General Seismic Architecture of the Southern San Andreas Fault Zone around the Thousand Palms Oasis from a Large-N Nodal Array

Pieter-Ewald Share*1, Hongrui Qiu2, Frank L. Vernon3, Amir A. Allam4, Yuri Fialko3, and Yehuda Ben-Zion5

Abstract
We discuss general structural features of the Banning and Mission Creek strands (BF and MCF) of the southern San Andreas fault (SSAF) in the Coachella Valley, based on ambient noise and earthquake wavefields recorded by a seismic array with >300 nodes. Earthquake P arrivals show rapid changes in waveform characteristics over 20–40 m zones that coincide with the surface BF and MCF. These variations indicate that the BF and MCF are high-impedance contrast interfaces—an observation supported by the presence of seismic reflections. Another prominent but more diffuse change in SSAF structure is found ∼1 km northeast of the BF. This feature has average-to-low arrival times (P and S) and ambient noise levels (at <30 Hz), and likely represents a relatively fast velocity block sandwiched between broader MCF and BF zones. The maximal arrival delays (P ∼0.1 s and S ∼0.25 s) and the highest ambient noise levels (>2 times median) are consistently observed southwest of the BF—a combined effect of Coachella Valley sediments and rock damage on that side. Immediately northeast of the MCF, large S minus P delays suggest a broad high Vp/Vs zone associated with asymmetric rock damage across the SSAF. This general overview shows the BF and MCF as mature but distinctly different fault zones. Future analyses will further clarify these and other SSAF features in greater detail.

Introduction
The southern San Andreas fault (SSAF) is estimated to pose one of the largest seismic risks in California (e.g., Weldon et al., 2005; Field et al., 2017). Clarifying the structural architecture and seismic properties of this major fault (Catchings et al., 2009; Lindsey and Fialko, 2013; Ajala et al., 2019) can improve the estimates of potential magnitudes and shaking of future large earthquakes. Several key structural characteristics of the SSAF in the Coachella Valley remain unclear. These include the fault dip at depth (e.g., Fialko, 2006; Fuis et al., 2017; Nicholson et al., 2017), the distribution of fault-related damage and slip on different fault strands (Fumal et al., 2002; Blinskiuk et al., 2021), and characteristics of velocity contrast interfaces in the fault zone (e.g., Qiu et al., 2019). To address these issues and better constrain SSAF structures, we analyze the seismic waveform from local earthquakes (Fig. 1a) recorded across a large-N array spanning a northern multistranded section of the SSAF near the Thousand Palms Oasis Preserve in the Coachella Valley (Fig. 1b). To compensate for the relative paucity of seismicity in the area, we also analyze data from...
regional and teleseismic earthquakes (Fig. 1a, inset) and ambient seismic noise.

The >300 node array transected both the Banning (BF) and Mission Creek (MCF) fault strands of the SSAF near the Thousand Palms Oasis, California, about 10 km northwest of Biskra Palms (e.g., Behr et al., 2010) where these two strands merge (Fig. 1a). Near the surface, the BF and MCF, together with other minor faults, comprise a >2 km wide fault zone embedded in Pliocene and Pleistocene stratified rocks, with thick Quaternary Coachella Valley sediments to the southwest and thinner sediments immediately northeast of the MCF (Rymer, 2000). Cretaceous plutonic rocks of the Peninsular Ranges underlie the Valley sediments (e.g., Matti et al., 1992; Ajala et al., 2019), whereas older metamorphic and igneous intrusive rocks of the Little San Bernardino Mountains likely abut the MCF beneath the shallow sediments to the NE (Catchings et al., 2009). Around the study site little is known about internal features of the SSAF structure, including the width and depth of core and broader damage zones, deeper attitudes of the BF and MCF, seismic velocity ($V_p$ and $V_s$) contrasts and variations across the fault zone, and the presence of crustal fluids. There is an evidence for a steep northeast dipping electrically conductive SSAF zone in the upper crust (~3 km wide at the surface) from electromagnetic imaging (Share et al., 2021). Active source seismic exploration and potential field results indicate similar near-surface geometries for the surrounding SSAF (e.g., Catchings et al., 2009; Fuis et al., 2017). To understand these and other features better, we apply a range of tools well-suited for imaging internal fault zone features (e.g., Ben-Zion et al., 2015; Share, Allam, et al., 2019; Qin et al., 2021; Qiu et al., 2021). In this article, we focus on general aspects of the recorded data such as the arrival time, frequency, amplitude, and phase variations across the array from direct, reflected, and transmitted seismic phases.

**Figure 1.** Study area in the context of the southern San Andreas fault (SSAF). (a) Major fault surface traces (black lines) and local earthquakes (circles; colors indicate depth) analyzed in this study. $P$ waveforms from all events are analyzed, whereas $S$ waveform data from only 14 events (thick-outlined circles) are used. Labeled example events are shown in Figures 2–3. The dense array location is indicated (red box) along with four subsidiary stations north of the array that are not used in this study (red triangles). The eastern California shear zone (ECSZ), the town of Palm Springs (PS), San Gorgonio Pass (SGP), and the San Jacinto fault zone (SJFZ) are shown for reference. The global map (inset) shows the eight regional (epicenters <30° away) and teleseismic (epicenters 30°–100° away) earthquakes employed in this study (circles), with example events labeled. (b) Zoom in on the red box in panel (a) showing the Banning (BF) and Mission Creek (MCF) faults and the nodal array (red triangles). Some nodes along the long linear profile are highlighted for reference, and the nearby Thousand Palms Canyon Road is also shown.
Data
The nodal array (Fig. 1) was deployed in the beginning of March 2020 and consisted of 322 Zland three-component 5 Hz geophones recording at 500 samples per second (see Data and Resources). Most of these nodes recorded into early April. The array configuration included two 100+ node 2D subarrays with apertures of 0.6–1 km centered on the BF and MCF (Fig. 1b). Each subarray included a grid of regularly spaced nodes, as well as several nodes scattered around that grid. In addition, we deployed a ~4 km long 100+ node quasi-linear profile crossing both the strands and connecting the two subarrays (Fig. 1b). The internode spacing varied from ~15 m around fault traces to ~100 m away from them. Four nodes were installed farther northeast (Fig. 1a) to help constrain local seismicity in the region, but data from these nodes were not analyzed in this general study.

During acquisition, the array recorded several local earthquakes of which we analyzed high-quality data from 27 M > 1 events that had diverse hypocentral distances (<100 km) and azimuths relative to the array (Fig. 1a; Table S1 available in the supplemental material to this article). In addition, we analyzed data from five moderate-to-large regional earthquakes (<30° away), including the M 5.7 Magna (Utah) and M 6.5 Stanley (Idaho) events, and three teleseismic events (30°–100° away) (Fig. 1a, inset; Table S2). Earthquake metadata were obtained from the Southern California Earthquake Data Center (see Data and Resources). Three-component waveforms associated with these events were extracted from the continuous recordings and detrended. In addition, a 2–20 Hz bandpass filter was applied to the local waveforms, and a 1 Hz lowpass filter was applied to the teleseismic and regional waveforms. The lower frequency data provide information about large-scale variations across the linear profile, and the higher frequency waveforms give insight into smaller scale complexities along this profile and across the two dense subarrays. The entire continuous dataset was used during ambient noise analysis.

Analysis and Results
To obtain a general sense of large-scale structural changes across the linear profile, we analyzed early waveform similarities and delay times of teleseismic and regional P arrivals. For a given event, the first P maximum or minimum coherent across vertical component recordings were picked and designated P arrivals (Fig. 2a). Each 3 s early P waveform (0.5 s before and 2.5 s after P pick) was then cross-correlated with every other 3 s P waveform in the linear array, and the maximum correlation coefficients were cataloged. The delays corresponding to the maximum correlation coefficients across the array were consistent with the delays obtained through manual picking. High correlation coefficients between neighboring stations imply similar large-scale structure beneath those stations. Low coefficients pinpoint a transition between two dissimilar media (e.g., Qin et al., 2021).

Cross-correlation results for three representative events are shown in Figure 2b. Given different azimuths and inclination angles of the incoming wavefronts, the cross correlations exhibit some variability between events. Nevertheless, taken together, Figure 2 suggests three locations of large-scale structural changes. One of these occurs over a 20–40 m wide zone near a mapped central strand of the BF (station 21 in Fig. 1b) and another around the MCF proper over a similarly narrow zone. A third region of rapid waveform changes is associated with a yet unknown large-scale structure around station 60 (Figs. 1b and 2). We, therefore, classify the site in terms of four crustal zones, one southwest of the BF, a second from the mapped BF to about ~1 km in the northeast, a third from that point to the mapped MCF, and the final zone extending from the MCF northeastward.

Next, we inspected the delay-time results from all teleseismic and regional events to help quantify the general P-wave velocity (V_p) variations across the profile. For each event, the picked arrival times were corrected for array geometry by subtracting the predicted travel times to individual stations using the TauP algorithm (Crotwell et al., 1999) and the IASP91 global velocity model (e.g., Share, Allam, et al., 2019; Qin et al., 2021; Qiu et al., 2021). A correction was then made for elevation above sea level by subtracting the excess travel times to elevated stations calculated with a reference near-surface V_p = 2 km/s (from Ajala et al., 2019). In this study, the choice of reference V_p has a minor effect on delay-time estimates, as the change in elevation from station 1 to 151 is only 138 m (see Data and Resources), which is a small fraction of the ~4 km profile length and probably below the length-scale resolution of the kilometer-long wavelengths of the teleseismic and regional waves. Finally, for each event, the median delay of the array was subtracted from individual station delays to provide a relative measure of delay (slowness) across the array. After corrections, the median of all event delay times shows a clear trend of increasing slowness (decreasing V_p) from northeast to southwest (Fig. 3a). This likely reflects a large-scale change in upper crustal structure from crystalline bedrock in the northeast (high V_p;
Catchings et al., 2009) to thick Coachella Valley sediments in the southwest (low $V_p$; Ajala et al., 2019) (Fig. 3a).

To provide information on smaller scale complexities, we analyzed next the higher frequency information provided by local earthquake arrivals. For the 27 high-quality events, we picked $P$ arrivals manually for linear array stations on vertical component recordings (Fig. 3b). For the 14 clear $S$ arrival producing events, we picked $S$ arrivals on both the horizontal components, and took the average of the two picks to produce a single $S$ pick per event and per station (Fig. 3b). These picks were corrected for their respective propagation paths using the

Figure 2. Teleseismic and regional earthquake analysis. (a) Vertical component $P$-wave recordings from three example teleseismic and regional earthquakes (e1–e3, locations in Fig. 1a, inset) and their manual picks (red triangles). The solid, dark-red lines at stations 21 and 98 indicate mapped BF and MCF locations, respectively. The dashed line at station 60 highlights another region of rapid changes in waveform characteristics. The zoom in on event e1’s early $P$ waves (−0.3 to 0.3 s) highlights the change in delay times from stations 1 to 100 for these low-frequency waveforms. (b) The maximum cross-correlation coefficients between early $P$ waveforms (3 s long) across the linear array for each event in panel (a). Note the change in color scale. Gray lines are missing stations for each event.
3D $V_P$ and $V_S$ models of Share, Guo, et al. (2019). Again, for each event and each phase, the median delay was subtracted from individual station delays to provide a relative measure of $P$ and $S$ delays across the array. The resultant median $P$ and $S$ delay times for all the events (Fig. 3a) reflect upper crustal features not captured by the regional 3D velocity models, that is, internal features of the SSAF structure at the site.

In addition to a general increase in delay (slowness), as shown by the teleseismic and regional earthquake results, the local delay-time curves reveal two other key characteristics; peak $P$ and $S$ delays (~0.1 and ~0.25 s) within 200 m southwest of the BF and, relative to $P$, rapidly increasing $S$ delays within a >500 m wide zone around the BF and immediately northeast of the MCF within a similarly sized zone. These are relatively high $V_P/V_S$ regions, as the $S$ arrivals Figure 3. Delay-time analysis and local earthquake waveforms. (a) Median delay-time curves for all teleseismic and regional (8 total, cyan), local $P$ (27 total, blue), and local $S$ (14 total, red) arrivals. Blue and maroon arrows show the peaks in local $P$ and $S$ delays. Error bars equal 1 standard error converted from median absolute deviations (MADs) for each station using $\sqrt{\pi/2} \times \text{MAD}/\sqrt{N}$, in which $N$ is the number of events. The black trendline indicates average slowness in the uppermost 1 km of the $V_P$ model of Ajala et al. (2019), a proxy value for increasing sediment thickness toward the southwest. Vertical, black-solid and dashed lines from left to right are the BF, station 60, and the MCF. (b) $P$ and $S$ waveforms from example local earthquakes (e4–e6) and the manual arrival picks made on them (red triangles). First three panels are $P$, the last is $S$. Instead of a 2–20 Hz bandpass ($P$), a 2–10 Hz filter is applied to $S$ for better waveform illumination. See Figure 1a for e4–e6 locations, and Figure S1 for $S$ waveforms and picks associated with both horizontal components of example events e5 and e6. Horizontal, dark-red lines have the same meaning as in Figure 2a. Yellow curves indicate seismic reflections off the MCF (e4) and BF (e6).
are disproportionately delayed compared with $P$. The cause of high $V_P/V_S$ southwest of BF is likely a combination of thick sediments (usually $V_P/V_S \gg 2$; Brocher, 2005), rock damage around that fault, and associated fluids. Sediments cannot be the only cause, as they gradually increase in thickness (Fig. 3a) up to ~2 km in the central Coachella Valley (west of the study site; Ajala et al., 2019) but the observed $P$, $S$, and $S$ minus $P$ delays peak around the BF. This suggests that rock damage along the BF is a contributing factor. There is significantly less and decreasing sediment coverage northeast of the MCF, and we associate most of the high $V_P/V_S$ there to asymmetric rock damage on that side, which is potentially infiltrated by the same fluids feeding the nearby oases. The asymmetric nature of brittle damage is supported by the presence of several mapped minor faults northeast of the MCF and none immediately to the southwest (Fig. 1b).

Assuming this high $V_P/V_S$ structure extends to a depth of 4 km implies a $V_P/V_S > 2.15$, which is the average $V_P/V_S$ in the uppermost 4 km of the local crust in the Share, Guo, et al. (2019) models. A depth of 4 km is approximately the extent of damage-related high electrical conductivities in the area (Share et al., 2021) and fault zone damage along other major faults (e.g., Qin et al., 2021; Qiu et al., 2021).

Local earthquake full waveforms provide additional information on fault zone characteristics (Fig. 3b). The region around station 60 acts as an inflection point in which incoming wavefronts undergo a sudden change in azimuth and/or apparent slowness. The supplemental material illustrates similar rapid changes in three-component motion around station 60 for an example teleseismic event (Movie S1). Also, $P$ and $S$ reflected phases emanate from the BF and MCF but not the region around station 60. These show that the BF and MCF are sharp, high-impedance contrast interfaces, whereas the region around station 60 likely represents a more diffuse change in structure.

Finally, we quantified variations in the background ambient noise levels across the array at different frequencies, which provides a multiscale measure of near-surface ground amplification and attenuation properties. In addition to the typical low-frequency ambient noise from remote natural sources (e.g., ocean waves), there are strong local sources producing noise at frequencies >1 Hz. These include primarily frequent daily road and rail traffic in the Coachella Valley (Brenguier et al., 2019), and traffic along the Thousand Palms Canyon Road (Fig. 1b). The latter is near-parallel to the linear profile along most of its length and is offset to the northwest by ~200 m, providing a relatively uniform source of noise to the entire array. For each of the frequency ranges <5, 5–10, 10–15, 15–20, 20–25, and 25–30 Hz, we applied a bandpass filter to each 24 hr recording for each station and calculated the median of the respective energy envelopes. This was done separately for the vertical, and each horizontal component recording after the latter were rotated by 38° E of north to align with mean fault-normal and fault-parallel coordinates.

Over a wide frequency range, the ambient noise results corroborate a change from relatively stiff (crystalline bedrock) in the northeast to relatively compliant (thick unconsolidated Coachella Valley sediments) in the southwest, with the most rapid increase in noise level toward the southwest around BF and the Valley sediments producing the maximum noise level (>2 times the medial level; Fig. 4a). The effects of local near-surface structures on ambient noise recordings are clearly greater than the proximity to local high-frequency noise sources, as, for example, the Thousand Palms Canyon Road is closest to the intersection of the linear profile and the MCF (Fig. 1b), but the region southwest of the BF, and not the MCF, consistently experiences the highest noise levels (Fig. 4a). At frequencies lower than 15 Hz, there are some elevated noise levels in a ~100 m wide zone immediately northeast of the MCF (Fig. 4b, Movie S2), potentially the surface manifestation of a core damage structure within the broader asymmetric damage zone (e.g., Fig. 3a). Higher median absolute deviations of the noise field around the BF (Fig. 4a) point to greater time-dependent near-surface structural changes in response to atmospheric and other sources (e.g., rainfall, air pressure, etc.) (Movies S3–S4). Moreover, the 5–10 Hz band shows rapid variations in fault-parallel minus fault-normal ground motions within this zone and especially around station 21 (Fig. 4b; Movie S5). This suggests that the broader BF zone has anisotropic properties. Finally, the region around station 60 has associated average-to-low ambient noise levels (Fig. 4), highlighting a lack of unconsolidated sediments and/or damage-related materials in this zone of diffuse structural change.

**Discussion and Conclusions**

At large scale, seismic velocities generally decrease from northeast to southwest, as near-surface crystalline bedrock (low ambient noise levels) transitions to unconsolidated Coachella Valley sediments (the highest ambient noise levels) (Figs. 3 and 4; Movies S3–S4). This general change is not smooth and includes abrupt changes in structure near a central
strand of the BF strand (station 21), the MCF proper (station 98), and in a region around station 60 (Figs. 2 and 3b; Movie S1). Also present are >500 m wide low velocity, high $V_p/V_S$ fault damage-related structures around the BF and immediately northeast of MCF (asymmetric) (Fig. 3a). The total contribution of fault damage to the elevated high $V_p/V_S$ around the BF is unclear, but it is likely the main cause northeast of the MCF. On that side of the MCF, the upper crustal fractures and pores within this region of active damage production may be infiltrated by crustal fluids from various sources (e.g., Catchings et al., 2009; Share et al., 2021). The region around station 60 is characterized by average-to-low slowness, $V_p/V_S$, and ambient noise levels, which we interpret as a relatively intact faster velocity block sandwiched between broader BF and MCF zones.

Overall, both the BF and MCF exhibit crustal structures characteristic of mature faults. The BF strand separates two crustal blocks of very different material properties, whereas the MCF is actively deforming (e.g., Blisniuk et al., 2021) and likely has strong asymmetric damage. The latter is a signature of persistent ruptures along a bimaterial interface with greater damage generation on the competent rock side (crystalline basement in northeast) of the interface (e.g., Ben-Zion and Shi, 2005; Dor et al., 2008; Share, Allam, et al., 2019). For a dextral strike-slip earthquake, this bimaterial contrast polarity may induce preferred sub-shear velocity propagation in the direction of motion of the slow compliant block (Shlomai et al., 2020), that is, toward the northwest in this case (e.g., Bombay

Figure 4. Ambient noise analysis. (a) Median ambient noise levels (vertical component) in different frequency bands over roughly a month-long period (Julian day 63–90). As an example, MADs for the 10–15 Hz band are also shown. This indicates median variability across the month-long window. Vertical black lines have the same meaning as in Figure 3a. Note the increase in MAD and variability around the BF and to the southwest of that fault. (b) Median ambient noise levels in map view for the same time window at 5–10 Hz for the entire array (left two panels) and at 10–15 Hz focused on the BF and MCF (right two panels). The short, dashed-red line in the left two panels indicates the location of station 60 and a region of relatively diffuse change in structure. Note the change in color scale between the panels. All the panels show vertical component motions, except the second, which indicates fault-parallel minus fault-normal motions. Movies associated with these four panels are in the supplemental materials (Movies S2–S5).
Beach M 7.8 earthquake scenario; Field et al., 2017). We observe that this velocity contrast polarity is opposite to that observed at ∼10 km depth along a major structure northeast of the mapped SSAF (Qiu et al., 2019; Share, Guo, et al., 2019), which implies depth-dependent contrast polarities if the deeper structure is a downdip continuation of the SSAF and subsequently more complex earthquake rupture dynamics. Moreover, relatively large (>500 m) widths of low-rigidity zones associated with the BF and MCF (Fig. 3) are consistent with a mature fault system impacted by multiple large seismic events (e.g., Mitchell and Faulkner, 2009), resulting in pervasive off-fault damage. Additional details on the seismic structure of the SSAF in the study area can be obtained by analyzing fault zone head and trapped waves (e.g., Qin et al., 2021; Qiu et al., 2021), reverse-time migration of fault-reflected phases, and migration of the scattered wavefield (Touma et al., 2021). Analyses of such signals are the subject of continuing research.

Data and Resources
The nodal array data and associated metadata are freely available for download from the Incorporated Research Institutions for Seismology Data Management Center (IRIS DMC) (Vernon et al., 2020). Local and teleseismic earthquake metadata obtained from the Southern California Earthquake Data Center (SCEDC, 2013) and used during analysis are summarized in Tables S1 and S2. The supplemental materials also include Movies S1–S5 and Figure S1. Movie S1 is a three-component motion animation of event e2 in Figure 2. Movies S2–S5 are time lapeses over a roughly month-long window of the same data used to produce Figure 4b (in this subfigure, the medians over that month are shown). Figure S1 shows S-wave arrivals and picks for local events e5 and e6 (Fig. 3b).

Declaration of Competing Interests
The authors declare no competing interests.

Acknowledgments
The authors thank Deborah Rogers and Ginny Short (Center for Natural Lands Management) and Brian Ousley (Bureau of Land Management) for help with permitting at the Thousand Palms site. They also acknowledge the IRIS Pascall Instrument Center and the University of Utah Seismograph Stations for providing nodal seismometers. Constructive comments from Editor Keith Koper and two anonymous reviewers helped to improve this article. The deployment and retrieval of nodes were made possible through the dedication and hard work of students, postdoctoral researchers, faculty, and staff from the University of Southern California, Southern California Earthquake Center, University of California San Diego (Scripps Institution of Oceanography), The University of Utah, San Diego State University, the U.S. Geological Survey (USGS), and the NASA Jet Propulsion Laboratory. This study was supported by the National Science Foundation (Grants EAR-1841315 and 1841273), and Pieter-Ewald Share was also supported by a Green Foundation Postdoctoral Fellowship (UCSD/Scripps).

References


