

Millisecond phonon lifetimes in piezoelectric nanomechanical resonators

Corresponding Author: Mr Oliver Hitchcock

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Comments for the Author)

This work demonstrates for the first time that piezoelectric nanoresonators made of lithium niobate (LN) can achieve phonon lifetimes in the millisecond range — a value previously reserved exclusively for non-piezoelectric materials such as silicon. The measured phonon lifetime of $T_1 = 4.9$ ms at 3.6 GHz corresponds to an inferred internal quality factor of $Q_i > 4 \times 10^8$ and a $f \cdot Q_i$ product of more than 10^{18} . The resonator quality factor $M = k^2_{\text{eff}} \cdot Q_i$ lies between 1.3×10^5 and 8.6×10^5 , thereby surpassing all previously published electromechanical platforms by more than an order of magnitude. These results represent an improvement of over three orders of magnitude compared to earlier LN nanoresonators. On the scientific level, the work refutes the widely held assumption in the field of a fundamental conflict of objectives between strong piezoelectricity and high mechanical coherence — an assumption that had pushed large parts of the research community toward hybrid architectures with non-piezoelectric resonators.

The central research question—that the material damage was caused during manufacturing ("[...]The implicit assumption has been that high material quality and strong piezoelectricity are incompatible. In this work we show that they are not.[...]") — could not be proven, as no identical structures with material damage were available for comparison.

The study compares three systems with one another (spacing 360 nm / 2800 nm / 4760 nm). The results are listed in table 1 and Appendix C (Table 2). The derived results for spacing 2800 nm are missing. If you enter them yourself, you will see that the three compared systems do not follow any explanatory logic. The phonon lifetime for spacing 2800 nm is higher, with $T_{1,\text{in}} = 26$ ms, than that of the system with 4760 nm. The results were intentionally not included. Why?

The paper's greatest weakness is the trade-off between the high manufacturing tolerance and the required precision of the spacings at the center of the device. Appendix G explains that this tolerance does not have a significant impact on the evaluation of the results. The claim that the alignment error is negligible is logically sound and well supported by measurement data for the key result regarding the 4760 nm device, since k_e is measured directly and the relative misalignment effect is small. For the 360 nm device, it is less convincing: the low yield suggests a strong selection mechanism, and the intrinsic loss parameters of this device could depend on an unknown effective geometry. However, due to the missing values for Design 2 (spacing 2800 nm), an external reader may question the validity of the argument.

A physically interesting finding, which is not discussed in the paper, lies in the dephasing rate:

Spacing $\Gamma \Phi / 2\pi \Gamma \Phi / k$

360 nm 74 kHz 2.3

2800 nm 30 kHz 127

4760 nm 17 kHz 523

As the spacing increases, dephasing increasingly dominates over energy decay—meaning the system becomes increasingly phonon-number-conserving but phase-destructive. This is consistent with TLS defects in the vicinity that generate frequency noise without directly dissipating energy. A physical analysis of this dependence is completely absent from the paper, although it would allow conclusions to be drawn about the defect density as a function of distance.

Furthermore, there is a lack of a statistically significant number of measurement results per design. It is therefore unclear whether the results are random values or reproducible. It would therefore be desirable to obtain a statistical distribution of relevant measurement parameters with a sample size (n) and a standard deviation. For derived measurement values, an

error tolerance corresponding to 1-sigma, 2-sigma, or 3-sigma is ideal.

The self-citation rate is high, at just under 30%, but since the work draws on the author's own previous research, this is acceptable.

Certain stylistic elements need to be revised. Captions should be placed above the tables (e.g., Table 1). A "no-line-break" should be inserted between measured values and units.

Personally, I feel that it weakens the line of reasoning to present the measurement results first (Section 2) and only then explain the manufacturing process (Section 3).

The drawings in Figure 2 are difficult for non-experts in the field of nanotechnology to understand.

In 2a), two of the figures appear identical to the viewer. In Figure 2b), there are layout issues. A number of parameters are not explained (e.g., T_e).

Given the novel nature of the findings and data, it is recommended that optional measurements related to the reviewer's open research questions be addressed in future work, rather than conducting new measurement series in this publication (for example, comparing identical designs manufactured using different methods to directly demonstrate the influence of material defects on quality factor).

The paper is academically high-ranking, innovative, and very well written. I recommend it for publication.

Reviewer #2

(Comments for the Author)

The manuscript offers a material agnostic fabrication process to extend phonon lifetimes piezoelectric materials wherein a shadow mask insures no contamination to the wafer after the SC1 clean ($\text{NH}_4\text{OH} - \text{H}_2\text{O}_2 - \text{H}_2\text{O}$). The process is demonstrated with lithium niobate (LN) and the phonon lifetime is measured at 10 mK by observing the energy ringdown of coherent states with respect to the intracavity phonon number. The authors achieve an internal quality factor (Q_i) of greater than 4×10^8 , a frequency-quality factor product of greater than 1018, and an effective electromechanical coupling (k_{eff}^2) of 5.2×10^{-4} measured or 3.5×10^{-3} simulated. The manuscript gives a detailed explanation of the fabrication processes and addresses concerns about misalignment and shadow mask distortion. The work is impactful for the fabrication and use of LN for acoustic resonators with opportunities to be applied to other materials as well. The manuscript is scientifically sound and presented well. Below a few questions and suggestions are offered to improve the clarity and completeness of the manuscript:

1. Figure 3 shows this work's Q_i and k_{eff}^2 compared to other materials, including other LN-PNC devices [5,6]. It is well explained how the fabrication process improves the Q_i by cleansing the surface of contaminants. Can the authors offer an explanation as to why the k_{eff}^2 is lower than the other LN-PNC devices?
2. Appendix E explains how k_{eff}^2 is estimated for works where it is not explicitly quoted. Did the authors test their estimation techniques against the works where it is explicitly quoted to ensure consistency across all the devices being compared?
3. The misalignment offset seems to be the largest tradeoff to the fabrication process, and it is limited to the mechanical tolerance of the bonding tool as described in appendix G. The yield is heavily affected by this misalignment and can be improved by using a tool with lower tolerance. Can the authors briefly discuss the availability of lower tolerance bonding tools and the minimum tolerance one might find commercially available?

Sunil Bhawe and Alaina Attanasio

Version 1:

Reviewer comments:

Reviewer #1

(Comments for the Author)

The authors have explicitly addressed all of Reviewer 1's comments and have in fact incorporated a large portion of the proposed changes into the manuscript. Overall, the revision is substantial and largely convincing.

The work demonstrates LN resonators with very long internal lifetimes and high Q_i .

This convincingly supports the claim that strong piezoelectricity and high mechanical coherence are fundamentally compatible.

The missing 2800 nm analysis was appropriately replaced by a threshold value. The reproducibility of the 2800 nm design was significantly improved through additional measurements.

Reviewer #2

(Comments for the Author)

the authors have satisfactorily answered our questions. I recommend this manuscript for publication.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Response to Referees

Jul 20, 2026

Millisecond phonon lifetimes in piezoelectric nanomechanical resonators

Oliver A. Hitchcock, Matthew P. Maksymowych, Kaveh Pezeshki, Takuma Makihara, Erik Szakiel, Chou-Wei Kiang, Daniel Rugar, and Amir H. Safavi-Naeini

Reviewer #1 (Comments for the Author):

This work demonstrates for the first time that piezoelectric nanoresonators made of lithium niobate (LN) can achieve phonon lifetimes in the millisecond range — a value previously reserved exclusively for non-piezoelectric materials such as silicon. The measured phonon lifetime of $T_1 = 4.9$ ms at 3.6 GHz corresponds to an inferred internal quality factor of $Q_i > 4 \times 10^8$ and a $f \cdot Q_i$ product of more than 10^{18} . The resonator quality factor $M = k^2_{\text{eff}} \cdot Q_i$ lies between 1.3×10^5 and 8.6×10^5 , thereby surpassing all previously published electromechanical platforms by more than an order of magnitude. These results represent an improvement of over three orders of magnitude compared to earlier LN nanoresonators. On the scientific level, the work refutes the widely held assumption in the field of a fundamental conflict of objectives between strong piezoelectricity and high mechanical coherence — an assumption that had pushed large parts of the research community toward hybrid architectures with non-piezoelectric resonators.

The central research question—that the material damage was caused during manufacturing ("[...]The implicit assumption has been that high material quality and strong piezoelectricity are incompatible. In this work we show that they are not.[...] ")—could not be proven, as no identical structures with material damage were available for comparison.

Thank you for raising this point. We will begin by clarifying our precise claim and then present additional data that addresses the requested comparison.

As stated in our introduction, fabrication processing damages the surfaces of lithium niobate: "Detailed materials studies identified several mechanisms [of damage] and tied them to specific fabrication steps. Ion milling...amorphizes exposed surfaces and creates a ~10 nm layer hosting a dense population of two-level system (TLS) defects." [L46-L51]

We therefore consider it established prior art that material damage is caused by manufacturing and that chemical treatments can strip damaged surfaces [Phys. Rev. Appl. 23, 064055 (2025), Applied Physics Letters 118, 123501 (2021)].

The central questions we seek to address are therefore: 1) *how* to apply these material science results to released nanomechanical devices, and 2) *the degree to which* phonon lifetimes improve as a result. This requires resolving the following challenge: "the wet chemistry required

for both device release and surface restoration is incompatible with metal contacts, and conventional resist-based liftoff metallization contaminates the very surfaces being cleaned” [L67-L70]. The key innovation we report that solves this challenge is the material-agnostic, contactless shadow mask metallization technique, which defers metallization until after release and surface restoration.

The degree of improvement is then quantified by our phonon lifetime measurements, which rise by several orders of magnitude relative to all previous piezoelectric resonators, until energy decay is dominated by external coupling rather than intrinsic material losses. We would also note that the sentence quoted by the reviewer, our claim that: high material quality and strong piezoelectricity *are compatible*, is a claim about compatibility, and it is demonstrated directly by the devices reported here: strongly piezoelectric resonators with internal quality factors comparable to the best non-piezoelectric platforms. This demonstration stands independently of any comparison against damaged structures.

Nonetheless, we agree with the reviewer that a direct, within-study comparison between damaged and restored structures strengthens the paper. The manuscript now provides this and we discuss the details here. Throughout this study, we fabricated devices on separate chips from the same lithium niobate on insulator (LNOI) wafer, with nearly identical device design and fabrication processing, except that some batches omitted the SC1 surface damage removal steps.

We have now additionally cooled down and measured lifetimes for two devices from our legacy, pre-shadow-mask fabrication process (400 nm and 1700 nm electrode gaps), containing known ion-mill damage and contamination from liftoff-resist metallization. We find for three distinct modes on the 400nm spaced devices:

Frequency [GHz]	T1 (std) [us]	Te (std) [us]	kappa_e/2pi (std) [kHz]
2.126221	2.1 (0.2)	3.652 (0.003)	43.58 (0.04)
2.155882	1.3 (0.1)	4.971 (0.006)	32.02 (0.04)
2.182663	2.8 (0.3)	6.588 (0.008)	24.16 (0.03)

These phonon lifetimes are *not* dominated by external coupling and are consistent with previous work [Nature 571, 537 (2019), Nature 604, 463 (2022)], where devices were limited to microsecond lifetimes.

For three distinct modes on the 1700 nm spaced devices:

Frequency [GHz]	T1 (std) [us]	Te (std) [us]	kappa_e/2pi (std) [kHz]
2.147424	25 (1)	85.4 (0.35)	1.86 (.01)

2.175234	13 (2)	120.2 (0.5)	1.32 (.01)
2.195269	2.6 (.4)	182 (2)	0.872 (.008)

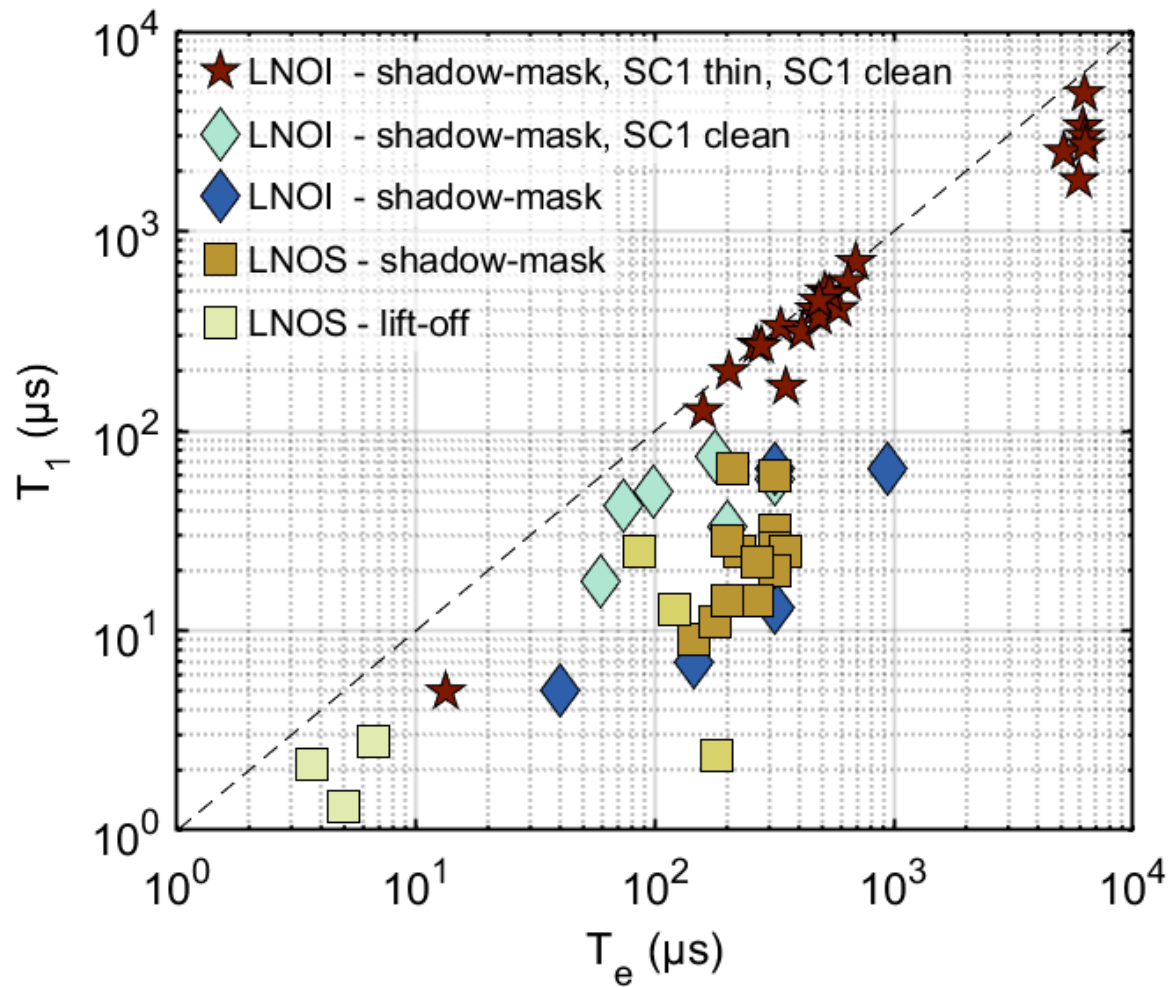
These devices are *also not* dominated by external coupling, in contrast to devices with 2800nm-gap made with the new fabrication process.

Furthermore, we have cooled down a previously unmeasured device array from the 2800nm-gap design batch to increase the statistics for the ‘Design 2’ group. The resulting data for these measurements has been added to Fig. 1c, the individual values are shown below:

Frequency [GHz]	T1 (std) [us]	Te (std) [us]	kappa_e/2pi (std) [kHz]
3.404642	166 (1)	352 (4)	0.45 (.01)
3.434473	311 (1)	411 (3)	0.387 (0.003)
3.461041	372 (6)	486 (4)	0.327 (0.003)
3.488115	408(23)	585 (6)	0.272 (0.003)
3.517525	556 (11)	641 (7)	0.248 (0.003)
3.537129	484 (3)	539 (5)	0.295 (0.003)
3.569299	447 (2)	484 (5)	0.328 (0.003)
3.586657	382.4 (0.6)	369 (3)	0.431 (0.003)
3.620333	266.1 (.3)	278 (2)	0.571 (0.004)
3.645697	198.4 (0.3)	204 (1)	0.782 (0.005)
3.682808	126.1 (0.7)	159 (1)	1.002 (0.007)

These results are consistent with our previous measurements of ‘Design 2’ devices: many of the phonon lifetimes are a significant fraction of the external coupling lifetimes, with some modes dominated by the external coupling lifetime. This is despite belonging to an entirely separate device array and undergoing multiple cooldown/warmup cycles.

A new Appendix H containing a new Fig. 11 and Table 4 summarizes the measurements of resonators across these multiple process variants. Devices that received the full SC1 surface restoration are external coupling dominated ($T1 \sim T_e$), while devices retaining surface damage fall further from the external coupling limit.



Together, the prior material studies [Phys. Rev. Appl. 23, 064055 (2025)], added comparison of fabrication process variants, and the systematic spacing of electrodes supports our claim that shadow-masks and the enabled surface restoration techniques have reduced phonon losses and resulted in the breakthrough to millisecond phonon lifetimes.

The study compares three systems with one another (spacing 360 nm / 2800 nm / 4760 nm). The results are listed in table 1 and Appendix C (Table 2). The derived results for spacing 2800 nm are missing. If you enter them yourself, you will see that the three compared systems do not follow any explanatory logic. The phonon lifetime for spacing 2800 nm is higher, with $T_{1,in} = 26$ ms, than that of the system with 4760 nm. The results were intentionally not included. Why?

You are correct that direct propagation gives $T_{1,in} = 26$ ms for the 2800 nm device, exceeding the 4760 nm value, and we agree that a blank entry was confusing. This entry was omitted for the following reason: the measured lifetime $T_1 = 674 \pm 4$ μ s lies within the uncertainty of the coupling limit $T_e = 690 \pm 30$ μ s, so the inferred internal loss rate $k_i = k - k_e = 2\pi (6 \pm 10)$

Hz is consistent with zero, therefore direct propagation of uncertainty yields a value of T_1 with an unbounded upper value. Our measurement simply could not resolve a difference between energy decay and the coupling limit and so instead of a value of T_1 with infinitely large error bars, we simply omitted it. We therefore limited our claim to: the devices are external coupling dominated, as stated in the main text [L138-L142]:

“For the two widest spacings, the measured T_1 is either consistent with or close to the external coupling limit T_e , indicating that the energy decay is dominated by leakage into the measurement circuit rather than by intrinsic material loss”

This result was precisely what motivated pushing the electrodes to 4760 nm spacing, where the smaller k_e makes the internal lifetime statistically resolvable. We have revised Tables 1 and 2 so that they now report one-sided bounds on the derived intrinsic loss values, with brief explanations in the captions. We note these bounds are consistent with the 21 ± 6 ms inferred for the 4760 nm device, supporting reproducibility across independent devices/designs for the new fabrication process.

The paper’s greatest weakness is the trade-off between the high manufacturing tolerance and the required precision of the spacings at the center of the device. Appendix G explains that this tolerance does not have a significant impact on the evaluation of the results. The claim that the alignment error is negligible is logically sound and well supported by measurement data for the key result regarding the 4760 nm device, since k_e is measured directly and the relative misalignment effect is small. For the 360 nm device, it is less convincing: the low yield suggests a strong selection mechanism, and the intrinsic loss parameters of this device could depend on an unknown effective geometry. However, due to the missing values for Design 2 (spacing 2800 nm), an external reader may question the validity of the argument.

Thank you for your question regarding the role of the 360 nm device; we understand the concern that low yield could act as a selection filter, leaving a single measured device that is unrepresentative. Several points addressing this:

1. The yield is deliberately limited to guarantee at least one working device per measurement – we are not post-selecting on performance. We intentionally offset the shadow-mask array alignment to ensure that some devices on the chip will almost always be measurable even with imperfect mask alignment. However, this also means that for a given global flipchip-bond misalignment, a deterministic subset of devices receives correctly placed electrodes, while the rest are visibly misaligned and never measured. This is true for all designs (360, 2800, and 4760 nm) but is especially pronounced for the smallest 360 nm gap devices. However, this is a purely geometric pass/fail, uncorrelated with phonon loss rates. Given the analysis of shadow mask misalignment in Appendix G, future devices will no longer require vernier offsets, increasing the yield.

2. The effective geometry is well constrained. An SEM image of the device is shown in Fig. 9 and the geometry is consistent with the design. Additionally, the measured k_e is roughly consistent with simulation (11.97 kHz measured vs 10.8 kHz simulated, Table 2), suggesting that we realized the designed geometry and target mode.
3. The central claim of the manuscript that phonon losses are limited by external coupling rests on the higher yield 2800 nm and 4760 nm designs which had consistent behavior across multiple devices (Fig. 1c, Fig. 11). The role of the 360 nm design is to show that the same shadow-mask process can produce functional ‘on-defect’ electrode devices, to reveal the metal loss channel that appears in such devices, and to illustrate the necessity to push the electrodes further from the defect in order to probe the intrinsic lithium niobate material losses.

We acknowledge that additional measurements of different devices within the 360 nm design group and the resulting statistics would be desirable, however, due to the low yield, this would require a dedicated fabrication campaign for devices which are not critical to support the claim of this paper (lifetime measurements that resolve intrinsic lithium niobate phonon losses). Consistent with the reviewer's recommendation that optional measurements be addressed in future work, we defer this, while noting that commercially available bonders with 200-300 nm placement accuracy (see response to Reviewer 2, point 3) would raise on-defect yield substantially and make such statistics practical.

A physically interesting finding, which is not discussed in the paper, lies in the dephasing rate:

Spacing $\Gamma\phi/2\pi$ $\Gamma\phi/k$

360 nm 74 kHz 2.3

2800 nm 30 kHz 127

4760 nm 17 kHz 523

As the spacing increases, dephasing increasingly dominates over energy decay—meaning the system becomes increasingly phonon-number-conserving but phase-destructive. This is consistent with TLS defects in the vicinity that generate frequency noise without directly dissipating energy. A physical analysis of this dependence is completely absent from the paper, although it would allow conclusions to be drawn about the defect density as a function of distance.

Your observation is correct, physically significant, and the subject of an in progress follow-up study of these devices. The dephasing persists even as the intrinsic energy lifetime improves by orders of magnitude, and depends only weakly on electrode placement.

This suggests that the slow dephasing measured in spectroscopy originates predominantly in the non-metallized lithium niobate of the central defect region rather than at the metal-LN interface. This is an interesting and unexpected result, since metal interfaces and amorphous oxides are natural suspects for hosting TLS. The behavior is consistent with residual TLS in the lithium niobate that are dispersively coupled to the mechanical mode, contributing negligibly to

energy relaxation rate ($\sim 1/\Delta^2$) while dominating dephasing rate ($\sim 1/\Delta$), where Δ is the detuning between the resonator and the TLS. A quantitative analysis requires experimental approaches beyond the ringdown and spectroscopy used here, and is in progress.

Consistent with the reviewer's recommendation that open questions be addressed in future work, we report the lifetime and fabrication results here and will present the dephasing study and our attempts at improvement/mitigation, separately.

Furthermore, there is a lack of a statistically significant number of measurement results per design. It is therefore unclear whether the results are random values or reproducible. It would therefore be desirable to obtain a statistical distribution of relevant measurement parameters with a sample size (n) and a standard deviation. For derived measurement values, an error tolerance corresponding to 1-sigma, 2-sigma, or 3-sigma is ideal.

We have worked to address the lack of a statistically significant number of measurements results per design. We performed new measurements from 2800nm-gap devices to increase the number of measurements for this design. We have added this data to Fig. 1c. For the 360 nm design, the vernier offsets that made the process robust to bonder misalignment limited the yield, so ensemble statistics are not attainable with the currently available devices and data. As the main results of the paper concern the larger gap devices, we did not find it within scope to fabricate new 360nm devices and defer this to future studies.

While the overall design and measured modes shown in this work are nearly identical –having the same crystal cut, targeting the same shear mode, suspended by a phononic bandgap structure– the precise geometries are perturbed within a device array and between devices to allow for frequency multiplexing. As a result, T_e and T_1 are naturally expected to vary between devices, therefore, the T_1/T_e (coupling limit saturation ratio) is the most informative statistic for us to quote within a given design.

Variant	Stack	Electrode gap	Metalization	SC1 thin	SC1 clean	T_1/T_e	n
1	LNOI	4760 nm	shadow-mask	Y	Y	0.50 ± 0.16	6
2	LNOI	2800 nm	shadow-mask	Y	Y	0.86 ± 0.15	16
3	LNOI	360 nm	shadow-mask	Y	Y	0.37 [†]	1
4	LNOI	1700 nm	shadow-mask	N	Y	0.36 ± 0.17	6
5	LNOI	1700 nm	shadow-mask	N	N	0.09 ± 0.06	6
6	LNOS	1850 nm	shadow-mask	N	Y	0.11 ± 0.07	13

7	LNOS	1700 nm	lift-off-resist	N	N	0.13 ± 0.14	3
8	LNOS	400 nm	lift-off-resist	N	N	0.42 ± 0.16	3

† single resonator, no uncertainty reported

We have revised the manuscript to include information related to the statistical distribution of relevant measurement results. Information about the sample size, mean and standard deviation have been placed in the new Appendix H, and referenced within the main text [L162-L164].

The self-citation rate is high, at just under 30%, but since the work draws on the author's own previous research, this is acceptable.

We appreciate the reviewer's understanding on this point; no changes were made. Though the NEMs community has been taking up phononic crystal devices in the last few years for low-temperature measurements and ultra-long lifetimes, the field is quite new. We have added an additional reference to recent work by the Purdy group.

Certain stylistic elements need to be revised. Captions should be placed above the tables (e.g., Table 1). A “no-line-break” should be inserted between measured values and units.

The captions of Tables 1 and 3 now appear above the tables. We have inserted non-breaking spaces between all values and units throughout the manuscript. Thank you for noting these.

Personally, I feel that it weakens the line of reasoning to present the measurement results first (Section 2) and only then explain the manufacturing process (Section 3).

This is a fair point, and we considered restructuring the manuscript along these lines. However, we have chosen to retain the original ordering for the following reasons:

1. The paper's headline result is the measurement of millisecond phonon lifetimes that are external coupling dominated. For a broad readership we prefer to present that evidence first, consistent with the results forward convention of Nature family journals. The shadow-mask fabrication section then explains how these results were achieved.
 2. The introduction already previews the process logic [L75-L89], and introduces the shadow mask technique before the Section 2 discussion of devices and measurements. It is therefore only the details of the fabrication that are left to section 3.
 3. We feel it may confuse some readers to present the shadow mask details before introducing the devices they are used to create.
-

The drawings in Figure 2 are difficult for non-experts in the field of nanotechnology to understand.

In 2a), two of the figures appear identical to the viewer. In Figure 2b), there are layout issues.

A number of parameters are not explained (e.g., T_e).

We have redrawn Figure 2 for clarity. The film-thinning step is now shown explicitly; the HF-undercut and surface-cleaning panels are condensed and given distinct surface coloring to emphasize that the wet etch removes the amorphous damaged layer; front/back labeling of the SiN wafer in panel b is added.

All parameters are now defined at first use;

- In particular, $T_e \equiv 1/k_e$, the external-coupling-limited lifetime, is defined at first use in the main text [L137] as well as in the caption of Table 1.

If specific layout issues remain, we would gladly address them.

Given the novel nature of the findings and data, it is recommended that optional measurements related to the reviewer's open research questions be addressed in future work, rather than conducting new measurement series in this publication (for example, comparing identical designs manufactured using different methods to directly demonstrate the influence of material defects on quality factor).

The paper is academically high-ranking, innovative, and very well written. I recommend it for publication.

We thank the reviewer for their time and effort in reviewing this work, for the constructive feedback, and for the recommendation to publish. Their suggestions have measurably improved the manuscript and its conclusions.

Reviewer #2 (Comments for the Author):

The manuscript offers a material agnostic fabrication process to extend phonon lifetimes piezoelectric materials wherein a shadow mask insures no contamination to the wafer after the SC1 clean ($\text{NH}_4\text{OH} - \text{H}_2\text{O}_2 - \text{H}_2\text{O}$). The process is demonstrated with lithium niobate (LN) and the phonon lifetime is measured at 10 mK by observing the energy ringdown of coherent states with respect to the intracavity phonon number. The authors achieve an internal quality factor (Q_i) of greater than 4×10^8 , a frequency-quality factor product of greater than 1018, and an effective electromechanical coupling (k_{eff}^2) of 5.2×10^{-4} measured or 3.5×10^{-3} simulated. The manuscript gives a detailed explanation of the fabrication processes and addresses concerns about misalignment and shadow mask distortion. The work is impactful for the fabrication and use of LN for acoustic resonators with opportunities to be applied to other materials as well. The manuscript is scientifically sound and presented well. Below a few questions and suggestions are offered to improve the clarity and completeness of the manuscript:

1. Figure 3 shows this work's Q_i and k_{eff}^2 compared to other materials, including other LN-PNC devices [5,6]. It is well explained how the fabrication process improves the Q_i by cleansing the surface of contaminates. Can the authors offer an explanation as to why the k_{eff}^2 is lower than the other LN-PNC devices?

The lower k_{eff}^2 is a consequence of the design choice (see L112), moving the metal electrodes further away from the central defect cell. For a given applied voltage, distant electrodes generate less strain in the resonator, reducing the effective electrical to mechanical conversion ($k_{\text{eff}}^2 \sim 1\text{e-}3$) when compared to a device with metal on or adjacent to the central defect ($k_{\text{eff}}^2 \sim 1\text{e-}1$). For LN-PNC devices, this coupling is a tunable knob we exploit in this work to reveal the intrinsic material losses, rather than a limit imposed by the platform.

2. Appendix E explains how k_{eff}^2 is estimated for works where it is not explicitly quoted. Did the authors test their estimation techniques against the works where it is explicitly quoted to ensure consistency across all the devices being compared?

We appreciate this comment as this is a useful check to confirm the validity of our estimations. Our estimation techniques follow those typically used in the field when k_{eff}^2 is quoted and we have tested them for the references with explicitly quoted values of k_{eff}^2 :

Ref. [22] [<https://doi.org/10.1038/nature08967>]

The paper quotes $k_{\text{eff}}^2=1.2\text{e-}2$, obtained from the frequency splitting of the measured transmission spectra, shown in Fig. 1d:

$$k_{\text{eff}}^2=(6.10^2 - 6.07^2)/(6.10^2)*\pi^2/8=1.2\text{e-}2.$$

This includes a factor of $\pi^2/8$ (thickness-mode factor) which is an alternative and slightly different convention for k_{eff}^2 than the one we use in our work (Eq. E2, consistent with IEEE convention), representing a correction of order unity. Interestingly, this paper also includes equivalent BVD circuit parameters C_0 and C_m , which when plugged into Eq. E1 yield $k_{\text{eff}}^2 = C_m/(C_0 + C_m) = 3.4 \times 10^{-3}$. This is a factor of ~ 3 lower than the paper's quoted frequency-splitting value, despite the equivalence of these two expressions (up to the thickness-mode factor). This internal inconsistency within Ref. [22] may result from the quoted numeric value of C_0 being analytically estimated from the physical geometry, (which the paper seems to imply). An overestimation of C_0 (including non-resonant electrode area) could account for this factor of ~ 3 . **To conclude, Ref. [22] uses the same technique we present in Appendix E (Eq. E2, frequency-splitting form of k_{eff}^2) to obtain its explicitly quoted value, and is consistent with our other estimations.**

Ref. [28] [<https://doi.org/10.1126/science.1257219>]

This reference does not have enough information to directly enter our appendix E expressions, which is why we defaulted to the generous material parameter upper bound for GaAs SAW devices ($k_{\text{eff}}^2 = k^2 = 7 \times 10^{-4}$). Nevertheless, a comparison between the equivalent-circuit and the transducer model of [28] permits an approximate consistency check, which supports the quoted value, as we show below.

Following the main text [28], we can write the acoustic admittance element of the transducer as $Y_{a,\text{tr}} = G_a(w) + i B_a(w)$. Where the conductance $G_a = G_a(w_{\text{IDT}})$ captures the acoustic emission of the qubit through the transducer IDT on resonance, and can be expressed as $G_a = 2c_g^2 K^2 N_{\text{tr}}^2 W \epsilon w_{\text{IDT}}$. The supplementary material also expresses the electrostatic capacitance of the transducer as $C_0 = C_{\text{tr}} = \sqrt{2} N_{\text{tr}} W \epsilon \sim 85$ fF. The ratio of these two expressions, at the resonance condition $w_{10} = w_{\text{IDT}}$, gives the acoustic relaxation rate of the qubit through the transducer IDT quoted in main text, $\Gamma_{10} = G_a/(2C_{\text{tr}}) = c_g^2 N_{\text{tr}} K^2 w_{10}/\sqrt{2}$, with fitted measurements yielding, $\Gamma_{10}/2\pi = 38$ MHz. We therefore estimate $G_a = 2C_{\text{tr}} \Gamma_{10} \sim 40$ uS. Furthermore, the supplement gives $G_a(w) = G_a \text{sinc}^2(x)$, with $x = N_{\text{tr}} \pi(w - w_{\text{IDT}})w_{\text{IDT}}$, which when integrated is:

$$\int G_a(w) dw = G_a w_{\text{IDT}}/N_{\text{tr}}.$$

From the modified BVD equivalent circuit, the integral over conductance is $\int G(w) dw = \pi/2 w_{\text{IDT}}^2 C_m$. Equating these expressions yields: $C_m = 2G_a / \pi w_{\text{IDT}} N_{\text{tr}}$. Inserting these into our Eq. E1: $k_{\text{eff}}^2 = C_m/(C_0 + C_m) \sim C_m/C_0 = 2G_a / \pi w_0 N_{\text{tr}} / C_{\text{tr}} \sim 5 \times 10^{-4}$, which is consistent within $\sim 30\%$ of the quoted material parameter upper limit of 7×10^{-4} for a GaAs SAW. **Therefore the quoted value from Ref. [28] is consistent with our estimation technique.**

Ref. [30] [<https://doi.org/10.1063/5.0125122>]

For the quartz SAWs in this reference, the electromechanical coupling coefficient is defined by $k^2 = 2(v_f - v_s)/v_f$, where v_f and v_s are the free and shorted/metallized velocities of a SAW propagating along the substrate, respectively. This definition is the first-order expansion of our Eq. E2. In a SAW device with fixed wavelength (lithographically defined IDT period), the frequency and velocities are proportional, while the admittance resonance (pole) and

anti-resonance (zero) correspond to the shorted and free velocity boundary conditions. Substituting and expanding for small velocity splitting, we obtain: $k_{\text{eff}}^2 = 2(v_f - v_s)/v_f + O((v_f - v_s)^2/v_f^2)$. The quadratic correction is negligible for small k_{eff}^2 , as is the case for a quartz SAW. **As such, the quoted value of Ref. [30] is consistent with our estimation technique.**

3. The misalignment offset seems to be the largest tradeoff to the fabrication process, and it is limited to the mechanical tolerance of the bonding tool as described in appendix G. The yield is heavily affected by this misalignment and can be improved by using a tool with lower tolerance. Can the authors briefly discuss the availability of lower tolerance bonding tools and the minimum tolerance one might find commercially available?

The reviewer is correct that yield is primarily driven by the bonding tool misalignment, rather than our shadow masks, and could be improved with a better bonding tool. Our results were obtained with a ~1 um-class research bonder (Finetech lambda; ~2 um practical), near the top of the commercial tolerance range. Production die bonders currently reach 200 - 300 nm (3 sigma) (ASMPT AMICRA NANO, Finetech femto 2, Accura Opto), which alone would raise our 360 nm design yield by shrinking the required vernier. Beyond tool tolerance, the shadow-mask approach may admit self-aligning techniques [<https://doi.org/10.1109/ECTC.2014.6897268>, <https://doi.org/10.1109/OMEMS.2012.6318818>], such as etched divots in the shadow mask chip that mate with LN bumps and would effectively decouple the ultimate achievable alignments from the bonder entirely and potentially reach below 100 nm. This brief discussion has been added in Appendix G [L648-L654] .

Response to Reviewer Comments - 2

Aug 27, 2026

Millisecond phonon lifetimes in piezoelectric nanomechanical resonators

Oliver A. Hitchcock, Matthew P. Maksymowych, Kaveh Pezeshki, Takuma Makihara, Erik Szakiel, Chou-Wei Kiang, Daniel Rugar, and Amir H. Safavi-Naeini

Reviewer #1:

Comments for the Author:

The authors have explicitly addressed all of Reviewer 1's comments and have in fact incorporated a large portion of the proposed changes into the manuscript. Overall, the revision is substantial and largely convincing.

The work demonstrates LN resonators with very long internal lifetimes and high Q_i .

This convincingly supports the claim that strong piezoelectricity and high mechanical coherence are fundamentally compatible.

The missing 2800 nm analysis was appropriately replaced by a threshold value. The reproducibility of the 2800 nm design was significantly improved through additional measurements.

[We would like to thank Reviewer #1 for their recommendation.](#)

Reviewer #2:

Comments for the Author:

the authors have satisfactorily answered our questions. I recommend this manuscript for publication.

[We would like to thank Reviewer #2 for their recommendation.](#)