Breakdown data analysis – preliminary report

Paarangat Pushkarna

January 1, 2025

Contents

1 Analysis aim

Visual inspection of dark current waveforms recorded by upstream and downstream faraday cups (FCs) is useful for planning future analyses of dark current. Previous RF test stand data (X-Box 3 2018-2019) is filtered to isolate breakdown events, and dark current signals from breakdown and pre-breakdown pulses are visualised. Additional filtering measures are identified to improve breakdown isolation. Typical dark current signals are obtained, and signal characteristics benefitting from further investigation are highlighted. It is found that breakdown dark current waveforms closely match pre-breakdown waveforms until the point of breakdown ignition, for both upstream and downstream FCs.

2 Breakdown candidate filtering

The dark current waveforms inspected must be those from true breakdowns. In previous datasets, it is possible that non-breakdown events are tagged as breakdowns due to conservative thresholds on power reflected, or structure outgassing, for example. Additionally, structure breakdowns must be separated from breakdowns occurring in the pulse compressor or load, as we are only interested in structure performance. A data filtration pipeline that reduces a pool of breakdown candidates to the true breakdowns has been implemented. Similar pipelines are well documented [Raj16; Luc18; Woo15; Pus22], so a brief summary is given alongside any notable differences from existing literature.

2.1 Boolean filtering and noise rejection

The most straightforward filtering is on boolean flag variables indicating tripping of thresholds on reflected power and FC signals. Labelling of RF pulses in previous test stand data is explained on page 17 of [Pus22]. For all pulses labelled with log type “2”, indicating a breakdown pulse, the following logic statement must be met by boolean flag variables:

$$\mathrm{BD}_{\mathrm{structure}} = ((\mathtt{DC\_UP}_{\mathrm{flag}} \lor \mathtt{DC\_DOWN}_{\mathrm{flag}}) \land \mathtt{PSR}_{\mathrm{flag}}) \land \neg\mathtt{PER}_{\mathrm{flag}}, \tag{1}$$

where $\land$ is the logical “and” operator, and $\neg$ is the logical “not” operator, and $\lor$ is the logical “or” operator. This indicates that the thresholds on the upstream or downstream FC can be tripped, but in addition the reflected power from the structure (PSR) must be tripped. Also, a tripped threshold on PER, power reflected from the load, indicates a load breakdown, which must be excluded. Equation 1 is more stringent than the logic of [Luc18] which reads:

$$\mathrm{BD}_{\mathrm{structure}} = (\mathtt{DC\_UP}_{\mathrm{flag}} \lor \mathtt{DC\_DOWN}_{\mathrm{flag}} \lor \mathtt{PSR}_{\mathrm{flag}}) \land \neg\mathtt{PER}_{\mathrm{flag}}, \tag{2}$$

and less stringent than the logic of [Pus22], which reads:

$$\mathrm{BD}_{\mathrm{structure}} = (\mathtt{PLR}_{\mathrm{flag}} \land \mathtt{PKR}_{\mathrm{flag}} \land \mathtt{DC\_UP}_{\mathrm{flag}} \land \mathtt{DC\_DOWN}_{\mathrm{flag}} \land \mathtt{PSR}_{\mathrm{flag}}) \land \neg\mathtt{PER}_{\mathrm{flag}}. \tag{3}$$

Equation 3 was made less stringent because a structure breakdown may not release enough charge to trigger reflected signal thresholds in the klystron and pulse compressor directional couplers. It was thought that abandoning Equation 3 would admit breakdowns of smaller 'intensity'. Equation 2 was made more stringent because triggering of PSR was thought a necessary BD criterion, but the $\vee$ operator between upstream and downstream FC thresholds was maintained, as it is not guaranteed that a breakdown will trip both thresholds.

Calculation of average input power to the structure, as well as the flat top pulse width, is defined in [Luc18] and [Pus22]. Average input power must be greater than 650kW, and flat top pulse width less than 300ns, for a pulse to be considered as part of normal test stand operation. Pulses not meeting this criteria are considered noise.

2.2 Filtering on total signal energy

The total pulse energies, $U_{\texttt{INC}}$, $U_{\texttt{REF}}$ and $U_{\texttt{TRA}}$, are determined in the following way:

$$U_{\texttt{SIGNAL}} = \int P_{\texttt{SIGNAL}}(t)\mathrm{d}t \text{ [Joules]}, \tag{4}$$

$$P_{\texttt{SIGNAL}}(t) = C_2(A_{\texttt{SIGNAL}}(t))^2 + C_1 A_{\texttt{SIGNAL}}(t) + C_0 \text{ [Watts]}. \tag{5}$$

Define two quantities $m_{\texttt{TRA}}$ and $m_{\texttt{REF}}$ using the total pulse energies, according to [Raj16]:

$$m_{\texttt{TRA}} = \frac{U_{\texttt{INC}} - U_{\texttt{TRA}}}{U_{\texttt{INC}} + U_{\texttt{TRA}}} \text{ [unitless]} \tag{6}$$

$$m_{\texttt{REF}} = \frac{U_{\texttt{INC}} + U_{\texttt{REF}}}{U_{\texttt{INC}} - U_{\texttt{REF}}} - 1 \text{ [unitless]}. \tag{7}$$

Better separation of breakdowns from non-breakdowns is achieved by putting thresholds on $m_{\texttt{REF}}$ and $m_{\texttt{TRA}}$. This can be due to overly conservative breakdown thresholds, or gassing from the structure in early conditioning periods. Call the 2D-space defined by $m_{\texttt{REF}}$ and $m_{\texttt{TRA}}$, "m-space". A breakdown implies that $U_{\texttt{REF}}$ is closer to $U_{\texttt{INC}}$, so the denominator in Equation 7 gets smaller, making $m_{\texttt{REF}}$ bigger. A non-breakdown implies that $U_{\texttt{REF}} \ll U_{\texttt{INC}}$ so the fraction in Equation 7 goes to 1, and $m_{\texttt{REF}}$ goes to zero. This is better visualised on a log axis in Figure 1.

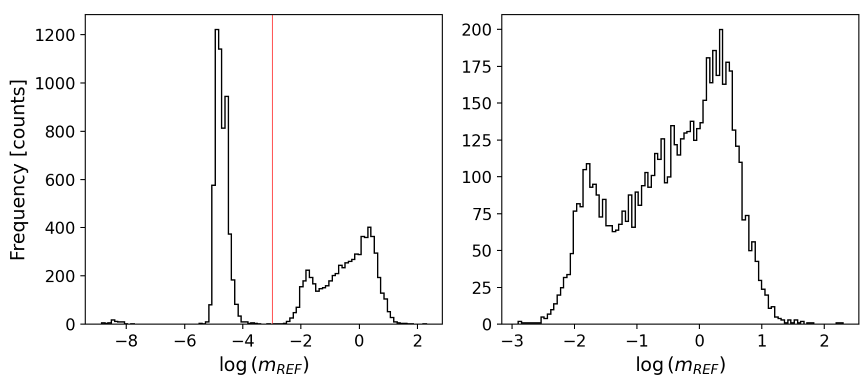

Figure 1: Left: the full space of $m_{\texttt{REF}}$, leftmost cluster showing breakdown candidates to reject due to low reflection. Right: distribution of $m_{\texttt{REF}}$ after rejecting cluster to left of red line.

Apply a threshold $\log(m_{\texttt{REF}}) > -3$, indicated by the red line in Figure 1. Since the leftmost cluster in Figure 1 (left) has $\log(m_{\texttt{REF}}) \in (-6, -4]$, we suspect this reflection coefficient is too low, and breakdown candidates in the leftmost cluster are rejected, modulo direct waveform inspection. It seems that applying this threshold automatically filters $m_{\texttt{TRA}}$ as indicated in Figure 2. This means the two peaks in $m_{\texttt{TRA}}$ and $m_{\texttt{REF}}$ spaces are related. Since both are well clustered at low values of $m_{\texttt{TRA}}$ and $m_{\texttt{REF}}$, we can justify their rejection.

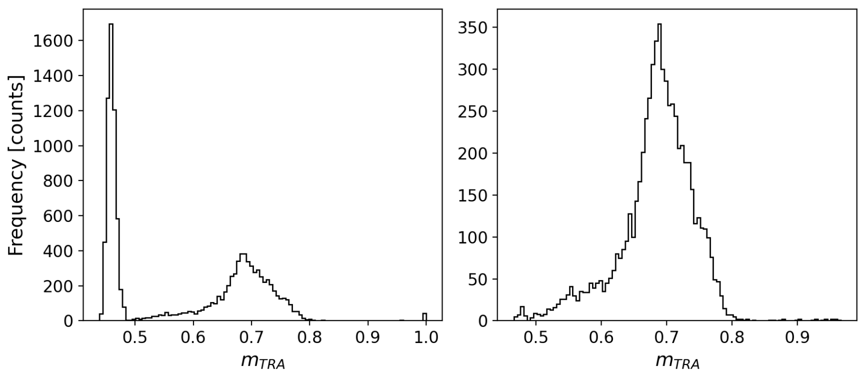

Figure 2: Left: the full space of $m_{\texttt{TRA}}$, leftmost cluster showing breakdown candidates to reject due to low reflection. Right: distribution of $m_{\texttt{TRA}}$ after rejecting cluster to left of red line.

As a sanity check to ensure that the discarded candidates are not breakdowns, we directly examine the pulses that make up the leftmost clusters in figures 1 and 2. We also put a minimum threshold of 10MW on the average power to enable better waveform visualisation. We are confident that pulses within this rejected cluster are not breakdowns because the reflected waveforms are much diminished in intensity compared to their peak power during a breakdown (c.f. existing test stand literature).

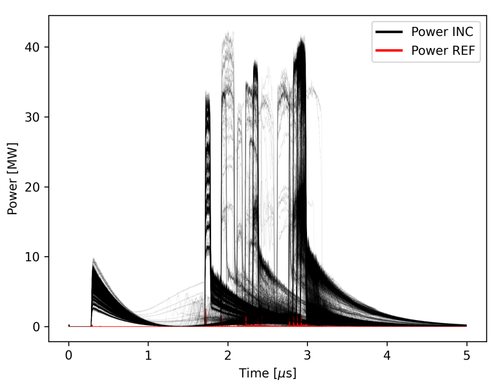

Figure 3: Waveforms belonging to the leftmost peak in figures 1 and 2

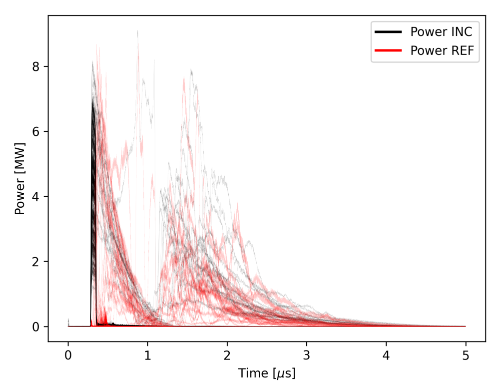

Figure 4: Non-standard pulses passing through average power filters.

Another straightforward filter is to exclude non-standard pulses that have slipped past the pulse width and average power filtering step. A subset is identified where the leading edge of the flat-top occurs less than 1μs into the waveform window. These pulses cannot be compared to other breakdowns because charging of the pulse compressor is absent, and no tail is observed on the falling edge, as with other (non-)breakdown pulses. We are therefore justified in discarding these pulses, which are visualised in Figure 4. The m_{REF} distribution after discarding these spurious pulses is shown in Figure 5.

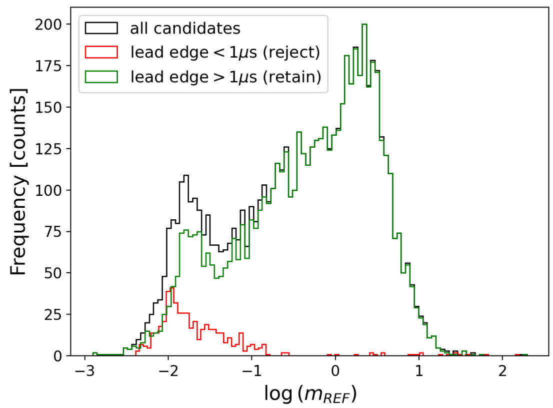

Figure 5: Rejection of spurious breakdown candidates with non-standard incident power, and lack of pulse compressor charging time.

2.3 TOP (Time-of-Propagation) exclusion in m_{REF} space

Despite filtering non-standard pulses in Figure 4, there still remains a protrusion on the left side of the distribution for remaining breakdown candidates in Figure 5 (in green), for $\log(m_{\text{REF}}) \in [-2, -1.3]$. A last, physically motivated filter can be instituted to investigate causes of this subsidiary peak.

Boolean flags triggered on reflected signal thresholds are used to isolate structure breakdowns. However, cell-by-cell breakdown location is discerned from the raw RF waveforms incident, reflected and transmitted through the structure, using the "Time-of-Propagation" (TOP) and "Edge Correlation" techniques [Woo15; Luc18; Raj16]. TOP is calculated by subtracting the falling edge time of the transmitted waveform from the rising edge time of the reflected waveform. Usually, the 90 percent falling edge, and 10 percent rising edge times are used. In our analysis, we use a 75 percent falling edge on the transmitted waveform, and 25 percent rising edge on the reflected waveform, for better robustness to pulse shape aberrations. The condition

$$t_{\text{TOP}} \in [-t_{\text{fill}}, t_{\text{fill}}] \text{ or } t_{\text{TOP}} + t_{\text{fill}} \in [0, 2t_{\text{fill}}] \tag{8}$$

must hold for breakdowns occurring within the structure. TOP cannot exceed structure fill time for any structure breakdown. We can use the bounding condition in equation 8 to exclude breakdowns occurring outside the structure, for example in the RF load. While these should have been excluded by boolean thresholds, we ensure their rejection by calculating TOP and structure fill time (64.55ns).

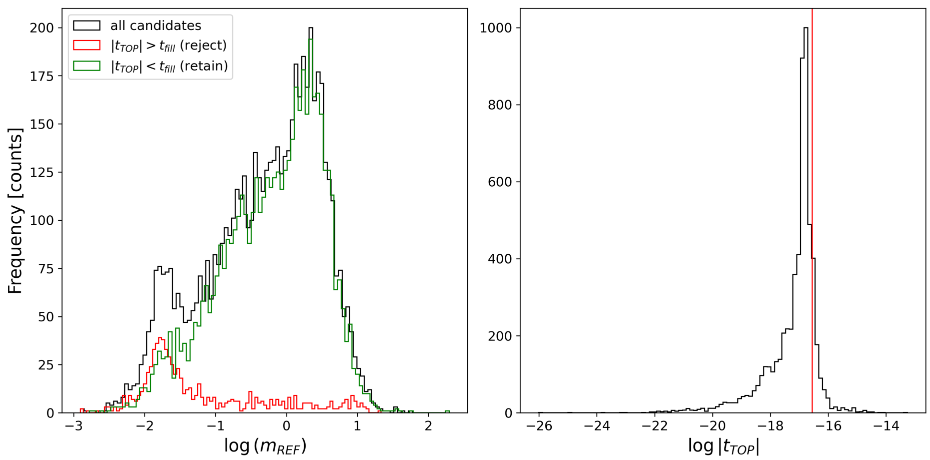

Figure 6: Rejection of spurious breakdown candidates with $|t_{\mathrm{TOP}}| > t_{\mathrm{fill}}$. Log-space used on $|t_{\mathrm{TOP}}|$ to aid visualisation. Red-line showing cut threshold on $|t_{\mathrm{TOP}}|$.

The subsidiary peak is well accounted for by pulses with TOP greater than the structure fill time (Figure 6). Imperfections in edge time calculations should mitigate impact on other parts of the distribution, which are already minimal when compared to filters based on pulse count, time stamp, and average power. Physical reasons for the red peak in Figure 6 relate to power reflection on timescales longer than t_{fill}, which points to load breakdowns and vacuum activity in the line between the hybrid board and structure as possible culprits.

There are two similar, separate lines in the XBOX3 dataset. Only data for line A has been shown, but the same filtering steps have been undertaken for line B. In what follows, calculations will be demonstrated mostly on line A, with the understanding that identical steps must be taken on data from line B. Differences between the two datasets, in filtering or otherwise, will be raised as they arise. As an example, the TOP filtering step is demonstrated in Figure 7 below, after leading edge, power, pulse width, flag and log type filtering has been undertaken. A summary of the filtering steps taken to remove spurious breakdown candidates is:

1. Boolean filtering by breakdown flags from triggered thresholds, and log types
2. Filtering noise and non-standard waveforms with average power and pulse width thresholds
3. Removing pulses with flat-top leading edges less than 1μs, and with absent pulse compressor charging times
4. Removing pulses with $\log(m_{\mathrm{REF}}) < -3$, as this reflection coefficient is too small to constitute breakdown
5. Removing pulses with breakdown TOP larger than the structure fill time.

Direct, numerical thresholds on average power and pulse width are avoided, other than those to remove obviously noisy pulses. Instead, cuts based on general physical principles, such as

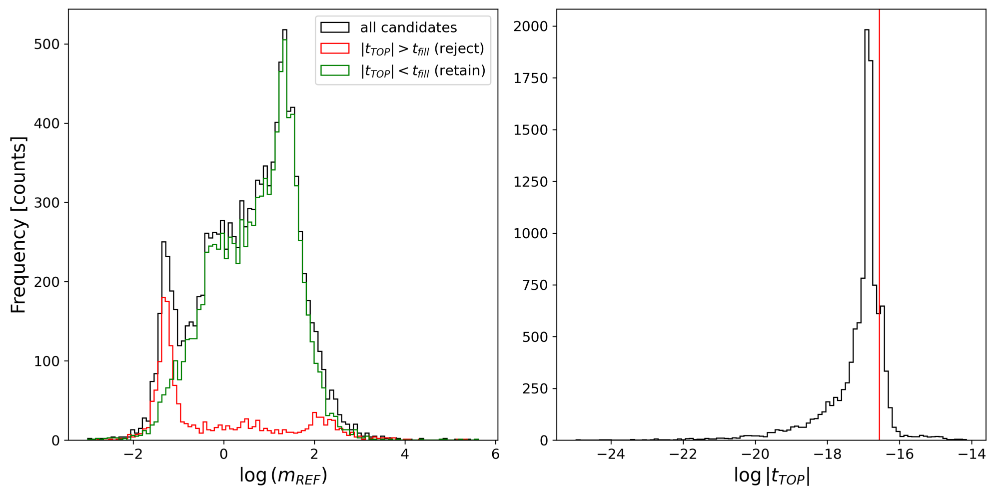

Figure 7: TOP rejection for line B dataset.

uniformity of incident pulse shape (to compare apples with apples), or structure fill time as a bound on breakdown location, have been adopted. Where doubts occur, the waveforms of the discarded pulses have been directly inspected to ensure that genuine breakdowns are not lost. This work is the first to exclude pulses by taking logs of the reflection coefficient, to expose non-breakdowns that in previous analyses would have been treated as breakdowns. Without putting $m_{\mathtt{REF}}$ on a logarithmic axis, the candidates rejected in step 4 above could not have been differentiated.

2.4 2d m-space visualisation

Having followed the steps outlined above, we can plot $\log(m_{\mathtt{REF}})$ against $m_{\mathtt{TRA}}$ to examine the relationships between scattering parameters. Additionally, the impact of filtering steps on clusters in the 2d-space can also be observed. The first filtering step however, that of discarding pulses with $\log(m_{\mathtt{REF}}) < -3$ has been omitted from the 2d plots in Figure 8 for clarity of visualisation. The line B dataset contains more breakdowns than line A, and a range of $\log(m_{\mathtt{REF}})$ values with a larger upper bound. If $\log(m_{\mathtt{REF}})$ can be thought of as a representation of breakdown magnitude, then line B appears to have more numerous breakdowns that are also more severe than line A. This is why the bottom right 2d histogram in Figure 8 appears to have more empty space – a select few events have large $\log(m_{\mathtt{REF}})$ values, creating isolated bins on the $\log(m_{\mathtt{REF}})$ axis. The cause of this discrepancy is not known. The DUT on line B was a TD24 cavity that had not been baked out, whereas line A hosted a baked out TD24 structure. A first reason could be increased robustness to breakdown after bake out.

Another note is on the linearity of $\log(m_{\mathtt{REF}})$ with $m_{\mathtt{TRA}}$. If $m_{\mathtt{TRA}}$ and $m_{\mathtt{REF}}$ are connected to magnitude of charge released during breakdown, then the exponential relationship between $m_{\mathtt{TRA}}$ and $m_{\mathtt{REF}}$ may be explained by trying to find a connection with Fowler-Nordheim theory. It might just be that $m_{\mathtt{TRA}}$ and $m_{\mathtt{REF}}$ are real parts of the complex valued RF reflection and transmission coefficients, and are as such exponentially related, which could be the less

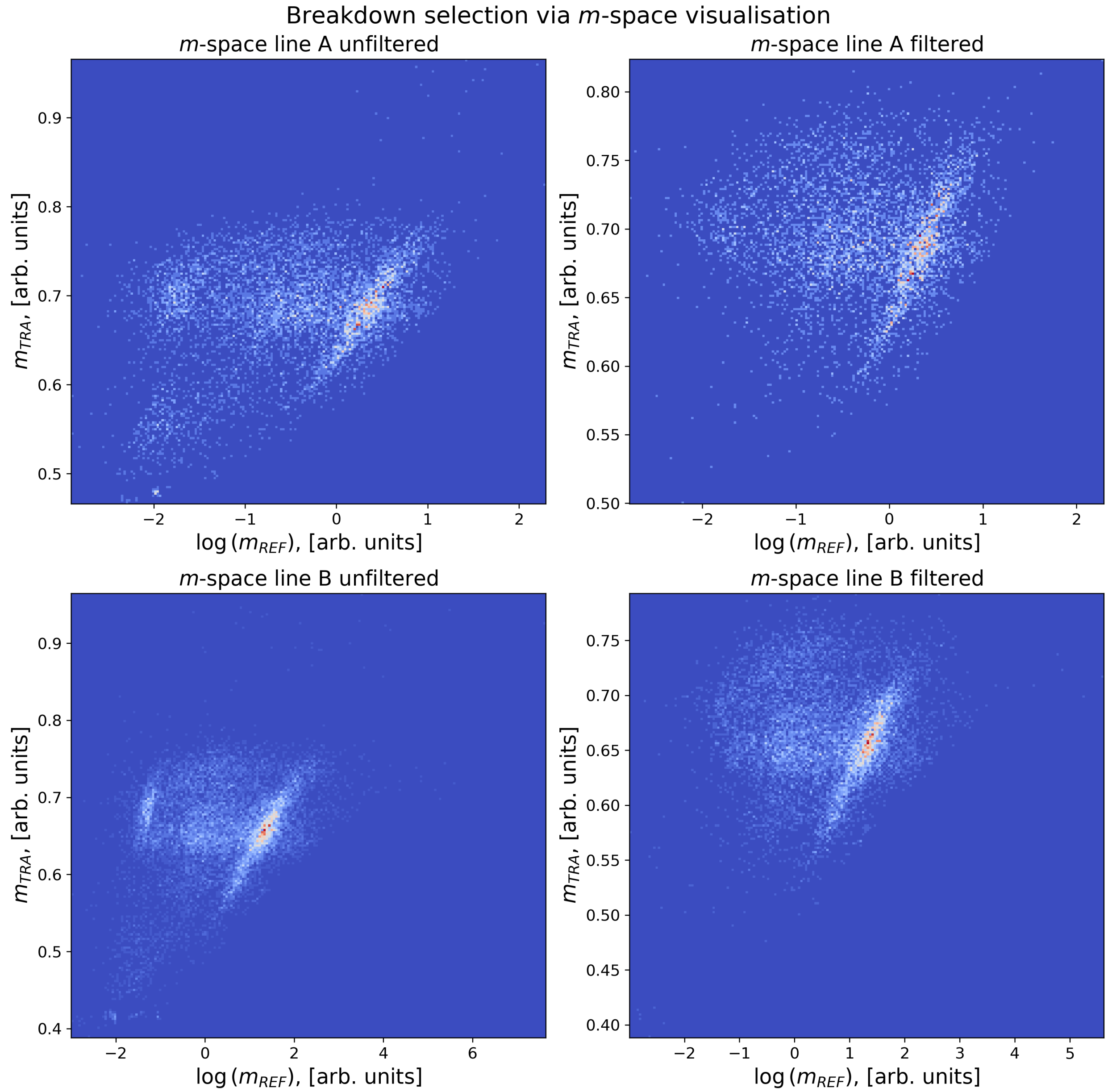

Figure 8: Final visualisation of scattering parameters after filtering steps.

interesting, and more likely, reason.

Finally, the clustering in the left column of Figure 8 can be explained partially by the filters applied. In the top left panel, at coordinates (-2, 0.7) and in the bottom left panel at coordinates (-1.75, 0.7), clustering due to pulses with TOP greater than the structure fill time is observed. It was found that after rejecting such pulses in the previous section, that this cluster was removed, as evidenced in the filtered 2d histograms in the right column of Figure 8. Similarly, a cluster due to pulses with leading edges less than one microseconds manifests at the rough coordinates (-2, 0.55) in the top left panel and at (-2, 0.45) in the bottom left panel. This too is removed after filtering, and two main clusters remain. One evidences the linear relationship between $\log(m_{\mathtt{REF}})$ with $m_{\mathtt{TRA}}$, and another more diffuse cluster lies just behind, roughly at coordinates (-0.5, and 0.65) in both top right and bottom right panels. The physical reason for the diffuse cluster is not yet known, and it's relation to the main cluster is currently being probed. The diffuse cluster does not depend on the leading edge time of the incident power flat top.

Visualisation of m-space in 2d exposes several interesting behaviours, and enables detailed global data set inspection. This work is novel in plotting $\log(m_{\mathtt{REF}})$ instead of just $m_{\mathtt{REF}}$ (as in [Raj16; Luc18]), and this has enabled finer isolation of breakdowns as compared to previous studies. This work is the first to reject breakdown candidates by TOP and structure fill time. m-space represents a useful projection of the full dataset onto two features, reducing dataset dimensionality, and allowing data of interest to be discriminated more conveniently for further analysis. This is preferable to an automated ML based dimensionality reduction, as adopted by [Obe+22], since the features extracted are transparent, closed-form, and physically motivated. An example of an extended analysis may be a linear fit of $\log(m_{\mathtt{REF}})$ against $m_{\mathtt{TRA}}$, and definition of a region of interest around the fitted line which captures data constituting the main cluster. Then, data falling inside and outside the region of interest can be examined separately, and physical reasons for the diffuse, secondary cluster narrowed down.

To highlight the utility of $\log(m_{\mathtt{REF}})$ in studying breakdown behaviour, the variation of TOP with $\log(m_{\mathtt{REF}})$ is shown in Figure 9. If $\log(m_{\mathtt{REF}})$ represents the energy reflected as a

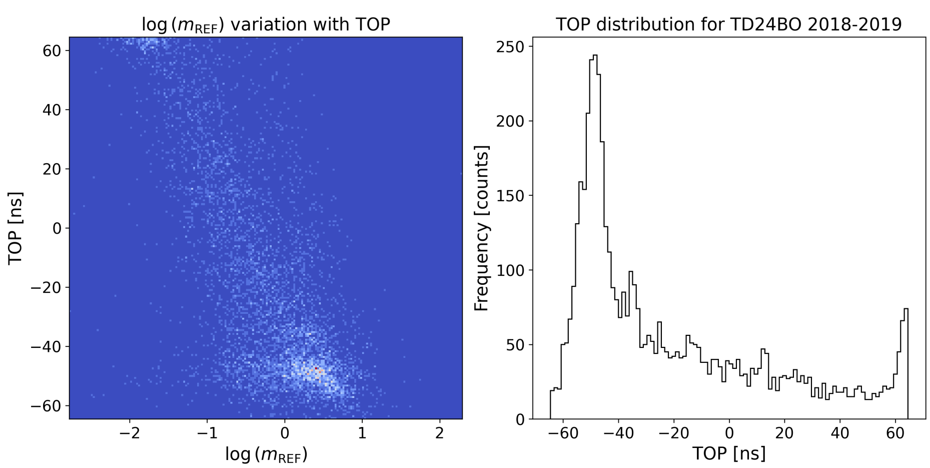

Figure 9: Variation of $\log(m_{\mathtt{REF}})$ with TOP.

result of the breakdown, relative to the energy incident on the structure, then Figure 9 shows that not only are breakdowns concentrated at locations more upstream in the structure (the peak at TOP $\in [-60, -40]$ns, Figure 9 right) but that these breakdowns also reflected a greater proportion of the input energy than breakdowns occurring at locations downstream. That breakdowns occur in hot cells concentrated at the start of the structure was known [Woo15; Raj16; Luc18], but that they have correspondingly larger "intensities" is also interesting. This shows that this calculation, alongside others mentioned in the paragraphs above, are worth checking and formalising.

An important next step would solidify the interpretation of m_{REF} as an indicator of breakdown intensity, by examining correlation of m_{REF} with charge measured on upstream and downstream faraday cups. Since a greater charge release should result in greater power reflection, m_{REF} should increase for pulses with large deflections in faraday cup signals. Evolution of breakdown intensity over the conditioning period, and association of breakdown intensity with 'hot' cell behaviour are possible questions that such a study could address.

3 Breakdown dark current waveform inspection

Once breakdowns are filtered, dark current waveforms can be retrieved. Input power variation, and change in flat-top leading edge time, make it difficult to draw any inference from 'persistence' plots (several dark current waveforms plotted one over the other, with opacity adjusted). Instead, each dark current waveform is inserted as a row in a 2d matrix, with each cell containing the ADC counts recorded for that time step. The time base is common for all waveforms: 1250 samples at 250MSPS, meaning a 4ns sample spacing with a full 5 μs window. Individualised time data for each pulse is not available. This matrix is visualised as a 2d histogram or heatmap, so that global trends can be more conveniently exposed. Dark current heatmaps are shown in Figures 10 and 11.

The first μs (250 samples) of data in each waveform is used as a background sample, and the average background subtracted from each full waveform. The absolute value of each waveform is taken to assist visualisation and heatmap colour scaling, just for Figures 10 and 11. Immediately clear is the leading edge time of dark current pulses in both upstream and downstream directions. This leading edge time is almost the same as the leading edge of the incident power flat top. Movement of the flat top within the 5μs window results in movement of the dark current leading edge. It is understandable that timing of dark current emission is correlated with incidence of RF power to the structure under test. Saturation of the 14-bit ADC in almost every breakdown dark current signal is apparent, with the maximum ADC count being 8000.

Most often, saturation occurs in the leading peak for the upstream signals, and then again in subsequent 'aftershocks', subsidiary peaks coming after the mean peak. A leading peak is common to almost all waveforms, whilst the nature of aftershocks seems to change with upstream versus downstream signals, as well as the leading peak time-of-arrival. Aftershocks are less pronounced for the downstream signals, especially for dark current waveforms with leading edges at $\sim 3\mu$s. A gap between aftershocks and leading peaks appears for $\sim 3\mu$s waveforms in the upstream faraday cup, while no gap is apparent for waveforms with leading

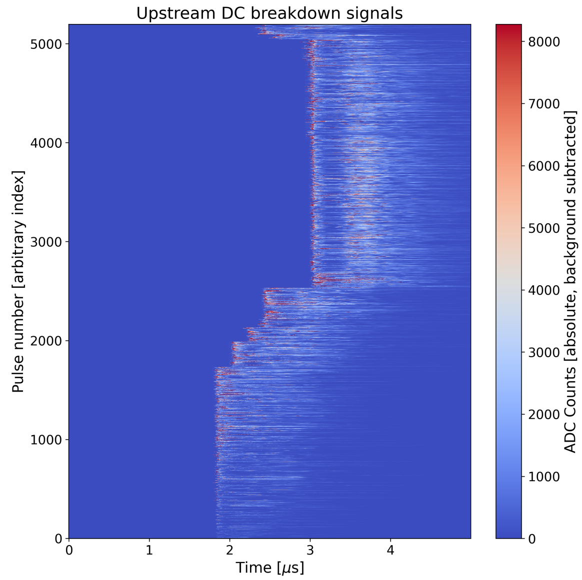

Figure 10: Upstream faraday cup waveforms for breakdown pulses.

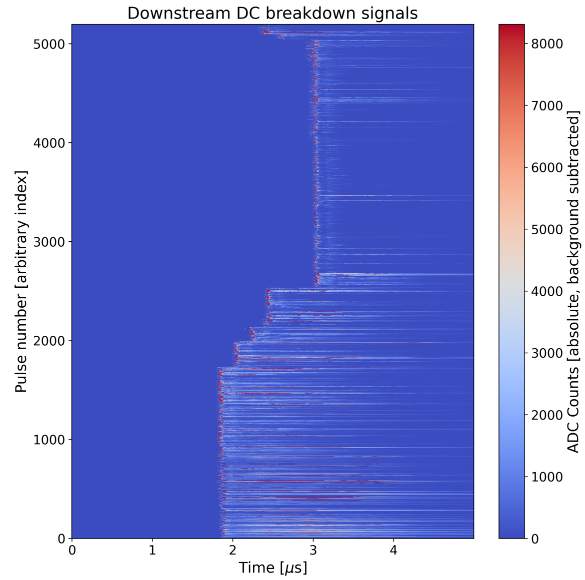

Figure 11: Downstream faraday cup waveforms for breakdown pulses.

edges $< 3\mu$s, or for waveforms in the downstream faraday cup. For waveforms with leading edge $< 3\mu$s, aftershocks in downstream signals do not seem to follow any specific pattern.

To better probe signal shape for waveforms with leading edge $\sim 3\mu$s, a 'typical' waveform was constructed by first amplitude normalising each waveform, then averaging waveforms. This is shown in Figure 12. To maintain uniformity, only waveforms with flat-top pulse width less than 60ns were selected. The same approach can be extended to other pulse widths and leading edge times. The gap between leading peaks and aftershocks for upstream signals is now more apparent. These aftershocks are also present in the downstream signals, but with a smaller separation. This could imply that charge travelling downstream may fall into RF buckets, and is accelerated and bunched, whereas upstream travelling charge is not. In both upstream and downstream cases, a short, high-amplitude leading peak is followed by a lower amplitude, broader peak. This could correspond to charge emission at different stages of the breakdown, which occur over few ns timescales in the onset phase, to sub-μs (few 100ns) timescales in the burning and extinction phases.

The assumption that downstream waveforms have greater activity and higher amplitude is not borne out by this data, since the upstream cup has secondary peaks of greater intensity. It could be that the first peak of the downstream is yet larger in amplitude, since more charge could be bunched into it by the RF, but may therefore be saturating the ADC.

Waveforms with leading edge $< 3\mu$s, are different to those with leading edge $\sim 3\mu$s. Providing a longer klystron input pulse, and therefore more average power to the structure, seems to change the nature of dark current emission as recorded by upstream and downstream faraday cups. This motivates correlating dark current signals with klystron input power, klystron output power, and incident power to the structure. The signal features exposed encourage comparison of signal artefacts to existing theory of breakdown arc mechanisms.

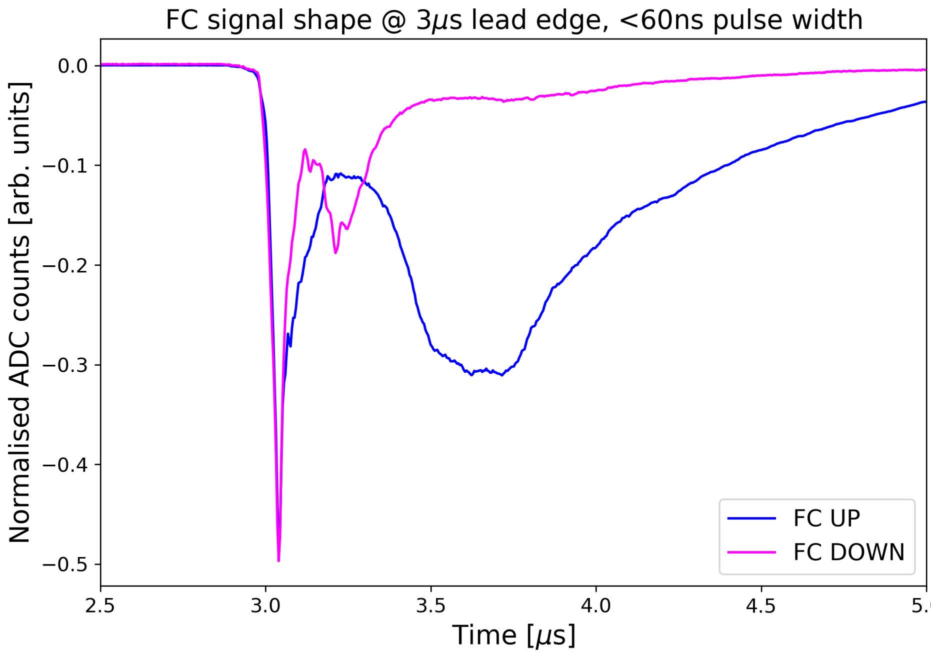

Figure 12: Typical breakdown pulse shapes, normalised and averaged.

3.1 Inspection of pre-breakdown waveforms

Just as breakdown waveforms have been inspected, pre-breakdown waveforms corresponding to those breakdowns can also be examined. Similar heatmaps and typical pulses can be created to expose the typical signal shape, and its variation over the chosen (pre-)breakdown waveforms.

Pre-breakdown activity is again related to incidence of the flat-top on the structure. Pre-breakdown leading edges follow the flat-top leading edge times (roughly 1.7μs and 2.9μs), as shown in Figures 13 and 14. Additionally, since the breakdown is not ignited, and only field emission occurs, the duration of field emission matches the flat-top pulse width. The increase in pulse width during conditioning can be observed in both the upstream and downstream dark current signals, from roughly 50ns to 150ns. Since less charge is emitted during the pre-breakdown pulses, the corresponding ADC counts are also lower, not reaching more than 500. The field emission behaviour is regulated by the square pulse "on-off" nature of the flat-top – field emission only occurs during the flat-top, and no 'tail' is observed in the dark current signals, as it was for breakdown waveforms.

Pre-breakdown signals collected at the downstream faraday cup differ from those collected at the upstream faraday cup, as shown in Figure 15. Again, all waveforms have been normalised and averaged, and the same criteria for leading edge and pulse width as for the breakdown waveforms have been applied. The upstream faraday cup resembles a squarer pulse, with a FWHM larger than that for the downstream faraday cup. This could again be due to bunching and acceleration by the RF. The downstream cup also exhibits the expected electronic behaviour from such a device – an initially sharp rise time followed by a longer

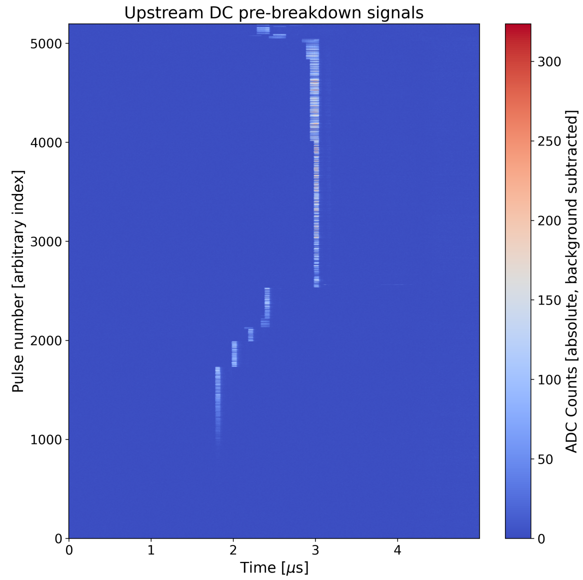

Figure 13: Upstream faraday cup waveforms for pre-breakdown pulses.

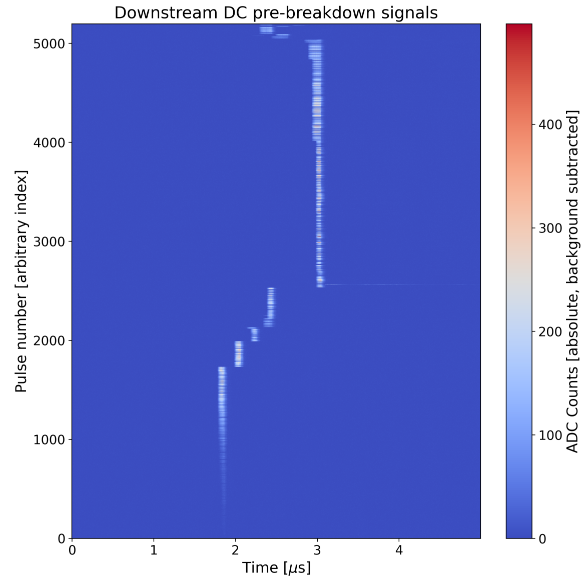

Figure 14: Downstream faraday cup waveforms for pre-breakdown pulses.

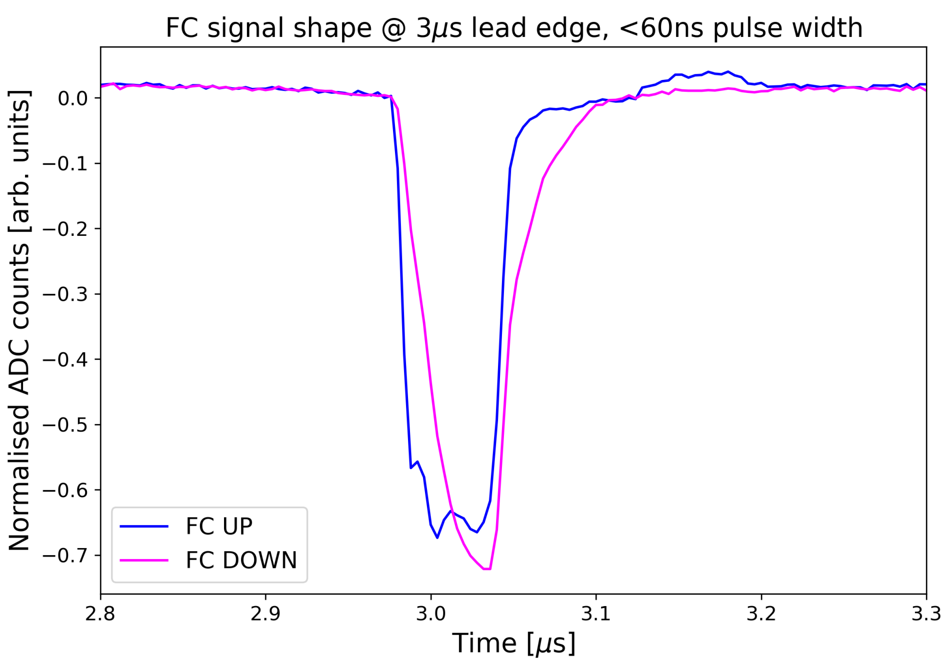

Figure 15: Typical pre-breakdown pulse shapes, normalised and averaged.

decay time, as produced by tests of the same faraday cup in single bunch at the synchrotron linac (single bunch faraday cup response test by Eugene earlier this year). It could be that due to bunching, the downstream cup is responding to an electron bunch resembling a 'single' event, whereas the upstream cup is the superposition of several smaller 'electron packets' that are dispersed longitudinally, having not been captured by the RF.

3.2 Breakdown ignition time

Searches for precursory phenomena giving predictive information on breakdown onset are ongoing [Obe+22; Pas20]. For this reason, pre-breakdown waveforms are of interest, since they provide information on the evolution of field emitter sites before they ignite into full breakdowns. Comparison of pre-breakdown waveforms with their "parent" breakdown waveforms is therefore the aim of this section. Proper matching of pre-breakdown pulses, labelled "0" in the X-box 3 dataset, with their parent breakdowns is important for drawing accurate inferences.

As an aside, using pulse counts is not possible for this matching task because of non-uniform incrementation of pulse count variables. It was expected that pre-breakdowns will have a pulse count of one less than the breakdown pulse, making pre-breakdown waveform retrieval straightforward. However this was not borne out by the dataset, and raw timestamps were instead used. From a database organisation perspective, UTC timestamp is the only truly unique key.

Once matched, pre-breakdown waveforms can be plotted on common axes with their parent breakdown waveforms to enable comparison. These plots are shown in Figures 16 and 17. An important note is that the upstream and downstream signals shown below are not from the same breakdown. The upstream signals show dark current waveforms from one breakdown event, and the downstream signals show dark current waveforms from another.

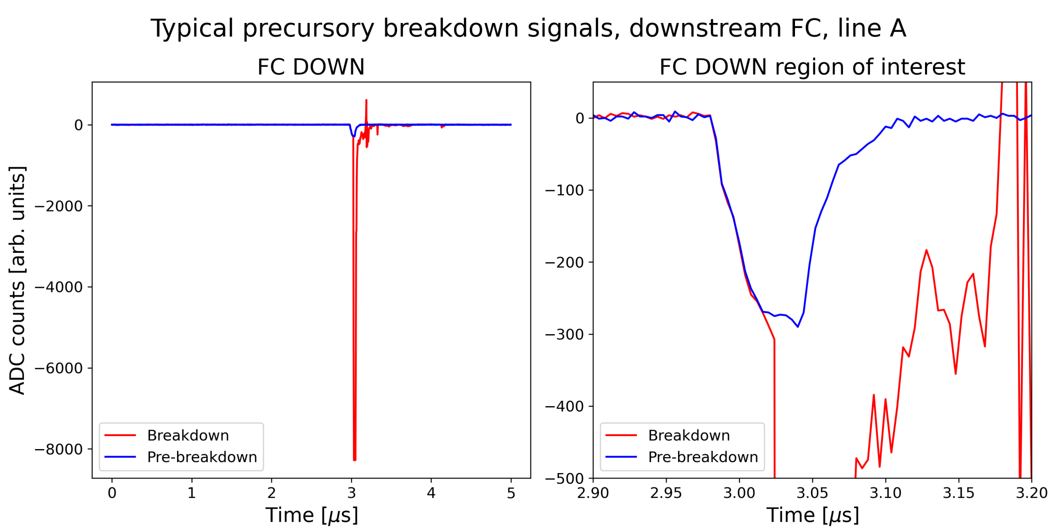

Figure 16: Precursory behaviour on the downstream faraday cup.

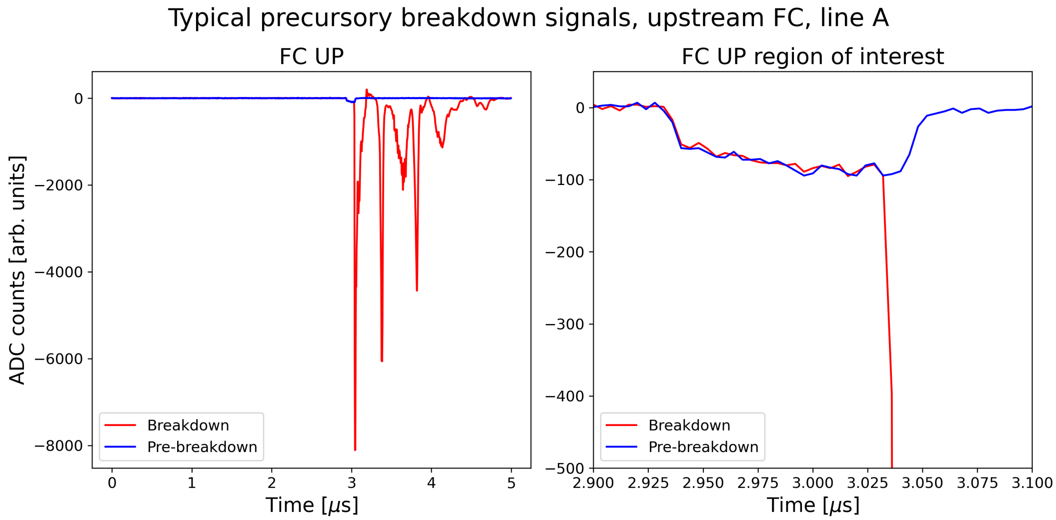

Figure 17: Precursory behaviour on the upstream faraday cup.

Firstly we observe the distinction between upstream and downstream dark current signals – leading peaks in both upstream and downstream signals saturate the ADC, but afterpulses in the upstream faraday cup are much more prominent than in the downstream cup. The width of the leading peak on the upstream signal is also larger than in the downstream signal. We reference the earlier discussion on waveform inspection regarding bunching of dark current due to incident RF as a possible explanation.

Once plotted on the same axes, it is apparent that the breakdown waveforms follow the same trajectory as the pre-breakdown waveforms up until the point of breakdown ignition, which is recorded on the faraday cup signals as a sharp, almost infinitely steep rising edge on the leading peak. The ignition point as implied by the rising edge is gated by the time for which the emitter site expels charge in the pre-breakdown dark current waveform. The waveforms are almost identical up until the ignition point, in both cases. This implies that the physical processes emitting and transporting charge to each faraday cup are the same until the point of breakdown "ignition". It is crucial also that such precursory behaviour does not occur for all pulses examined; pre-breakdown pulses with only noise are also common.

It is possible that the width of the field emitter "gate" depends on the incident flat top pulse width. Then, the location of the ignition point within the gate becomes the first quantity of interest. Do breakdowns ignite predominantly in the earlier, or latter part of the gate? Can a distribution be created for the breakdown ignition time within this gate? Does the length of the gate matter, or is the distribution of ignition times unaffected by the gate length, and therefore the incident flat top pulse width? For the two different breakdowns shown in Figures 16 and 17, one ignition point occurs just before the gate midpoint (downstream signal, Figure 16) while the occurs later in the pulse (upstream signal, Figure 17). The expected distribution of breakdown ignition times is not yet known, so existing breakdown and field emission literature will be consulted. The number and nature of pulses without precursory behaviour also needs to be determined.

3.3 Consequences for future studies

The most immediate application of this result is to the development of data acquisition frameworks for recording dark current waveforms with higher bandwidths and sample rates. Design of a cyclic buffer to store the previous N pulses before a breakdown pulse is ongoing, but more information is needed to know how large N should be. Extra storage space used up by storing N pre-breakdown pulses is also unknown, as long as N is unknown. This question can be addressed further by examining X-Box 2 data, which stores not only pre-breakdown pulses, but pulses immediately prior to the pre-breakdown pulse. This would allow the evolution of the field emitter site to be monitored not just over two, but three waveforms. If the emitter is absent in the penultimate, or pulse before pre-breakdown pulse, then a cyclic buffer of $N = 10$ will not yield any extra information. The extra data wrangling effort of examining X-Box 2 data, which follows different naming conventions and organisation to X-Box 3 data, is therefore justified.

Should more data be required, the timing system of the PXI to store "normal", non-breakdown pulses can be exploited. X-Boxes 2 and 3 save a "log" or normal pulse every 60 and 40 seconds respectively, without consideration to whether or not a breakdown occurred. It is therefore possible that a "log" pulse is saved within a few seconds of a breakdown occurring. Due to the short time separation, it would be possible to frame such pulses as part of the "chain" of pulses leading up to breakdown, and use them to shed light on emitter development.

Another application of the previous section is to the design and calibration of more sensitive diagnostics for dark current measurement. By accurately calibrating the ADCs with a DC input, and then measuring the ADC counts recorded, the ADC counts can be converted to Voltages. The vertical dynamic range and offset can therefore be adjusted so that saturation of the ADC is avoided, and leading peaks can be recorded with good fidelity. This is important because the relative charge contained underneath the leading peak of downstream signals, relative to the charge contained in upstream signals, is currently unknown. The dynamic range of the NI-5761 250MSPS 14-bit ADCs on the PXI has to be adjusted to allow the higher amplitude signals to be recorded.

To target pre-breakdown dark current waveforms, multimode optical fibers coupled to Silicon Photomultipliers (SiPMs) are being considered as a more sensitive means of dark current detection. Emission of secondary charge from the structure bulk will be incident on the silica fiber core, creating Cherenkov radiation, which will then be detected by Silicon Photomultipliers. If faraday cups address the upper end of the dark current dynamic range, then optical fibers are well suited to record the lower amplitude end (e.g. the pre-breakdown waveforms) with greater fidelity, as SiPMs can detect single photon events. Suitability of optical fibers in detecting the pre-breakdown waveforms shown in this work, with improved frequency response and time-resolution, should be assessed.

4 Conclusion

This document explained how breakdowns are identified in an RF test stand dataset. Tagging, filtration and cleaning of breakdown candidates was summarised, and useful tools for further analysis, such as m-space, were highlighted. The expected dark current signal shapes were explained, for both pre-breakdown and breakdown waveforms. It was found that breakdown and pre-breakdown waveforms are identical up until the point of breakdown ignition, providing impetus for studies of field emitter evolution using a combination of historical (X-box 2) data, new data acquisition frameworks under design (both hardware and software), and novel, fiber optical breakdown diagnostics.

References

[Woo15] Benjamin J. Woolley. "High Power X-band RF Test Stand Development and High Power Testing of the CLIC Crab Cavity". PhD thesis. Lancaster University, 2015. (Visited on 05/18/2022).

[Raj16] Robin Rajamäki. "Vacuum arc localization in CLIC prototype radio frequency accelerating structures". MA thesis. 2016.

[Luc18] Thomas Geoffrey Lucas. "High Field Phenomenology in Linear Accelerators for the Compact Linear Collider". PhD thesis. University of Melbourne, 2018. (Visited on 05/18/2022).

[Pas20] Jan Paszkiewicz. "Studies of Breakdown and Pre-Breakdown Phenomena in High Gradient Accelerating Structures". PhD thesis. St. John's College, Oxford, 2020. (Visited on 05/18/2022).

[Obe+22] Christoph Obermair et al. *Explainable Machine Learning for Breakdown Prediction in High Gradient RF Cavities*. 2022. URL: `https://arxiv.org/abs/2202.05610`.

[Pus22] Paarangat Pushkarna. "Examination of breakdown precursors in high gradient accelerating structures". PhD thesis. Melbourne: The University of Melbourne, Oct. 2022. 50 pp.