)}_{pq}$ and covariance matrices $V^{(n)}_{pqst}$ and $V^{(g)}_{pqst}$ were computed from equations (4.15), (4.46), (4.16) and (4.48), respectively, using these snr estimates and the values of the heterogeneity parameters discussed in Section 4.2 ($\xi_\omega = 0.5$, $\xi_\alpha = 0.2$, $\xi_A = 0.4$).

Time-domain transfer functions were generated from the 1066A model of Gilbert & Dziewonski (1975), the recently derived Q model of Masters & Gilbert (1982), and the location parameters of Table 2. These time series were processed exactly as the observed seismograms to yield the matrices $\mathbf{H}_{pq}$.

6.1 EFFECTS OF CONSTRAINTS ON THE SOURCE MECHANISM

The estimates of scalar seismic moment can be constrained by imposing a likelihood function on the source mechanism $\hat{\mathbf{m}}$ of the sort discussed in Section 4.3. To investigate the effects of such constraints on $\tilde{M}_\mathrm{T}$ we performed a series of numerical experiments with the Honshu event. Synthetic seismograms were computed from the moment-rate tensor of

Masters & Gilbert (1981, private communication) listed in Table 3 ($M_T = 0.644 \times 10^{27}$ dyne cm $= 0.644 \times 10^{20}$ N m), and the resulting cross-spectra were integrated across the frequency band 4–5 mHz to give theoretical values of U_{pq} for the nine IDA stations recording this event. The estimation algorithm of Section 5.1 was applied to these synthetic data using the various likelihood functions described in Section 4.3.

Table 4 summarizes the results for four basic types of likelihood functions. Each type characterizes a different manifold of source mechanisms and is parametrized by a dimensionless constant κ, whose inverse measures the dispersion about a most probable mechanism $\hat{\mathbf{m}}_0$. Type HA corresponds to a hyperspherical density function for a general asynchronous source; its bias and variance are given by equations (4.60) and (4.61). Type HS is also hyperspherical, but the mechanism is constrained to be synchronous (equations 4.57 and 4.58). Type HSD incorporates the further restriction that the mechanism be purely deviatoric; its statistics are constructed by the methods outlined at the end of Section 4.3. Finally, type GSD represents the Gaussian approximation to type HSD (equations 4.64–4.66). In all these distributions $\hat{\mathbf{m}}_0$ was taken to be the synchronous, deviatoric mechanism of Table 3, the same mechanism used to generate the data.

For each of the experiments in Table 4 the coefficients $\{a_{pq}; p, q = 1, 2, \ldots, 9\}$ were computed assuming the snrs and heterogeneity parameters assigned to the real data. However, because the synthetic data were not, in fact, contaminated by such observational errors, the values of $\tilde{M}_T$ listed in the table were not corrected for the relative observational bias $\hat{\beta}_O$.

At a fixed value of κ restricting the source to be synchronous and deviatoric generally decreases the errors in $\tilde{M}_T$ by a small amount. At $\kappa = 18$, for example, the relative error $\rho_T = (\hat{\sigma}_m^2 + \hat{\sigma}_O^2)^{1/2}$ decreases from about 10 per cent for the HA distribution to 8 per cent for HSD. This modest gain is a manifestation of the heuristic notion that imposing *a priori* constraints on the source mechanism should permit an improved estimate of scalar moment (conditional, of course, on the constraints being correct!). The exception occurs at $\kappa = 0$, where going from an HA to an HS distribution actually increases the estimated error. The

Table 4. Statistics of total-moment estimates for the Honshu event derived from synthetic data in the band 4–5 mHz for various source-mechanism likelihood functions.

Type	Probability density function	Source description	κ	$\hat{\beta}_m$	$\hat{\sigma}_m$	$\hat{\beta}_O$	$\hat{\sigma}_O$	$\tilde{M}_T$* (10^{27} dyne cm)
HA	Hyperspherical	Asynchronous	0	−0.058	0.120	0.042	0.095	0.696
			9	−0.044	0.105	0.035	0.087	0.709
			18	−0.029	0.081	0.024	0.075	0.707
			50	−0.014	0.047	0.010	0.062	0.682
HS	Hyperspherical	Synchronous	0	−0.072	0.126	0.051	0.110	0.677
			9	−0.038	0.093	0.030	0.084	0.694
			18	−0.021	0.066	0.018	0.072	0.684
			50	−0.012	0.038	0.007	0.059	0.663
HSD	Hyperspherical	Synchronous, deviatoric	0	−0.045	0.082	0.042	0.108	0.644
			9	−0.028	0.075	0.026	0.080	0.675
			18	−0.018	0.054	0.016	0.069	0.672
			50	−0.011	0.032	0.007	0.059	0.658
GSD	Gaussian	Synchronous, deviatoric	9	−0.012	0.085	0.029	0.088	0.692
			18	−0.011	0.058	0.016	0.072	0.677
			50	−0.010	0.032	0.007	0.059	0.659

*Correct value is $M_T = 0.644 \times 10^{27}$ dyne cm.

explanation for this apparently anomalous behaviour is subtle: restricting the mechanism to be synchronous will increase the expected error if the error matrix $\mathbf{E}$ is nearly real, since the elements of $\hat{\mathbf{m}}$ will then be in phase and the quadratic form $\hat{\mathbf{m}}^* \cdot \mathbf{E} \cdot \hat{\mathbf{m}}$ will, on the average, be larger. This is indeed the case at $\kappa = 0$, where the power-spectral integrals dominate the sum more.

Increasing κ implies increasing the probability that the actual mechanism lies near $\hat{\mathbf{m}}_0$; for all the distributions in Table 4 this reduces the estimation errors, since $\tilde{M}_T$ can be optimized accordingly. In the limit $\kappa \to \infty$ the error assigned to $\hat{\mathbf{m}}_0$ goes to zero, and the optimal estimate yields $\hat{\sigma}_m = 0$.

The Gaussian approximation is expected to be valid in the large κ limit. In fact, at $\kappa = 50$ the HSD and GSD distributions give almost identical estimates of M_T and its errors. The agreement deteriorates as κ is decreased, but is still adequate for values as low as $\kappa = 9$. The expected error in $\hat{\mathbf{m}}_0$ scales approximately as $\kappa^{-1/2}$; i.e. a value of $\kappa = 9$ corresponds to an error of about 33 per cent, which exceeds the uncertainty in most focal mechanisms obtained from first-motion analysis. The Gaussian approximation should therefore be applicable to the study of the many events for which good mechanisms exist, including the examples considered here.

6.2 ESTIMATION OF TOTAL MOMENT

M_T was estimated from real data using GSD distributions to constrain the source mechanism. The parameters $\hat{\mathbf{m}}_0$ and $\hat{\mathbf{V}}_m$ are displayed in Table 3. For the Honshu event, $\hat{\mathbf{m}}_0$ was taken to be the synchronous, deviatoric mechanism of Masters & Gilbert (1981, private communication), and $\hat{\mathbf{V}}_m$ was constructed according to equation (4.67) from the variance matrix obtained by their regression analysis. For Oaxaca, the double-couple mechanism of Stewart *et al.* (1981) was adopted. The uncertainties in the strike, dip and slip of the auxiliary plane were assigned standard errors of $\pm 7°$, $\pm 7°$ and $\pm 15°$, respectively, which are consistent with the constraints imposed by their study. We assumed these errors to be uncorrelated and derived $\hat{\mathbf{V}}_m$ by first-order perturbation theory. To allow for the possibility of a small, non-double-couple, deviatoric component in the source we augmented this variance matrix by a matrix of the form (4.64) with $\kappa = 100$.

The bias-corrected estimates of M_T derived from actual data are shown in Figs 2 and 3. The points correspond to integrals over disjunct 1-mHz bands spanning the frequency interval 1–11 mHz. Despite the narrowness of the integration bandwidth, the precision of the estimates is quite good; their relative standard deviations range from 10 to 22 per cent.

Our estimates of total moment are generally consistent with other studies. Moment-rate tensors for the Honshu earthquake have been computed by Masters & Gilbert (1981, private communication) from IDA data in the band 2–5 mHz and by Dziewonski *et al.* (1981) from SRO data peaked at approximately 16 mHz; these yield total moments of 0.64 and 0.40×10^{27} dyne cm, respectively. We obtain values that vary from $0.52 \pm 0.071 \times 10^{27}$ dyne cm (3–4 mHz) to $0.30 \pm 0.048 \times 10^{27}$ dyne cm (10–11 mHz). Averaging $\tilde{M}_T^2$ over the ten frequency points yields a root mean square (rms) $\tilde{M}_T$ of 0.43×10^{27} dyne cm, essentially identical to that found by applying the estimation algorithm to power and cross-spectra integrated across the entire 1–11 mHz interval.

A similar averaging of our results for the Oaxaca earthquake gives an rms value of 2.8×10^{27} dyne cm, with individual estimates ranging from $2.2 \pm 0.45 \times 10^{27}$ dyne cm (7–8 mHz) to $3.2 \pm 0.45 \times 10^{27}$ dyne cm (3–4 mHz). In comparison, Stewart *et al.* (1981) computed a moment of 3.2×10^{27} dyne cm from their double-couple mechanism by fitting surface waves recorded by the World-Wide Standardized Seismographic Network (WWSSN), and Reichle *et al.* (1982) got 2.9×10^{27} dyne cm using a somewhat different mechanism and

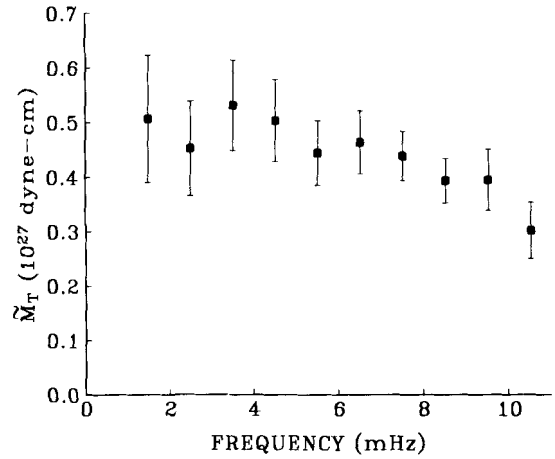

Figure 2. Estimates of total moment averaged over 1-mHz bands for the Honshu event of 1978 March 7. Error bars represent one standard deviation.

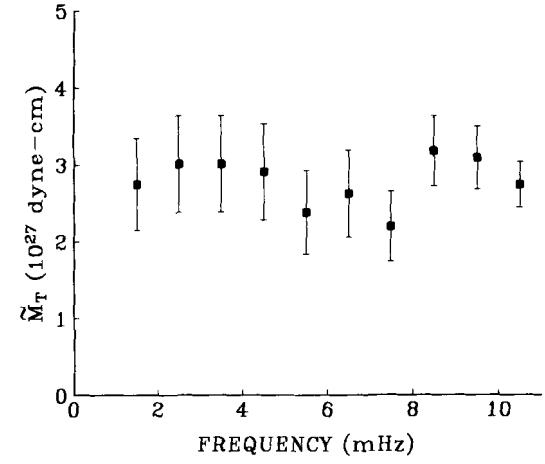

Figure 3. Estimates of total moment averaged over 1-mHz bands for the Oaxaca event of 1978 November 29. Error bars represent one standard deviation.

surface waves recorded by IDA and SRO instruments. Both data sets had predominant frequencies near 5 mHz. Ward's (1980b) analysis of higher-frequency (~ 30 mHz) body waves produced estimates ranging from 1.4 to 2.8×10^{27} dyne cm.

The precision attainable by our algorithm should permit the resolution of a fairly weak frequency dependence. The Oaxaca event yields an essentially flat moment spectrum – a result consistent with the short source times ($\leq$ 20 s) obtained by Stewart *et al.* (1981) and Reichle *et al.* (1982). In contrast, there is evident in the plot for Honshu a systematic decrease in M_T with frequency, resulting in a drop of about 25 per cent across the band. A regression fit to a Brune-type source spectrum yields an apparent corner frequency of 15 ± 4 mHz, which is anomalously low for an event of this magnitude.

However, we have excluded from our analysis two potential contributors of frequency-dependent bias: source mislocation and an incorrect model of radial Q structure. To investi-

gate the former we generated noise-free synthetic seismograms for the hypothetical sources of Table 3 using the hypocentres of Table 2; we then inverted these synthetic data assuming various source depths. The results are given in Fig. 4. Perturbing the source depth of the Honshu event by 20 km (~ 5 per cent) introduced an error in $\tilde{M}_T$ that averages only 5 per cent and is nowhere greater than 10 per cent. For the shallow-focus Oaxaca event, comparable errors are generated by a 6-km displacement of the hypocentre. The detailed studies of Dziewonski *et al.* (1981) and Stewart *et al.* (1981) suggest that focal-depth errors much in excess of these values are unlikely. Hence, the marginal uncertainty in M_T due to an incorrect depth is only on the order of a few per cent.

The effect of an incorrect radial Q model on $\tilde{M}_T$ can be illustrated by considering a shallow-focus event such as Oaxaca, where the fundamental spheroidal modes dominate the seismogram. Over an integration bandwidth of only 1 mHz the ratio $\Delta\bar{\alpha}_k/\alpha_k$ of equation (4.35) can be approximated by $-\Delta Q/Q$, where $\Delta Q/Q$ is the average relative bias in the quality factors of fundamental modes contained in the band. Referring to equations (4.25) and (4.26) and ignoring other sources of bias we obtain $\langle \Delta U_{pq}^{(g)} \rangle \simeq (\Delta Q/Q) U_{pq}$. This implies that the relative bias in $\tilde{M}_T^2$ due to errors in Q is approximately $\Delta Q/Q$. Given the quality and size of the fundamental-mode data set used by Masters & Gilbert (1982), their model, which has been adopted to generate the transfer functions, should be able to predict

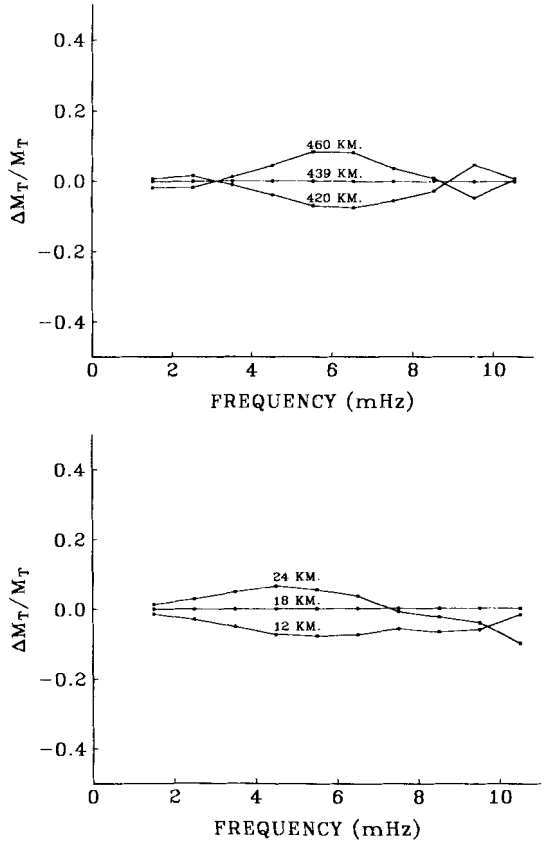

Figure 4. Relative bias in total moment caused by displacement of the assumed source depth from its true value. Curves computed from synthetic data for: (a) the Honshu source at 439 km, and (b) the Oaxaca source at 18 km assuming focal depths of 420, 439, 460 km and 12, 18, 24 km, respectively.

the average fundamental-mode Qs to 5 per cent (their prediction uncertainties are only a few per cent or less) and, thus, would not lead to an appreciable error in the values of $\tilde{M}_T$ for a typical shallow-focus event. In the case of the deep-focus Honshu event, however, the magnitude of the error cannot be so strictly bounded, because the overtones contribute significantly to the seismogram, especially in the band 6–11 mHz, and their Qs are less well known. We do not feel warranted, therefore, to exclude the hypothesis that the true value of corner frequency for Honshu is much greater than 15 Hz.

6.3 TEST FOR AN ISOTROPIC PART

The detection algorithm formulated in Section 5.2 was applied to the Honshu and Oaxaca data to test for the existence of an isotropic moment M_I. In Figs 5 and 6 the bias-corrected estimates $\tilde{M}_I^2$ given by equation (5.18) are compared with the 'detection thresholds' computed for the 90 and 95 per cent confidence levels. These thresholds are defined by the inequality (5.20). All values of $\tilde{M}_I^2$ have been normalized by our best estimates of M_T^2 ($\tilde{M}_T = 0.43 \times 10^{27}$ dyne cm for Honshu, 2.8×10^{27} dyne cm for Oaxaca).

The detection thresholds for the two events are quite different, being generally higher for Honshu, especially in the band 3–6 mHz. Indeed, as Fig. 5 shows, the data for Honshu at 3–4 mHz are effectively incapable of distinguishing a purely isotropic mechanism from a purely deviatoric one, even at the 90 per cent confidence level.

The mode-excitation efficiency of an isotropic source relative to a deviatoric source of equal M_T can be measured by the ratio of the mode's compressional energy density, E_C, to its shear energy density, E_S, at the source depth (Gilbert 1973b). At the 18-km hypocentral depth of the Oaxaca event, the ratio E_C/E_S increases gradually with frequency for the fundamental spheroidal modes. Since these modes dominate the seismograms, the detection thresholds decrease with frequency. Despite their low values (< 0.1 in the band 7–11 mHz for the 95 per cent level), only one estimate of M_I^2/M_T^2 (9–10 mHz) is significantly positive at the 95 per cent level. Taking the ten measurements together, the hypothesis that the Oaxaca event is purely deviatoric cannot, therefore, be rejected.

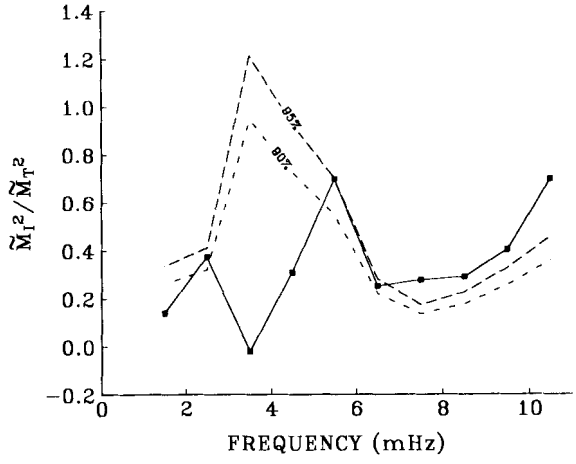

Figure 5. Test for the detection of an isotropic component of the Honshu mechanism. Points connected by solid line are estimates of M_I^2/M_T^2 averaged over 1-mHz bands. Points lying above the light and heavy dashed lines are significantly greater than zero at the 90 per cent confidence levels, respectively.

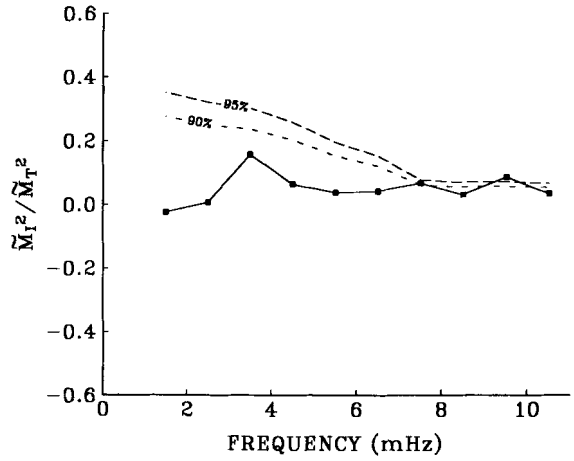

Figure 6. Test for the detection of an isotropic component of the Oaxaca mechanism, following the conventions of Fig. 5.

At the much greater focal depth of the Honshu event (439 km), E_C/E_S for the fundamental modes falls off rapidly beyond 2 mHz, causing a sharp rise in the detection thresholds plotted on Fig. 5. Based on fundamental-mode excitation alone, it would be very difficult to detect an isotropic component of a deep-focus earthquake, a fact recognized by Okal & Geller (1979). Fortunately, at frequencies above about 4 mHz, the detection thresholds for our data set are controlled primarily by the excitation of overtones, not fundamentals, and the detectability of an isotropic part improves, reaching an optimum at 7–8 mHz.

Except for the points at 1–2 and 3–4 mHz the estimated values of M_I^2/M_T^2 for Honshu are much more positive than those for Oaxaca, varying between 0.25 and 0.70 with an average value over all 10 points of 0.34. All six points at frequencies greater than 5 mHz show an isotropic part that is significant at 90 per cent confidence level, and four of the six are significant at the 95 per cent level. Formally, therefore, we can reject the hypothesis that the Honshu mechanism is purely deviatoric.

The sensitivity of our conclusions to focal-depth mislocation was investigated by re-inverting the data set with ±20 km displacements in the Honshu hypocentre; these produced only minor shifts in the curves of Fig. 5 and did not alter the outcome of the significance test.

Other potential sources of extraneous bias are more difficult to evaluate. Both an incorrect radial Q model and mode coupling by lateral heterogeneities could alter the distribution of compressional and shear energies in the displacement field and thus affect the estimates of M_I^2/M_T^2. Until additional experiments can be conducted to pin down the magnitude of these effects, the positive findings of Fig. 5 must be considered provisional.

7 Discussion

Our tentative identification of an isotropic component to the Honshu mechanism constitutes an interesting result, lending credence to the conjecture of previous seismologists (e.g. Gilbert & Dziewonski 1975) that isotropic compression accompanies some deep-focus earthquakes. Our hypothesis-testing algorithm has been rigorously formulated for an attack on this difficult observational problem and is being applied to the study of additional events recorded by the IDA network and other digital arrays.

The techniques for scalar-moment estimation developed here offer many advantages. Theoretical transfer functions are employed by our analysis, but the algorithms avoid the matched filtering of observed seismograms by synthetics and are, therefore, fairly robust with respect to errors in the assumed earth model. Careful attention has been given to the estimation statistics and to schemes for optimizing the estimation errors, including those associated with incorrect transfer functions. As the experiments of the previous section demonstrate, the total moment M_T can be obtained with a precision of better than 20 per cent by averaging the data from a sparse network over bandwidths of only 1 mHz. This precision should be sufficient to identify 'slow earthquakes' such as those discussed by Kanamori, Sacks and their co-workers and to investigate other anomalous aspects of source time functions. In some cases a significant improvement in the estimation statistics, especially in the detectability of an isotropic part, may be achievable by judiciously choosing time windows lagged to reduce the contributions of the fundamental modes (Dratler *et al.* 1971).

Our method for estimating scalar moment requires no *a priori* information about the source mechanism, but, if such information is available, it can be specified in terms of the various likelihood functions discussed in Section 4.3. Imposing constraints on the source mechanism derivable from, say, first-motion analysis allows total moment to be obtained from very few seismic recordings. For example, consistent estimates of M_T for both the Honshu and Oaxaca events have been computed using only two stations. Moreover, specifying a likelihood function allows us to eliminate the negative bias in moment estimates for shallow focus events caused by the degeneracy of the strain tensor at the free surface (Kanamori & Given 1981).

Acknowledgments

We are grateful to T. G. Masters and F. Gilbert for providing us with edited time series, transfer functions and moment-tensor estimates for the Honshu and Oaxaca events, and we thank G. Backus, J. Brune and J. B. Minster for helpful discussions. T.H.J. is an Alfred P. Sloan Foundation Fellow. This research was sponsored by the National Science Foundation under grants EAR79-23779 and EAR80-07823.

References

Aki, K., 1966. Generation and propagation of *G* waves from the Niigata earthquake of June 16, 1964. Part 2. Estimation of earthquake moment, released energy, and stress-strain drop from the *G* wave spectrum, *Bull. Earthq. Res. Inst. Tokyo Univ.*, **44**, 73–88.

Aki, K. & Richards, P. G., 1980. *Quantitative Seismology, Theory and Methods,* Vol. I, W. H. Freeman, San Francisco.

Backus, G., 1977. Interpreting the seismic glut moments of total degree two or less, *Geophys. J. R. astr. Soc.*, **51**, 1–25.

Backus, G. & Gilbert, F., 1970. Uniqueness in the inversion of inaccurate gross Earth data, *Phil. Trans. R. Soc. A*, **266**, 123–192.

Backus, G. & Mulcahy, M., 1976a. Moment tensors and other phenomenological descriptions of seismic sources – I. Continuous displacements, *Geophys. J. R. astr. Soc.*, **46**, 341–361.

Backus, G. & Mulcahy, M., 1976b. Moment tensors and other phenomenological descriptions of seismic sources – II. Discontinuous displacements, *Geophys. J. R. astr. Soc.*, **47**, 301–329.

Benioff, H., 1963. Source wave forms of three earthquakes, *Bull. seism. Soc. Am.*, **53**, 893–903.

Buland, R. & Gilbert, F., 1976. Matched filtering for the seismic moment tensor, *Geophys. Res. Lett.*, **3**, 205–206.

Burridge, R. & Knopoff, L., 1964. Body force equivalents for seismic dislocations, *Bull. seism. Soc. Am.*, **54**, 1875–1888.

Dahlen, F. A., 1979. The spectra of unresolved split normal mode multiplets, *Geophys. J. R. astr. Soc.*, **58**, 1–33.

Dratler, J., Jr, Farrell, W. E., Block, B. & Gilbert, F., 1971. High-Q overtone modes of the Earth, *Geophys. J. R. astr. Soc.*, **23**, 399–410.

Dziewonski, A. M., Chou, T. A. & Woodhouse, J. H., 1981. Determination of earthquake source parameters from wave-form data for studies of global and regional seismicity, *J. geophys. Res.*, **86**, 2825–2852.

Dziewonski, A. M. & Gilbert, F., 1974. Temporal variation of the seismic moment tensor and the evidence of precursive compression for two deep earthquakes, *Nature*, **247**, 185–188.

Erdélyi, A., 1953. *Higher Transcendental Functions, II*, chapter 11, McGraw-Hill, New York.

Evison, F. F., 1963. Earthquakes and faults, *Bull seism. Soc. Am.*, **53**, 873–891.

Fisher, R. A., 1953. Dispersion on a sphere, *Proc. R. Soc. A*, **217**, 295–305.

Gilbert, F., 1971a. Excitation of normal modes of the Earth by earthquake sources, *Geophys. J. R. astr. Soc.*, **22**, 223–226.

Gilbert, F., 1971b. The diagonal sum rule and averaged eigenfrequencies, *Geophys. J. R. astr. Soc.*, **23**, 119–123.

Gilbert, F., 1973a. Derivation of source parameters from low-frequency spectra, *Phil. Trans. R. Soc. A*, **274**, 369–371.

Gilbert, F., 1973b. The relative efficiency of earthquakes and explosions in exciting surface waves and body waves, *Geophys. J. R. astr. Soc.*, **33**, 487–488.

Gilbert, F., 1976. Representation of seismic displacements in terms of travelling waves, *Geophys. J. R. astr. Soc.*, **44**, 275–280.

Gilbert, F. & Buland, R., 1976. An enhanced deconvolution procedure for retrieving the seismic moment tensor from a sparse network, *Geophys. J. R. astr. Soc.*, **50**, 251–255.

Gilbert, F. & Dziewonski, A. M., 1975. An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra, *Phil. Trans. R. Soc. A*, **278**, 187–269.

Hanks, T. C. & Kanamori, H., 1979. A moment magnitude scale, *J. geophys. Res.*, **84**, 2348–2350.

Jenkins, G. M. & Watts, D. G., 1968. *Spectral Analysis and its Applications*, chapter 6, Holden-Day, San Francisco.

Johnson, N. L. & Kotz, S., 1970. *Continuous Univariate Distributions – II*, chapter 33, Wiley, New York.

Jordan, T. H., 1978. A procedure for estimating lateral variations from low-frequency eigenspectra data, *Geophys. J. R. astr. Soc.*, **52**, 441–455.

Jordan, T. H., 1980. Earth structure from seismological observations, Physics of the Earth's interior, *Proc. int. School of Physics 'Enrico Fermi'*, Course LXXVIII, pp. 1–40, eds Dziewonski, A. M. & Boschi, E., North-Holland, Amsterdam.

Jordan, T. H. & Silver, P. G., 1981. Parameters of normal-mode multiplets and their relationship to lateral heterogeneity, *Eos, Trans. Am. geophys. Un.*, **62**, 336 (abstract).

Kanamori, H., 1970a. Synthesis of long period surface waves and its application to earthquake source studies – Kurile Islands earthquake of October 13, 1963, *J. geophys. Res.*, **75**, 5011–5027.

Kanamori, H., 1970b. The Alaska earthquake of 1964: radiation of long-period surface waves and source mechanism, *J. geophys. Res.*, **75**, 5029–5040.

Kanamori, H., 1977. The energy release in great earthquakes, *J. geophys. Res.*, **82**, 2981–2987.

Kanamori, H. & Cipar, J., 1974. Focal process of the great Chilean earthquake May 22, 1960, *Phys. Earth planet. Int.*, **9**, 128–136.

Kanamori, H. & Given, J. W., 1981. Use of long-period surface waves for fast determination of earthquake source parameters, *Phys. Earth planet. Int.*, **27**, 8–31.

Kanamori, H. & Stewart, G. S., 1979. A slow earthquake, *Phys. Earth planet. Int.*, **18**, 167–175.

Knopoff, L. & Randall, M. J., 1970. The compensated linear-vector dipole: a possible mechanism for deep earthquakes, *J. geophys. Res.*, **75**, 4957–4963.

Masters, G. & Gilbert, F., 1982. Attenuation in the Earth at low frequencies, *Phil. Trans. R. Soc.*, submitted.

McCowan, D. W., 1976. Moment tensor representation of surface waves, *Geophys. J. R. astr. Soc.*, **44**, 595–599.

Mendiguren, J. A., 1977. Inversion of surface wave data in source mechanism studies, *J. geophys. Res.*, **82**, 889–894.

Okal, E. A. & Geller, R. J., 1979. On the observability of isotropic seismic sources: the July 31, 1970 Colombian earthquake, *Phys. Earth planet. Int.*, **27**, 393–446.

Olver, F. W. J., 1970. Bessel functions of integral order, *Handbook of Mathematical Functions*, chapter 9, eds Abramowitz, M. & Stegun, I. A., Dover, New York.

Patton, P. & Aki, A., 1979. Bias in the estimate of seismic moment tensor by the linear inversion method, *Geophys. J. R. astr. Soc.*, **59**, 479–495.

Randall, M. J., 1971. Elastic multipole theory and seismic moment, *Bull. seism. Soc. Am.*, **61**, 1321–1326.

Randall, M. J., 1972. Multipolar analysis of the mechanisms of deep focus earthquakes, *Meth. comp. Phys.*, **12**, 267–298.

Randall, M. J. & Knopoff, L., 1970. The mechanism at the focus of deep earthquakes, *J. geophys. Res.*, **75**, 4965–4976.

Reichle, M., Orcutt, J. A. & Priestly, K., 1982. A comparison of source mechanisms of three Mexican earthquakes: teleseismic *P* waves, *J. geophys. Res.*, in press.

Riedesel, M., Jordan, T. H. & Masters, G., 1981. Observed effects of lateral heterogeneity on fundamental spheroidal modes, *Eos, Trans. Am. geophys. Un.*, **62**, 332 (abstract).

Sacks, I. S., Linde, A. T., Snoke, J. A. & Suyehiro, S., 1981. A slow earthquake sequence following the Izo-Oshima earthquake of 1978, *Earthquake Prediction: an International Review*, eds Simpson, D. W. & Richards, P. G., American Geophysical Union, Washington.

Sacks, I. S., Suyehiro, S., Linde, A. T. & Snoke, J. A., 1978. Slow earthquakes and stress redistribution, *Nature*, **275**, 599–602.

Silver, P. G. & Jordan, T. H., 1981. Fundamental spheroidal mode observations of aspherical heterogeneity, *Geophys. J. R. astr. Soc.*, **64**, 605–634.

Stewart, G. S., Chael, E. P. & McNally, K. C., 1981. The 1978 November 29 Oaxaca, Mexico earthquake – a large simple event, *J. geophys. Res.*, **86**, 5053–5060.

Strelitz, R. A., 1980. The fate of the downgoing slab: a study of the moment tensors from body waves of complex deep-focus earthquakes, *Phys. Earth planet. Int.*, **21**, 83–96.

Stump, W. S. & Johnson, L. R., 1977. The determination of source properties by the linear inversion of seismograms, *Bull. seism. Soc., Am.*, **67**, 1489–1502.

von Mises, R., 1918. Über die "Ganzzahligkeit" der Atmogewichte und verwandte Fragen, *Phys. Z.*, **19**, 490–500.

Ward, S. N., 1980a. Body wave calculations using moment tensor sources in spherically symmetric, inhomogeneous media, *Geophys. J. R. astr. Soc.*, **60**, 53–66.

Ward, S. N., 1980b. A technique for the recovery of the seismic moment tensor applied to the Oaxaca, Mexico earthquake of November 1978, *Bull. seism. Soc. Am.*, **70**, 717–734.

Watson, G. S., 1966. The statistics of orientation data, *J. Geol.*, **74**, 786–797.

Appendix A: formulation of $V^{(n)}_{pqst}$

The covariance matrix for ambient seismic noise is

$$V^{(n)}_{pqst} \equiv \mathrm{Cov}[\Delta U^{(n)*}_{pq} \Delta U^{(n)}_{st}] = \langle \Delta U^{(n)*}_{pq} \Delta U^{(n)}_{st} \rangle - B^{(n)}_{pq} B^{(n)}_{st}. \tag{A1}$$

We assume the endpoints of the integration are $\omega_a = J_a\omega_0$ and $\omega_b = J_b\omega_0$, where ω_0 is the fundamental Fourier frequency $2\pi/T$ and J_a and J_b are integers, and we approximate the frequency integrals by sums over discrete Fourier points,

$$\Delta U^{(n)}_{pq} = \omega_0 \sum_{j=J_a}^{J_b} [u^*_p(j\omega_0)\,\bar{n}_q(j\omega_0) + \bar{n}^*_p(j\omega_0)\,u_q(j\omega_0) + \bar{n}^*_p(j\omega_0)\,\bar{n}_q(j\omega_0)]. \tag{A2}$$

The covariance matrix involves the first four moments of $\bar{n}_p$. All odd-order moments are zero for a Gaussian white-noise process with zero-mean. The required even-order moments are

$$\langle \bar{n}_p(j\omega_0)\,\bar{n}_q(k\omega_0) \rangle = 0 \tag{A3}$$

$$\langle \bar{n}^*_p(j\omega_0)\,\bar{n}_q(k\omega_0) \rangle = T\nu_p^2\,\delta_{pq}\,\delta_{jk} \tag{A4}$$

$$\langle \bar{n}_p(j\omega_0)\,\bar{n}_q^*(j\omega_0)\,\bar{n}_s^*(k\omega_0)\,\bar{n}_t(k\omega_0)\rangle = T^2(\nu_p^2\,\nu_s^2\,\delta_{pq}\,\delta_{st} + \nu_p^2\,\nu_q^2\,\delta_{ps}\,\delta_{qt}\delta_{jk}). \tag{A5}$$

The resulting expression for the covariance matrix is

$$V_{pqst}^{(n)} = T\nu_p^2\,U_{qt}\delta_{ps} + T\nu_q^2\,U_{ps}^*\delta_{qt} + \Delta\omega T^2\nu_p^2\nu_q^2\delta_{ps}\delta_{qt}. \tag{A6}$$

Appendix B: moments of the hyperspherical normal distribution

In this section we derive general expressions for the moments of the hyperspherical normal distribution $S_n(\hat{\mathbf{m}})$ defined by equation (4.55). We also examine the form of these moments when $\kappa = 0$, corresponding to a uniform distribution on the n-dimensional unit sphere, and given formulae asymptotically valid in the Gaussian limit of small dispersion ($\kappa \to \infty$).

Let $\hat{\mathbf{m}}^p$ denote the pth-order outer-product tensor

$$\underbrace{\hat{m}_i\hat{m}_j\ldots\hat{m}_l}_{p\text{ times}}\,. \tag{B1}$$

The pth moment of S_n is

$$\mu^{(p)} \equiv \langle \hat{\mathbf{m}}^p \rangle = C_n^{-1}(\kappa)\int_{\Omega_n} \hat{\mathbf{m}}^p \exp(\kappa\,\hat{\mathbf{m}}_0\cdot\hat{\mathbf{m}})\,d\Omega_n(\hat{\mathbf{m}}). \tag{B2}$$

The normalization factor C_n (equation 4.56) is an integral of the general form

$$C_n(\kappa r) = \int_{\Omega_n} \exp(\kappa\,\mathbf{r}\cdot\hat{\mathbf{m}})\,d\Omega_n(\hat{\mathbf{m}}), \tag{B3}$$

where $r \equiv \|\mathbf{r}\| = (\mathbf{r}\cdot\mathbf{r})^{1/2}$. To the function (B3) we apply ∇_r, the gradient operator with respect to the vector field r, and evaluate the results at $\mathbf{r} = \hat{\mathbf{m}}_0$:

$$[\nabla_r C_n(\kappa r)]_{\mathbf{r}=\hat{\mathbf{m}}_0} = \kappa\int_{\Omega_n} \hat{\mathbf{m}}\exp(\kappa\,\hat{\mathbf{m}}_0\cdot\hat{\mathbf{m}})\,d\Omega_n(\hat{\mathbf{m}}). \tag{B4}$$

This surface integral is proportional to $\mu^{(1)}$. Hence, the pth moment is generated by applying the gradient operator p times,

$$\mu^{(p)} = \kappa^{-p}\,C_n^{-1}(\kappa)\,[\nabla_r^p C_n(\kappa r)]_{\mathbf{r}=\hat{\mathbf{m}}_0}. \tag{B5}$$

Expressing the surface element $d\Omega_n$ in hyperspherical coordinates (Erdélyi 1953, chapter XI) and using the integral representation of the modified Bessel function of the first kind I_ν (Olver 1970, equation 9.6.18) we can reduce (B3) to the simple expression

$$C_n(\kappa r) = (2\pi)^{\nu+1}(\kappa r)^{-\nu}I_\nu(\kappa r),$$
$$\nu \equiv \frac{n}{2} - 1. \tag{B6}$$

The gradient of C_n is thus related to the derivative of the function $z^{-\nu}I_\nu(z)$; a formula for this derivative (*op. cit.*, equation 9.6.28) yields

$$\nabla_r C_n(\kappa r) = \mathbf{r}\frac{\kappa^2}{2\pi}\,C_{n+2}(\kappa r). \tag{B7}$$

The repeated application of (B7) in equation (B5) generates a finite series of $\mu^{(p)}$ involving the functions C_n only through the coefficients

$$a_\nu^{(q)}(\kappa) \equiv (2\pi)^{-q}C_n^{-1}(\kappa)\,C_{n+2q}(\kappa)$$
$$= \kappa^{-q}I_{\nu+q}(\kappa)/I_\nu(\kappa), \qquad q \in (1, 2, \ldots, p). \tag{B8}$$

We give explicit expressions for the Cartesian components of the first four moments:

$$\mu_i^{(1)} = \kappa a_\nu^{(1)}(\kappa)\hat{m}_{0i} \tag{B9}$$

$$\mu_{ij}^{(2)} = a_\nu^{(1)}(\kappa)\delta_{ij} + \kappa^2 a_\nu^{(2)}(\kappa)\hat{m}_{0i}\hat{m}_{0j} \tag{B10}$$

$$\mu_{ijk}^{(3)} = \kappa a_\nu^{(2)}(\kappa)\,(\hat{m}_{0i}\delta_{jk} + \hat{m}_{0j}\delta_{jk} + \hat{m}_{0k}\delta_{ij}) + \kappa^3 a_\nu^{(3)}(\kappa)\hat{m}_{0i}\hat{m}_{0j}\hat{m}_{0k} \tag{B11}$$

$$\begin{aligned}\mu_{ijkl}^{(4)} = {} & a_\nu^{(2)}(\kappa)\,(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \\ & + \kappa^2 a_\nu^{(3)}(\kappa)\,(\hat{m}_{0i}\hat{m}_{0j}\delta_{kl} + \hat{m}_{0i}\hat{m}_{0l}\delta_{jk} \\ & + \hat{m}_{0k}\hat{m}_{0l}\delta_{ij} + \hat{m}_{0j}\hat{m}_{0k}\delta_{il} + \hat{m}_{0j}\hat{m}_{0l}\delta_{ik} + \hat{m}_{0i}\hat{m}_{0k}\delta_{jl}) \\ & + \kappa^4 a_\nu^{(4)}(\kappa)\,\hat{m}_{0i}\hat{m}_{0j}\hat{m}_{0k}\hat{m}_{0l}.\end{aligned} \tag{B12}$$

In the case $\kappa = 0$, corresponding to a uniform distribution on the unit hypersphere, the coefficients are easily calculated from the closed-form expression

$$a_\nu^{(q)}(0) = [2^q(\nu+1)(\nu+2)\ldots(\nu+q)]^{-1}, \tag{B13}$$

and the low-order moments reduce to the isotropic tensors

$$\mu_i^{(1)} = 0, \tag{B14}$$

$$\mu_{ij}^{(2)} = n^{-1}\delta_{ij}, \tag{B15}$$

$$\mu_{ijk}^{(3)} = 0, \tag{B16}$$

$$\mu_{ijkl}^{(4)} = [n(n+2)]^{-1}(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{kl}\delta_{jk}). \tag{B17}$$

To examine the limit of small dispersion (large κ) we observe that the probability density function $S_n(\hat{\mathbf{m}})$ is proportional to

$$\exp(\kappa\,\hat{\mathbf{m}}_0\cdot\hat{\mathbf{m}}) = \exp(\kappa)\exp[-(\kappa/2)\,(\hat{\mathbf{m}}-\hat{\mathbf{m}}_0)\cdot(\hat{\mathbf{m}}-\hat{\mathbf{m}}_0)]. \tag{B18}$$

When κ is large, $S_n(\hat{\mathbf{m}})$ is appreciably different from zero only in the immediate vicinity of $\hat{\mathbf{m}}_0$, where $\hat{\mathbf{m}}_0\cdot\hat{\mathbf{m}} \simeq 1$. The domain of $S_n(\hat{\mathbf{m}})$ can thus be approximated by the hyperplane tangent to Ω_n at $\hat{\mathbf{m}}_0$, on which the density function (B18) is a Gaussian distribution with a mean $\langle\hat{\mathbf{m}}\rangle = \hat{\mathbf{m}}_0$ and a variance matrix $\hat{\mathbf{V}}_\mathrm{m} \equiv \mathrm{Var}[\hat{\mathbf{m}}] = \kappa^{-1}\,(\mathbf{I} - \hat{\mathbf{m}}_0\hat{\mathbf{m}}_0)$. The moments of this distribution are:

$$\mu_i^{(1)} = \hat{m}_{0i} \tag{B19}$$

$$\mu_{ij}^{(2)} = (\hat{V}_\mathrm{m})_{ij} + \hat{m}_{0i}\hat{m}_{0j} \tag{B20}$$

$$\mu_{ijk}^{(3)} = \hat{m}_{0i}(\hat{V}_\mathrm{m})_{jk} + \hat{m}_{0j}(\hat{V}_\mathrm{m})_{ik} + \hat{m}_{0k}(\hat{V}_\mathrm{m})_{ij} + \hat{m}_{0i}\hat{m}_{0j}\hat{m}_{0k} \tag{B21}$$

$$\begin{aligned}\mu_{ijkl}^{(4)} = {} & (\hat{V}_\mathrm{m})_{ij}(\hat{V}_\mathrm{m})_{kl} + (\hat{V}_\mathrm{m})_{ik}(\hat{V}_\mathrm{m})_{jl} + (\hat{V}_\mathrm{m})_{il}(\hat{V}_\mathrm{m})_{jk} \\ & + \hat{m}_{0i}\hat{m}_{0j}(\hat{V}_\mathrm{m})_{kl} + \hat{m}_{0i}\hat{m}_{0l}(\hat{V}_\mathrm{m})_{jk} \\ & + \hat{m}_{0k}\hat{m}_{0l}(\hat{V}_\mathrm{m})_{ij} + \hat{m}_{0j}\hat{m}_{0k}(\hat{V}_\mathrm{m})_{il} + \hat{m}_{0j}\hat{m}_{0l}(\hat{V}_\mathrm{m})_{ik} + \hat{m}_{0i}\hat{m}_{0k}(\hat{V}_\mathrm{m})_{jl} \\ & + \hat{m}_{0i}\hat{m}_{0j}\hat{m}_{0k}\hat{m}_{0l}.\end{aligned} \tag{B22}$$