APPENDIX

Impurity Ionization

The influence of impurities on electrical conduction is largely determined by their ionized concentration in comparison to the intrinsic carrier density. The intrinsic carrier density of any non-degenerate semiconducting mineral n_i is simply:

$$n_i = \sqrt{N_c N_v}\exp\left(-E_g/2k_bT\right) \qquad (1)$$

where N_c and N_v are the CB and VB density-of-states (DOS) (cm^{-3}) respectively, k_b is the Boltzmann constant (J/K) and T is the lattice temperature (K) (Sze 1985). Dasbach et al. (1994) determined the DOS for pyrite using Seebeck and Hall-Effect measurements on synthetic p-type material combined with current-voltage measurements on Schottky diodes. Values obtained at 300 K were N_c $\sim 8.5\times10^{19}$ cm^{-3} and N_v $\sim 3\times10^{18}$ cm^{-3} which are in approximate agreement with theoretical band structure estimates by Zhao et al. (Zhao et al. 1993). For the best-measured bandgap of 0.95 eV (Schiek et al. 1990) the intrinsic carrier density using equation (1) is therefore close to a value of $2\times10^{12}cm^{-3}$ as used by Altermatt et al. (2002) for simulating the photovoltaic response of pyrite. Importantly, for pyrite, the majority carrier concentration, resistivity and semiconducting type are all sensitive to ionized extrinsic impurities greater than this value (Sze 1985).

The fraction of impurities ionized $N_{d,a}^{+/-}/N_{d,a}$ at a particular temperature depends on the activation energy $\Delta E_{d,a}$ of each level which is simply the difference between the respective Fermi level $E_F^{n,p}$ and dopant level $E_{d,a}$ as shown in **Figure 3**. Although measurements for most sulfide levels remain unknown, the observation of p and n-type conductivity associated with certain impurities such as As and Co means they must be close to the band edge to be thermally ionized at room temperature nearly all measurements are made. The ionized impurity fraction is simply:

$$N_{d,a}^{+/-} = N_{d,a}/(1 + g_{d,a}\exp\left(\Delta E_{d,a}/k_bT\right)) \qquad (2)$$

where $N_{d,a}$ are the dopant and acceptor impurity concentrations here measured with the electron probe, $g_{d,a}$ are the level degeneracies and k_bT is the lattice thermal energy (eV) which is 0.026 eV at room temperature (Sze 1985). Inherent in the above statement is the rather gross assumption that the total concentration of impurities is substitutional but without recourse to models or an elaborate series of experiments this will have to suffice. A literature review suggests relevant dopant and acceptor energy levels as well as their degeneracies remain largely unknown in sulfides. Accurately calculating the ionized fraction is unfortunately not possible. However, for the large donor-acceptor concentrations in the arsenian pyrite studied here, activation energies

also shrink by an amount $\alpha_{d,a}(N_a + N_d)^{1/3}$ where α is a constant. This mechanism is called Band Gap Narrowing (BGN) and is related to the tail of the distribution of states associated with the impurity reaching into the relevant band (Sze 1985). The higher the impurity concentration the more $N_{d,a}^{+/-} \rightarrow N_{d,a}$ allowing elemental maps coupled with knowledge of their impact on doping, to be used to estimate effective doping. This simple assumption does of course break down when the impurity in question approaches its solubility limit. Our inability to measure S and Fe vacancies is another large source of error. Within this simplified schemata the overall dopant density $D\ (x, y)$ is simply:

$$D(x,y) = \sum_{i=0}^{N} N_a^-\ (x,y) - \sum_{i=0}^{M} N_d^+\ (x,y) \qquad (3)$$

where N and M are the number of ionized p and n-type impurities whereas N_a^- and N_d^+ are their respective densities at room temperature. A positive difference results in *p*-type behavior whilst a negative one gives overall *n*-type conduction. This equation together with several assumptions based on (1) and (2) are used to predict extrinsic doping in the Otago assemblage as shown in the main text.

Photon Absorption in Pyrite

Above band-gap photon absorption in pyrite occurs concurrently with the promotion of a free electron-hole pair; the general trend is known to follow the general quadratic rule for an indirect bandgap (Abraitis et al. 2004;Kou et al. 1978). Optical measurements by Karguppikar et al. (1988) suggest the indirect gap has an absorption coefficient α for phonon-assisted transitions given by:

$$\alpha hv = \frac{A\left(hv - E_g + E_p\right)^2}{e^{\theta_D/T} - 1} + \frac{B\left(hv - E_g - E_p\right)^2}{1 - e^{-\theta_D/T}} \qquad (4)$$

where A and B are coefficients, E_p is the single phonon energy associated with the transition across the gap and θ_D is the Debye temperature of 610 K. The 1st and 2nd terms represent phonon absorption and emission with respective coefficients of A=1.134×10^3 and 2.07×10^4 with an E_p value of 0.08 eV. For above gap laser excitation used here $hv > E_g$ and phonon emission dominates. As an aside, the extremely high absorption coefficient for pyrite in the visible spectrum makes it an ideal candidate for thin-film solar cells (Ennaoui et al. 1985, 1986; Altermatt et al. 2002; Tributsch et al. 2003).

Using the above formulae and the Gaussian optical transport model of Melinger et al. (Melinger et al. 1994) laser absorption profiles were calculated to estimate the depth d to which LBIC can in theory detect junctions. For comparison the free electron concentrations for two laser wavelengths (633 nm and 1064 nm) are shown in **Figure 6**. At λ=1064 nm the absorption length 1/α is roughly twice that at 633 nm as noted in the figure. For this work however a 633 nm laser was selected to probe close to the surface. For both calculations a 10 nm oxide was included on the surface to illustrate its transparency to the LBIC technique. Other field mapping methods such as Kelvin microscopy would suffer

interpretation issues in the presence of these "dirty" highly heterogeneous oxides and would probably require careful surface treatment prior to analysis (Laird and Johnson et al. 2012).

Modulation Dispersion Analysis

Dispersion analysis attempts to use the frequency dependence of LBIC as a means of increasing the certainty with which we positively identify micro-junctions above the thermoelectric background. Consider the various plots in **Figure 7**. According to a theoretical model by Niu et al. (1976) for the Lateral Photovoltaic Effect (LPE) (the core principle behind remote-LBIC), the signal amplitude from a micro-junction increases almost exponentially as f_m is reduced as shown by the dashed green line. This signal must pass through the equivalent circuit of **Figure 3** in the main text before being acquired by an external circuit. The transfer function of the micro-junction circuit (dashed blue in the same figure) acts as a "high-pass" filter [1] resulting in an overall response curve (purple curve) that peaks at an optimal modulation frequency, f_m^{opt} (solid purple). Capacitive elements tend to block DC potentials but pass higher-frequency signals via a displacement current[4]. Note that whilst this frequency give the maximum amplitude in $|P|$ it still contains a thermoelectric component.

Finally consider the thermoelectric (TE) current induced by a step change in temperature *between* the two probes. As seen in **Figure 5** in the main text, all lower frequency images are prone to large spatial artifacts which will depend on numerous complicated interplaying factors including laser power, thermal conductivity, Seebeck coefficients, as well as semiconductor type and majority carrier concentration. Heat build-up during laser dwell can generate a lateral thermoelectric current due to the temperature gradient near the laser spot as seen by the two probes. At higher f_m^{hi} the temperature cannot fluctuate in harmony with the power dissipated due to thermal mass. Furthermore, the proposed TE artifact behaves differently with modulation compared to LBIC signals from micro-junctions as its effective circuit may be very dissimilar. The TE contribution increases with proximity to either electrode because of the larger sustained temperature differential between them. In Figure the electrodes can be seen as a result of shadowing and reflection of the laser back into the large grains. These two lobes essentially form a background over the entire area that can mask subtle alteration in the field distribution if not removed or at least understood. With respect to the collected LBIC data this effect manifests as an exponential like response close to either electrode, the sign of which depends on the majority carrier. For example, excitation close to the left probe results in a negative X component. As f_m increases however, the TE response reduces in magnitude and spatial extent. The reason for this is that the not inconsiderable thermal mass of the pyrite grain imposes thermal impedance to fluctuations of the local temperature. The thermal time constant of a region τ_{RC} is simply the product of the thermal resistance, R_t and heat capacity, C_t. For a stepwise change in power, the temperature increase ΔT in response to an input

[1] A high-pass filter allows the passage of high-frequency signal whilst a low-pass does the opposite.

positive step in dissipated power has a dependence: $\Delta T \sim \Delta P(1-\exp(-t/\tau_{RC}))$ where ΔP is the step in Watts. If ΔP is replaced by a sinusoidal function $P_0 \sin(2\pi f_m t)$ then regions with large "thermal mass" and correspondingly long τ_{RC}, will dampen all but low-frequency temperature oscillations. This "thermal mass" dampens high frequency components of the TE acting as seen by the external circuit. At higher modulation rates the temperature increase since the root-mean square (RMS) power dissipated into the sample is in fact higher, but the temperature difference between the electrodes and resultant thermoelectric currents cannot ramp up and down in sync with the applied sinusoidal power. At a sufficiently high frequency f_m^{hi} there will be no residual TE component left in the signal as shown in **Figure 7**. Regions close to either probe where TE effects are obviously highest exhibit a low-frequency response represented by the solid red line with a guesstimate drop off with frequency. Calculating the form of the drop-off is beyond the scope of this paper; fortunately it can be measured from the data and appears exponential in nature close to either electrode as assumed in the figure. Note that the shape of the LBIC response function at any position will depend on the composite of these effects and whether a junction does exist or not. Close to the mid-point between the probes the TE background is almost zero assuming thermal properties are the same on either of the spot resulting in no temperature difference between the two probes. Generally this will not be the case for a complex mineral assemblage with highly asymmetric thermal properties on either side of the probe midpoint. For this sample however, there is an approximate symmetry about the HAGB.

With this simple model in mind, non-linear fitting of the complete 7-d array proceeded using an equation to model the overall dispersion relationship. By modeling the response we aim to separate micro-junction related components from the TE artifact. An *empirical* equation for the LBIC dispersion curve is simply:

$$L_{X,Y,|P|}(x,y,f_m) = L^0_{X,Y,|P|} + L^a_{X,Y,|P|}e^{-f_m/\tau_a} + L^b_{X,Y,|P|}e^{+f_m/\tau_b} \qquad (5)$$

where τ_a and τ_b are both constrained as positive values i.e L^a and L^b will mostly measure thermoelectric and photocurrent components, respectively. Although this lacks theoretical rigor it does separate mineral regions where the thermoelectric response complicates interpretation. Fitting used a non-linear Levenberg-Marquardt algorithm to minimize the following quantity:

$$\sum_{i=0}^{N-1} w_i\{y_i - L_{x,y,|P|}(f_m^i)\}^2 \quad (6)$$

with N being the number of frequency points and w_i being a weight for each measurement assumed here to be 1. The above equation was fitted to the $L_{x,y,|P|}(f_m)$ data-sets with the results being the series of images shown in **Figure 8**. Fitting across the image is generally good except close to regions where the signal is low or non-existent. The $L^a_{|P|}$(x,y) image shown in **Figure 8** represents the thermoelectric component; it clearly demarcates two distinct regions; a high region near-electrodes sandwiching a low region where its presence is negligible. Likewise the $L^b_{|P|}$ image displays strong positive peaks suggesting dispersive signatures unrelated to thermoelectric currents. Importantly,

anything with the characteristic LBIC bipolar signature within this low region is definitely associated with micro-junction fields. For illustrative purposes, raw data and fits in the electrode regions as well as those in left (*L*) and right (*R*) hot spots in the cp-py junction are shown in **Figure 9**. The response close to either probe seems to peak well below 20 Hz. Since it is questionable as to whether these thermoelectric currents pass through any junction capacitance at all, the high-pass filter may be absent in these locations or at least much reduced in efficacy due only to parasitic elements such as probe cable capacitance etc. As thermoelectric currents are largely absent at the highest modulation frequency dataset of 500 Hz, its complete LBIC dataset is best able to illustrate the spatial complexity in micro-junctions related to mixed sulfide heterojunctions as well as those inherent to the Arsenian pyrite.

Figure 6:

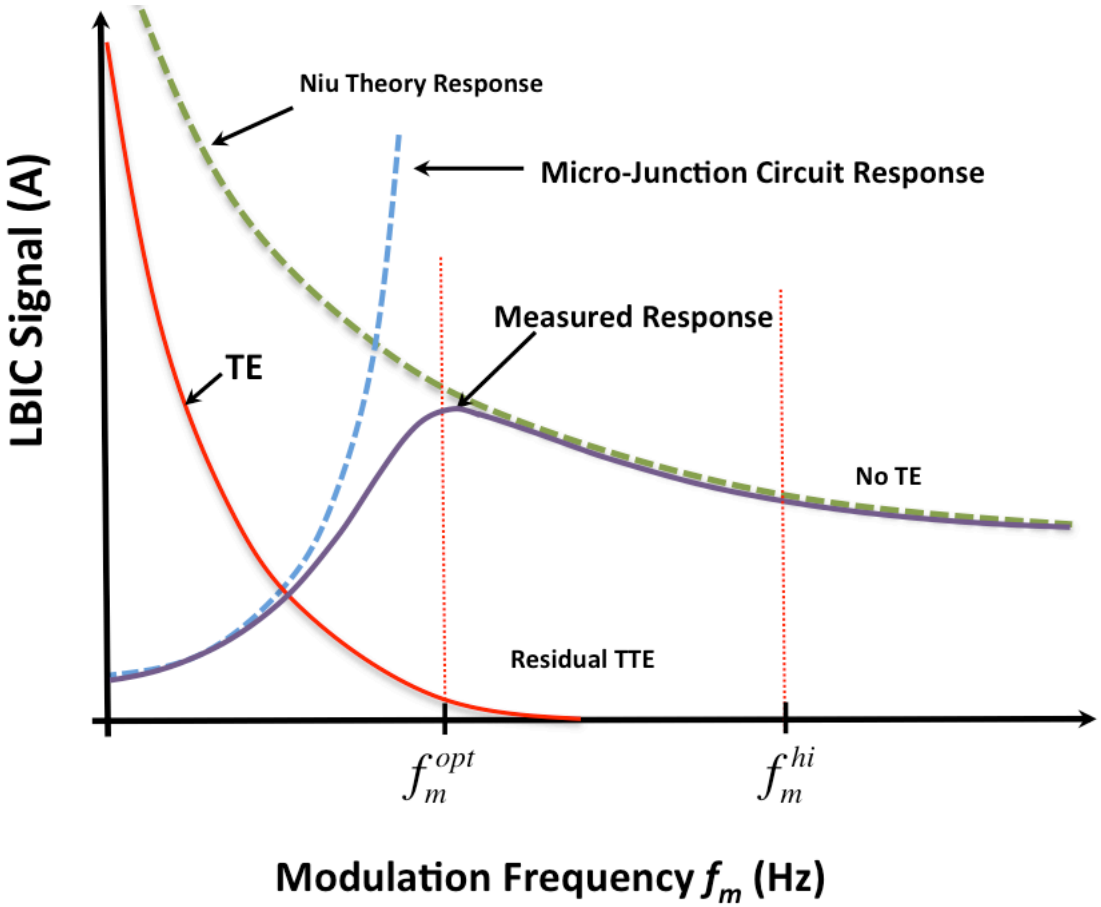

Figure 7:

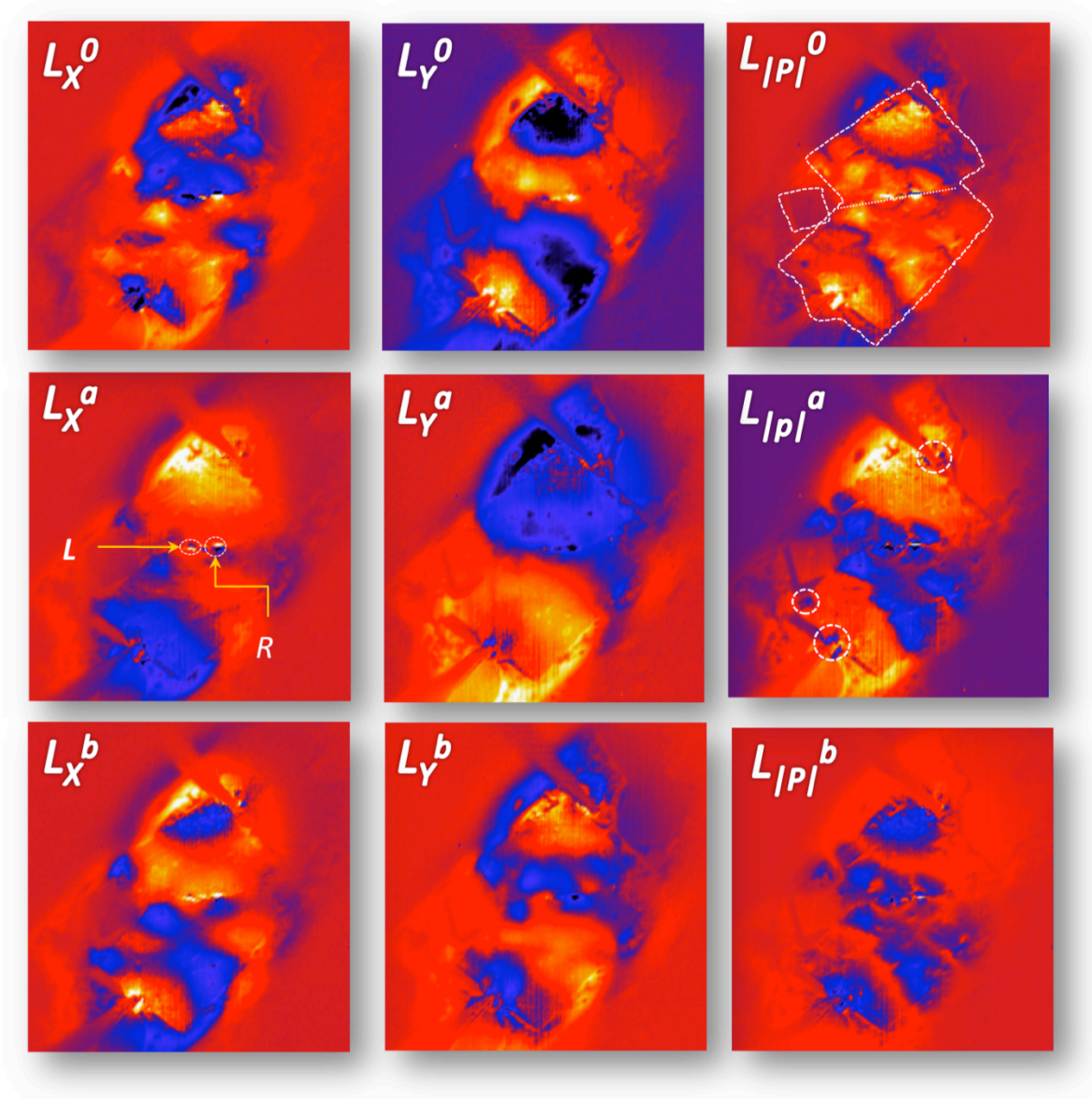

Figure 8:

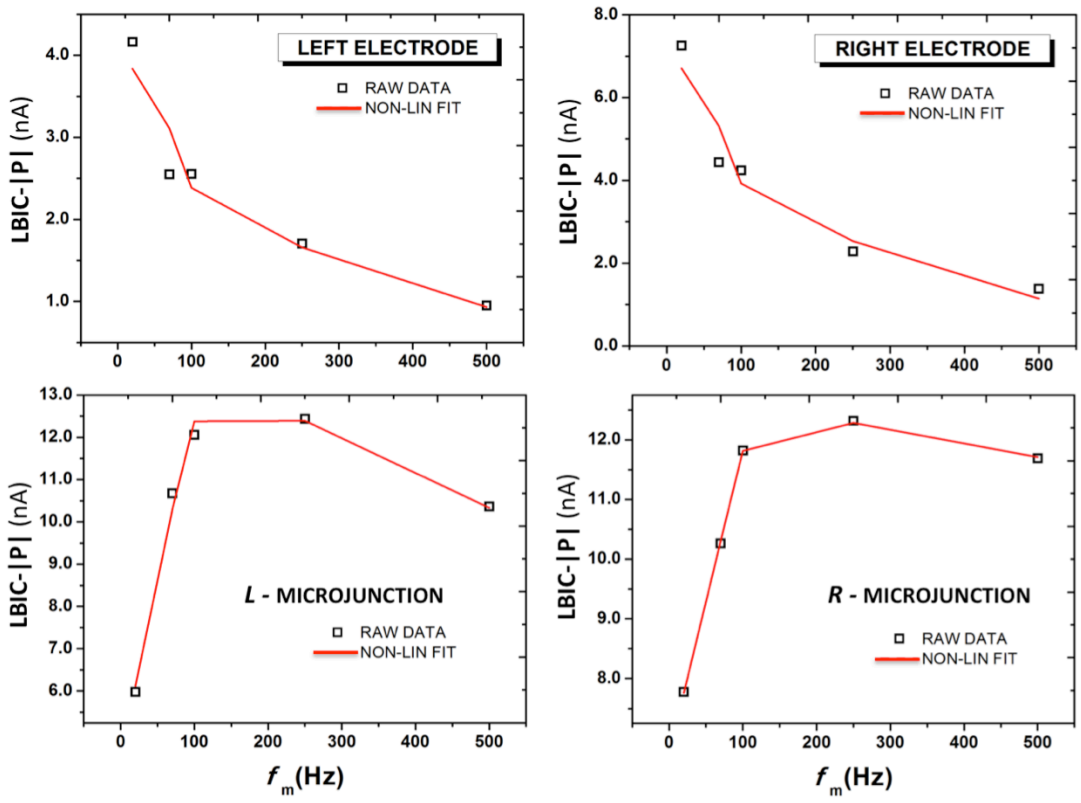

Figure 9:

Figure 10:

Figure 6: Focused laser absorption profiles in pyrite for two common laser wavelengths. A 10 nm wide-bandgap oxide (such as hematite) has been included to illustrate it has almost zero influence on probing underlying structure. For this work 633 nm is preferred as it probes the near-surface. Absorption lengths ($1/\alpha$) are indicated by the dashed white line.

Figure 7: Basic model of the overall measured frequency dependence for LBIC in the presence of a thermoelectric background (red line). The dashed green line is the exponential like response expected from the theory of Niu whilst the dashed blue line is the typical curve for a high-pass filter response expected from the equivalent circuit.

Figure 8: Fitted maps for the LBIC data-set using the non-linear sum of exponentials discussed in the text. In the top image the border region from the optical micrograph has been superimposed to guide the reader. All five components of the fit are shown. The $L_{|P|^a}$ image indicates a clear demarcation between regions close to electrodes where TE currents dominate the broad response. Regions identified as junctions (L and R) are enclosed with dashed circles in the $L_x{}^a$ image including junctions associated with cp in the R_2 region along the HAGB B-B'.

Figure 9: (Top) The magnitude of the peak LBIC-|P| signal in the left and right thermoelectric lobes (black squares) as a function of modulation frequency together with a non-linear numerical fit (dashed red line) using the equation described in the text. (Bottom) The same fitting process applied to the peak response in the two high signal regions L and R within the R_2 zone associated with a cp-py heterojunction.

Figure 10: The 500Hz LBIC dataset including (Top left) real component X of the LBIC dataset illustrating the probe positions, (Bottom Left) imaginary component Y of the LBIC phasor, (Top Right) magnitude $|P|$ of the phasor, (Bottom Right) the phase ϕ of the phasor with respect to laser excitation. Note the current scales here differs to that shown in the main text. Also shown on the phase map are the location of both ohmic probes.

Appendix References cited

Altermatt, P.P., Kiesewetter, T., Ellmer, K., and Tributsch, H. (2002) Specifying targets of future research in photovoltaic devices containing pyrite (FeS2) by numerical modelling. Solar Energy Materials and Solar Cells, 71, 181–195.

A M Karguppikar. (1988) Electrical and optical properties of natural iron pyrite (FeS2), 109, 549–558.

Dasbach, R. (1994) PhD Thesis, Basic investigations on Pyrite thin-film solar cells. (University of Konstanz).

Ennaoui, A., and Tributsch, H. (1986) ENERGETIC CHARACTERIZATION OF THE PHOTOACTIVE FES2 (PYRITE) INTERFACE. Solar Energy Materials, 14, 461–474.

Ennaoui, A., Fiechter, S., Goslowsky, H., and Tributsch, H. (1985) PHOTOACTIVE SYNTHETIC POLYCRYSTALLINE PYRITE (FES2). Journal of the

Electrochemical Society, 132, 1579–1582.
Kou, W.W., and Seehra, M.S. (1978) Optical-Absorption in Iron Pyrite (Fes2). Physics Review B, 18, 7062–7068 (1978).
Melinger, J. S. et al. (1994) Critical Evaluation of the Pulsed Laser Methods for Single Event Effects Testing and Fundamental Studies. IEEE Transactions on Nuclear Science, 41, 2574–2584.
Niu, H., Matsuda, T., Yamauchi, K., and Takai, M. (1972) Lateral photovoltaic effect in nitrogen-implanted p-type silicon. Appl. Phys. Lett, 21, 423–424.
Zhao, G. L., Callaway, J. and Hayashibara, M. (1993) Electronic-Structures of Iron and Cobalt Pyrites. *Physics Review B* 48, 15781–15786 (1993).