Single-seismometer Focal Mechanism and Uncertainty Estimation from Body Waves

Victor Agaba * [1], **Suzan van der Lee** [1], **Madelyn Sita** [2], **Caio Ciardelli**[3], **Paula Babirye**[4]

[1]Department of Earth and Planetary Sciences, Northwestern University, Evanston, IL, USA, [2]Department of Chemistry, University of Virginia, Charlottesville, VS, USA

Author contributions: *Conceptualization*: SL, MS.. *Methodology*: VA, SL.. *Software*: VA.. *Formal Analysis*: VA, SL.. *Writing - Original draft*: VA, SL.. *Visualization*: VA, SL.. *Supervision*: SL..

Abstract Work on later, 200-word max.

Non-technical summary Work on later, shorter than abstract.

1 Introduction

[Suzan] Insert words here...
Writing tip: maximize information content of words

2 Cosine similarity misfit criterion

2.1 Background

When a quake event happens, a single seismometer reads P, SV and SH body waves through which the respective amplitudes A^P, A^{SV} and A^{SH} along with their noise levels σ_P, σ_{SV} and σ_{SH} (representing standard deviation of an assumed normal distribution centered at the amplitude value) can be extracted. Given a ray path's azimuth ϕ and take-off angles i, j of P and S waves respectively, we model the amplitudes as functions of the source mechanism's fault strike ψ, fault dip δ and slip rake λ, further accounting for the P- and S-velocities α_h, β_h at source depth h (Aki and Richards, 1980). The expressions are:

$$\begin{aligned} A^P &\sim \left(s_R(3\cos^2(i)-1) - q_R\sin(2i) - p_R\sin^2(i)\right)/\alpha_h^3 \\ A^{SV} &\sim \left(1.5s_R\sin(2j) + q_R\cos(2j) + 0.5p_R\sin(2j)\right)/\beta_h^3 \\ A^{SH} &\sim \left(q_L\cos(j) + p_L\sin(j)\right)/\beta_h^3 \end{aligned} \tag{1}$$

where

$$\begin{aligned} s_R &= 0.5\sin(\lambda)\sin(2\delta) \\ q_R &= \sin(\lambda)\cos(2\delta)\sin(\psi_r) + \cos(\lambda)\cos(\delta)\cos(\psi_r) \\ p_R &= \cos(\lambda)\sin(\delta)\sin(2\psi_r) - 0.5\sin(\lambda)\sin(2\delta)\cos(2\psi_r) \\ p_L &= 0.5\sin(\lambda)\sin(2\delta)\sin(2\psi_r) + \cos(\lambda)\sin(\delta)\cos(2\psi_r) \\ q_L &= -\cos(\lambda)\cos(\delta)\sin(\psi_r) + \sin(\lambda)\cos(2\delta)\cos(\psi_r) \end{aligned} \tag{2}$$

Our goal is two-fold:

i) Find estimates $\widehat{\psi}, \widehat{\delta}, \widehat{\lambda} \in [0°, 360°) \times [0°, 90°] \times [-180°, 180°)$ for the strike, dip and rake that *best* describe the collected data.

ii) Estimate the joint noise distribution of fault parameters to more accurately constrain the set of *acceptable* estimates for the source mechanism.

We thus need a metric against which to evaluate what the best and acceptable estimates are. The misfit function used by Sita and van der Lee (2022) is defined as the angle between the vector of observed amplitudes and that produced from a forward simulation of synthetic parameters. In this paper, we shall use a misfit Φ of negative cosine similarity between the two vectors which has a 1-to-1 correspondence with angle:

$$\Phi(\boldsymbol{m}) = \frac{-\boldsymbol{A}_o^\top \boldsymbol{A}_s}{\|\boldsymbol{A}_o\|\|\boldsymbol{A}_s\|} \tag{3}$$

where $\boldsymbol{A}_o = (A_o^P, A_o^{SV}, A_o^{SH})^\top$ are observed amplitudes, $\boldsymbol{A}_s = (A_s^P, A_s^{SV}, A_s^{SH})^\top$ are synthetic amplitudes, and $\boldsymbol{m} = (\psi, \delta, \lambda)^\top$ are the underlying parameters.

2.2 Justification

We are interested in situations where absolute amplitudes are either unknown or unnecessary, for instance if using data collected from NASA's Insight mission (Sita and van der Lee, 2022). In this case, we cannot know the absolute velocities because they depend on an understanding of Mars' interior that is good enough to build a reliable velocity model. Alternative misfit functions can account for the relative amplitudes, such as in equations 4 and 5:

$$\Phi(\boldsymbol{m}) = \sqrt{\left(\Delta\frac{A^P}{A^{SV}}\right)^2 + \left(\Delta\frac{A^{SV}}{A^{SH}}\right)^2} \tag{4}$$

*Corresponding author: victoragaba2025@u.northwestern.edu

$$\Phi(\boldsymbol{m}) = \sqrt{\log^2\left(\Delta\frac{A^P}{A^{SV}}\right) + \log^2\left(\Delta\frac{A^{SV}}{A^{SH}}\right)} \tag{5}$$

First, we note that using cosine similarity is more numerically stable in cases where one of the observed or synthetic amplitudes is close to zero since the norm of the associated vector in equation 3 may still be large enough to make division feasible.

Secondly, we have chosen cosine similarity over angle because it involves one less operation, which is more computationally efficient and easier to differentiate for directed search algorithms (see Section 3). The Gaussian function in equation 6 is similarly easy to differentiate, but it does not account for relative amplitudes so it would exclude otherwise acceptable solutions.

$$\Phi(\boldsymbol{m}) \sim |\Sigma|^{-1/2}\exp\left(-\frac{1}{2}(\boldsymbol{A}-\boldsymbol{A}_o)^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{A}-\boldsymbol{A}_o)\right) \tag{6}$$

2.3 Tolerance derivation

Given vectors $\boldsymbol{A}_s$ and $\boldsymbol{A}_o$, we would like to know the largest cosine similarity for which $\boldsymbol{A}_s$ is considered an acceptable fit with respect to the recorded amplitudes while accounting for the asymmetry of error levels for different amplitudes. Visualized in 3D space, we have a confidence ellipsoid centered at $\boldsymbol{A}_o$ (equation 7) such that acceptable fits are vectors which fall inside the elliptical cone tangent to it with vertex at the origin.

$$E(\boldsymbol{m}) = (\boldsymbol{A}-\boldsymbol{A}_o)^\top\tilde{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{A}-\boldsymbol{A}_o) = 1 \tag{7}$$

where $\boldsymbol{A} = (A^P, A^{SV}, A^{SH})^\top$ is arbitrary and $\tilde{\boldsymbol{\Sigma}} = \mathrm{diag}(\sigma_P^2, \sigma_{SV}^2, \sigma_{SH}^2)/\chi^2_{2,\alpha}$ for some $\alpha \in (0,1)$. The elliptical cone intersects with the ellipsoid at every boundary point for which

$$(\nabla_{\boldsymbol{A}}E)^\top\boldsymbol{A} = 2(\boldsymbol{A}-\boldsymbol{A}_o)^\top\tilde{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A} = 0 \tag{8}$$

Combining equations 7 and 8 gives the plane containing every point of tangency:

$$\frac{1}{2}(\nabla_{\boldsymbol{A}}E)^\top\boldsymbol{A} - f(\boldsymbol{A}) = \boldsymbol{A}_o^\top\tilde{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{A}-\boldsymbol{A}_o) = -1 \tag{9}$$

To account for asymmetry, we recognize that the orientation of $\boldsymbol{A}_s$ relative to $\boldsymbol{A}_o$ has an effect on the tolerance because the associated "thickness" of the ellipsoid changes. This information is captured by the plane containing vectors $\boldsymbol{A}_o$ and $\boldsymbol{A}_s$:

$$(\boldsymbol{A}_s\times\boldsymbol{A}_o)^\top\boldsymbol{A} = 0 \tag{10}$$

To find the constrained points of tangency, we can intersect the original ellipsoid with planes 9 and 10. The line of intersection is such that

$$\boldsymbol{A}(t) = (1-{}^1\!/_{c_1})\boldsymbol{A}_o + t\boldsymbol{D} \tag{11}$$

where c_1 (equation 12) is a location parameter projecting $\boldsymbol{A}_o$ onto plane 9 and $\boldsymbol{D}$ (equation 13) is the direction perpendicular to planes 9 and 10.

$$\begin{aligned}\boldsymbol{A}_o^\top\tilde{\boldsymbol{\Sigma}}^{-1}((1-{}^1\!/_{c_1})\boldsymbol{A}_o - \boldsymbol{A}_o) &= -1\\ \Rightarrow c_1 = \boldsymbol{A}_o^\top\tilde{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}_o &= \|\boldsymbol{A}_o\|^2_{\tilde{\boldsymbol{\Sigma}}^{-1}}\end{aligned} \tag{12}$$

$$\begin{aligned}\boldsymbol{D} &= (\boldsymbol{A}_o^\top\tilde{\boldsymbol{\Sigma}}^{-1})^\top\times(\boldsymbol{A}_s\times\boldsymbol{A}_o)\\ &= c_1\boldsymbol{A}_s - c_2\boldsymbol{A}_o\end{aligned} \tag{13}$$

where $c_2 = \boldsymbol{A}_o^\top\tilde{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}_s = \langle\boldsymbol{A}_o^\top, \boldsymbol{A}_s\rangle_{\tilde{\boldsymbol{\Sigma}}^{-1}}$.

Combining equations 7 and 11 produces a quadratic parameterized by the scalar t, from which we can obtain a quadratic equation for the points of tangency:

$$\begin{aligned}(t\boldsymbol{D} - {}^1\!/_{c_1}\boldsymbol{A}_o)^\top\tilde{\boldsymbol{\Sigma}}^{-1}(t\boldsymbol{D} - {}^1\!/_{c_1}\boldsymbol{A}_o) &= 1\\ (\boldsymbol{D}^\top\tilde{\boldsymbol{\Sigma}}^{-1}\boldsymbol{D})t^2 - {}^2\!/_{c_1}(\boldsymbol{A}_o^\top\tilde{\boldsymbol{\Sigma}}^{-1}\boldsymbol{D})t + ({}^1\!/_{c_1} - 1) &= 0\end{aligned} \tag{14}$$

Denote $c_3 = \boldsymbol{A}_s^\top\tilde{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}_s = \|\boldsymbol{A}_s\|^2_{\tilde{\boldsymbol{\Sigma}}^{-1}}$. Then by substitution we have $\boldsymbol{D}^\top\tilde{\boldsymbol{\Sigma}}^{-1}\boldsymbol{D} = c_1(c_1c_3 - c_2^2)$ and $\boldsymbol{A}_o^\top\tilde{\boldsymbol{\Sigma}}^{-1}\boldsymbol{D} = 0$. Solving then yields the necessary points (denoted $\boldsymbol{A}_b$) as

$$\boldsymbol{A}_b = \left(1-\frac{1}{c_1}\right)\boldsymbol{A}_o \pm \left(\frac{1}{c_1}\sqrt{\frac{c_1-1}{c_1c_3-c_2^2}}\right)(c_1\boldsymbol{A}_s - c_2\boldsymbol{A}_o) \tag{15}$$

Equation 15 shows that for the points to be well-defined, the following conditions must hold relative to the inner-product space defined by $\tilde{\boldsymbol{\Sigma}}^{-1}$:

- Strict Cauchy-Schwarz inequality: $c_1c_3 > c_2^2$. This also implies that $\tilde{\boldsymbol{\Sigma}} \succ 0$ and $\boldsymbol{A}_s \nparallel \boldsymbol{A}_o$. Practically, it means that all uncertainties are necessary and inversion methods that depend on forward simulation should treat parallelism as an edge case.
- Confidence ellipsoid cannot contain origin: $c_1 \geq 1$. Practically, we should expect errors to be a small fraction of observed amplitudes.

Thus, given 2 nonparallel vectors $\boldsymbol{A}_o$ and $\boldsymbol{A}_s$ and positive definite covariance matrix $\tilde{\boldsymbol{\Sigma}}$, we accept the simulation if $\hat{\boldsymbol{A}}_o^\top\hat{\boldsymbol{A}}_s \geq \mathrm{tol} = \min_b \hat{\boldsymbol{A}}_o^\top\hat{\boldsymbol{A}}_b$. The condition, more formally, is

$$\max_b \hat{\boldsymbol{A}}_o^\top(\hat{\boldsymbol{A}}_b - \hat{\boldsymbol{A}}_s) \leq 0 \tag{16}$$

where $\hat{\boldsymbol{A}}_o$, $\hat{\boldsymbol{A}}_s$ and $\hat{\boldsymbol{A}}_b$ are unit vectors.

3 Search Methods

3.1 Misfit gradient

The vector $\boldsymbol{A} = (A^P, A^{SV}, A^{SH})^\top$ can be treated as a continuously differentiable nonlinear function of parameters $\boldsymbol{m} = (\psi, \delta, \lambda)^\top$, defined by equation 1. We can thus apply directed search methods to minimize the following cost function:

$$\Phi(\boldsymbol{m}) = -\hat{\boldsymbol{A}}_o^\top\hat{\boldsymbol{A}} \tag{17}$$

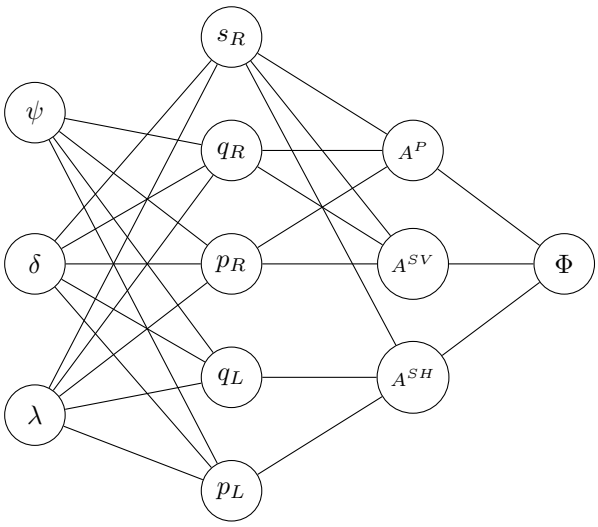

Figure 1 Computational graph of forward model (edges are non-zero partial derivatives).

To calculate the gradient, first note that $\boldsymbol{A}$ is a linear combination of nonlinear functions of $\boldsymbol{m}$ (equation 2).

Define a vector $\boldsymbol{\eta} = (s_R, q_R, p_R, q_L, p_L)^\top$. Then we can encode the misfit as

$$\Phi(\boldsymbol{m}) = \tilde{\Phi}(\boldsymbol{A}(\boldsymbol{\eta}(\boldsymbol{m}))) \tag{18}$$

This composition is visualized as a computational graph in figure 1. Each stage of the composition has an associated Jacobian, which we can demonstrate layer by layer:

i) From $\boldsymbol{m}$ to $\boldsymbol{\eta}$, we can use equation 2 to define the Jacobian as follows:

$$\begin{aligned}
&a = \sin(\lambda) && b = \cos(2\delta) && c = \cos(\lambda)\\
&d = \sin(2\delta) && e = \sin(\delta) && f = \cos(\delta)\\
&g = \sin(\psi - \phi) && h = \cos(\psi - \phi) && p = cos(2(\psi - \phi))\\
&q = \sin(2(\psi - \phi))
\end{aligned}$$

$$\boldsymbol{J_p}(\boldsymbol{\eta}) = \begin{bmatrix} 0 & ab & 0.5cd \\ abh - cfg & -2adg - ceh & cbg - afh \\ 2cep + adq & cfq - abp & -aeq - 0.5cdp \\ -cfh - abg & ceg - 2adh & afg + cbh \\ adp - 2ceq & abq + cfp & 0.5cdq - aep \end{bmatrix} \tag{19}$$

ii) From $\boldsymbol{\eta}$ to A, we have a linear transformation whose Jacobian is the matrix of coefficients from equation 1:

$$\boldsymbol{J_\eta}(\boldsymbol{A}) = \begin{bmatrix} \frac{3\cos^2(i)-1}{\alpha_h^3} & \frac{-\sin(2i)}{\alpha_h^3} & \frac{-sin^2(i)}{\alpha_h^3} & 0 & 0 \\ \frac{1.5\sin(2j)}{\beta_h^3} & \frac{\cos(2j)}{\beta_h^3} & \frac{0.5\sin(2j)}{\beta_h^3} & 0 & 0 \\ 0 & 0 & 0 & \frac{\cos(j)}{\beta_h^3} & \frac{\sin(j)}{\beta_h^3} \end{bmatrix} \tag{20}$$

iii) From $\boldsymbol{A}$ to Φ, we have a scalar function whose Jacobian is the gradient transposed.

$$\begin{aligned}
\nabla_{\boldsymbol{A}}\tilde{\Phi} &= \frac{\|\boldsymbol{A}\|\nabla_{\boldsymbol{A}}(-\hat{\boldsymbol{A}}_o^\top \boldsymbol{A}) - (-\hat{\boldsymbol{A}}_o^\top \boldsymbol{A})\nabla_{\boldsymbol{A}}(\|\boldsymbol{A}\|)}{\|\boldsymbol{A}\|^2} \\
&= \frac{-1}{\|\boldsymbol{A}\|}(\hat{\boldsymbol{A}}_o + \Phi\hat{\boldsymbol{A}})
\end{aligned} \tag{21}$$

Combining 19, 20 and 21 gives the overall gradient:

$$\nabla_{\boldsymbol{m}}\Phi = \boldsymbol{J_p}(\boldsymbol{\eta})^\top \boldsymbol{J_\eta}(\boldsymbol{A})^\top \nabla_{\boldsymbol{A}}\tilde{\Phi} \tag{22}$$

3.2 Optimization algorithms

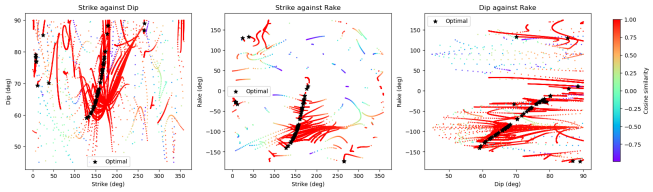

Figure 2 Results of steepest descent algorithm after a systematic sweep of starting points.

4 Uncertainty Quantification

4.1 Kagan angle distribution

An alternative way of quantifying model deviations.

4.2 Error propagation

Say something here, composed Jacobian is important » Earth 353, IEMS 401.
Emphasize the interest in modality.

4.3 Monte Carlo methods

Place to talk about different seeding stategies and their effect on uncertainty. Why hybrid search was used, why systematic over random.

5 Testing

5.1 Synthetic data

Say something here about the Monte Carlo

5.2 Earth data

From Uganda (Paula)

5.3 Mars data

From Insight (Suzan/Maddy)

6 Conclusions

6.1 Performance overview

Put some words here...

6.2 Growth avenues

Put some more words here...

Acknowledgements

[Suzan] Insert words here...

Data and code availability

Create a new repo with beautiful code, READMEs and everything. One function/algorithm at a time.

Competing interests

The authors have no competing interests.

References

Aki, K. and Richards, P. G. *Quantitative Seismology, Theory and Methods*, volume Vol. 854. Freeman & Co., 1980.

Sita, M. and van der Lee, S. Potential Volcano-Tectonic Origins and Faulting Mechanisms of Three Low-Frequency Marsquakes Detected by a Single InSight Seismometer. *Journal of Geophysical Research: Planets*, 127(10):e2022JE007309, 2022. doi: 10.1029/2022JE007309.