Phase Pilot Recovery: A Foundation for Mitigating Speckle Scattering Noise in Seismic Data

Akshika Rohatgi, Andrey Bakulin, and Sergey Fomel, Bureau of Economic Geology, University of Texas at Austin*

Summary

The complex near-surface scattering introduces significant distortions in deep reflection data, necessitating effective noise mitigation strategies. In this study, we investigate the recovery of signal phase in the presence of multiplicative noise which is a crucial step for subsequent despeckling algorithms. Leveraging a stack-based approach, we introduce the concept of a "phase pilot trace", which represents the trace closest to the signal phase using the minimum number of traces possible. Our analysis focuses on evaluating the quality of the phase pilot by assessing the standard deviation of the residual phase and its enhancement with increasing stack size. Through numerical experiments, we demonstrate a consistent reduction in phase spread across all frequencies, adhering to the 1/√N rule, indicating a uniform reduction in normalized values for all the frequencies. We propose utilizing the standard deviation of the residual phase as a metric for assessing the quality of phase pilot traces, providing insights into the necessary ensemble size for desired phase recovery quality. However, the pronounced frequency dependency adds complexity to determining the necessary trace count for broadband data.

Introduction

The study of small-scale scattering, recognized in the acoustic and ultrasonic fields (Abbott and Thurstone, 1979; Fink and Derode, 1998; Goodman, 2007), has predominantly focused on analyzing and mitigating noise in intensity images at a fixed frequency. However, seismic imaging operates with broadband signals, presenting unique challenges. Complex near-surface scattering can introduce significant distortions in deep reflection data (Bakulin et al., 2022). As a result, even after standard processing, land seismic data in complex areas exhibit weak and distorted pre-stack reflections with low coherency. Furthermore, both phase and amplitude exhibit huge variability. To model and explain these effects, a multiplicative random noise model based on the speckle mechanism of small-scale scattering has been proposed (Bakulin et al., 2022).

While the multiplicative model employed by surface-consistent deconvolution (Taner and Koehler, 1981; Cary and Lorentz, 1993) has proven useful in correcting for some phase variations in seismic data, it fails to address strong random-like phase variations caused by small- and medium-scale near-surface scattering. Measuring the local phase is a valuable attribute for analyzing seismic data. Non-stationary phase correction helps in identifying significant horizons and increasing their resolution in time (van der Baan and Fomel, 2009). However, local phase measurements are significantly impacted by speckle scattering noise, necessitating despeckling.

In our research, we focus on the crucial task of isolating a clean signal phase, which we identify as phase pilot trace recovery. This process is fundamental for any despeckling or speckle noise reduction algorithm in processing and imaging. Our study delves into a simple, stack-based approach for retrieving the phase pilot. Although this idea was previously mentioned, there was a lack of detailed investigation into how the pilot's quality is affected by the stacking volume, specifically the number of traces used. We evaluate the phase pilot's quality by analyzing the standard deviation of the residual phase and how this indicator of quality enhances as the stack size increases. Our numerical experiments demonstrate a consistent reduction in phase spread, adhering roughly to the 1/√N rule across all frequencies, indicating a uniform decrease in normalized values, independent of frequency. This finding illuminates the necessary trace count to attain the desired quality for the phase pilot. Moreover, we highlight the profound effect of frequency on phase spread within multiplicative noise, underscoring the added complexities. Our frequency-dependent analysis of the residual phase effectively characterizes multiplicative noise behavior across the seismic data's full frequency range, providing a basis for devising multiplicative noise mitigation strategies.

Multiplicative noise model

This model applies multiplicative noise to seismic traces within the Fourier domain as,

$$X_k(\omega) = S(\omega) * R_k(\omega), \tag{1}$$

where $S(\omega)$ is the clean signal and $R_k(\omega)$ denotes the random multiplicative noise. Equation (2) can model two types of multiplicative noise: random phase perturbations and random time shifts, with both assumed to follow a normal distribution akin to that observed in seismic field data,

$$R_k(\omega) = e^{i(\varphi_k + \omega\tau_k)} . \tag{2}$$

Here, φ_k represents phase perturbations varying independently across frequencies within a trace, leading to complex signal change of the ballistic arrivals in the time domain consistent with the field observations in complex regions. In contrast, residual statics introduce linear phase shifts proportional to a constant time delay τ_k for each trace. We emphasize that both types of noise solely perturb the phase of the signal in distinct ways while preserving the amplitude, which is crucial. In this study, our focus is on investigating the phase characteristics of seismic signals affected by multiplicative noise.

Figure 1a illustrates the ensemble of 50 clean traces containing three flat events each represented by a Klauder wavelet. These traces are subjected to random phase perturbations with a standard deviation of $\pi/3$ and random residual statics with a spread of 4 ms. The traces with multiplicative noise are illustrated in figure 1b.

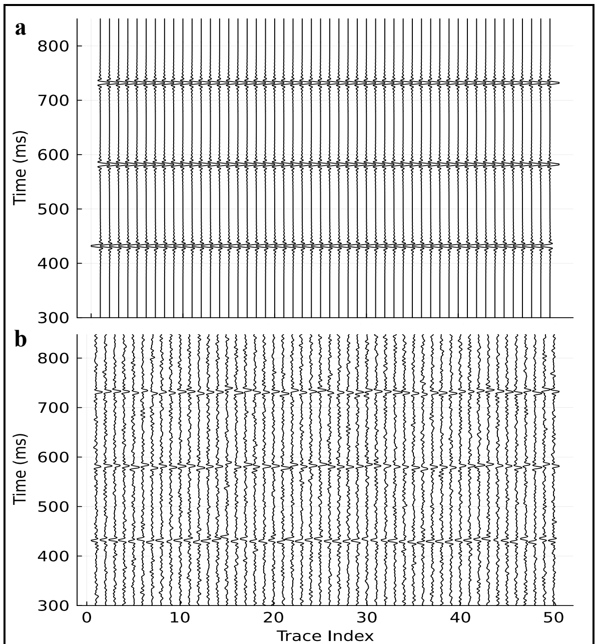

Figure 1. (a) Synthetic gather with three flat events each represented by Klauder wavelet. (b) Same as figure 1(a) but after perturbation with random multiplicative noise containing phase perturbations (spread of $\pi/3$) and residual statics (spread of 4 ms).

Transformation of multiplicative noise while stacking

For multiplicative noise, local stacking of traces with the same underlying signal but different noise is represented by the equation,

$$\hat{S}(\omega) = \frac{1}{K}\sum_{k=1}^{K}\{S(\omega) * R_k(\omega)\} . \qquad (3)$$

Considering that τ_k and $\varphi_k(\omega)$ are independent of each other and both random normally distributed with standard deviations σ_φ and σ_τ, the mathematical expectation of the stack can be written as (Bakulin et al., 2022),

$$E\left[\hat{S}(\omega)\right] = |S(\omega)| e^{i\varphi_s(\omega)} e^{-\frac{\omega^2\sigma_\tau^2}{2}} e^{-\frac{\sigma_\varphi^2}{2}} . \qquad (4)$$

We conclude that stacking recovers the clean signal phase while attenuating the amplitudes due to two factors: residual statics and phase perturbation. Berni and Roever (1989) previously derived exponential loss $e^{-\frac{\omega^2\sigma_\tau^2}{2}}$ by analyzing intra-array residual statics without employing a multiplicative noise model and without analyzing the phase. We executed stacking on ensembles with multiplicative noise at varying scales: 10, 100, 1000. Figures 2(a), 2(b), and 2(c) present these stacked traces alongside the clean signal in the time domain, where we notice a swift decrease in noise. This reduction, however, comes with the expected attenuation of higher frequencies as outlined by equation (4).

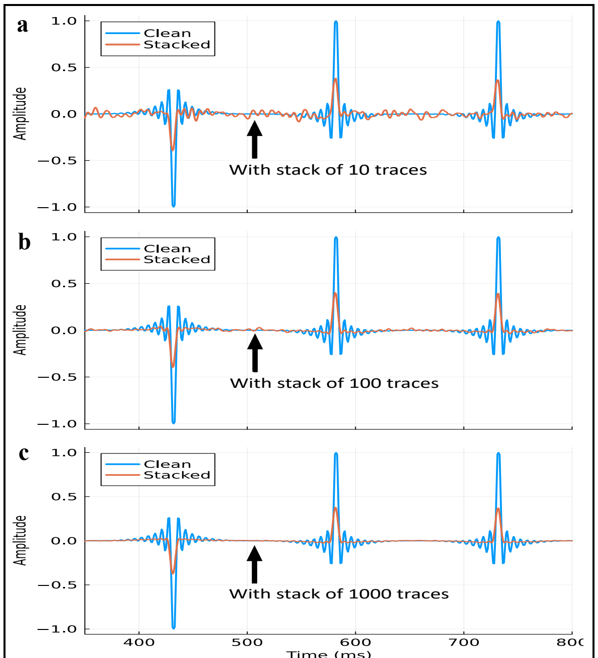

Figure 2. Stacked signal compared with clean signal in the time domain with the number of traces varying as (a) 10, (b) 100, and (c) 1000 in stack ensemble.

Phase pilot recovery

The multi-channel and redundant nature of pre-stack seismic data presents an opportunity to mitigate multiplicative noise. This process can greatly benefit from accurately recovering the correct signal phase, even in cases where the amplitude is diminished (Bakulin et al., 2023b). Khalil and Gulunay (2011) demonstrate the stack-based pilot is useful even if one wants to derive intra-array statics in single-sensor land data when analyzing direct arrivals. Bakulin et al (2021) explained that the phase derived from local stacking indeed provides an accurate estimation of the signal phase in the presence of random multiplicative noise caused by near-surface scattering. As a result, we present the task of creating a broadband "phase pilot trace", which is a trace with a phase that closely matches the signal phase, derived from the fewest possible traces. Then, we adopt the difference between the noisy and either the clean or pilot phase as our metric for evaluation and quantify with a simple standard deviation σ_{φ}^{pilot}. Equation (3) indicates that stacking is the most straightforward method for acquiring a phase pilot, highlighting that the average residual phase converges towards the signal phase. However, this approach does not clarify how the quality or standard deviation of the residual phase is influenced by the trace count. Thus, we undertake a numerical analysis to determine these relationships.

Focusing on our aim of phase pilot recovery, we proceed to assess the stacked phase. We executed stacking on ensembles with multiplicative noise at varying scales: 10, 100, 1000. Figure 3 demonstrates that increasing the number of traces gradually improves the phase accuracy. As the number of traces increases, there emerges the potential for phase recovery of the seismic signal, attaining the unbiased phases through stacking.

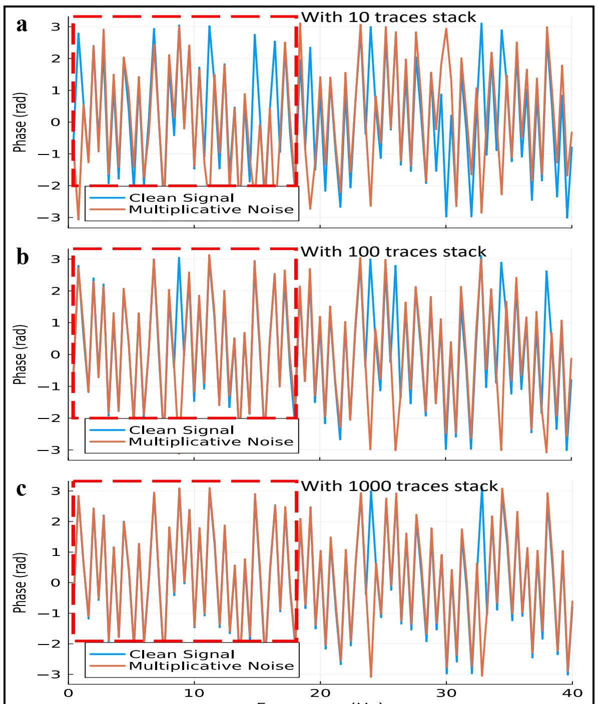

Figure 3. Phase spectra of stacked traces overlaid on the phase spectra of the clean signal. The number of traces stacked varies as (a) 10, (b) 100, and (c) 1000.

Yet, it's important to understand that figure 3 depicts just a single instance of the stacked ensemble. To calculate our targeted metric of σ_{φ}^{pilot}, a series of numerical experiments is required to produce various stacking outcomes for each ensemble size (*N*), from which we can then derive the mean and standard deviation. By repeatedly stacking, we mitigate the influence of random fluctuations and outliers. We conducted 10,000 iterations of the stack with the ensemble of traces varying from a single trace to 100 traces and examined how the σ_{φ}^{pilot} varies with the increase in the number of traces (Figure 4).

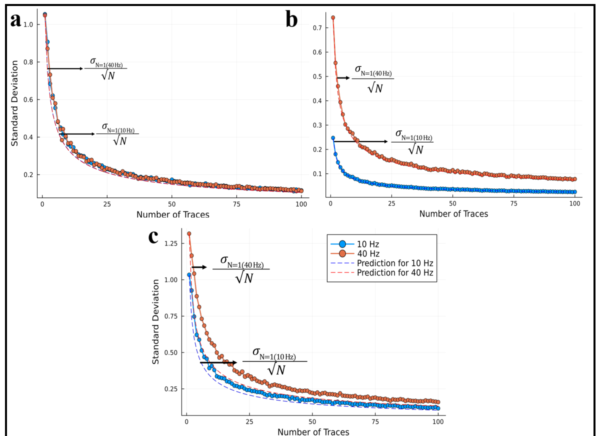

Figure 4. The standard deviation of residual phases post-stack vs number of traces at 10 Hz and 40 Hz for the noise with (a) random phase perturbations only, (b) random residual static only, and (c) both types of multiplicative noise.

To understand the two distinct types of multiplicative noise, comprising random phase perturbations (independent of frequency) and random residual statics (frequency-dependent), we initially examined these two noises separately before integrating them. Figure 4a shows the variation in standard deviation with respect to the number of traces, where the consistency of the curve across varying frequencies supports our claim regarding the frequency independence of random phase perturbations. Conversely, for traces with random residual statics only, the standard deviation varies with varying frequencies (Figure 4b). Subsequently, figure 4c presents the combination of both types of noises.

From figure 4, it's evident that the phase spread due to multiplicative noise diminishes roughly following the 1/√N rule. This law could be analytically derived for small-phase perturbations. By estimating the σ_{φ}^{pilot} of the raw data (N=1), we can apply this principle to determine the number of traces needed for stacking to meet the desired σ_{φ}^{pilot} threshold. This method offers a straightforward approach to defining the necessary ensemble size for reaching the desired quality of the phase pilot recovery. Nonetheless, the influence of σ_{φ}^{pilot} on frequency is significant and warrants further examination.

Phase pilot quality vs frequency

Having demonstrated the reduction of σ_{φ}^{pilot} with an increase in the number of traces at a constant frequency, we now explore how σ_{φ}^{pilot} varies across different frequencies within broadband seismic data.

With phase perturbations, we initially assume a constant standard deviation across all frequencies, leading to a uniform improvement in the quality of the pilot across the bandwidth (Figure 5a). On the other hand, residual statics exhibit a linearly increasing standard deviation σ_{φ}^{pilot} with frequency (Figure 5b), resulting in a linear relationship in the stacked phase but with a reduced slope proportional to ~ 1/√N. When both types of noises are present, the observed behavior represents a superposition of the two effects, yet the phase spread increases due to the impact of residual statics (Figure 5c).

Let us assume we aim for σ_{φ}^{pilot} to be less than 0.4 rad (22 degrees) as shown in figure 5c. This level of accuracy could be achieved with a stack-based phase pilot utilizing 10 traces but only for frequencies up to 20 Hz. Achieving a similar standard at 60 Hz would necessitate 100 traces. We can also see from figure 3 that while lower frequencies demonstrate phase recovery with as few as 10 traces, higher frequencies require a greater number of traces.

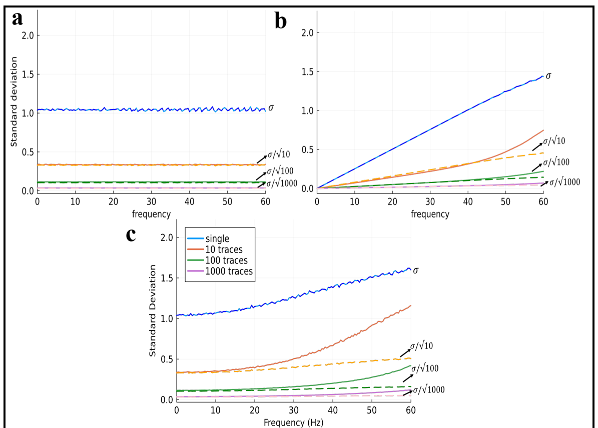

Figure 5. Standard deviation of residual phases post-stack (σ_{φ}^{pilot}) variation with frequencies for the stack of 10, 100, and 1000 ensemble of traces with (a)- random phase perturbations only, (b)- random residual static only, and (c)- Both types of multiplicative noise.

Given that the standard deviation is anticipated to increase with frequency (Bakulin et al., 2023b), there's a call for frequency-dependent processing approaches like those described by Retaillue et al. (2014). Exploring alternative machine-learning algorithms might offer a superior quality improvement compared to traditional stacking, particularly for higher standard deviations where phase enhancement becomes less effective than 1/√N (notably at higher frequencies in figure 5c).

Discussion

The specific threshold for phase recovery σ_{φ}^{pilot} varies by application and frequency band, necessitating dedicated research. Applications like beamforming and despeckling employ hundreds of traces as noted by Bakulin et al. (2023a), but perhaps do so to condition the amplitude. Cary and Nagarajappa (2013) highlighted that surface-consistent deconvolution is compromised by high phase instability, underscoring the need for data conditioning to achieve a certain level of phase stability for time processing. Different applications, such as pre-stack seismic inversion, might demand varying degrees of phase stability. Additionally, failing to correct for small-scale velocity variations can impair migration, as Xe et al. (2016) observed, suggesting that depth migration might need its own threshold for phase errors. Holt and Lubrano (2020) identified phase instability as a significant challenge in seismic processing and interpretation, particularly onshore. By defining acceptable phase error thresholds for each application and mitigating speckle noise within those bounds, we aim to facilitate the processing and inversion of complex data in scattering geological settings.

Conclusions

We have shown that it is possible to achieve an unbiased estimation of the signal phase amidst multiplicative noise using straightforward stacking techniques. Establishing phase pilot traces with accurate phases across all frequencies lays the groundwork for subsequent despeckling or denoising algorithms aimed at speckle noise. We suggest utilizing the standard deviation of the residual phase as a metric for assessing the quality of these phase pilot traces. This metric, measurable directly from the original data, is anticipated to reduce in proportion to 1/√N, provided the perturbations remain manageable. With an initial understanding of the phase spread, we can determine the necessary number of traces for stacking to attain the desired stacked phase quality. Yet, the pronounced frequency dependency introduces complexity in this determination for broadband data and might necessitate excessively large stacking ensembles to achieve our objectives. While stacking represents the most fundamental method for signal phase recovery, exploring alternatives, including nonlinear and machine learning (ML) approaches, could potentially offer improved phase recovery efficiency with fewer traces required. This underscores the importance of adopting frequency-specific seismic processing methods to address the challenges of multiplicative noise, thereby enhancing the precision of subsurface imaging and reservoir characterization efforts.