Post-publication review of ‘Flux-tuned topological superconductivity in full-shell nanowires’ Vaitiekenas et al. Science 2020

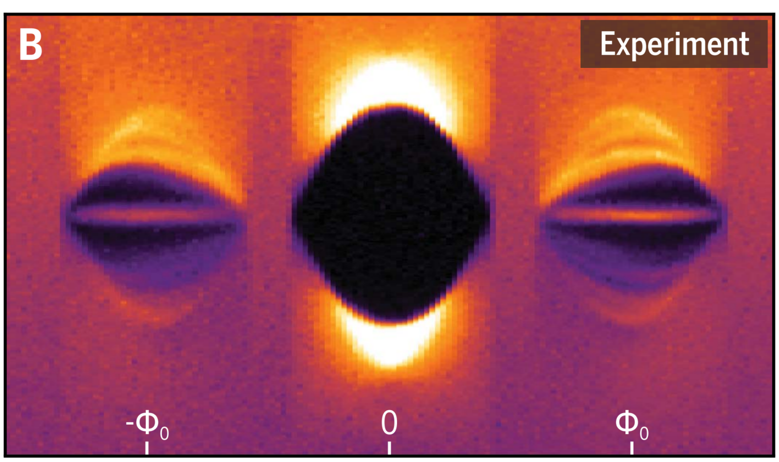
Sergey Frolov, University of Pittsburgh
Vincent Mourik, Forschungszentrum Jülich
March 11, 2022

Summary of Our Findings

In March 2020 a research group at Microsoft / University of Copenhagen published a paper in *Science* titled “Flux-induced topological superconductivity in full-shell nanowires” https://science.sciencemag.org/content/367/6485/eaav3392 (in what follows Science-2020). The authors claimed (both through explicit statements in the text and through their choice of what data to present) that some nanowires they have studied exhibit robust zero-bias peaks (i.e. peaks in conductance at zero voltage) coinciding with the odd lobes of a multi-lobed phenomenon known as the Little-Parks Oscillation. They interpret this as evidence for topological superconductivity, i.e. Majorana particles.

In October 2020 the authors shared additional data from those experiments on Zenodo http://doi.org/10.5281/zenodo.4263106. Having carefully studied the paper, as well as additional data, we conclude that data in a key figure, Figure 2, are not representative of the full experimental data obtained. Contradictions are found not only in analogous data from other devices, but even over larger parameter ranges from the same device.

Contrary to statements in the text, we find evidence that the authors’ supposedly robust zero-bias peaks fail to coincide consistently with the odd lobes of the Little-Parks Oscillations. We also find evidence that some supposed zero-bias peaks are split peaks, and that zero-bias peaks are gate-dependent, and not robust throughout the tunneling regime. Again these findings are contrary to explicit statements in the paper.



Left: Paper figure in the Research Article Summary shows a coincidence of a zero-bias peak and Little-Parks oscillations, but additional data contain examples where this coincidence is not present.

With respect to data as shown in a key figure, Figure 2, we found multiple statements throughout the paper to be inaccurate or misleading as applied to the full data available. These statements, identified below, concern both the description of the data, as well as the interpretation of the data as relevant to ‘flux-induced topological superconductivity’ - the topic of the paper. Thus the core conclusions of the paper are invalidated. We do not see how this paper can remain in the present form, nor how the problem can be addressed by publishing any form of clarification such as a correction.

Selection of data for publication is an inevitable part of the research process. Hundreds and thousands of datasets are obtained for each study in our field. Yet only 5-10 become figures in a given paper. Here we identify a situation in which the selection of data for this paper is not representative of the total data obtained. Our field of study is vulnerable to this problem. Data chosen for the figures tell a compelling story, while full data present contradictions. Sometimes the contradictions are so serious that the paper is retracted. This was the case for ‘Quantized Majorana Conductance’ by Zhang et al., Nature 2018.

On July 2 2021, Science published an experiment from a group at IST Austria, which uses the same nanowires and the same measurement technique used for Figure 2 of Science-2020, but arrives at opposite conclusions (Valentini et al., Science 2021 https://www.science.org/doi/10.1126/science.abf1513). The analysis by the IST group presents a fuller body of data. While Niels Bohr Institute has stated that they conducted an internal investigation into Science-2020, we believe external and independent expertise is required to further assess this issue. We continue to request that the corresponding authors share full data obtained.

The Editor-in-Chief of *Science* issued an Editorial Expression of Concern on Science-2020 on July 30, 2021.

Table of Contents

1. Background information on Majorana physics

It is not required to be an expert on Majorana research to understand our findings. In fact, little to none of the Majorana physics is involved in our arguments because the issue we identify is the non-representative character of the regimes shown in the figures. However, it is helpful to know several well-established and non-controversial facts about Majorana states and their relationship to the specific subject of the Science-2020 paper.

Names used for this phenomenon are: Majorana fermions, Majorana bound states (MBS), Majorana zero modes (MZM). A regime in which Majorana modes are present is commonly referred to as 'topological regime' or 'topological superconductor'. The title of the Science-2020 paper contains the words 'topological superconductivity', and the authors describe their observation of Majorana modes in the paper.

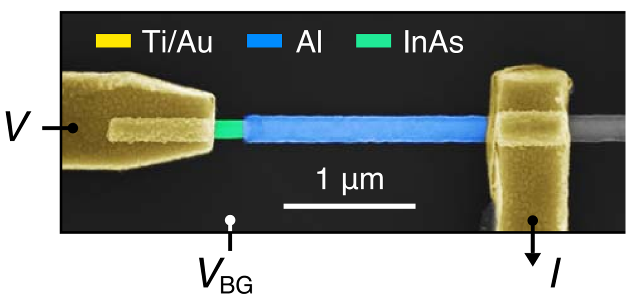
Majorana states have been predicted to produce conductance peaks in electronic transport measurements. These peaks are expected at <u>zero</u> source-drain bias voltage, hence Majorana <u>Zero</u> Modes. The peaks are referred to as zero-bias peaks (ZBP) or zero-bias conductance peaks (ZBCP). They are expected to not be transient, but rather remain at zero voltage while electric and magnetic fields are varied by experimentalists over some range.

It is also important to know that Majorana is not the only phenomenon that results in non-transient zero bias peaks. Another well-studied effect is called Andreev bound states and such states are expected to form much more easily than Majorana states, meaning they can exist over much wider ranges of experimental parameters. It is known, at least since 2014, that under certain conditions Andreev bound states can generate non-transient ZBCPs that can mimic Majorana states. Andreev states are sometimes called 'trivial', which is understood as 'not a Majorana state' and/or 'not topological'.

Having provided this background, we emphasize that the conclusions of the present analysis and its findings do not depend on the currently ongoing scientific debates about whether evidence supporting the existence of Majorana modes in any device or materials platform has been already obtained. We do not rely on the present scientific understanding of this question, and we do not rely on any research that has taken place after 2018, the date of the initial posting of results on arXiv. Instead, our analysis is confined to checking for consistency between written statements and the figures presented by the authors in their Science-2020 paper, and the larger dataset now available.

1.2 Measurement techniques used

The technique used to obtain data in Figure 2 is electrical measurements whereby a small bias voltage is applied across a sample and current, or conductance (the change in current as bias voltage varies), is measured. This is the technique used to search for a ZBP in nanowire devices. Devices studied this way are referred to by the authors as ‘NIS’ (normal-insulator-superconductor). This is because a separate voltage on a nearby gate electrode is used to deplete a segment of the bare semiconductor nanowire in a field effect transistor-like manner. This creates an insulating segment located between superconductor and non-superconductor (normal) metal contacts. Electrons have to quantum-mechanically tunnel across the insulating segment, so a zero-bias peak is obtained in a tunneling or a tunneling spectroscopy measurement.



A segment of Figure 2A. Source-drain voltage bias V is applied to normal contact (Ti/Au) on the left. Current I is measured after passing through Al superconductor (blue shell on top of nanowire). InAs (green) is the bare nanowire that becomes an insulator when gate voltage V_{BG} is applied to the doped silicon layer underneath (dark background).

Once a device is fabricated and set up for measurements in a cryostat, the primary control knobs available to the authors are source-drain bias voltage between normal and superconducting contacts, gate voltage(s), external magnetic field (or magnetic flux) and temperature.

The two additional measurement techniques used in the Science-2020 paper are supercurrent through the shell (Figure 1) and Coulomb blockade (Figures 6 and 7). Figure 1 does not show Majorana measurements and is redundant with respect to Figure 2 for the purposes of our analysis. Figures 6 and 7 (Coulomb blockade measurements) are secondary to data in Figure 2, meaning that the weight of the discussion in Science-2020 falls heavily on Figure 2. The non-representative nature of Figure 2 cannot be rectified by data in Figures 6 and 7. We make further comments on the Coulomb blockade technique below after presenting our analysis of Figure 2. Figures 3, 4 and 5 are devoted to theory, which is not relevant to how representative the tunneling data in Figure 2 are. We comment on the theory later.

2. Analysis of Figure 2

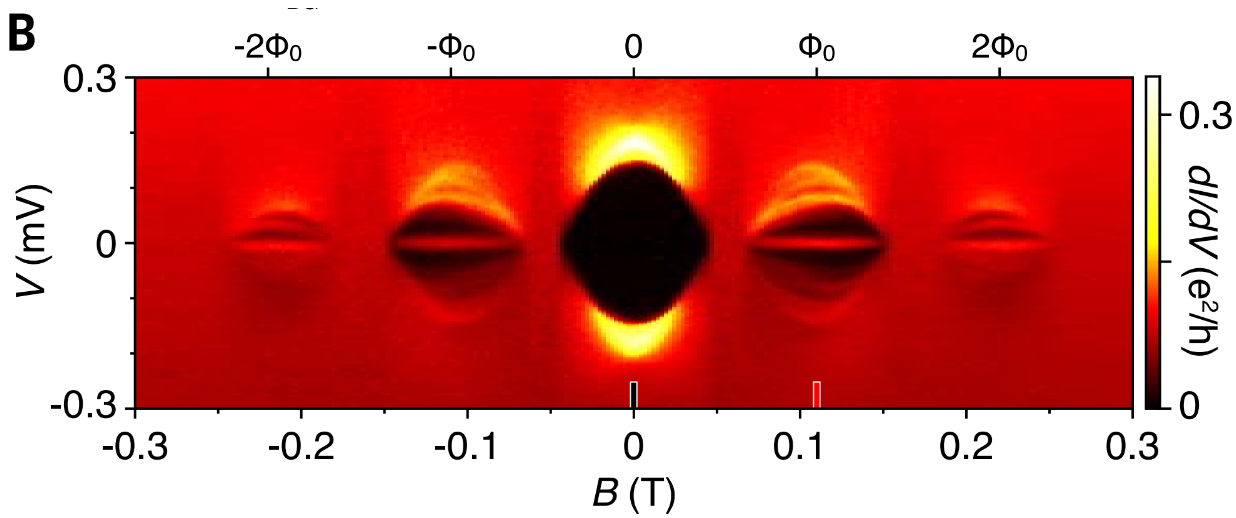
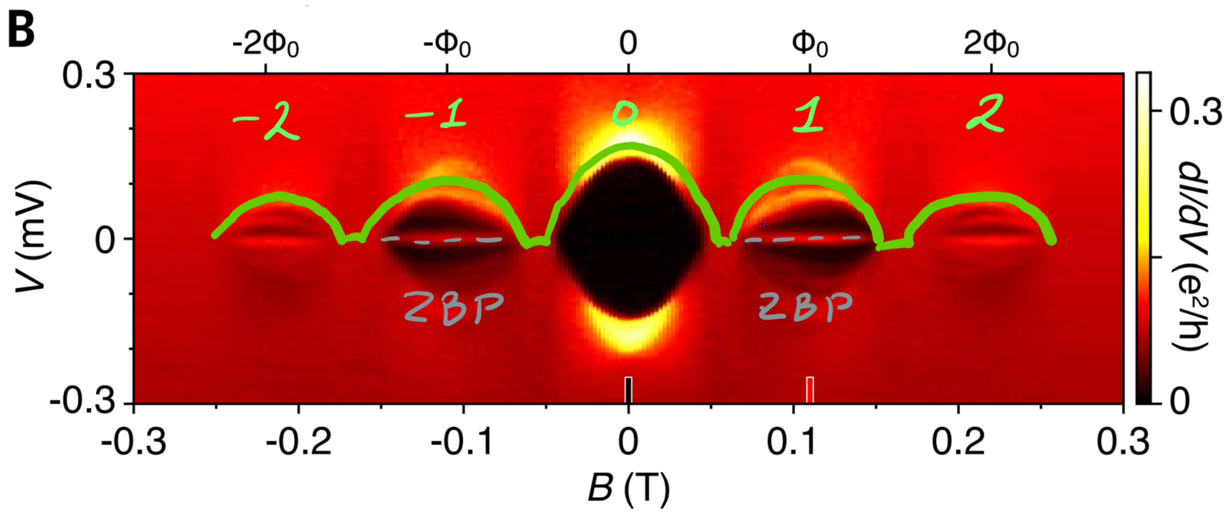


Figure 2 is composed based on three separate datasets, panels B, C and E, all presented as colormaps of conductance dI/dV, given in the units of e^2/h. Panel B is the most illustrative and the most discussed in the paper, it is a magnetic field B(T) evolution of conductance as function of source-drain bias voltage V(mV).

Panel 2B presents two key phenomena. The first of them is the zero-bias peaks that we mark with gray dashed lines below. They are visible at V = 0 mV and around B = 0.1 T and B = -0.1 T as two narrow and long red horizontal streaks surrounded by a dark background. The second is the oscillations of the dark, low conductance region around zero voltage bias that we marked by the green trace. They are well-known Little-Parks oscillations in superconducting rings. The lobes of oscillations correspond to the number of flux quanta threading the nanowire which we label with green numbers -2, -1 0, 1, 2. The top horizontal axis labels magnetic flux in units of flux quantum Φ_0. Important for understanding the authors' statements is that each flux quantum corresponds to the winding of the superconducting wavefunction's phase of 2π, a flux of $2\Phi_0$ is 4π winding, etc. This language of 'winding' is used in the paper.

The relationship between the two phenomena, zero-bias peaks (ZBP) and Little-Parks oscillations (LP) is the topic of the Science-2020 paper. The authors use Little-Parks oscillations as a counter for how many flux quanta are threading the wire. This number of flux quanta they connect to topological superconductivity (MZM):

“We show experimentally and theoretically that the winding of the superconducting phase around the shell induced by the applied flux gives rise to MZMs at the ends of the wire.” From the Structured Abstract, also referred to as the Research Article Summary.



Also from the main text:

“The topological phase sets in at relatively low magnetic fields (~0.1 T), is controlled discretely by moving from zero to one phase twist around the superconducting shell” Science-2020, p1.

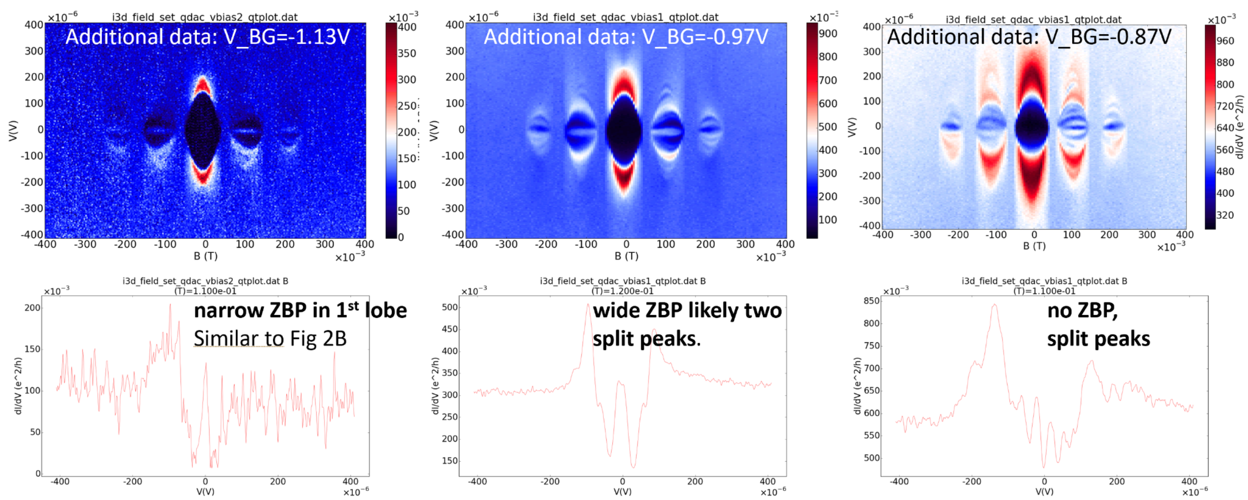
And here is an example of how the authors describe this relationship in more technical terms, relating zero-bias peaks to the first lobe of Little Parks oscillations (LP-1):

“In the superconducting regions around one quantum of applied flux, corresponding to phase twists of $\pm\ 2\pi$*in the shell, tunneling spectra into the core show stable zero-bias peaks, indicating a discrete subgap state fixed at zero energy.” Science-2020, p1.*

In Appendix A we list multiple instances where the relationship between ZBP and LP is stated in Science-2020, including 7 that discuss the significance of the coincidence of ZBP and the first lobe of the LP (LP-1). In what follows we analyze the ZBP/LP-1 coincidence itself and do not contribute any of our own interpretation of its relationship to the topological/Majorana physics. (We also analyze several other statements that the authors make about there being no subgap features in LP-0 or the even/odd pattern. Those statements are important but secondary to the coincidence of ZBP and LP-1.)

Figure 2 presents data from Device 1, and shows a clear coincidence of the ZBP with LP-1. To understand how robust this is, we first turn to additional data from the same Device 1; these data were not published with the Science-2020 paper. We replot these data from Zenodo files, hence the different colorscale.

We see three additional datasets, the difference between them is the setting of back gate voltage V_BG. Only one of the new datasets, V_BG = -1.13 V, shows a narrow ZBP comparable to panel 2B. At V_BG=-0.87V we see no ZBP in LP-1. Instead, peaks at finite V (also referred two as split peaks) are observed.



At V_BG=-0.97V (center) we see a ZBP which is thicker in voltage bias. Continuing the trend from data at V_BG=-0.87, this thickening of the peak can be a consequence of having not a single ZBP but two closely spaced and overlapping split peaks. This is also confirmed by the additional back gate voltage dependence data shown in the next section.

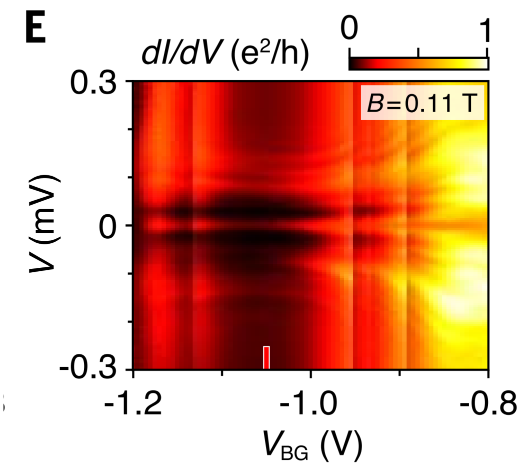
2.2 Gate-dependent zero-bias peaks

Additional data reveal that the presence/absence of ZBP is related to how the back gate voltage is chosen. This is seen in the magnetic field dependences presented above, and in the additional gate voltage dependences plotted below. The fact that gate voltage can control ZBP should be surprising to a reader of the paper, if one attributes its origin to the inner part of the nanowire that is protected by the full metallic shell. Lack of gating of the inner core of the wire is stated in the paper.

“Although full-shell wires do not allow for direct gating of the electron density in the semiconducting core, we demonstrated that, using a careful design of the wire

properties—for example, by choosing the appropriate radius— it is possible to obtain wires that harbor MZMs at a predictable magnetic field." Science-2020, p6.

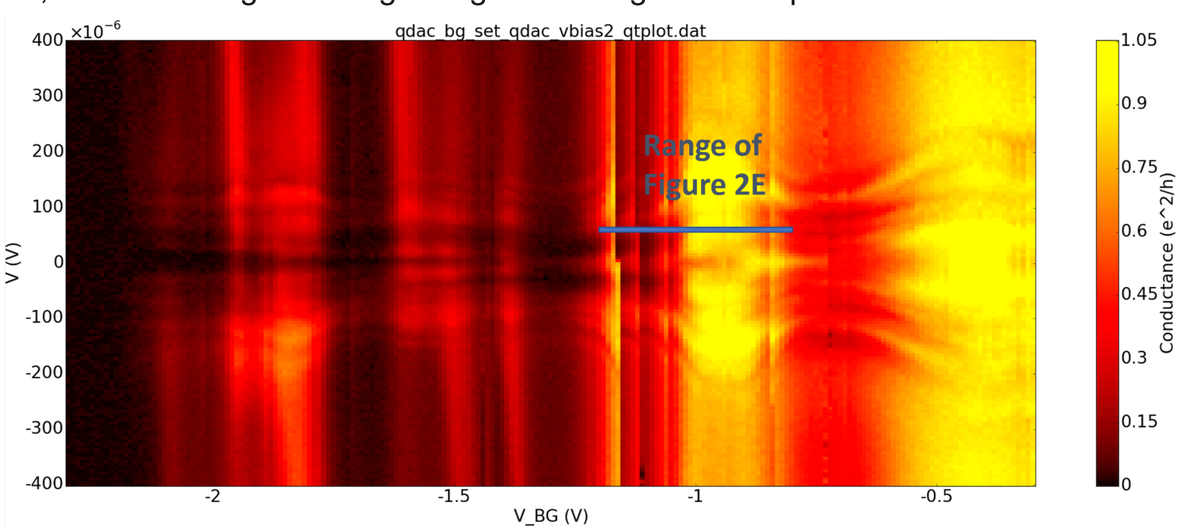
Zero-bias peaks due to these MZM would originate from a wire section fully covered by metal - and therefore should not be sensitive to gate voltages, because the electric fields from the gates are electrostatically screened and do not reach the semiconductor core clad in the full shell. The actual measurements related to gate sensitivity of ZBP are shown in Figure 2E, which is the back gate evolution of conductance, now at a fixed magnetic field. This is how data in panel 2E are described in the text:



"In the first lobe, at B = 110 mT, the sweep of VBG showed a zero-energy state throughout the tunneling regime (Fig. 2E). [...] As the tunnel barrier is opened, the zero-bias peak gradually evolves into a zero-bias dip." Science-2020, p2.

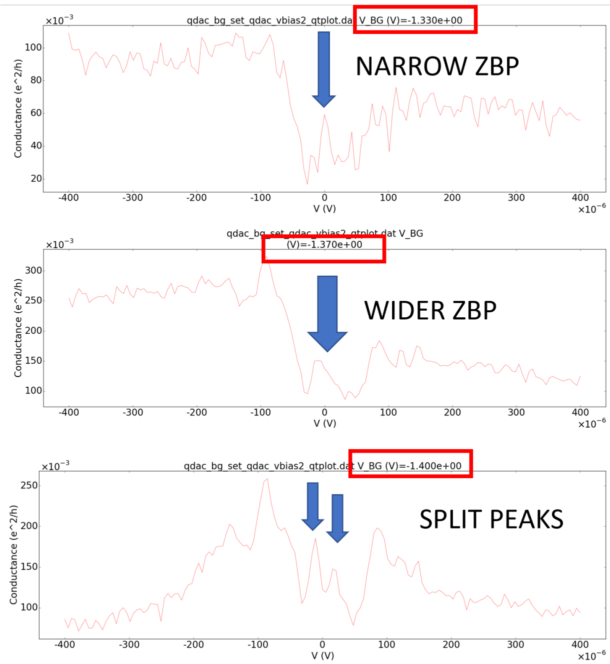
Indeed, in Figure 2E we see a zero-bias peak that spans most of the image. The 'dip' observed around V_BG=-0.8V is also interpreted by the authors in the context of MZM (more on this later). Thus, the entire panel E is associated in the paper with the topological regime.

However, in contrast with Figure 2E, additional data contain a dataset over a larger range of V_BG with a ZBP region similar to that in Figure 2E, but which covers only about 18% of the full gate voltage range. To establish some correspondence with Figure 2E, we mark the gate voltage range of the Figure in the plot of additional data.



After replotting and analyzing these data, we find the following regimes:

1) Noise without features between V_BG = -2.25 V and -2.19 V
2) Split peaks between V_BG = -2.19 V and -1.38V (does not appear in Figure 2E)
3) ZBP between -1.38V and -1.035V (similar to Figure 2E)
4) Split peaks between -1.03V and -0.90V
(similar to Figure 2E, and also where the magnetic field dependence shown in the previous section at V_BG=-0.97 falls, though gate voltages may not correspond to each other due to irreproducible charge jumps)
5) ZBP between -0.90 V and -0.70 V (does not appear in Figure 2E)
6) More split peaks between -0.7 V and -0.3 V (does not appear in Figure 2E)



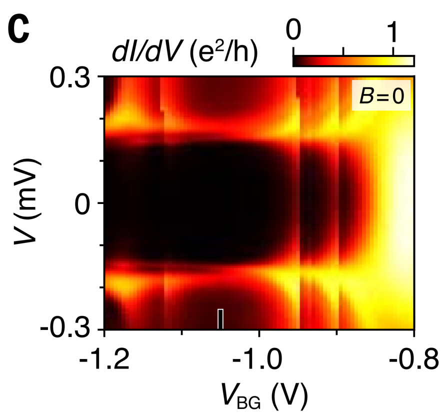
Thus, additional data of gate voltage dependence in LP-1 demonstrate regions of split peaks that exceed in gate voltage range the regions of zero bias peaks, and a second region of zero bias peaks at more positive gate voltages. This contradicts statements in the paper regarding the gate voltage dependence of zero-bias peaks.

We observe that the zero-bias peak never approaches the noise level in this additional dataset. It remains visible until it is replaced by a pair of split peaks. The splitting is gradual, as far as linecuts show. Below we explain how the limited amount of additional data, which is most likely not all of the data obtained, invalidates the claims of this paper and leaves little to no room for the claims to be accurate. We use the example of the extended range gate data. These data do not show 1:1 correspondence to the data in Figure 2E, presumably due to irreproducible charge jumps. Yet, no data acquired over a wider range of gate voltage than Figure 2E, but corresponding to Figure 2 within the figure's gate voltage range, have been provided. If further data do not exist, that is regrettable (knowing that ZBP existed only over part of the range in one dataset, the authors should have investigated whether that behavior repeated or not as part of determining whether ZBP is gate-tunable). It would also mean that the additional data are the best evidence of the behavior of device 1 over a larger gate voltage range, and should certainly have been discussed in the paper.

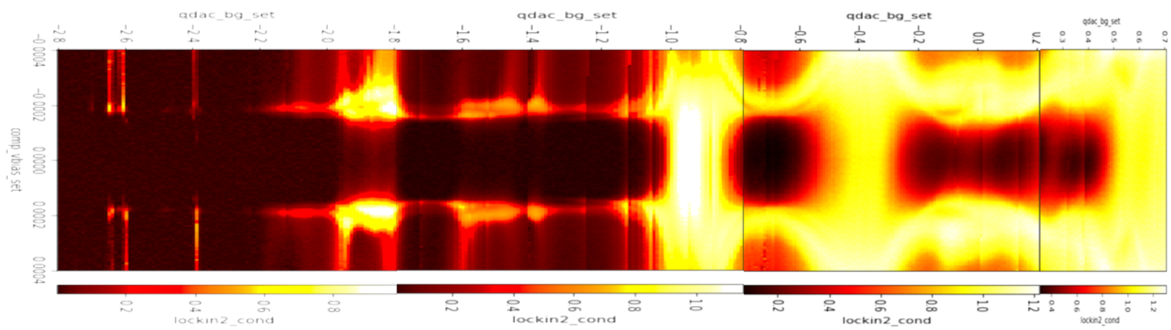
If data that correspond to Figure 2 even closer than the available additional data exist but are being withheld, they should be shown by the authors, and there are then two possibilities. One is that these withheld data show ZBP over the entire extended gate range. However unlikely, this would seemingly be in line with the claim of the paper of a "*zero-energy state throughout the tunneling regime".* However, taken together with the additional data we show here, this would demonstrate that device 1 can change between states of limited ZBP and "total ZBP", over time, and likely as a result of charge instabilities. This in itself suggests that the ZBP must be tunable by gates, which emanate electric fields.

The other possibility is that, like the available additional data, the rest of the withheld gate voltage data show ZBP only over the narrow range, i.e., they contain further evidence that the ZBP is gate tunable.

The description of Figure 2E contains the phrase '*throughout the tunneling regime'*. Science-2020, p2. We understand this phrase as the range of V_BG in Figure 2E (and 2C) being the only relevant range. The 'tunneling regime' is the regime where conductance at finite bias, e.g. V>0.3, is below 2 at zero magnetic field, and below 1 at large magnetic field. Figure 2E is composed in such a way that at the left boundary of the panel the overall conductance is low, 0.1-0.3, and at the right boundary of the Figure conductance is high, approaching 1. The same impression is given by panel 2C which is the same regime as 2E but at zero magnetic field.



The statement that the entire tunneling regime has been presented in Figures 2E and 2C can be scrutinized by considering the additional data obtained at zero magnetic field (this is a composite of four separate datasets, the range of Figure 2C is between -1.2 V and -0.8 V):



From these data, not published with the Science-2020 paper, we see that:

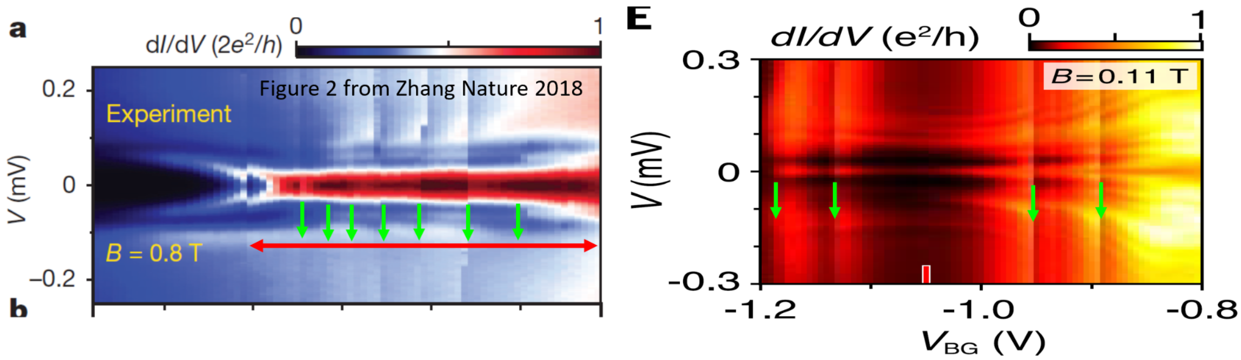
1) Conductance is below 1 at high V (e.g. top edge) for back gate voltages as high as V_BG=+0.2V, outside of high current resonances (vertical yellow streaks). At zero magnetic field, tunneling takes place for conductances below 2 in the basic picture.
2) The superconducting gap feature (horizontal peaks at V = +/-0.002) are present for back gate voltages as low as V_BG=-2.7V, indicating that electronic tunneling is still taking place.

From this we conclude that a much larger back gate voltage range -2.7V < V_BG < +0.2, or larger, corresponds to the tunneling regime. The data in Figures 2E and 2C, over the gate voltage range between -1.2 V and -0.8 V occupy only a fraction of the tunneling regime. Consequently, the statements in the paper that zero-bias states are observed throughout the tunneling regime are incorrect, since in other parts of the range split peaks are observed.

We stress that the argument we make does not boil down to a debate about the appropriate gate voltage extent of the tunneling regime. The need to have presented a larger range of gate voltages within the paper, so as to make it clear that gate voltage does have a significant effect on the apparent zero-bias peak, is only made stronger by the uncertainty in determining which range corresponds to the tunneling regime. The authors did not present these data, but instead put statements in their paper to suggest the absence of gate effect over a relevant gate range. The possibility that this is not the full relevant gate range was not presented.

The gate voltage range with ZBP that was shown in Figure 2 is extended by ‘charge jumps’ which are sharp discontinuities in the colormap. As the back gate voltage increases the signal shifts back to a previous value, in a manner similar to a skipping record, but due to a shift in the position of a small charged particle near the nanowire.

Charge jumps in the recently retracted ‘Quantized Majorana Conductance’ Nature paper, originally published in 2018 by a different set of authors, extended the gate voltage range of ZBP (green arrows, but see footnote[1]). In Science-2020, the fact that Figure 2 does not show an extended peak but a pair of peaks that come together and immediately disperse, becomes more apparent if you go through the exercise of removing segments that appear to repeat themselves upon charge jumps. In both cases, the main problem, in our view, is showing non-representative data affected by charge jumps.

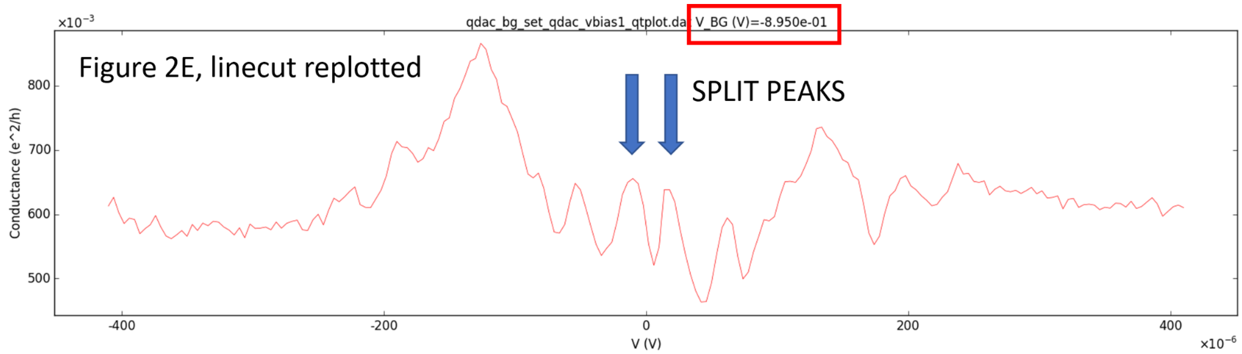


We furthermore find that the following sentence from Science-2020 is inaccurate. It is important to point out this inaccuracy because the statement connects a feature of the data to a ‘theory supporting MZMs’:

“As the tunnel barrier is opened, the zero-bias peak gradually evolves into a zero-bias dip. The increase of finite-bias conductance compared with zero-bias conductance as tunnel barrier decreases is in qualitative agreement with theory supporting MZMs” p2.

The inaccuracy occurs where the authors discuss the evolution from a zero-bias peak to a zero bias dip. In fact, the zero-bias peak splits and becomes two peaks. We replotted Figure 2E ourselves to verify this, and we can also see from additional magnetic field sweep data (e.g. top right panel of the figure on p7 of this analysis), as well as from additional gate voltage sweep data, that these are not dips but rather a pair of peaks,

[1] In ‘Quantized Majorana Conductance’, Zhang et al. Nature 2018 (retracted), in addition to choosing a dataset which extends the longevity of the ZBP in gate parameter space due to charge jumps, several unfavorable charge jumps that could have cast doubt on the ZBP longevity were removed by cutting a piece of data from the total dataset and stitching the remaining data together. This data manipulation was not disclosed. We explicitly do not allege that the same happened in Science-2020, on the contrary, as far as we can tell, no selective removal of charge jumps took place. However, we do allege that the charge jumps in Science-2020, as in the retracted Nature paper, fulfill the same role of artificially creating an impression of a long-lived ZBP in gate parameter space, whereas additional data shows that the ZBP is absent in most of the gate parameter space. We find this a more significant problem.



split away from zero bias voltage to a finite positive and finite negative voltage, symmetric around zero bias voltage.

At more positive gate voltages an apparent dip can indeed be seen at zero voltage bias. However, where the authors make a link to an MZM theory they do not discuss or explain how split peaks fit into this comparison between experiment and theory. We find this omission to be problematic, based on our understanding of the cited and other MZM theories.

<u>To summarize our analysis of available tunneling data for Device 1</u> (Figure 2 vs. additional data) we conclude that both the magnetic field dependence (panel 2B) suggestive of correlation between ZBP and LP-1, and the gate voltage dependences (panels 2C and 2E) suggestive of zero-energy states present throughout the tunneling region of back gate voltage are not representative of the total data. This leads to statements in the text of the paper, where these data are described and generalized, to be invalid.

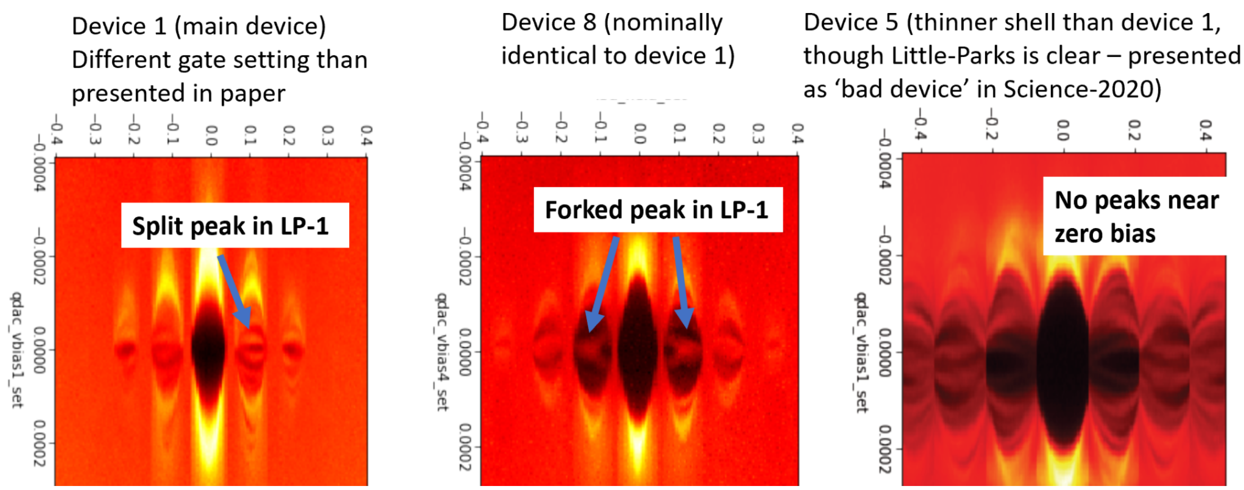
3. Device-to-device reproducibility

An important part of the argument that the authors make in their Science-2020 paper is reproducibility of these observations from device to device. They write:

“Two device geometries, measured in three devices each, showed similar results.” Science-2020, p1.

“Moreover, the full shell naturally protects the semiconductor from impurities and random surface doping, thus enabling a reproducible way of growing many wires with essentially identical electrostatic environments.” Science-2020, p6.

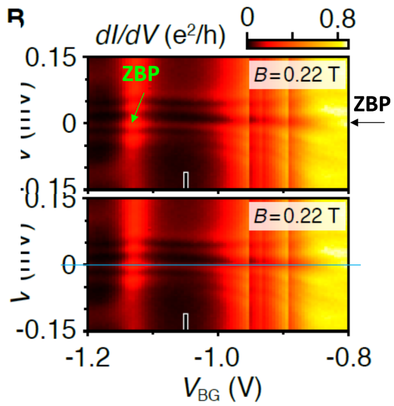
The total number of devices for the tunneling geometry was subsequently disclosed by the authors to be 9. We can compare data from those devices to Figure 2.



On the left we repeat data from Device 1, not shown in Science-2020, where no ZBP is observed in LP-1. In the center, we see data from Device 8, not shown in Science-2020, which authors describe as ‘nominally identical’ to Device 1. We see that within LP-1 a pair of split peaks merge into a ZBP, a behavior which suggests that there is no connection between the phenomena of LP and ZBP in this case. On the right, we see data from Device 5 which the authors describe as different from Device 1, a statement that does not find strong support in additional data (more on that later). In this device, while LP oscillations are clear, we see no ZBPs.

Some datasets from other devices, 3, 4 and 9, do show a coincidence between ZBP and LP-1 in magnetic field sweeps, in line with Figure 2B. However, as we demonstrate above for Device 1, this can be a result of acquiring data at an atypical setting of gate voltage where such coincidence appears. Recall that from each device we only see a few datasets on Zenodo. Taken together, additional data and paper data provide plenty of examples where the relationship between ZBP and LP-1 is either present or absent. However, Figure 2 of the Science-2020 paper itself only shows that the two phenomena coincide.

3.2 Even/odd pattern



The authors extend and generalize their experimental statements beyond the LP-1 coincidence. They suggest that whenever the number of flux quanta is even, 0 or 2, there is no ZBP. But when the number of quanta is odd, 1 or 3, there is a ZBP. This fits with their MZM narrative, and they make multiple statements throughout the paper, where they indicate that their data support this 'even/odd' pattern. We list statements in Appendix A. We do not find support for the existence of such an even/odd pattern in the total data we considered.

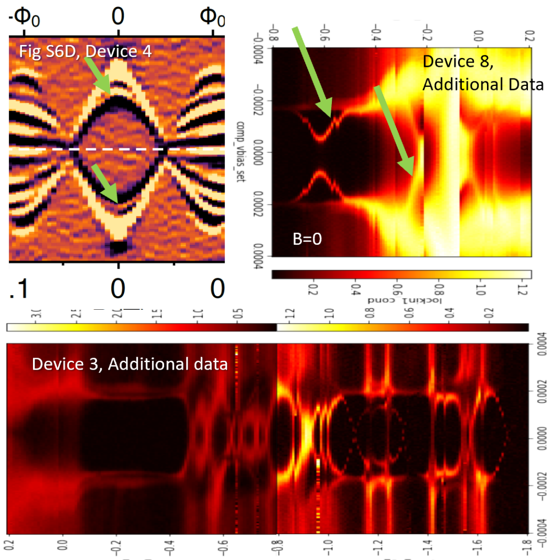
The data to support the entire even/odd ZBP/LP pattern in a single graph is limited to one dataset from Device 4 that shows a pattern up to LP-3. However, in Device 3, the key verification device, and in Device 12, nominally identical to Device 1, there are no ZBPs in LP-3. In fact, the only dataset from Device 12 shows no peaks in any lobes. In Device 5 we see broad ZBPs in LP-3. We term Device 5 a 'bad' device, by which we simply mean a device that has no ZBPs in LP-1. In other devices LP-3 is not visible.

The even/odd pattern would imply that no ZBP is present in LP-2. However, we find for Device 1 a zero-bias peak in the gate voltage dependence within LP-2. This contradicts this statement in the supplementary materials:

'In the second lobe, with even number of phase windings, the spectrum for device 1 features an asymmetric superconducting density of states with the lowest energy subgap state centered around ~5 ueV, see Fig. S3.' SM, p3.

To illustrate this statement, the authors present a magnetic field dependence focused on LP-2 at the back gate set to V_BG=-1.05 V. The zero-bias peak in Figure S3B is at V_BG=-1.15 V, a different back gate voltage.

3.3 Empty gap within LP-0 (near zero magnetic field)

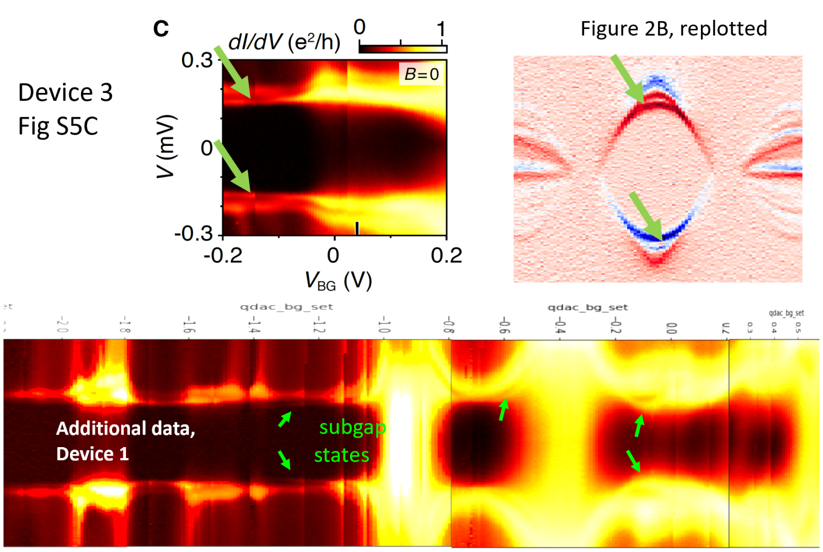


A statement regarding LP-0, repeated with variations several times in Science-2020, is stronger than the even/odd pattern would imply (i.e. no ZBPs in LP-0). The statement is that no subgap states at any bias are present in LP-0:

‘*Our measurements reveal that tunneling into the core in the zeroth superconducting lobe, around zero flux, yields a hard proximity induced gap with no subgap features.*’ Science-2020, p1.

The importance of the statement of ‘no subgap features’ in LP-0 for the narrative of the paper is high. If there are subgap features in LP-0, it leaves the possibility open that these features shift to zero energy in magnetic field due to its Zeeman/orbital effect purely within the semiconductor, rather than ‘discrete phase winding’ in the full-shell of a superconductor, an aspect that is presented as unique to this paper. The notion that ZBP in LP-1 appears ‘out of nowhere’ would then tilt a reader towards believing that a discrete value of magnetic flux, within LP-1, is what generates that zero-bias peak.

Contrary to the statement above, we find in devices 1, 3, 4, 5, 8, 9 subgap features at zero magnetic field. The features are marked with green arrows for devices 1, 4 and 8.



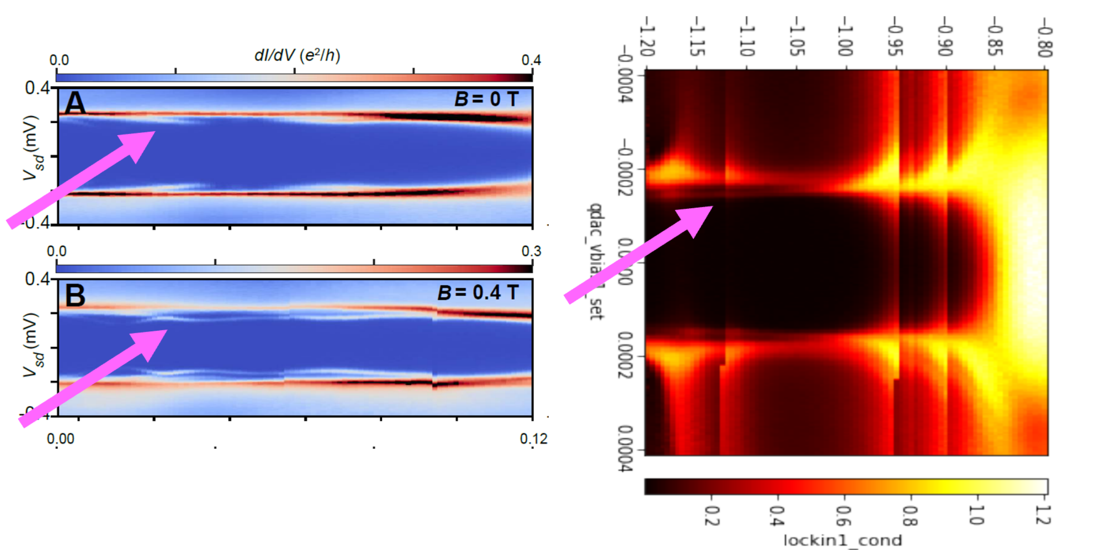
In Device 3, the states are so numerous that it is not practical to mark them all.

In Device 4, and in the main Device 1, which is the basis for Figure 2, we identify subgap features that happen to be located very near the edge of the gap. As a reminder, ‘the gap’ is a horizontal peak in the back gate voltage dependence at approximately +/- 0.0002 V.

We notice that these subgap features exist at and around those gate voltages that are used in Figure 2. In Figure 2 these are not apparent, and only can be identified if an additional data processing step is taken. In the replotted Figure 2B we take a numerical derivative in the vertical direction to identify a feature near zero field. This numerical derivative is the same procedure that the authors applied in Figure S6D for Device 4, which also shows similar features right below the gap.

In the additional gate-dependence data for Device 1, not shown in the paper, we see subgap states that are distinguishable from the gap without extra data processing. They appear as peaks at different V values than the gap. These subgap features appear either as peaks that track along the gap for a range of gate voltage, or as half-loop resonances. While device 1 does experience frequent charge jumps, meaning that data taken at different times come with differences, the overall features are comparable.

We find a comprehensive discussion of such states in an earlier Science paper from the same group that lists some of the same key authors, Deng et al., Science 2016 “Majorana bound state in a coupled quantum-dot hybrid-nanowire system” (Science-2016). Science-2016 in Figure S4 shows states that at zero magnetic field share appearance with those present in devices 1 and 3 in Science-2016 (see magenta arrows below).



Left: Deng et al, Science 2016 Right: Vaitiekenas et al, Science 2020

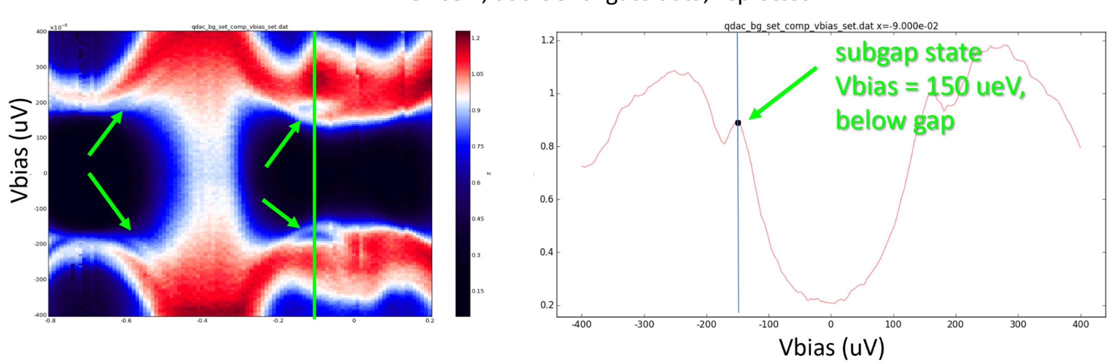
The caption of figure S4 in Science-2016, starts with a bold sentence “**Longitudinal ABS modes from the same subband**”. (‘ABS’ stands for Andreev Bound States, which are an example of subgap states). This figure focuses on the resonances highlighted by magenta arrows. The caption also provides a theoretical model that compares these resonances to ‘lowest-energy and second-lowest-energy ABSs’. In the supplementary text, these are described like this:

“At zero magnetic field, the induced superconducting gap is hard, with a few weak subgap states close to the bulk (continuous) gap edge. As the magnetic field increases, these subgap states (corresponding to one of the Zeeman-split branches) move towards lower energy. ”
Science-2016, Supplementary Material, p. 4

Indeed, Science-2016 Figure S4 does demonstrate how these ‘weak subgap states close to the bulk (continuous) gap edge’ move under the influence of Zeeman magnetic field towards zero bias. The nanowires used in Science-2016 do not have full Al shells therefore the movement of states could not have been induced by discretized phase winding. We also note that in Figure S6 of Science-2016, ABS merge into a zero-bias peak at a low field of B = 80 mT, comparable to the onset of LP-1 in Science-2020. Other papers, such as Albrecht Nature-2016, show zero-bias peaks onset at a variety of magnetic fields including low fields, merging from ABS at different zero-field energies, also without invoking ‘discrete phase winding’.

For additional data from device 1, the half-loop visible at Vgate = -0.1V reaches down to a voltage bias of 0.00015 V, below the value of the gap. Thus this constitutes a subgap resonance. The gate voltage range of this half-loop is greater than 0.3 V, a range comparable to the entire gate voltage range of Figure 2C, 2E. The half-loop shape is often encountered when ABS are studied. In those studies the gate range gives a characteristic range for tuning from even to odd occupation of ABS states in a quantum dot.

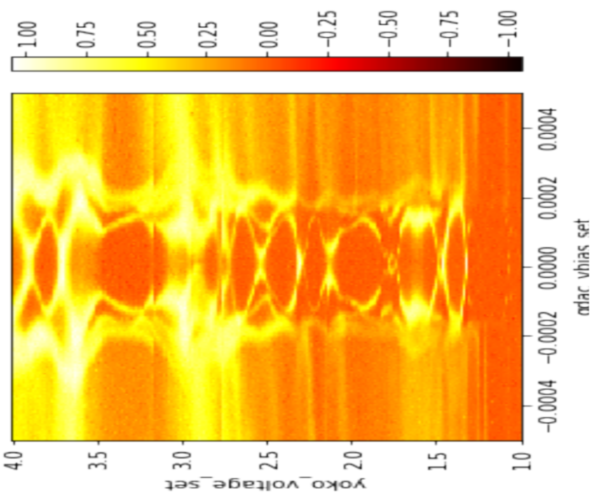
Device 1, additional gate data, replotted

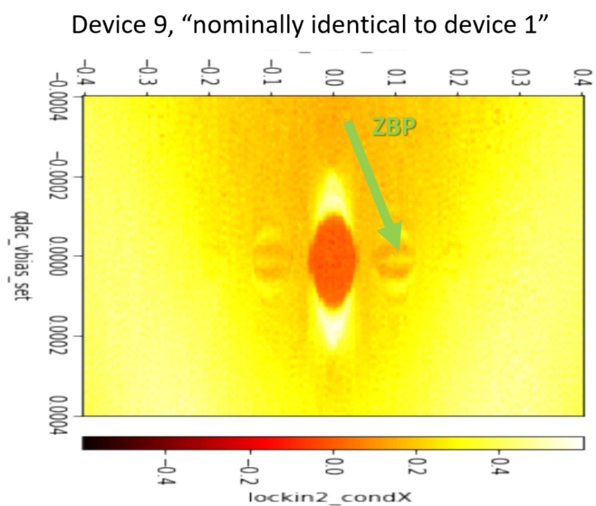


The fact that the existence of subgap features was not disclosed or discussed in the paper is problematic, regardless of their ultimate origin which cannot be unambiguously elucidated based on the available data. Statements about LP-0 made in the paper lose part of their support and become either inaccurate or misleading. Despite having studied this question earlier, in Science-2016, the real possibility that the resonances are ubiquitous and have trivial, non-topological explanations, has not been investigated through extensive data acquisition. At least, such data have not been shared with us despite multiple requests.

3.4 More examples illustrate the large variety of observed phenomena

Device 9, revealed with additional data and referred to as '*nominally identical to device 1 lithography*' (Zenodo repository, additionaldata_summary.pdf, pg. 118-121), shows that it is possible to see a ZBP correlated with LP-1 in a device that, at zero magnetic field, contains numerous states that are not confined near the gap edge, but instead cross V=0 throughout the gate voltage range. Thus, additional data also demonstrate that a zero-bias peak can be observed in the presence of ubiquitous subgap features. This raises further questions about the correlation of LP-1 and ZBP and how the data might look in other regimes not presented in the paper or in the additional data.





Device 8 offers another illustration that a wide range of behaviors has been observed in the authors' experiments, and the coincidence of ZBP and LP-1 is not a representative behavior. At zero magnetic field, Device 8, referred to as 'nominally identical to device 1 lithography' (Zenodo repository, additionaldata_summary.pdf, pg. 114-117), exhibits two features, one a resonance that does not cross V=0, and appears at gate voltage V_BG = -0.3. Another one is a resonance that does cross V=0 at V_BG=-0.2 (a charge jump at the same gate voltage makes it harder to see). The authors set V_BG in between the two resonances (blue line), presumably in order to avoid them. However, when the magnetic field is turned on we see a fork-like resonance in LP-1: a pair of peaks merging into ZBP only half way through LP-1.

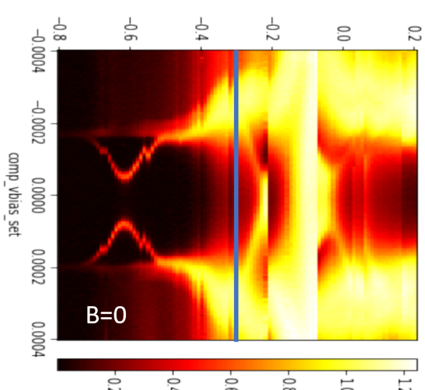
This is another example of possible behavior of subgap resonances. It is likely that if more magnetic field sweeps datasets were provided to us, e.g. at different gate voltages, more examples would be collected showing a rich variety of behavior, but not a correlation between ZBP and LP-1. It is possible that the authors already have more data like that, in which case they should share these data. At the very minimum, additional data demonstrate a range of phenomena that greatly exceeds those presented in the Science-2020 paper.

3.5 Device 5: “bad” device with no ZBP but with Andreev Bound States

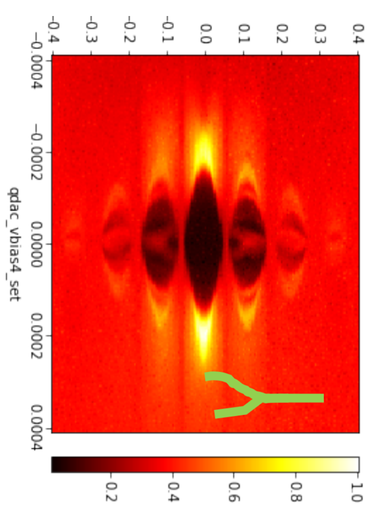
The authors present a device in the Science-2020 paper, Device 5, for which they do not show data with ZBP in LP-1. They explain the difference between Device 5 and other devices (1, 3, 4) as due to a thinner shell of Al, and/or thicker nanowire diameter in Device 5. We refer to Device 5 as ‘bad’ and to Devices 1, 3, 4 as ‘good’, with the meaning that ‘bad’ devices do not show ZBPs in LP-1.

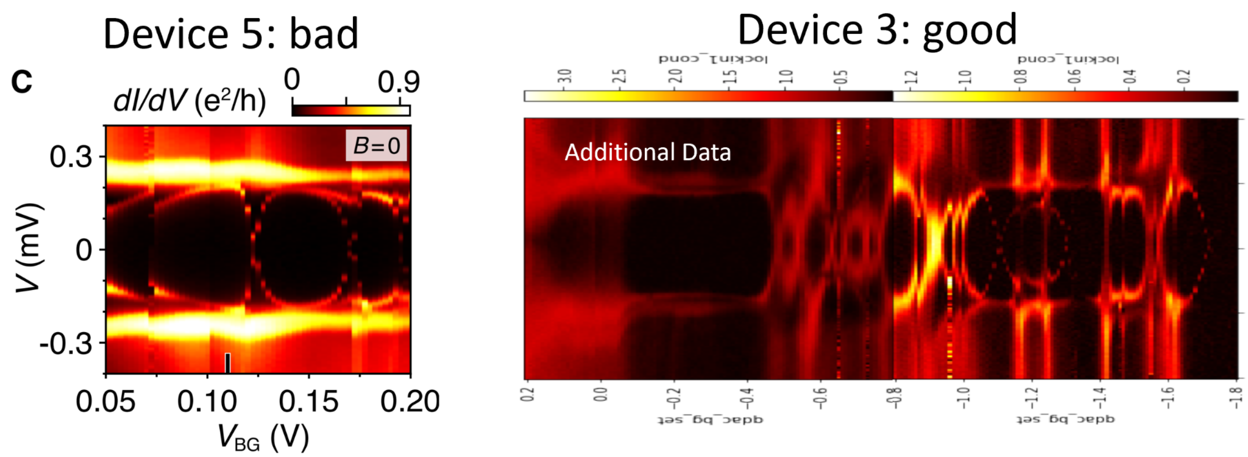
There is no evidence provided that the absence of ZBP is due to differences in the nanowire geometry. But we find even more significant the fact that Device 5, for which ZBP is not shown, is the only device for which the paper acknowledges the presence of subgap resonances in LP-0. This is how the authors describe them:

“For device 5, a discrete state crosses zero-energy around VBG = 0.12 V and then again at 0.17 V, resembling a proximitized quantum dot state, similar to the one previously studied in Ref. [84], see Fig. S7. We usually associate such state with a resonant level in the barrier and if possible avoid it in the measurements.” SM, p3.



Device 8,
“nominally identical to device 1”
From Additional data





A *'proximitized quantum dot state'* is another name for an Andreev Bound State (ABS). As a reminder, Andreev Bound States are a widely known explanation for zero-bias peak, which is potent because these states can appear over parameter ranges much larger compared to MZM. A reader might infer that Andreev Bound States were only found in this 'bad' Device 5. But as we already demonstrated, subgap features, some of which are likely Andreev Bound States, are visible in many Devices.

To illustrate this, we show side-by-side with data from Device 5, an analogous dataset from Device 3. In contrast to Device 5, which is a 'bad' device with ABS and no ZBP, this Device 3 plays a role of the verification device in the paper. Meaning, this is a 'good' device that is used to show that the findings from Device 1 can be reproduced. As we see from the data shown above, Device 3 is actually similar to Device 5. Both exhibit multiple subgap features that cross V=0. This can only be seen in the additional data for Device 3, and not in the Science-2020 paper, where the authors select regions with only subgap features near the gap edge, and not the ones that reach V=0, for supplementary Figure S5.

4. Non-Majorana interpretation of the same data

After the paper's figures, the main text of the paper is the most important factor in how the scientific claims are perceived by readers of the published paper, and before them by referees and editors when weighing the paper for publication. We already presented a case that individual statements from the main and supplementary text are incorrect. Here we discuss the paper as a whole.

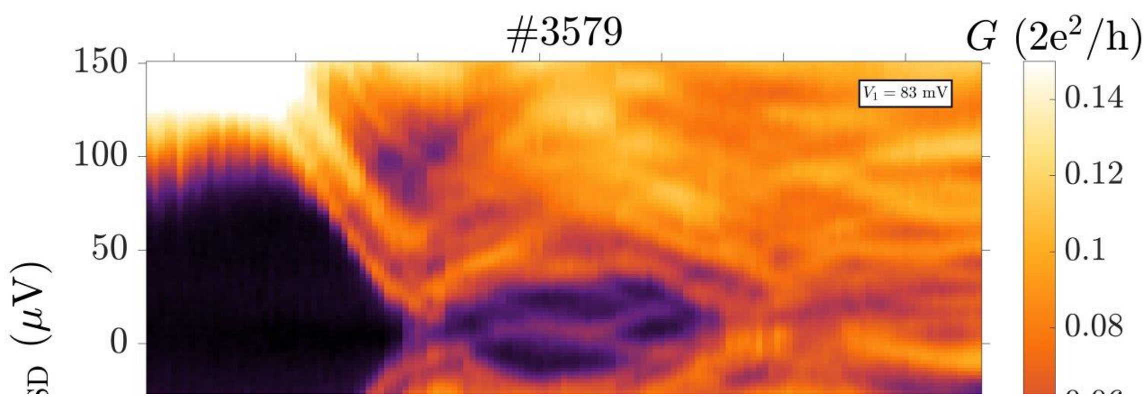
In the experimental part of Science-2020, data descriptions as well as data interpretation are presented in simple and non-nuanced fashion. They project confidence in conclusions and stand in contrast with the complexity of phenomena revealed in additional data.

It is especially striking that an alternative explanation of ZBP, in terms of Andreev bound states or related phenomena, is absent from the main paper text entirely. No mention of this possibility is made in the text. Out of 82 references, only a single theoretical paper (Ref 75) deals with Andreev ZBPs, but the paper is not cited in this context. No experiments that demonstrate a ZBP due to non-Majorana origins are referenced. Contrarily, 6 experimental articles are referenced interpreting zero bias peaks through the Majorana hypothesis.

A reader is not informed of the alternative explanation, and if they know about it, their mind is directed away from it. An example of such misdirection is the description of Device 5 which shows Andreev bound states in supplementary materials, where this device is explained as distinct from Device 1 due to different nanowire cross-section. It is dismissed using confident language but without evidence, based on speculation.

Distinguishing the Majorana hypothesis from the Andreev hypothesis is the main task in this research community when conducting experiments. The authors are well aware of widely discussed trivial zero-bias peaks due to Andreev bound states. They know about this from literature (first published in 2014), conferences (first presented in 2013) and many personal interactions. For example, in the data repository linked to their new arXiv publication (Bannerjee et al. arXiv 2022 https://arxiv.org/abs/2201.03453) members of the Copenhagen group refer to a plot of zero-bias peak in magnetic field as a ‘Silvano’, in an apparent reference to Silvano De Franceschi, the corresponding author of the 2014 paper that first discovered ZBP due to ABS. The 2016 Science paper from the Copenhagen group (Deng et al, Science 2016) with many of the same key authors discusses Andreev and Majorana states together. The corresponding authors confirmed their awareness of the Andreev interpretation directly to us in connection with this post-publication review.

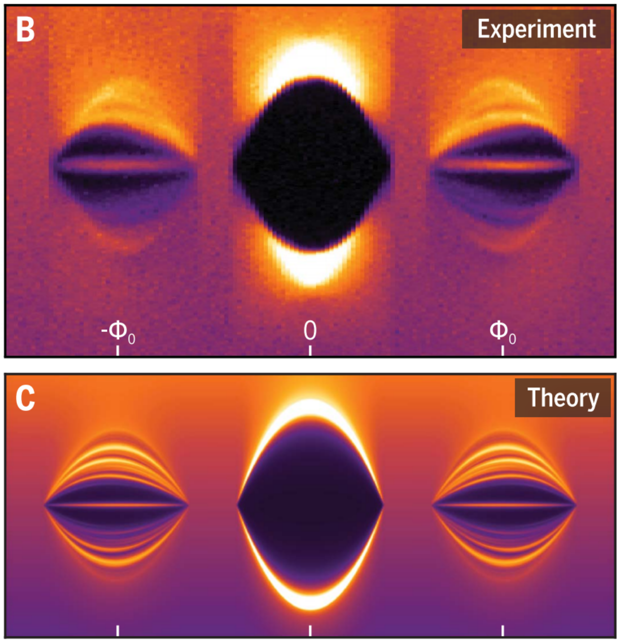
	High resolution phase scans to reconstruct a Silvano at V1 = +83 mV and V1=86mV



If there is an alternative explanation, established for many years, at the very least this urges caution and extra rigor in making an ambitious Majorana claim in a high-impact journal. It should not be enough to offer a mere possibility that experiments may have a connection to Majorana, and ignore substantial evidence that the optimism is not based on reality. It should not be enough to show only a few datasets to substantiate your conclusion. Any Majorana theory provides plenty of opportunities to check and cross-check the experiment. After obtaining an interesting zero-bias peak, it is possible to further verify whether any given ZBP it is due to Majorana or if it is just an example of fine-tuned Andreev states by e.g. repeating magnetic field sweeps for a large number of different fixed gate voltages. This type of cross-check is at the core of our analysis, and it shows that multiple claims within the paper are inconsistent with the data.

4.2 Comments on the theory part

Science-2020 paper contains a theoretical contribution, to which Figures 3,4,5 are dedicated. Originally, the theory part and the experimental paper were separate papers: they were posted on arXiv under separate identifiers (1809.05513 and 1809.05512). There was no overlap in the list of authors between the theory paper and the experimental paper. A version published in Science-2020 combined the author lists. The content of the two papers was also combined. Its two cores remain similar, meaning that the arXiv posting of the theory paper has strong overlap with the theory part of the joint Science-2020 paper. The parts of the paper are written in different style: the



non-nuanced style of the experimental part is in contrast with the more nuanced style of the theory part.

We contend that the fact that the experimental part dedicated to the tunneling technique contains multiple incorrect statements, and its experimental conclusions where data are concerned are not supported by additional data, makes the Science-2020 paper invalid, in the form it was published (i.e. joint with the theoretical part).

There is therefore no need to consider the theory part in great detail. We limit our discussion to only a few remarks.

First, the theoretical models of nanowires are approximate and do not include all effects present in real devices. For example, the authors revealed that the nanowire cross-sections are not regular hexagons. This means that theory plays a mostly supporting role in an experimental paper such as Science-2020 and cannot carry the weight of proof. To illustrate this point, Valentini et al. Science 2021 also presents a theoretical model of the same nanowire devices. That theory does not use identical assumptions. Yet it produces the same data, notably without Majorana.

Second, the theoretical model in Science-2020 actually shows that a ZBP correlated with LP-1 can appear in the non-topological regime (Figure S23). Whether this observation has any relationship to the experimental part is unclear, since theory considers an idealized model of the device. Nevertheless, the discussion of Figure 23 contained in the supplementary material appears to be in direct contrast with the main text discussion of Figure 2 where no such possibility that a ZBP correlated with LP-1 need not be related to a topological state, is presented.

Third, and despite the fact that the theory is weak at associating ZBP with MZM, Figure 5 shows numerical simulations that closely resemble Figure 2. They are highly suggestive of a good match between experiment and theory. These simulations serve to convince a reader that theory does, in fact, provide strong support for the MZM explanation of the experimental data. As stated above, this is not the case.

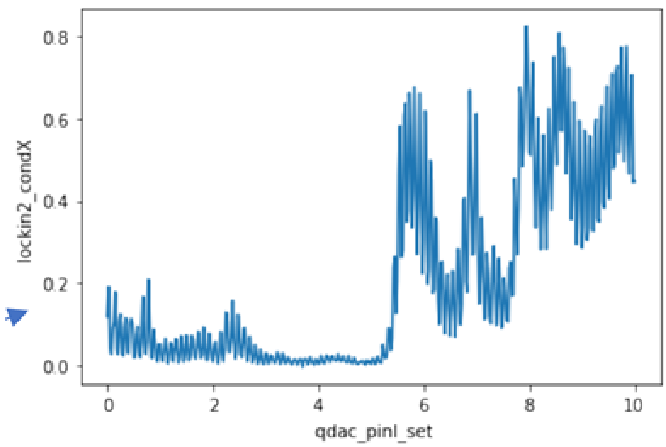
“As in the experimental observations in Fig. 2B, the zeroth lobe shows a hard gap with no subgap states. The first and second lobes, on the other hand, show multiple subgap states (58). The first lobe has a gap with a zero-bias peak owing to Majorana end states.” Science-2020, p.5, discussion of theory Fig. 5B

Fourth, the theory part also makes a connection to the experimental Coulomb blockade part, which we comment on below. The theory suggests that an ‘exponential scaling’ of the parameter extracted from the Coulomb blockade (peak oscillation amplitude) is indicative of the topological regime and is a better predictor of this regime than the coincidence of ZBP and LP-1. This conclusion appears to have been reached within a particular model and based on limited simulation results. It may be that by trying more realizations of the system the distinction between topological and trivial regimes becomes blurred. For example, unintended quantum dots can be added to the model and their effect on exponential scaling can be investigated. Here, again, the existence of another theory by Valentini et al. illustrates that conclusions reached in theoretical analysis are not model-independent.

4.3 Comment on Coulomb blockade section

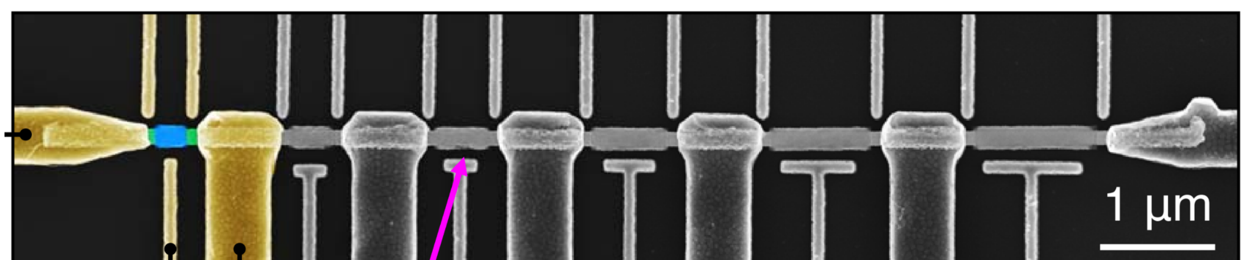
The authors present another experimental technique based on measuring spacings of the so-called Coulomb peaks. Figures 6 and 7 are dedicated to this technique. As with the theoretical part, the content of the Coulomb Blockade part cannot rectify problems in the NIS part of the paper. Since Figure 2 and its description misrepresent full data, any arguments contributed by an additional technique do not change the fact that the paper is invalid. However, we will make several remarks about the Coulomb blockade technique.

Coulomb Blockade (CB) measurements on small superconducting islands are not widely used in Majorana studies. Other groups studied such devices. But only the Copenhagen-Microsoft group made Majorana observation claims based on this measurement technique, in papers such as Albrecht et al. Nature 2016, Sherman et al. Nature Nanotechnology 2017. At the crux of the method is studying arrays of narrow peaks in gate voltage and monitoring the oscillations in peak spacing, when the magnetic field is turned on. The ‘peak oscillation amplitude’ is then plotted as a function of the length of a Coulomb island - a segment of a full-shell nanowire separated by gates from the rest of the wire. When a plot shows an exponential dependence of the extracted peak amplitude on the island length, the Copenhagen group refers to this as ‘exponential protection’, and they state it is a property unique to Majorana physics.



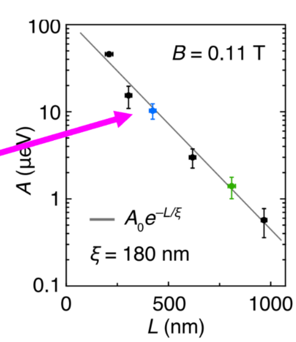
We find several aspects of this technique questionable. For example, in the papers listed in the previous paragraph only a small number of peaks are presented, while Coulomb islands can yield many more. For instance, in Science-2020, to extract peak distances the authors use positions of between 6 and 22 peak pairs. But in a wider range of gates, shown in additional data, we see of order 100 peaks, perhaps more.

What is not known is whether patterns shown in the Science-2020 paper persist over larger peak ranges? Do energy oscillations survive if one averages over 100 peaks? What about 40 peaks? What if the averaging window changed by a few peaks? Do oscillations look the same if different sets of peaks are averaged, a moving averaging window? The authors show in Figs S12, S15 that they can also change the back gate, and that it is set to a large negative value. Since NIS devices are highly sensitive to back gate, how does the 'peak oscillation amplitude' change in the full range of back gate?



The shell of Island #3 of Device 2 is damaged
Other inhomogeneities are visible along the nanowire

Island #3 falls perfectly on the
'exponential protection' plot



Are there examples where scaling is not exponential? Are there small islands with small 'peak oscillation amplitudes', and are there large islands with large amplitudes? In the electron microscope image of Device 2 shown above, island #3 appears to have a damaged shell. Other inhomogeneities are visible along the 8-micron long wire. How come Figure 7F, the 'exponential protection', remains insensitive to these issues, while the same paper argues that fine details of the shell make ZBP disappear in device 5?

There is also a more basic argument that illustrates the skepticism in the Coulomb technique. Figures 6 and 7 present data from Device 2, which is in fact 6 separate Coulomb islands of 6 different lengths. The results from all 6 islands are analyzed together as a length dependence of the topological regime. That means, each of the 6 nanowire segments contains MZM. Without looking at the data, let's cross-check this with what we already established from the tunneling technique.

We know that ZBPs, and hence Majorana modes, do not appear in every nanowire. Furthermore, finding a ZBP requires fine tuning of gate voltages. Based on this, it appears to be an unusually good device if it exhibits twelve Majorana zero modes, all at expected positions along the nanowire, at the ends of each Coulomb island. This can be compared to the NIS technique in the following way. The total length of Device 2 is around 8 microns. There has not been a demonstration of correlated zero-bias peaks on two ends of an 8 micron wire. There has also been no demonstration, with the better established NIS method, of two segments of the same nanowire exhibiting two sets of ZBP on each of the four ends.

What is the likelihood of realizing a device like device 2? If we assign a probability that a given segment has MZM to be 0.5, which we consider a high number, then the probability of having 6 segments is 1.5%. It would require, on average, 64 nanowires to find a single one like Device 2. If Majorana probability per device is 7/9 (which matches the authors' count of devices with ZBPs vs. total working devices), then the probability of 6 topological segments in one wire is 22%. The authors present not one, but three such nanowires, two more in additional data. To find so many nanowires with multiple Majorana would require testing tens, but more likely hundreds and even thousands of nanowires.

The coincidence appears too good to be true.

To address the concerns outlined above, we urge the Copenhagen group to release full data for all of its papers that rely on the Coulomb techniques as evidence that Majorana have been observed. We have requested data from Albrecht et al., Nature 2016, Sherman et al. Nature Nanotechnology 2017, and full data, including full Coulomb blockade data, from Vaitiekenas Science-2020. Professor Marcus has refused our requests.

5. Reproduction studies from IST Austria

Shortly after the Science-2020 paper was published, a preprint on arXiv from the IST group in Austria reported that in the exact same nanowires no Majorana modes are observed, but only trivial Andreev bound states (“Flux-tunable Andreev bound states in hybrid full-shell nanowires”, https://arxiv.org/abs/2008.02348). This work is now published as Valentini et al., Science 2021 (https://www.science.org/doi/10.1126/science.abf1513). An obvious contradiction between the Copenhagen Science 2020 paper and the IST Science 2021 paper was what motivated us to perform this post publication analysis. However, we do not base our analysis of the Copenhagen paper on the IST work. We find that problems identified in Figure 2 and its discussion make the Science-2020 paper invalid. However, we discuss the IST results given that it is a relevant and negative reproduction attempt.

Our conclusion so far is that the contradiction in the claims of the two papers does not extend to the actual measurements performed by the two groups. The small amount of data from Copenhagen and the large amount of data from IST do not contradict each other, if the physics interpretation of the data are set aside (i.e. Majorana vs Andreev). To put it another way, all of the data shown by Copenhagen can also be found in the large IST data library, of course with minor variations (A version of IST data repository is now open access here: https://research-explorer.app.ist.ac.at/record/9389) .

The IST experiments show that the ZBP/LP-1 coincidence is a result of fine-tuning of the gate voltage settings. We conclude this based on a large volume of data, over 500 datasets, shared with us by the IST group. Within these data, it is possible to find data that recreate the key features of Figure 2 of Science 2020. Namely, zero bias peaks that onset within LP-1. However, looking at full data from IST it becomes clear that only through selective presentation of data it is possible to argue that this ZBP/LP-1 connection is significant.

The full-shell nanowires for both studies are grown by the same grower using the same growth equipment. Device fabrication including positioning of nanowires on silicon substrates, and deposition of electron-beam lithography patterned metal contacts (and gates) is generally similar. Though different cleanrooms were used to make devices, the processes appear to be not different in substantive ways. One way to see this is from the data themselves: devices from IST and Copenhagen exhibit similar low-temperature characteristics such as the induced superconducting gap, including its magnitude, hardness and its persistence in applied magnetic field; as well as similar Little-Parks oscillations that are specific to full superconducting shell nanowires.

The overall data quality is comparable, meaning that datasets of the same quality are present in the full IST data and in the Science-2020 paper, though they do not necessarily correspond to completely equivalent regimes. This however does not leave much room for claiming that the two studies are different in any significant way. Given the larger variety of data we found in additional Copenhagen data compared to Science-2020 paper itself, we expect that full libraries of data show equivalent phenomenology. Again, this does not address the contradictory physics interpretations.

The total volume of the IST study is substantial and is in line with our expectation for this type of experiment. Approximately 100 devices on 8 separate silicon chips were characterized at low temperature. Close to 50 of them were explored in detail at zero applied magnetic field. And of those fifty 18 were studied with magnetic field applied (magnetic field is crucial for inducing Majorana modes in full-shell nanowires hypothetically by threading flux through the core of the wire). The IST group has shared with us, and with the Copenhagen group, the original source data from these 18 devices. The total number of raw unprocessed data files shared is exceeding 500. Each file corresponds to a separate instance of an executed measurement script.

We already discussed that the volume of data in terms of the number of devices studied and individual datasets is orders of magnitude larger from IST than from Copenhagen. But another important way in which it is larger is that at IST they have presented larger ranges of experimental parameters - for example larger gate voltage ranges. Furthermore, they show that by narrowing down gate voltage ranges it is possible to obtain data that are in qualitative agreement with the Science paper. We also see that within the limited additional data from Copenhagen, as discussed above.

The IST group has also shared a pdf summary with all of these data plotted alongside brief comments explaining their thinking while taking data. This additional pdf helps understand the principles they used for including data in their manuscript. We are satisfied that they have investigated a significant number of devices, drew generalities from their observations and then presented high resolution data that illustrate their conclusions that are confirmed by the full volume of data from 18 nanowire devices.

We would like to emphasize that this discussion is not about competing interpretations or schools of thought, or about a difference in opinion about the meaning of data. It is about experimental practices and in particular - which data were obtained to support the publication and about the degree of transparency that is needed to understand or reproduce the results. The IST team has done careful experiments with sufficient volumes of data, large parameter sweeps, re-sweeps and cross-checks. For every zero-bias peak they found they concluded that it is a coincidental phenomenon due to

Andreev states and not due to Majorana. They used the same nanowires and performed the same experiments as in Science-2020.

5.2 Status of data requests from Niels Bohr Institute, University of Copenhagen

In August 2020 we approached the corresponding authors of both IST and Copenhagen experiments with requests to share data. The purpose of those requests has been to compare their results. The question that we tried to find an answer to was the following: how exactly did the two research groups arrive at opposite conclusions using the same nanowire materials and similar experimental methods?

We asked for data that went beyond what was in their papers, ideally for the full original source data from all devices published and unpublished upon which conclusions of the papers are based. The IST group provided access to such significant data in early September 2020, even though their work was not yet published in a peer-reviewed journal and they were not required by any ethics codes or journal agreements to share their data.

The Copenhagen authors initially refused to provide any data, insisting that all data necessary to reproduce their manuscript were available within the paper. They shared a limited amount of data after the *Science* editor intervened. On finding that the additional data was not consistent with the paper, we asked the authors further questions. We did receive a response in May 2021. However, the response was not extensive or satisfactory. An earlier version of our analysis was shared with *Science* and it led to the publication of an Editorial Expression of Concern on July 30, 2021.

On July 26 2021 Jan Thomsen of Niels Bohr Institute informed *Science*, and shortly after a journalist at *Retraction Watch*, that the data shared by the authors on Zenodo constitute the entire body of experimental data and that ‘No additional data is left out’.

We could not believe this statement. The volume of data for the tunneling (NIS) component of the study is 52 datasets total over 9 devices, or between 5 and 6 datasets per device on average. If this were all the data, then no scientific conclusion could be justified by such an insufficient study. Given the number of experimental variables to explore, the temporal instabilities in devices and the richness of phenomena present in those data made available, the required volume of study would need to be between 1 and 2 orders of magnitude larger.

<u>Why do we need more data?</u> The volume of transport data that has been released to date, is inadequate; it does not permit any scientific conclusions to be reached. If this

were all the data, publication would not have been justified. Throughout this post-publication review, we pointed at examples where seeing more data is essential. In January 2022, Charles Marcus, the corresponding author of Science-2020, corrected the statement by Thomsen. Marcus said that the data shared so far were ‘not all the data’. He insisted that the data shared was ‘sufficient’. Marcus refused to share any more data.

5.3 What is a typical volume of data for a study like this

It typically takes several hours to obtain these 5 sets of data from one device. It takes multiple days to create and cool down these elaborate nanowire devices. Why would somebody go through the effort of creating on-chip markers and pads, positioning nanowires, fabricating contacts and gates, not counting the nanowire growth itself, - only to spend one afternoon taking data? It is common sense that devices such as this, with potential for discovery publishable in *Science*, are studied extensively. When a discovery is likely, even more data is obtained to verify, cross-check and confirm the findings. All these data should exist at the University of Copenhagen, somewhere, from at least the nine most relevant nanowire devices studied in Science-2020. If they exist, the easiest is to show them, so that the discovery of ‘full-shell Majorana’ can be verified.

‘You try once, and go with whatever the screen shows’ is not consistent with the scientific method. The richness of the obtained features would compel any experimental physicist to look further, not simply to repeat the same measurements but to vary parameters. For example, magnetic field dependences like those in Figure 2 should be studied for many different gate voltages, in the regions that do not show zero-bias peaks. It is necessary to repeat gate voltage sweeps and many different magnetic fields, and to push the gate settings to establish where the true boundaries of the tunneling regimes are. Repeating these standard methods for each device of the 9 studied would result in a ten-fold or more increase in the volume of data, required to reach conclusions at the level of those presented in the Science paper.

A comparison can be made with those experiments for which the volume of study is known. For instance, the volume of study for the IST-Austria Science paper is at least 500 datasets. The Copenhagen group shared significant volumes of data (in gigabytes) from two papers: Puglia et al PRB 2021, and Bannerjee et al, arXiv 2022.

The following considerations help to illustrate what we mean by full data. Once the nanowire device is fully created in the cleanroom, it is then tested in a dilution refrigerator. From the moment it reaches the lowest temperature, data acquisition starts.

At some later point in time, a decision is made to terminate data acquisition and warm up the sample. Between those two times, data are being obtained. Full data means - all of that data. If the same specimen was cooled down several times, then full data includes all cooldowns with indication of which cooldown data belong to.

Laboratory journals, data summaries are also part of the research process by which conclusions are reached, so they can also be requested under the Danish Code of Conduct for Research Integrity. These records contain information about data selection. Data selection can be active (i.e., during the measurement) decisions on which data to acquire. Or, they can be taken after acquisition. These decisions are part of the data analysis and processing and help understand how the conclusions were reached.

In the case of Science-2020 the most interesting full data at present is from those 9 nanowires where ‘zero-bias peaks’ were seen. However, more data may be necessary in the near future for example to unpack the Coulomb Blockade claims made in Science-2020, and other papers that use this ‘Coulomb’ technique - thus we have requested full data from all similar devices, and experimental journals and data summary powerpoints or pdf files that can help us to reconstruct how the conclusions were reached.

6. Who we are

Sergey Frolov (SF) is an associate professor at the University of Pittsburgh specializing in experimental condensed matter physics, in particular on transport experiments on low dimensional systems at cryogenic temperatures. Vincent Mourik (VM) is a junior team leader at Forschungszentrum Jülich (until recently postdoctoral scholar at UNSW), focusing on spin physics in silicon and related materials. Majorana research has been the topic of the PhD work of VM. SF has been focusing on Majorana in his research for the past 10 years. Both VM and SF have written articles for *Science* and are familiar with editorial policies and standards. Neither SF nor VM are driven by personal or other disqualifying motives to misrepresent our findings in any way to the scientific community.

7. Sources used

Our analysis is based on publicly available materials and can be verified and reproduced by third parties. The available data that are relevant to Figure 2 of the

Science-2020 paper is 52 datasets from tunneling measurements. The corresponding author has admitted that more data exist, but not shared these with us. We perform our analysis under the assumption that the data publicly shared are the best available in support of the claims of the paper. This is because the authors shared the additional data in response to a request from us that clearly stated our goal was to re-analyze their results. Furthermore, responses from the authors to our analysis did not incorporate any additional data.

Data can be found at: http://doi.org/10.5281/zenodo.4263106 and can be split in three groups:

•**The *Science-2020* paper main figures**: In March 2020, *Science* published a paper 'Flux-induced topological superconductivity in full-shell nanowires'.

•**SOM:** Supplementary Online Materials published with the *Science-2020* paper.

•**Additional data**: On August 18 2020 we asked the corresponding authors of the *Science-2020* paper to share more data. They uploaded more data on November 8 2020. These data can be found in additional_data.zip on Zenodo.

Typically many more samples are studied that inform a given paper than those presented in the paper. The Zenodo repository gives the number of samples that yielded substantive results within the tunneling technique as 9. All those studied samples and data obtained are part of the full volume of data. The conclusions of the paper are based upon those even if the data are not presented within the paper. Data obtained at a later moment in time, after the publication, or by different researchers, are not part of the full volume of data for the published study.

A dataset is a single graph, usually derived from executing a measurement script on a measurement computer that controls the measurement equipment. In the original form, a dataset is a spreadsheet of numbers that can be plotted as a curve or a two-dimensional colormap. 'Original' refers to the datafile, a settings file and a script as recorded at the moment of the measurement, without any further alteration. Sometimes 'raw data' is used which some people use to describe original data and others to describe any spreadsheet data including those already processed, e.g. for publication. To us, 'raw data', or 'full data' imply what we describe here as the original data.

APPENDIX A - Relevant quotations from the Science-2020 paper

A.1 Statements about Majorana in the paper

1) '*We show experimentally and theoretically that the winding of the superconducting phase around the shell induced by the applied flux gives rise to MZMs at the ends of the wire.*' Research Article Summary.

2) '*The topological phase […] is controlled by moving from zero to one phase twist around the superconducting shell.*' Research Article Summary.

3) '*Two device geometries, measured in three devices each, showed similar results.*' p1.

4) '*As the tunnel barrier is opened, the zero bias peak gradually evolves into a zero-bias dip. The increase of finite-bias conductance compared with zero-bias conductance as tunnel barrier decreases is in qualitative agreement with theory supporting MZMs...although the crossover from a peak to a dip occurs at lower conductance than expected.*' p2.

5) '*Moreover, the full shell naturally protects the semiconductor from impurities and random surface doping, thus enabling a reproducible way of growing many wires with essentially identical electrostatic environments.*' p6.

6) '*Although full-shell wires do not allow for direct gating of the electron density in the semiconducting core, we demonstrated that, using a careful design of the wire properties—for example, by choosing the appropriate radius— it is possible to obtain wires that harbor MZMs at a predictable magnetic field.*' p6.

A.2 On tunneling spectroscopy at zero magnetic field:

1) '*Our measurements reveal that tunneling into the core in the zeroth superconducting lobe, around zero flux, yields a hard proximity induced gap with no subgap features.*' p1.

2) '*At zero field, a hard superconducting gap was observed throughout the zeroth superconducting lobe (Fig. 2, B and D).*' p2.

3) '*In a weak tunneling regime, for VBG < −1 V, a hard gap was observed, with D = 180 meV (Fig. 2, C and D). For VBG ~ −0.8 V, as the tunneling barrier is decreased, the subgap conductance is enhanced owing to Andreev processes. The increase in conductance at VBG ~ −1.2 V is likely caused by a resonance in the barrier.*' p2.

4) ‘*A hard gap is seen in the zeroth lobe,…*’ and ‘*The zeroth lobe shows a hard superconducting gap,…*’ Fig. 2, commenting on data from Device 1

5) ‘*For all four tunneling-spectroscopy devices (1, 3, 4 and 5) the zeroth lobe, where the winding number is 0, shows a hard gap and no subgap states are visible.*’ SOM p3, commenting on all devices.

6) ‘*For device 5, a discrete state crosses zero-energy around VBG = 0.12 V and then again at 0.17 V, resembling a proximitized quantum dot state, similar to the one previously studied in Ref. [84], see Fig. S7. We usually associate such state with a resonant level in the barrier and if possible avoid it in the measurements.’* SOM p3, commenting on device 5.

7) *‘The zeroth lobe shows a hard superconducting gap,…*’ Fig S6, commenting on data from Device 4

8) ‘*The zeroth lobe shows a hard superconducting gap,…*’ Fig S7, commenting on data from Device 5.

A.3 On tunneling spectroscopy within the first lobe of the Little-Parks oscillations

1) ‘*In the superconducting regions around one quantum of applied flux, corresponding to phase twists of ±2π in the shell, tunneling spectra into the core show stable zero-bias peaks, indicating a discrete subgap state fixed at zero energy.*’ Research Article Summary.

2) ‘*Upon reopening, a narrow zero-bias conductance peak was observed throughout the first gapped lobe (Fig. 2, B and F). Several flux-dependent subgap states are also visible, separated from the zero-bias peak in the first lobe.*’ Research Article Summary.

3) ‘*In the first lobe, at B = 110 mT, the sweep of VBG showed a zero-energy state throughout the tunneling regime (Fig. 2E). The cut displayed in Fig. 2F shows a discrete zero-bias peak separated from other states by a softened gap, presumably owing to finite temperature and level broadening in the junction. As the tunnel barrier is opened, the zero-bias peak gradually evolves into a zero-bias dip.*’ p2.

4)‘…*the first lobes show a zero-bias peak,*…’ Fig. 2, commenting on data from Device 1.

5) '*In the first lobe, with the phase winding of 2π, the spectrum for devices 1, 3 and 4 (all with 30 nm Al shell) displays a discrete, zero-energy state (see main-text Fig. 2, and Figs. S5 and S6),…*' – SOM p3, commenting on data from devices 1, 3 and 4.

6) '*…the first lobe shows zero-bias peak,…*' Fig. S5, commenting on data from device 3.

7) '*…the first lobes show subgap states including a zero-bias peak,…*' – Fig. S6, commenting on data from device 4.

A.4 On tunneling spectroscopy within the second lobe of the Little-Parks oscillations:

1) '*A second gapped lobe centered around |B| = 220 mT then appeared, containing several subgap states away from zero energy...*' p2, commenting on data from Device 1.

2) '*…and the second lobes show nonzero subgap states.*' Fig. 2, commenting on data from Device 1.

3) '*In the second lobe, with even number of phase windings, the spectrum for device 1 features an asymmetric superconducting density of states with the lowest energy subgap state centered around ~5 eV, see Fig. S3; For devices 3 and 4, multiple subgap states can be identified at finite voltage, but no zero-bias peak, see Figs. S5 and S6; For device 5, a qualitatively similar to the first lobe spectrum with several finite-energy states is observed, see Fig. S7.*' SOM p3, commenting on all devices.

4) '*The spectrum shows subgap states away from zero energy.*' Fig. S3, commenting on data from Device 1.

5) '*…the second lobe shows non-zero energy subgap states.*' Fig. S5, commenting on data from Device 3.

6) '*…the second lobes show non-zero energy subgap states.*' Fig. S6, commenting on data from Device 5.

7) '*…the higher-order lobes show multiple discrete states away from zero energy.*' Fig. S7, commenting on data from device 5, '*…higher-order lobes…*' meaning first lobe and onward.

A.5 On tunneling spectroscopy within the third lobe of the Little-Parks oscillations:

1) '*Device 4 with slightly bigger diameter, displays the third lobe, with odd number of phase windings. The spectrum features subgap states and a peak at zero bias, see Fig. S6.*' SOM p3, commenting on data from Device 4, the only device shown with a third lobe.

2) '*The third lobes show subgap states again with a peak at zero bias.*' Fig. S6, commenting on data from Device 4.

A.6 On device 5 where no ZBP was shown (some quotes already included earlier):

1) '*In the first lobe, with the phase winding of 2π,… …for device 5 (with 10 nm Al shell) the spectrum consists of multiple discrete, but finite energy subgap states, see Fig. S7.*' SOM text p3, commenting on device 5.

2) '*For device 5, a qualitatively similar to the first lobe spectrum with several finite-energy states is observed, see Fig. S7.*' SOM text p3, commenting on tunnelling spectroscopy in the second lobe for device 5.

3) '*For device 5, a discrete state crosses zero-energy around VBG = 0.12 V and then again at 0.17 V, resembling a proximitized quantum dot state, similar to the one previously studied in Ref. [84], see Fig. S7. We usually associate such state with a resonant level in the barrier and if possible avoid it in the measurements.*' SOM text p3, commenting on subgap resonances in device 5

4) '*Tunneling spectroscopy without zero-bias peaks in device with thinner Al shell (device 5).*' Fig. S7, caption title.

5) '*The zeroth lobe shows a hard superconducting gap, the higher-order lobes show multiple discrete states away from zero energy. No destructive regime is present in the thinner-shell device.*' Fig. S7, commenting on data from device 5.

6) '*…whereas for device 5 (with 10 nm Al shell) the spectrum consists of multiple discrete, but finite energy subgap states, see Fig. S7*.' SOM text p3, commenting on data from device 5 in comparison to devices 1,3,4.