

Second report of Referee A, DF12830/Hanson, December 6th, 2021

The authors have adequately addressed my concerns and I can see the paper is significantly improved with the reorganization and updates to the simulations. My only comment is the the signals in Fig. 6 (right panels) are clipped and the authors should consider expanding the y-axis.

We thank Referee A for all the advice and guidance with the manuscript. We recongize how valuable your time is and we appreciate it.

Report of Referee B, DF12830/Hanson, December 6th, 2021

We would like to provide some context about the manuscript that Referee B might not know. This manuscript was first submitted on June 10th, 2021. Referee A filed a report and we received it on July 29th, 2021. In that initial report Referee A wrote the following about the topic and scope of the paper:

The manuscript derives a fully analytic model of Askaryan radiation in ice in the time domain. **This manuscript will make an excellent addition to the literature** on modeling Askaryan signals in ice ...

Referee A then presented a list of five areas where the manuscript could be improved. These remarks involved stylistic and structural critiques. None involved fundamental physics questions about any of the derivations or approximations. One question was about the analysis approach in the second half of the paper. Referee B refers to this section as the phenomenological section. We performed this analysis again, with the modifications suggested by Referee A. The results were improved slightly or unchanged, depending on the case. By October 4th, 2021, we re-submitted our manuscript. The editor noted that he was searching for a second referee, and by December 4th, 2021, we received a new referee report. Referee A wrote the following (from above):

The authors have adequately addressed my concerns and I can see the paper is significantly improved with the reorganization and updates to the simulations.

*We assumed that the report from Referee B would be of a similar nature: a concrete list of suggestions to improve the manuscript, especially six months into the process. What we received from Referee B, to put it mildly, tests the boundaries of professionalism. Telling authors that “major rewriting of the article is required for it to be published in any journal” comes across as an attempt to block the paper on any grounds one can find, and not a good faith attempt to improve the paper. How can Referee B know if the manuscript would or would not be published in **any** journal? We did, in fact, rewrite and reorganize sections of the paper for Referee A in ways that run counter to the demands made by Referee B. For example, Referee A specifically requested that we introduce the appendices as a way to store all of the details:*

I recommend [appendices] to keep the main body of the text shorter, while maintaining the authors’ preference for presenting the complete derivation (which I appreciate).

The original remark may be found below. We have shared our concerns about this report with the editor, who agreed to bring in an adjudicating referee at the end of the process if we request it. However, the editor has asked me to first submit a formal response.

I have made an attempt to categorize the critiques from the report of Referee B and address them systematically. First, there appear to be questions about whether the manuscript represents work that is of sufficient novelty to merit publication in PRD. Second, there are questions about distinctions between review material and new material. Third, there are questions about the meaning of citations. Finally, there are questions about notation, variables, and the details of derivations. There are also recurring themes, like the relationship between the mathematical and phenomenological parts of the paper, and the overall length. To make ourselves clear when responding to the four areas, we should develop some keywords.

*In our responses below, we will use the phrase “**apparently not**” to respond to critiques in which our response involves invoking the outstanding match our analytical model achieves when it is compared to a model derived in a completely different way. Semi-analytic models such as citations [19] and [30] in the manuscript received by Referee B have both a Monte Carlo component and an analytic component. Further, the analytic component is derived in the time domain. Our work from 2017 in citation [20] of the manuscript received by Referee B was derived in the Fourier domain. From*

the Fourier domain model, we converted it to the time-domain with great care and effort, attempting to preserve validity for any viewing angle. This was no easy task, given the complexity of the original equations. Correlation coefficients between our new time-domain model and the semi-analytic examples exceed 95%, with fractional power differences on the order of a few percent. That cannot be an accident, and we see that Referee B acknowledged that in the report.

*In our responses below, we will use the phrase “**citation misinterpretation**” to refer to the following problem. Referee B is rightly concerned with properly citing sources. When we provide a citation, we are NOT claiming credit for the implied discovery or finding. We are providing one example for a skeptical reader who would not otherwise believe the claim or assumption we are making. Suppose we write: “The lateral ICD is distributed exponentially near the cascade maximum [ab].” The paper we associate with citation [ab] does not necessarily have to represent the earliest paper ever written containing this finding. It just has to be an example. If every citation had to be the earliest possible work, then any physics article would have infinite citations! To address the concerns of Referee B, we are willing to add citations and rearrange them. We are also willing to include the specific citations Referee B suggests. But Referee B is not justified in raising this issue to the level of gravity discussed in the report.*

Remarks about Novelty, and distinctions about review material and new material

To our knowledge, the literature on modeling the Askaryan effect does not yet contain a fully analytic model of the Askaryan field $\vec{E}(t, \theta)$ in the time-domain. In the introduction section, we place our model in the appropriate context by describing the development of full MC and semi-analytic work. We articulated four distinct advantages a fully analytic model would provide to radio detectors like IceCube Gen2 (radio) in part B of the conclusion section. According to Referee A, and in our professional judgement, there is no question about whether this work is new enough to merit publication in a journal. What follows is our response to individual comments made by Referee B about this topic.

“However as one one gets down to the details and compares the results with those of earlier publications, the article starts rising many doubts, in particular about what is really new in the theoretical calculation presented,

and about what are the authors actually checking when they compare their analytic expressions to the results obtained with another ‘semi-analytical’ approximation.”

This model presents both a re-working of the on-cone derivation ($\theta = \theta_C$) of $\vec{E}(t)$ from citation [20] Section V, and a derivation of the off-cone ($\theta \neq \theta_C$) field $\vec{E}(t, \theta)$. Both calculations are new in that they both demonstrate the effect of the complex poles on the actual function for $\vec{E}(t, \theta)$. Further, to our knowledge, no one has published an off-cone fully analytic model. Finally, the effect of the poles on the form of $\vec{E}(t, \theta)$ is validated against the output of NuRadioMC based on citations [23] and [30] and for both the on and off-cone cases, and for electromagnetic and hadronic cascades. We are completely open about how this comparison is done. We note that the authors of [19] and [30] had access to the the ZHS MC code, but we do not, and it is not public. Thus, we cannot compare our model to ZHS MC. However, the ARVZ models in [19] and [30] match the full ZHS MC, and our model matches the ARVZ one. We decided to include the derivation of the form factor $\tilde{F}(\omega, \theta)$ from citation [20] only as a matter of convenience for the skeptical reader who would not want to take our word for it, or be forced to read the paper from 2017 just to accept the details of the current manuscript. While Referee B had the time, patience, and curiosity to read the background papers, we do not think leaving out details is wise. Referee A is in agreement. We want the reader to see the connection between the form factor, and the longitudinal development, with the final product.

“As a result I have been drawn into comparing the results of this paper to earlier work in some detail to come to a reasonable conclusion with respect to what is the actual reach of this paper. I have found that the paper has many repetitions of earlier work and that it is extremely difficult to say what is truly new in the analytical calculation. The theoretical results are basically the same as those presented in ref [31], only that they are now given in the time domain and the form factor is evaluated with simple approximations of the shower model, a gaussian for the depth development and an exponential decay for the lateral distribution. Moreover the exercise of converting the frequency domain to the time domain is already the main content of ref [20], so perhaps the new contribution is to give an explicit expression in the time domain for the radio pulse away from the Cherenkov cone. ”

In the Section I: Introduction, we specifically describe what is new in this manuscript, and place it in the appropriate context of the development of

the sub-field. In Section II: Units Definitions, and Conventions, we describe Eq. (5) as containing the five essential pieces required to build the model before taking the inverse Fourier transform. The goal is to build an all- θ , all- t analytic model by converting Eq. (5) into the time-domain. We credit citation [31] with \vec{E} , but strongly disagree that the results of this work are the same as reference [31]. The authors of [31] do not provide an all- t , all- θ analytic model, nor do they match a form factor to data or examine the effect in the time-domain. In fact, John Ralston himself was the one who first suggested we produce the form factor and use it to obtain a model like this. The GEANT4 studies that justify our form factor were presented in citation [20], and we complete the model in this paper by examining the effect of the form factor (and other factors) on $\vec{E}(t, \theta)$.

Moving on, Section III A provides the form factor to be entered into Eq. (5). Section III B provides the reader with an understanding of the longitudinal length parameter, which is necessary to evaluate \vec{E} . We discuss below the remarks Referee B makes about Section III B in particular. In Sections IV and V, we assemble the pieces for Eq. (5) and transform the model into the time-domain while making the assumptions necessary to keep the model analytic. There is no question that Section VI is also brand new, since no other paper to our knowledge performs this type of comparison (fully analytic to semi-analytic). Some works in this genre compare the semi-analytic models to the ZHS MC (citations [19] and [30] are good examples).

In Section VII: Conclusion, part A, we summarize new results in Table IX. In the text, we explain why we think each result in this table represents something new. We explain that the form factor was first described in [20], but how its effect on $\vec{E}(t, \theta)$ is now presented in this work. We explain how Ralston and Buniy introduced the length parameter a , but how we have finally connected it to $\vec{E}(t, \theta)$ and presented a novel approach for energy reconstruction. In part B of Section VII, we justify our approach by listing four different cases in which a fully analytic time-domain model provides advantages to detectors like IceCube-Gen2. Thus, at no point in the paper is the thread of reasoning for how this work is novel ever dropped. At the end of the quote above, Referee B finally acknowledges that “perhaps” the new contribution is the explicit expression for $\vec{E}(t, \theta)$. We state what is new all throughout this paper, in that in no uncertain terms.

“[In ref [20] the explicit expression of the pulse in the time domain is only given in the Cherenkov cone. There is a small difference between the

two expressions but no comment is made on this difference. Is this a correction? Moreover reference [20] treats LPM showers incorrectly, because the frequency spectrum of the pulses has the wrong low frequency limit (it should be the same as that of a shower of the same energy and no LPM effect). In this article the authors should discuss differences wrt ref [20] and address errors as well.]”

*The above remarks come after the prior quote. They do not even pertain to this manuscript, but work in citation [20]. The work in citation [20] did not demonstrate many aspects of this calculation because $\vec{E}(t, \theta)$ was not the focus of that work. Citation [20] does not address the uncertainty principle for on-cone signals, nor does it explain how the poles that lead to its shape connect to **both** the lateral ICD **and** longitudinal development. Referee B states that the treatment of low frequency modes is wrong in [20], when both reviewers of [20] were satisfied with the treatment of it. Further, Referee B states that the modes should match a cascade with no LPM effect. We presented results that demonstrate that they do, so this criticism is invalid. Not only is it invalid, it is not even about this manuscript. Referee B should understand that we avoided the LPM effect in this analytic model by using cascade energies of 10 PeV for EM cascades and 100 PeV for hadronic ones. Finally, does our model disagree with others’ models that do include the LPM effect? **Apparently not.** Please consult Section VI to observe the excellent agreement between a semi-analytic model that would have demonstrated the LPM effect if it were relevant at these energies in ice, which it is not.*

“Leaving aside the inadequacy of the presentation of the results (to be addressed below), the immediate question that crops up is, is this result of enough new content to justify a new article in PRD? The answer in my opinion is that the result could be of interest for the field if the presentation of this is clear and the calculation leads to a compact expression (it appears to be so) which describes the pulses in this region with sufficient accuracy.”

If Referee B is acknowledging that a compact expression of an all- θ all- t model is of interest for the field - “it appears to be so” - then why is Referee B raising concerns about novelty at all in this report? My colleagues in group meetings, collaboration meetings, and workshops have always made time for me to present findings along these lines. If there was not sufficient interest, we would not have embarked on a six-month effort to get these findings published.

“There is a phenomenological part of the article that compares these

analytical results in the time domain to an alternative calculation. This comparison is used to find the fit parameters for the physical scales of the showers. This, I believe, is new and maybe the authors are intending to justify the approximations made in the new analytical expression by making a detailed comparison of the result with a reference calculation. Putting aside all the doubts that one can have about the procedure used for such a “validation,” the same question comes up, is this sufficient to be addressed in a PRD paper? In my opinion the answer is that maybe it is enough, given the outstanding effort that is being made to exploit the radio technique in ice in the last decade, and the current interest that the technique has for neutrino astronomy. It is true that simple models for radio pulses could be of much use at different levels in such experiments, as the authors state. ”

This comparison is brand new in the literature. At no point, to our knowledge, has anyone compared a fully analytic all- t all- θ model to another one and succeeded. Indeed, the match is of high quality, as quantified by correlation coefficients and fractional power differences. Referee B shares that one has to put aside doubts about the validity of the technique, but cross-correlation and amplitude or power are literally the first ways one establishes a match between any two waveforms. Further, one obtains a -values, ω_0 -values, and ω_G -values, and values for ϵ that all make sense in light of Section III B. Referee B acknowledges that we state reasons for why such a validated model would be of use to the community given the outstanding effort underway to observe neutrinos via the Askaryan effect. Unless our stated justifications for the interest in such a model are unjustified, why is there any question at all about the novelty and merit of this work?

“There is a lot of spurious text. Parts of the text are clearly to be completely removed. I was particularly surprised to find out that section III B, that takes nearly two pages and seems an elaborate calculation, is simply an estimate of the width of the well known Greisen and Gaisser parameterizations, as obtained approximating them with gaussian functions. This is efficiently described in ref [31] already (which is the basis of this work) with a reference to Bruno Rossi. Most people working with showers have approximated the shower development with gaussians so it seems too trivial to be reported in detail. Moreover the derivation is particularly cumbersome. I do not even think it should be put in an appendix, the final result can just be displayed after a brief and clear description of what is actually being done. Incidentally references could be made to [31] and to the work of Andringa, S. et al. Astrop Phys 34 (2011) 360 describing both the gaussian approxima-

tion, fluctuations in the longitudinal spread (as named in [31], this is a better choice than "longitudinal length") and corrections to it. Incidentally the authors have payed very little attention to fluctuations in the shower profile which are relevant. It is particularly relevant to understand how the fluctuations in the longitudinal spread relate numerically to the small changes as the energy is varied. The sections on the uncertainty principle are in my opinion a distraction, it is not clear why they are introduced (nor is it clear that the principle is always satisfied, which is suspicious)."

*Section III B requires 1.5 pages out of a 23 page manuscript - 6% of the work. While it is true that citation [31] states the energy-dependence of the longitudinal length, the authors do not show the derivation, but do describe how they confirm it numerically. We believe typical experimentalists and NuRadioMC developers in our field would never accept a simple formula for a in the EM or hadronic cases without proof. Further, the standard Referee B is using to justify removing content from the paper is inconsistent. Referee B raises "alarms" regarding **citation misinterpretations** in other sections, motivated by the concern that they are misleading to a non-expert reader. However, Referee B also advocates cutting content because the average reader is an expert in cascade physics who knows automatically the energy dependence of the longitudinal length. The reader cannot be both an expert and a non-expert. As we stated earlier, these remarks just come across as an attempt to find any reason to block the paper, no matter the logic. Besides, the formulas for the longitudinal length are useful for energy estimation. At the end of Section III B we calculate a typical a -value of ≈ 4 meters for a 10 PeV event. Sure enough, this is exactly the scale we find when fitting a completely different model to our own. Experimentalists in our field need proof that the result from cascade theory matches the result from our analytic waveform.*

*We are willing to include the reference given by Referee B from 2011, but re-state our position that the LPM effect is irrelevant for the energies considered in this manuscript. Do our models demonstrate disagreement with semi-analytic calculations that include the LPM effect? **Apparently not.** The implementation in NuRadioMC incorporates shower libraries that have the LPM effect activated, and we have validated our model against this NuRadioMC output. NuRadioMC even produced a hadronic cascade with a sub-shower (waveform 5 in Fig. 10, and entry 5 in Table VIII). Our model still has a correlation coefficient of 95.5% and a power difference of 8.76%, at 100 PeV. Why does Referee B require a discussion of how our model does nothing*

to account for the LPM effect when the LPM effect is shown to be irrelevant? We even state at the end of Section I that we are considering energies for which the LPM effect has a tiny influence. Finally, these energies represent the energy threshold region for IceCube Gen2 designs. All models of UHE- ν predict a flux that decreases with energy. Thus, the threshold energies are the most important.

The sections on the uncertainty principle are not a distraction, but a mandatory check that we have done our transformation from the Fourier domain to the time domain correctly. If an approximation we made affected the shape of the waveform incorrectly, the uncertainty principle would not hold. In the off-cone case, we use the width of the field in the Fourier domain from citation [20], and multiply by the time-domain width found in this manuscript. This product is greater than or equal to $1/(2\pi)$ for $\eta \rightarrow 0$, a clear validation. Referee B does not even specify how this result is not always satisfied. Nor does Referee B specify what is wrong with the on-cone uncertainty principle check. In that case, we find Eq. 49, which requires a pair of ratios to be greater than $\Delta\theta$, which approaches zero. Why is that controversial? In the on-cone case, $\Delta\theta \rightarrow 0$, but a/r and l/a are both small and non-zero.

Citation Misinterpretation

We discuss here the “problem” Referee B raises with respect to citations. The first such remarks from Referee B begin here:

“The authors discuss in much detail the behavior of the results obtained in the time domain, relating them to shower properties. A non-expert reader could be misled into thinking these are new findings, but most of them are actually well documented in the abundant literature on this topic starting with the Monte Carlo simulation of ref [13] and practically in all the articles that have followed up to date. Stating these properties is not a problem of course, moreover it is a plus because they serve as a check giving confidence that the calculation gives reasonable results, however proper references should be made to articles that have actually found the same properties with similar or different techniques.”

This criticism comes across, again, as just an effort to delay or block the paper for any reason that may be found. Referee B does not even give a specific list of locations in the manuscript where improvements can be made. We will add any citations to the manuscript that Referee B feels are necessary.

However, there are only a few places where Referee B even states a specific example. How can we respond to a referee report when it does not contain a list of examples? The bulk of the discussion from Referee B begins here:

“More careful referencing is compulsory. The article has many issues with the references which are not acceptable. These are actually the most striking alarms on a first read. The first one is part of the first general comment, the way the article references earlier work on theoretical calculations is misleading, because the authors are not clear enough about what is new and what is not. The article should make a fair summary of what is being done in relation to what was done in the past, particularly in ref [31], what ref [20] added to it and finally limit the first part of this article to the minimal amount of material that is needed to present the time expression that follows from the work of [20] and the approximations made for the form factor.”

Referee B is asking for a “fair summary of what is being done in relation to what was done in the past.” Did Referee B not read the final three paragraphs of Section I? We dedicated half of a page to giving credit to examples of past work. We specifically describe the contributions by the authors of [13], [19], [20], and [31], and place each in context. Please read Section I.

“I will mention a few others but the list is not complete. In several places an inadequate reference is given, for instance the authors cite Hanson et al [20] for the Greisen model, they cite again [20] for showers with many maxima because of the LPM effect but the showers described in [20] have no multiple peaks, and again refer to [20] when they say that the lateral cascade width has the effect of a “low-pass filter,” while [20] may be the first use of such a “term”, such a reference is somewhat misleading in the sense that the effect of the cutoff in the frequency spectrum related to the lateral spread of the shower is described in multiple works on this topic, including the early ZHS MC calculation.”

First, if Referee B is going to raise the issue, and take the time to read the background literature, why not take the time to provide an exact list of proposed changes so that we can make them? Once again, it comes across as an attempt to delay or block the paper for any reason that may be found. Remarkably, the first mention of the Greisen distribution in Section II does NOT trigger a critique from Referee B, even though the citation for the ARA collaboration and the associated Monte Carlo code, AraSim. The reader is smart enough to know that we are citing the use of the Greisen distribution within AraSim, and not that the ARA collaboration discovered this distribution. Referee B states that we cite [20] for showers with multiple maxima,

but does not say where. Referee B could mean, for example, the last sentence of the first paragraph of Section VI, but that would be **citation misinterpretation**. We are merely citing the fact that the LPM changes the shape of $n(z')$, and that [20] accounted for LPM elongation, not multiple sub-showers. Similarly, in the last paragraph of part A of Section VII, we remark that the longitudinal length, a grows faster than $\sqrt{\ln(E_C/E_{\text{crit}})}$, and cite [20] as an example. If the remark of Referee B about the LPM effect refers to this example, that would again be **citation misinterpretation**. In that instance, however, we removed the citation because it is redundant. While it is true that citation [20] was the first to treat the effect of the form factor as a low-pass filter with physical poles, we have removed a reference to citation [20] in the sixth paragraph of Section I, just in case it is too ambiguous.

“The citation of the semi-analytic approach, ref [19], states that the method was “introduced ... to account for non-Gaussian fluctuations in the charge excess profile.” This is simply not true, the method was introduced to provide a fast and accurate method to describe radio pulses in general (similar to the motivation behind the analytic calculation the authors are discussing). The list again is longer. Such a careless citation style is not acceptable.”

Referee B makes a good point that citation [19] was not created just to account for non-Gaussian fluctuations in the charge excess. It was created for a variety of reasons. However, that model does account for a charge excess profile $n(z')$ produced by numerical simulation. We have changed the language in that paragraph to clarify this. Referee B repeats the mantra that “the list again is longer,” but does not provide the list.

Questions about notation, variables, and the details of derivations.

Original Remarks of Referee B

General discussion about suitability of the addressed topic:

This article is difficult to evaluate because it appears to be a thorough analytical calculation of radio pulses in ice in the time domain which seems to put in perspective and build from a couple of earlier articles on this topic, namely references [20] (involving one of the authors) and the original one [31] on the same subject (from 2001), and it provides expressions for the pulses in terms of parameters that relate to the geometry and to physical scales such as the longitudinal and lateral spread of the shower. In its second part the

article seems to validate these results by a detailed comparison with an alternative calculation based on ref [19], which is used to fit the pulses to those obtained in the alternative approach and to get the best fit parameters. As such it could be both classified as a theoretical paper on the mathematical modelling of radio pulses in ice and also as a phenomenological article in which a given analytic parameterization is compared to a more precise calculation. However as one gets down to the details and compares the results with those of earlier publications, the article starts raising many doubts, in particular about what is really new in the theoretical calculation presented, and about what are the authors actually checking when they compare their analytic expressions to the results obtained with another "semi-analytical" approximation.

As a result I have been drawn into comparing the results of this paper to earlier work in some detail to come to a reasonable conclusion with respect to what is the actual reach of this paper. I have found that the paper has many repetitions of earlier work and that it is extremely difficult to say what is truly new in the analytical calculation. The theoretical results are basically the same as those presented in ref [31], only that they are now given in the time domain and the form factor is evaluated with simple approximations of the shower model, a gaussian for the depth development and an exponential decay for the lateral distribution. Moreover the exercise of converting the frequency domain to the time domain is already the main content of ref [20], so perhaps the new contribution is to give an explicit expression in the time domain for the radio pulse away from the Cherenkov cone. [In ref [20] the explicit expression of the pulse in the time domain is only given in the Cherenkov cone. There is a small difference between the two expressions but no comment is made on this difference. Is this a correction? Moreover reference [20] treats LPM showers incorrectly, because the frequency spectrum of the pulses has the wrong low frequency limit (it should be the same as that of a shower of the same energy and no LPM effect). In this article the authors should discuss differences wrt ref [20] and address errors as well.]

The authors discuss in much detail the behavior of the results obtained in the time domain, relating them to shower properties. A non-expert reader could be misled into thinking these are new findings, but most of them are actually well documented in the abundant literature on this topic starting with the Monte Carlo simulation of ref [13] and practically in all the articles that have followed up to date. Stating these properties is not a problem of course,

moreover it is a plus because they serve as a check giving confidence that the calculation gives reasonable results, however proper references should be made to articles that have actually found the same properties with similar or different techniques. Leaving aside the inadequacy of the presentation of the results (to be addressed below), the immediate question that crops up is, is this result of enough new content to justify a new article in PRD? The answer in my opinion is that the result could be of interest for the field if the presentation of this is clear and the calculation leads to a compact expression (it appears to be so) which describes the pulses in this region with sufficient accuracy.

There is a phenomenological part of the article that compares these analytical results in the time domain to an alternative calculation. This comparison is used to find the fit parameters for the physical scales of the showers. This, I believe, is new and maybe the authors are intending to justify the approximations made in the new analytical expression by making a detailed comparison of the result with a reference calculation. Putting aside all the doubts that one can have about the procedure used for such a "validation," the same question comes up, is this sufficient to be addressed in a PRD paper? In my opinion the answer is that maybe it is enough, given the outstanding effort that is being made to exploit the radio technique in ice in the last decade, and the current interest that the technique has for neutrino astronomy. It is true that simple models for radio pulses could be of much use at different levels in such experiments, as the authors state.

Due to issues addressed above and a number of important issues that I will address below, it is clear to me that the article cannot be published in its current form and that major rewriting of the article is required for it to be published in any journal. This would imply adjusting the article to address my general comments above and to a number major issues that are very diverse and I will try to summarize in the text to follow. These problems are of sufficient severity and number that the final decision on whether the article is adequate for being published in PRD would depend on how they are finally addressed by the authors and I will not make a definitive statement.

Major comments:

The length of the article. This article is far too long for the amount of information given. More importantly it is written in a cumbersome style so that it is very tedious to find out what is exactly being done. It should be

drastically shortened and simplified.

There is a lot of spurious text. Parts of the text are clearly to be completely removed. I was particularly surprised to find out that section III B, that takes nearly two pages and seems an elaborate calculation, is simply an estimate of the width of the well known Greisen and Gaisser parameterizations, as obtained approximating them with gaussian functions. This is efficiently described in ref [31] already (which is the basis of this work) with a reference to Bruno Rossi. Most people working with showers have approximated the shower development with gaussians so it seems too trivial to be reported in detail. Moreover the derivation is particularly cumbersome. I do not even think it should be put in an appendix, the final result can just be displayed after a brief and clear description of what is actually being done. Incidentally references could be made to [31] and to the work of Andringa, S. et al. *Astrop Phys* 34 (2011) 360 describing both the gaussian approximation, fluctuations in the longitudinal spread (as named in [31], this is a better choice than "longitudinal length") and corrections to it. Incidentally the authors have payed very little attention to fluctuations in the shower profile which are relevant. It is particularly relevant to understand how the fluctuations in the longitudinal spread relate numerically to the small changes as the energy is varied. The sections on the uncertainty principle are in my opinion a distraction, it is not clear why they are introduced (nor is it clear that the principle is always satisfied, which is suspicious).

A large number of variables is used together with several unfortunate choices of notation. Although I would usually classify such a comment under "minor issues about style", this article is particularly exaggerated in this respect and more importantly it becomes an obstacle to clarity, obscuring the reasoning and the expressions presented and contributing to the confusion of the reader. Here are some examples, again the list is not exhaustive: The authors use unnecessary changes of variables introducing names such as $p, q, u, x, x_c, \omega_0, \omega_{CF}, \omega_1 \dots$ that either giving little intuition to interpret the expressions, or I cannot see the need to introduce them all, and that in my opinion obscure the interpretation. In places the authors introduce redundant variables, for instance there is a cutoff frequency ω_{CF} and a second one ω_0 which is simply $\omega_0 = \sqrt{2/3}\omega_{CF}$. One could question the need to introduce two frequency scales which are so close, when a single one would do the job. The introduction of $\sigma = \omega/\omega_{CF}$ is another unnecessary new variable. For the lateral distribution a simple exponential is used as an approximation. The radius is labeled "rho" in cylindrical coordinates and a length scale is

introduced as $1/\rho_0/\sqrt{2\pi}$ which is extremely confusing because here ρ_0 is used as an inverse radial distance. $1/\rho_0$ is related linearly to the Moliere radius (incidentally the definition given of the Moliere radius is incorrect). It is unnecessary to introduce two scales for the lateral spread given the simple approach given in this article, one of the two is enough to set the scale and using two adds to the confusion of variables, particularly with the unnatural definition of ρ_0 . Another example of the latter type is introducing $u = 1 - i\eta$, which is not justified because afterwards an expansion is made in powers of $(1 - u) = i\eta$, when u tends to 1. Clearly u is redundant and the natural variable to use is η , the limit corresponding to η going to zero. These problems include a few examples of bad choices of variables and labels, but the list is longer. The article should reduce the number of variables, change names where they lead to confusion such as the previous case, and eliminate spurious definitions that obscure the mathematical steps. Incidentally what I am demanding about improving the notation would end with a notation very close to that used in reference [31], which is an excellent article and has a rational labeling of variables. It should be used by the authors as a guide of style. Moreover, using the same conventions would make it a lot simpler to relate their work to the original work by Buniy and Ralston.

More careful referencing is compulsory. The article has many issues with the references which are not acceptable. These are actually the most striking alarms on a first read. The first one is part of the first general comment, the way the article references earlier work on theoretical calculations is misleading, because the authors are not clear enough about what is new and what is not. The article should make a fair summary of what is being done in relation to what was done in the past, particularly in ref [31], what ref [20] added to it and finally limit the first part of this article to the minimal amount of material that is needed to present the time expression that follows from the work of [20] and the approximations made for the form factor. I will mention a few others but the list is not complete. In several places an inadequate reference is given, for instance the authors cite Hanson et al [20] for the Greisen model, they cite again [20] for showers with many maxima because of the LPM effect but the showers described in [20] have no multiple peaks, and again refer to [20] when they say that the lateral cascade width has the effect of a “low-pass filter,” while [20] may be the first use of such a “term”, such a reference is somewhat misleading in the sense that the effect of the cutoff in the frequency spectrum related to the lateral spread of the shower is described in multiple works on this topic, including the early ZHS

MC calculation. The citation of the semi-analytic approach, ref [19], states that the method was “introduced ... to account for non-Gaussian fluctuations in the charge excess profile.” This is simply not true, the method was introduced to provide a fast and accurate method to describe radio pulses in general (similar to the motivation behind the analytic calculation the authors are discussing). The list again is longer. Such a careless citation style is not acceptable.

The expression of the pulse in the time domain away from the Cherenkov angle could be of interest. Yet the presentation of this result is again not clear. The authors have split the calculation into the text and the appendix in a way that is not adequate because the text cannot be followed without reading the appendix. The expression involves several approximations, firstly an expansion for small $\eta = \omega/\omega_{CF}$ (more clear than in $(1 - u)$). The result is an integral which is claimed to be related to the line-broadening function. The comparison involves the definition of $\omega_1 = t_r/(2p)$, which depends on the retarded time, t_r . this is a variable and therefore it is not a number. So the condition for approximating the integral obtained by the symmetric approximation $\omega_1 < \omega$ is clearly not satisfied for p small, for t_r large and obviously for low values of ω that are unavoidable in the ω integral from infinity to -infinity. The justification of this approximation is non-existent and deserves at least a careful discussion.

I am doubtful that this is a good approximation. In any case the authors could argue that the expression obtained this way is a reasonable approximation to the pulse. In this respect the latter part of the article, which compares these results to a reference calculation, takes some relevance. I did not find such reasoning though and if this is the intention of the authors it should be clearly stated. Indeed when one looks at the comparisons made for the off-cone case it is apparent that the agreement is a lot worse than for the on-cone case. Given the number of approximations made, another question crops up. What are the advantages of this expression relative to those of ref [13] that are intended to provide a good and rapid description of the pulses? This should be addressed for completeness. Moreover the approach used is clearly not valid for the LPM effect (this effect was also addressed in ref [20] with flaws), while the approach of [13] can deal with fluctuations of the shower profile and with the completely different profiles of showers fully affected by the LPM effect. This issue deserves a discussion.

Incidentally a comparison is not made relative to a detailed simulation but rather to a calculation based on [13] that uses the simulation of the

radio pulse in the Cherenkov cone to calculate the radio pulse away from it. This presumably has been done because [13] is fast but it is nevertheless an indirect way to compare. Some discussion of this is missing and multiple references are given to a simulation package that includes the approach of [13], what again leads to confusion because the reader may doubt whether they are comparing the parameterization to a MC simulation or to another model. If this is the case it needs to be discussed.

There are many other issues of the paper that I am not willing to spend more time in writing because the condition for me to accept this article for publication is that the authors should rewrite it according to all these recommendations and concerns.

Report of Referee A, DF12830/Hanson, September 12, 2021

The manuscript derives a fully analytic model of Askaryan radiation in ice in the time domain. This manuscript will make an excellent addition to the literature on modeling Askaryan signals in ice after some revisions are made. I have five larger comments and then several more minor ones.

1. Section V: Off-cone field equations is very hard to follow in part because the derivation starts from an assumed form for $\vec{\mathcal{E}}$ (line 4 in Table III) that is not presented earlier in the paper, but also because there are several changes of variables including reusing x , u and swapping their order.

We now provide one appendix for the on-cone derivation, and one appendix for the off-cone derivation. We tried to discern what details were important enough to keep in the main body of the text. For the on-cone case, we felt that stating the problem, then introducing the inverse Fourier transform were important enough to keep in the main body. For the inverse Fourier transform, we chose to explain the nature of the poles, and the two cases of the sign of retarded time in words in the main body. We still conclude with the verification of the uncertainty principle for the on-cone case. The rest of the details are now in Appendix B, along with the full definition of \mathcal{E} . For the off-cone case, we re-introduce the main field equation, but rely on the Appendix A for the full specification of $\vec{\mathcal{E}}$. We start the expansion and explain in

words how to reach the line-broadening function (related to the Voigt function). From there, we explain in words the symmetric approximation, but relegate the details to Appendix C. We then quote the result and verify the uncertainty principle. Appendix B contains enough detail for the skeptical reader to follow the reasoning from start to finish. Of course, we welcome any further suggestions from the reviewer. We went a step further and included the form factor details in Appendix A as well.

2. The comparison between NuRadioMC and the analytic model using digitized waveforms should be improved. There are several discussions of artifacts found in the analysis that would be eliminated if the authors just asked the waveforms directly from the authors of Ref. [23] or running NuRadioMC for example cases themselves.

Our response to this general comment is why we have taken so long to respond. We thank the reviewer for patience and understanding. We decided to install and learn NuRadioMC tools, and to create scripts that produced output fields. Further, since the goal was to compare MC truth values from NuRadioMC to values derived from field equations, we had to modify NuRadioMC to pass derived quantities like the longitudinal width parameter, a . We began this effort right at the end of the summer during preparations for the Fall 2021 semester. We are happy to report several major improvements to the waveform comparison analysis. We chose to stay with a small number of explicit waveforms, hoping to keep the analysis repeatable for another user. We find that the electromagnetic and hadronic on-cone waveforms from NuRadioMC (ARZ 2020 model) strongly agree with our on-cone formula. In particular, the higher energy hadronic waveforms showed improvement in correlation coefficient and fractional power difference. Further, electromagnetic fits to on-cone waveforms seem to produce agreement between the waveform a -value and the a -value from NuRadioMC. The a -value agreement is not as good for the hadronic case, but still at the 10-50% level. As for the off-cone cases, both the electromagnetic and hadronic waveforms improved in their correlation and power match with respect to NuRadioMC ARZ2020. We find the correct viewing angle in all cases, and we even show that if we know θ in advance, the correlation between a -values from NuRadioMC and the waveforms is 97 percent. This sug-

gests an energy reconstruction that is robust to changes in comparison waveforms. Finally, we find that our ϵ parameter appears to be greater than 1 for on-cone electromagnetic (10 PeV) events, while it is less than 1 for hadronic (100 PeV) events. This is intriguing and requires further study. We also highlight in several places where we intend to perform large-scale statistical analysis of our analytic model parameters in comparison to semi-analytic parameterizations. We hope these improvements satisfy the reviewer's concerns.

3. Given that the manuscript derives on- and off-cone equations, it seems appropriate to consider the limit of the off-cone equations as they approach θ_C . What are the boundaries between using the on- and off-cone equations?

We now provide further clarification surrounding the symmetric approximation, and connect that idea to a new subsection, VB. Section VB is entitled Usage of the On-Cone versus Off-Cone Fields. We note that taking the limit $\theta \rightarrow \theta_C$ means $p \rightarrow 0$, and that the pulse amplitude (and therefore power) goes to zero for the off-cone case. At first, it seems like the amplitude should diverge, but the pulse width goes to zero. Conversely, as pulse width increases, amplitude decreases. If one tries to make the amplitude of the off-cone model similar to the on-cone model, this fails because the pulse width goes to zero. Observed power should be maximized at $\theta = \theta_C$, and decrease with increasing $\Delta\theta$. Thus, the pulse width of the off-cone field should be greater than or equal to the on-cone pulse width. We do not prove that this is the only requirement one can impose, but it yields a physical result in the form of an equation for $\Delta\theta_{\min}$. We also estimate the minimum appropriate angle to evaluate the off-cone field to be 1 degree off-cone. We hope this argument makes sense, and welcome any suggestions from the reviewer.

4. This is really a stylistic comment, but it strikes me that several aspects of the derivations are too detailed and could be removed or moved into an appendix. I recommend the later to keep the main body of the text shorter, while maintaining the authors' preference for presenting the complete derivation (which I appreciate). The advantage of a shorter main text is that the primary results can come through more cleanly. But again this is a style comment and I leave it to the authors what

their preference is. This could perhaps help with comment #1.

We hope our changes discussed in response to main comment #1 address this point as well. We have added two appendices that contain all the details not in the main body. We are always open to further suggestions.

5. This treatment is valid under the assumption that $\eta < 1$, which is mentioned at the beginning of the manuscript and used in the derivation. The conclusions and summary should discuss the implications of this assumption or whether there will be followup work for $\eta > 1$.

We have provided a paragraph at the end of Sec. VII A that describes the limitations imposed by requiring $\eta < 1$. First, we review the several reasons $\eta < 1$ was necessary for the calculations. Second, we estimate the minimum observer distance given $\eta = 1$ for typical parameters within η . We find that the restriction is a mild one. According to NuRadioMC, most UHE- ν Askaryan signals arrive from distances greater than the result implied by $\eta \leq 1$. Third, we describe how future work will test the limits of $\eta > 1$. Since this work already reveals a potential energy reconstruction, we must embark on a study to develop the probability distributions of parameters like a , ω_C , and ω_0 when our model is forced to match the other models in NuRadioMC. In these statistical comparisons it will be straightforward to determine in parallel how necessary it is to assume $\eta < 1$. If the reviewer has any insight for future directions, we'd be happy to consider those as well.

Minor Comments by page:

- Page 1: Abstract: “a simulation being used to design IceCube-Gen2”
→ “a simulation being used to design the radio component of IceCube-Gen2”

We adopted this change (see abstract).

- Page 1, Para 2: Clarify that the charge that radiates is the summed charge of the shower

We adopted this change (see paragraph 2, page 1) by adding “the net charge in the cascade”

- Page 1, Para 2: The discussion of the flux constraints strikes me as out of place. I recommend making a new paragraph prior to this one that describes the flux constraints across the energy band of interest to the radio component of Gen2. Furthermore, why not use the IceCube and/or Auger constraints here which are more constraining than ANITA at EeV energies? Or perhaps add them together.

We adopted this change, in that we introduce the IceCube EHE differential limit. Instead of creating a new paragraph, we decided to allow the reasoning to “flow down” from the Askaryan effect, to the low UHE- ν flux, to the conclusion that large in situ detectors are being developed.

- Page 1, Para 4: What can be approximated as a Gaussian? The radiation or Xmax?

Technically we meant the charge excess in the cascade. We deemed this phrasing unnecessary, so we decided to cut it since it was confusing (see current paragraph 3).

- Page 2, Last Para Section I: Seems odd to compare waveforms to “NuRadioMC” rather than to the ZHS or ARVS models. In principle, the authors are using NuRadioMC but the comparison is to the specific semi-analytical model used in Ref. 23.

We adopted this by changing “by NuRadioMC” to “by the semi-analytic parameterizations in NuRadioMC.”

- Section II Para 1: “index value” \rightarrow “refractive index for solid ice”

We adopted this change verbatim.

- Page 4: Typo in this line (subject / verb agreement): “As presented in Section II the form factor for the cascade ICD be the 3D Fourier transform of $f(z', \vec{\rho}')$.”

This is now fixed.

- Eqns 8-11 could be written more succinctly by just saying in words that the normalization factors are chosen such that integrating the form factor integrating over a sphere is normalized to 1.

We adopted this change. We now write in words that f is normalized to 1.

- Notation is inconsistent in the format of the vector in the dot product ($\vec{q} \cdot \vec{x}$) going from Eqn (12) to (13). (Eqn 12 has arrows over the vectors while Eqn 13 typesets the vectors in bold.)

This is now fixed.

- Paragraph between Eqns (15) and (16): Isn't γ the exponential scale factor of the ICD rather than the slope?

We adopted the change. Sometimes we use slope and scale factor interchangeably. It now reads exponential scale factor.

- Page 5: Eqns 24 through 27 seem unnecessary. I could imagine that you could skip directly to Eqn (28) or move several aspects of these derivations into an appendix. But I suppose this is somewhat of a stylistic choice.

We extra-adopted this one by creating Appendix A for the form factor. We felt quoting the integral tables will give future readers the ability to check for themselves.

- Fig. 2: I recommend adding the original ZHS parameterization to this plot for comparison.

If we adopted this edit, the ZHS curve would lie right on top of the gray curve of Fig. 2. That is the reason we have (what is now) Eq. 11 (Eq. 31 in the first draft). We did, however, discover a bug in the plotting script that creates Fig. 2. The curves are now more aligned, as they should be.

- Page 6: Eqn (49) I would revise so that the sqrt is over the entire equation, but this is a style choice. It also might be useful to point out that since $f(x)$ is a normalized function < 1 , $\ln(f)$ will always be negative and a is real valued.

We revised the format to keep the $\sqrt{\ln(x)}$ energy dependence by itself, but used the same symbol for both roots. We added a sentence reminding the reader that a will be real-valued.

- Page 7: Eqns (60) and (61) The approximations are valid if the approximations are valid if $\omega \ll \omega_C$ and $\omega \ll \omega_{CF}$ rather than only less than.

Note that these formulas are now in Appendix B. We chose to keep the language “such that the following approximations of the factors in the denominator are valid.” For example, the reader can see from Fig. 2 how close to 1 the parameter σ can be.

- Page 8: For clarity, I recommend introducing the fact that you will change to the time domain between (63) and (64)

We adopted this change by adding “...converts the field to the time domain.”

- Page 10, Fig. 3: The differences in the curves are hard to discern, particularly at the peaks and the top and bottom curves are almost indistinguishable. Perhaps use different colors or grey scales? Or offset them in time? Explain why you chose these parameters as examples either in the text or in the caption.

We adopted this edit by altering Figure 3 in the following ways: we now use gray scale for three cases, changed the normalization, explain in the text that the values are motivated by fit comparisons in Sec. 6, and the two plots are the same data. The second plot is a zoomed in version of the first. The data in the graphs are normalized to the black curve peak.

- Page 10: First paragraph: Is ω_C missing an n ?

Normally, yes, equations should have n ’s. In this work, in the units and vocabulary section, we state that we take things like k and c to be the wavenumber and speed “in the medium.” In this case, the n ’s would just cancel, so we chose to leave it as is.

- Section V: 2nd paragraph: Where is $\vec{\mathcal{E}}(\omega, \theta)$ defined? If it’s already derived in a reference, please cite it.

Fixed: the full version of \mathcal{E} is given in Appendix B.

- Eqn 85 and subsequent eqns in Section V: The equations are being evaluated at $u = 1$, but $x \rightarrow 1$ in the argument of $\mathcal{E}(u, x)$ rather than u . I would expect it to be written $\mathcal{E}(u, 1)$.

Fixed. The equations are now in Appendix B.

- Page 12, Para 1: Several variable (e.g. x) are reused multiple times in this manuscript.

What's funny is if you recite the alphabet, we have used almost every letter at some point in the paper. We might as well stay with x for the ratio of cascade energy to critical energy. We think given the tables of definitions, it's clear what we mean in the different cases.

- Page 12, Para 1: Voigt function \rightarrow Voigt function, $U(x, s)$. (I assume.)

We adopted this change, though we recast it as the line-broadening function (the relevant case of the Voigt function). We write “the line-broadening function, $H...$ ” and then use H subsequently.

- Page 13, Section VI: Para 1: While NuRadioMC includes semi-analytic models, they are all based on other work (eg. ZHS, AVRZ, etc) so you might as well cite the specific model that they used in Ref. 23 rather than the wrapper class.

We adopted this change. We cite the most recent ARZ result, as the waveforms are compared to that semi-analytic model using NuRadioMC.

- Eqn (108): $(\Delta f)^2$ seems to be a poor choice of variable name since it has units of power, but other variables in the fit are named “ f ” and represent frequencies.

Fixed: we can change to capital E to avoid confusion, as in $(\Delta E)^2$

- Pages 13-19, Figs 6-9: The white space around the waveforms is too large and the black curve hides the peak of the gray curve. I suspect

that some differences are hidden to the reader.

We decided to zoom in for the x-axis to reduce whitespace. For the y-axis, it's not that differences are being hidden, it's that the theoretical fit is minimizing $(\Delta E)^2$ and maximizing ρ . Thus, the waveforms should lie on top of each other as much as possible.

- Tabs V-VI, IX: Where appropriate please add the Monte Carlo truth to these tables. I also don't understand why "PeV" is listed in bold.

We adopted this change by changing PeV to regular face instead of bold, and adding the new MC truth values for a and θ to these tables when possible.

- Tab VII, IX: The description of the columns in the caption and the columns in the table are not consistent.

Fixed, the captions are now consistent with the columns for the off-cone tables.

- Tab VIII: Why is waveform #1 missing from this table?

Fixed. One waveform in the NuRadioMC paper for that EM energy is cut off visually, so we could not digitize it completely. We wanted others to be able to repeat our analysis, so we digitized from the paper. Now that we have transitioned to generating NuRadioMC waveforms, this is no longer an issue.

- Page 20, Section VII, Last Para: I agree that the potential for reconstruction is really intriguing. I would add a comment that this could be comprised of a study using the full statistical comparisons to the semi-analytical models.

We adopted this change by adding a sentence that future work will build the probability distributions of parameters like a and the cutoff frequencies. The hope is to have an energy reconstruction independent of waveform normalization (in the y-axis).