Reviewers' comments:

Reviewer #1 (Remarks to the Author):

The manuscript presents a two-photon comb source for quantum communications complemented with a wavelength converter.

The subject is interesting and timely, but we regret to say we found the presentation confusing at many points. More in detail:

- There is some ambiguity in the distribution distance. We understand that each photon travels a distance of 10km, thus covering an overall separation by 20km, and allowing quantum communications over that length. However, this should not be phrased as a 20km transmission, as this describes a single photon travelling that far.

- The introduction starts from the recent demonstration of quantum supremacy by Google. However much we may like their result, Shor's algorithm is still far from posing a believable threat to our crypto-systems. We think that a more solid argument can be drawn, but, admittedly, this is just our personal taste.

- On line 81, there is some claim that the bow-tie configuration is less demanding in terms of maintenance. The authors should also bear in mind that locking to a specific line should be put in place for coupling to a quantum memory. Can the authors explain if this advantage would remain even in these conditions.

- The discussion about dispersion on lines 96-97 is confusing; please revise.

- The fidelity of the output state to a maximally entangled one is high, but this decreases when the some control is inserted to manipulate it to a specific target. Can the authors account for this decrease in fidelity? Can it be attributed to dispersion effects of the waveplate?

- The polarisation behaviour of the WC is not at all discussed in the main text.

- The discussion on Fig. 4 in lines 140-141 is very terse, and only focusses on the enhancements. Likewise, the comment on the single photon conversion is excessively concise.

- The abstract reports a value of 5 for the g2, higher than 1 by 15 standard deviations (we assume). These quantitative considerations tend not to be reported in the main text, however. Furthermore, error bars should be reported in Fig.4, or a statement about them being smaller than the points should be made.

- On line 157 loss is identified as a reducing factor of g2, however this figure is notoriously resistant with respect to loss. We suspect that the authors are being once again too concise in their treatment; can they spell out how loss affects the g2? Is it a matter of signal-to-noise ratio?

- Lines 173-175 are confusing; the authors should expand this argument.

- When a short comparison with the bandwidth of a Pr quantum memory is carried out, it is not understood whether the source is compatible with existing memories. In detail: is the overall bandwidth sufficiently compact for an AFC protocol to address it? is the separation between the teeth compatible with that of an AFC? if so, what would the rephasing time be? If not, what more is needed?

- The discussion on the correlation filter perhaps needs some revision and extensions.

We do not feel able to provide a recommendation before these clarifications are take into account.

Reviewer #2 (Remarks to the Author):

The manuscript presents an entanglement source targeted at fiber-based quantum internet. This work, while timely, requires substantial revision. Publication in current form is premature.

Given that the introductory text/context is all about converting entangled photons to the correct wavelength for quantum memories, one would expect that there would be some measure of the fidelity of these converted photons, but instead, all they present is Fig 4c where the comb sort of looks like the comb before conversion. They should have done tomography on the converted photons, not just the photons from the two-photon comb.

Let's say the intro were re-written to make it clearer that this is a very technical paper in experimental optics and not necessarily something very useful to the quantum information community. Then their major claims are that 1) they developed a way of creating a two-photon comb that is more robust to external conditions, setup, etc. and 2) they couple it to a wavelength converter.

Regarding 1), they claim their method gives linewidths an order of magnitude shorter than the versatile entanglement source for QI applications they cite in Ref. 19, but then if you look at the table in SI Table 1 (in note 6) comparing their results to other photon sources that were not explicitly developed as versatile entanglement sources, it seems that their line widths are not substantially smaller, so not sure if this work is significant for experimental optics community either.

Regarding 2), the authors don't provide much context on how hard it is to convert photons or whether they did any innovation to increase efficiency/fidelity (not even in SI notes 4 where they talk about experimental details of wavelength conversion), so it's difficult for me to say whether this is a big deal.

Paper writing clarity/style could be substantially improved. While this is not a significant factor in deciding whether a paper should move forward, I would like to suggest changes to the presentation; the current version was not an easy read.

Reviewer #3 (Remarks to the Author):

Please see attached file.

The manuscript I have been asked to review presents a spontaneous parametric down conversion based versatile entanglement source (VES) that operates at telecom wavelengths and is capable of generating photon pairs in a narrow linewidth. While telecom regime operation is attractive from a fiber-optic communication point of view, narrow linewidths are a requirement for being able to interface the photons with the narrow atomic transitions of quantum memories (QMs). The first part of the experiment describes the operation and characterization of the VES, e.g., via Hanbury-Brown-Twiss (HBT) measurements for two-photon correlations and Bell state tomography for entanglement fidelity, etc. Effects such as dispersion arising from transmitting the entangled pairs through an optical fiber channel are also analysed. In the second part, the authors present a wavelength conversion (WC) stage: as typical QMs operate in the visible regime, the generated photon pairs need to have their wavelength converted to be compatible to the QM. This WC is realized through the process of sum frequency generation (SFG) after which, they compare the results obtained in the telecom regime with the visible regime intensively, e.g., through normalized second-order signal-idler correlation coefficients.

Overall, this is a beautiful and fairly complex/challenging experiment and the obtained results – from the first part at least – are definitely nice. However, I have a rather serious concern about the second part. In line 126, the authors claim, "Two photons in the telecom regime will be regenerated as two QM-compatible photons after passing through the PPLN waveguide without losing their original quantumness". Such a lossless quantum frequency conversion[1] (QFC) however happens only if a certain condition – depending on the interaction length of the nonlinear medium and the pump power along some other factors – is met. This is indeed also discussed in the Supplementary material: see the equation after line 101. From Supplementary Fig. 4a, I would infer that QFC happens at a SFG pump power well above 200 mW, whereas, the main results (presented in Fig. 4 in the manuscript) seem to be at a much lower power of 50 mW. In light of that, I am not entirely sure what to make of claims such as "...clearly showed that the low SNR in the signal from the two photons in the telecom range can be increased by passing it through the wavelength converter." in line 140. Could this be an instance of the (severe) limitations of the photon counting approach[2]? Or is this some other effect? I think the authors should clarify this further.

As far as I understand, the reason that they do not present the final results at a pump power requisite for QFC is most likely because of the very high noise-photon creation rate; see Fig. 6c in Supplementary. But perhaps the quantumness of the wavelength-converted photon pair is not so critical for the actual QM operation anyway? It would be great if the authors could elaborate this.

Apart from that, I have some other minor concerns.

1. Overall, this work is hard to read (Text in the Supplementary material is especially quite hard to understand). Apart from grammatical errors, strange phrases and sentence structures, the authors also use new technical terms such as *frequency mode-locked multimode* that I had never heard of before (and could also not find with a web search). Perhaps they could try to find a native speaker to improve the overall level of English? And try to use well-established technical terms, wherever possible?

2. Is the transmission fiber length 10 km or 20 km? For instance, they write 20 km in the abstract, while in Fig. 1 it is 10 km. Maybe when they say 20 km, they mean the *collective* distance traveled by the photon pairs? In any case, would be nice to clarify this a bit.

3. The name 'single-photon counting module' (SPCM) is used in the same sense as a single photon detector, which is technically incorrect and leads to confusion. I assumed they used a Si based avalanche photodiode for photon detection at visible wavelengths and superconducting detector at telecom wavelengths; and obviously, both these detectors were interfaced to some kind of counting apparatus. It would be good to differentiate this properly throughout the paper.

4. Line 96: "When dispersion less than the time interval is occurred, the width of peak after long-fiber transmission reflects the distribution of frequency". This is somewhat arcane, are they alluding to dispersion in the 10 km channel?

[1] Kumar, P. (1990). Quantum frequency conversion. Optics Letters, 15(24), 1476

[2] Page 302 of Lvovsky, A. I., & Raymer, M. G. (2009). Continuous-variable optical quantum state tomography. Reviews of Modern Physics, 81, 1–33.

We thank all the three reviewers for careful reading of our manuscript, and we are greatly helped by their thoughtful comments. After rethinking and improving the manuscript based on the comments, we believe that our revised manuscript satisfies the requirements for the publication in *Communications Physics*. Following is our reply to the reviewers.

We made this document in the following manner:

The comments from reviewers: **Black bold**

Our reply: Blue

The previous manuscript: indicated by [old] (indented)

Tea revised manuscript: indicated by [new] (indented)

Reply to the comments of reviewer #1 :

We are very grateful to reviewer #1 for the appropriate and constructive comments. We have taken all the comments into account in the revised version of our paper. Here we describe our response to each of them.

The manuscript presents a two-photon comb source for quantum communications complemented with a wavelength converter.
The subject is interesting and timely, but we regret to say we found the presentation confusing at many points. More in detail:

- There is some ambiguity in the distribution distance. We understand that each photon travels a distance of 10km, thus covering an overall separation by 20km, and allowing quantum communications over that length. However, this should not be phrased as a 20km transmission, as this describes a single photon travelling that far.

We thank reviewer #1 for pointing out this issue. We noticed that 20-km distribution and 10-km transmission were mixed in abstract and summary part. To make the description clear, we have now replaced “transmission” with “distribution” in the following parts.

[old] line 12: “transmission”, line 229: “transmission”

[new] line 13: “distribution”, line 268: “distribution”

- The introduction starts from the recent demonstration of quantum supremacy by Google. However much we may like their result, Shor’s algorithm is still far from posing a believable threat to our crypto-systems. We think that a more solid argument can be drawn, but, admittedly, this is just our personal taste.

We appreciate the opinion for improving the introduction. According to this comment, we changed sentences aiming at not only preparing for decryption but also developing quantum internet.

> [old] line 19: Recently, a quantum processor with 53 qubits was demonstrated by Google [1], and they insisted that quantum supremacy is being verified experimentally. If larger numbers of qubits become controllable in the future, Rivest–Shamir–Adleman (RSA) cryptosystems could be deciphered using quantum computers [2]. Therefore, unconditionally secure communication needs to be developed. Quantum communication can achieve unconditional security [3], [4], and use of the existing global optical fiber network is highly desirable.

> [new] line 22: Recently, quantum technologies have been intensively developed globally. For instance, a quantum processor with 53 qubits was demonstrated by Google [1], and they reported that quantum supremacy was verified experimentally. If even larger numbers of qubits become controllable in the future, Rivest–Shamir–Adleman (RSA) cryptosystems could be decrypted using quantum computers [2]. Quantum communication needs to be developed for the purposes of unconditional security [3][4] and of efficient quantum computing by connecting quantum computers through quantum communication network, or quantum internet [5].

- On line 81, there is some claim that the bow-tie configuration is less demanding in terms of maintenance. The authors should also bear in mind that locking to a specific line should be put in place for coupling to a quantum memory. Can the authors explain if this advantage would remain even in these conditions.

We thank reviewer #1 for the thoughtful comment. We have already locked the cavity length in these experiments for aiming at a narrow-line memory frequency. Accordingly, we can stabilize the relative phase of a quantum state to realize higher fidelity, while remaining the advantage of the long bow-tie configuration. In the process of locking, there is no particular difficulty once the cavity is aligned.

Furthermore, compared with the Fabry-Perot type scheme, there is another advantage; because the cavity here has extra mirrors, we can separate the SPDC pump input from the locking laser, and the beam path adjustment is more convenient. Compared to a small cavity, the advantage of the setup here is free of maintenance aiming at a narrow memory absorption line of ~4 MHz in Pr^{3+}:YSO. To explain this, we added the following sentence.

> [new] line 98: Cavity locking was achieved using the Pound–Drever–Hall technique to target the absorption line of QM and stabilize the relative phase of entanglement.

- The discussion about dispersion on lines 96-97 is confusing; please revise.

We are very sorry for the confusing statement. To make the description clear, we modified the corresponding sentences as follows:

> [old] line 96: We observed ~2 ns of broadening per photon, and we concluded that the TPC had

a frequency range of 1–2 THz by assuming a Gaussian spectrum (see Supplementary Information).

[new] line 113: We observed ~2 ns broadening per one channel of TCSPC, and we estimated that TPC had a comb range of approximately 1–2 THz from the fiber dispersion of 15 ps·nm^{-1}·km^{-1} (please see Supplementary Information).

(Supplementary note 3) line 89: T_{FWHM} is the FWHM of one timing peak which is approximated by a Gaussian-shape peak. Although a Voigt function should be used for the strict estimation of the spectrum, we used a Gaussian function to reduce complexity.

- The fidelity of the output state to a maximally entangled one is high, but this decreases when the some control is inserted to manipulate it to a specific target. Can the authors account for this decrease in fidelity? Can it be attributed to dispersion effects of the waveplate?

We thank reviewer #1 for the thoughtful comment. We agree with reviewer #1 and consider that this fidelity decrease is mainly caused by the dispersion effects of the waveplate, namely wavelength dependency of retardance. In the experiments, the fidelity became lower by increasing the number of waveplates. For extra waveplates not involved in tomography, the number of them are 0, 1, 2 for the maximally entangled state in Fig. 2, $|\Phi^{\pm}\rangle$, $|\Psi^{\pm}\rangle$, respectively. It is possible that tomographic result becomes worse because of these effects.

We also think there is another small reason in the process of fidelity calculation. The fidelity to a Bell state includes a slight phase gap from 0 or π: Supposing ρ_{meas} is an obtained state, the fidelity to the maximally entangled state is calculated as $\langle\psi_{near}|\rho_{meas}|\psi_{near}\rangle$, where $|\psi_{near}\rangle$ is the nearest state to ρ_{meas}, while the fidelity to a Bell state is $\langle\psi_{Bell}|\rho_{meas}|\psi_{Bell}\rangle$, where $|\psi_{Bell}\rangle$ is one of the Bell states. The relative phase of $|\psi_{near}\rangle$ is not always 0 or π, while that of $|\psi_{Bell}\rangle$ is always 0 or π.

The revised sentences now read,

[new] line 221: Based on the tomographic results, we consider that the dispersion or wavelength-dependent retardance of the waveplates affects the tomographic reconstruction. To realize desired states, additional waveplates unrelated to the tomography are utilized: one for $|\Phi^{\pm}\rangle$, and two for $|\Psi^{\pm}\rangle$. Although estimating the magnitude of the effect is difficult, the obtained fidelity does appear lower when this is taken into account.

- The polarisation behaviour of the WC is not at all discussed in the main text.

We are sorry for the lack of detailed information. The three polarizations interacting in the WC (pump, signal and idler) are all vertical polarization, that is, type-0 phase matching condition. To explain this situation, we added a sentence below.

[new] line 311: The PPLN for the wavelength converter was of the waveguide type with an area

of 9.9×11 μm and length of 48 mm (NTT electronics) with a type-0 phase matching condition (all three interacting lights are vertically polarized).

- The discussion on Fig. 4 in lines 140-141 is very terse, and only focusses on the enhancements. Likewise, the comment on the single photon conversion is excessively concise.

We thank reviewer #1 for pointing out this issue. We revised the corresponding sentences including the single photon conversion, and also added the detailed explanation in the Discussion section as follows:

[new] line 154: The variation in $g_{s,i}^{(2)}(0)$ clearly showed that the low SNR in the signal from the two photons in the telecom range can be increased by putting the two photons through the wavelength converter (see Discussion for details). When we tried to demonstrate only single-photon WC, which corresponds to a correlation between the telecom photon and the visible-range photon, we did not observe a clear correlation function, despite extensive adjustment of the SPDC/SFG pump power and spectral filters; we observed $g_{s,i}^{(2)}(0) \approx 1$, i.e., almost the noise floor (data not shown).

(discussion part) line 240: This effect can also explain the result of very low SNR in one-photon WC: in the telecom range, there were many noise photons consisting of other pairs generated by next SPDC and, consequently, although the WC had high loss and additional noise, the correlation of the two photons reduced.

- The abstract reports a value of 5 for the g2, higher than 1 by 15 standard deviations (we assume). These quantitative considerations tend not to be reported in the main text, however. Furthermore, error bars should be reported in Fig.4, or a statement about them being smaller than the points should be made.

We appreciate his/her careful reading and suggestion.

The g2 values in the abstract included the data in the supplementary information, so that we deleted the corresponding sentences to avoid confusion from the main text.

We added the statement for error bar in Fig. 4 caption, (it is smaller than the points in figure.) and it reads,

[new] (Fig. 4 caption) line 363: The error bars are smaller than the points (or some can be seen in lower power range; see Supplementary Information) and are therefore omitted.

- On line 157 loss is identified as a reducing factor of g2, however this figure is notoriously resistant with respect to loss. We suspect that the authors are being once again too concise in their treatment; can they spell out how loss affects the g2? Is it a matter of signal-to-noise ratio?

We thank reviewer #1 for the kind attention. Because the noise almost remained constant despite fiber loss while signal decreased, g2 therefore became lower. The loss caused by the 10-km fiber was measured as 62% and when comparing the Fig.4 (a) with (c), the noise floor was not affected by this fiber loss. This is because the noise is mainly produced by the pump laser for the WC. According to the comment by reviewer #1, we modified the corresponding sentences as follows:

[old] line 155: The factors decreasing $g_{s,i}^{(2)}(0)$include: fiber loss depending on length (including connection loss), polarization rotation by birefringence due to polarization-sensitive WC, and pulse spreading by wavelength dispersion over the time-bin size of the time-correlated single-photon counting module.

[new] line 173: The factors decreasing $g_{s,i}^{(2)}(0)$ do not affect the noise floor (but only weaken the coincidence counts) and include fiber losses depending on its length (including a connection loss), polarization rotation induced by birefringence due to the polarization-sensitive WC, and pulse broadening by wavelength dispersion over the time-bin size of the TCSPC.

- Lines 173-175 are confusing; the authors should expand this argument.

We thank reviewer #1 for pointing out this issue. We modified the corresponding sentences and also emphasized that quantum memory does not demand very high finesse thanks to short FSR. Because of this, we were released from labor exhausting alignment in every experiment. The revised sentences now read,

[old] A narrow linewidth is guaranteed by the long cavity length (or short FSR). Hence, the highest finesse is not required for fine tuning for QM as the linewidth will stay narrower than the absorption line of QM if the finesse reduced somewhat. Therefore, TPC is free of frequent cavity alignment, resulting in good compatibility with the troublesome Bell-state generation and QM. At low finesse, it was reported that the intensity of side clusters of the main frequency comb increased [28], but because of spectral separation, side clusters can be removed using a spectral filter or can even help increase the number of modes without filtering.

[new] line 189: A narrow linewidth is guaranteed by the long cavity length (or short FSR), even with low finesse. Therefore, the highest finesse in the current setup is not required because the linewidth will remain narrower than the absorption line of QM, even if the finesse reduces slightly due to either a shift in alignment or something else; in our case, the photon with 1-MHz linewidth has a redundancy and will couple to 4.6-MHz window of Pr^{3+}:YSO QM even finesse becomes less than a half of the present value. Accordingly, TPC is free of frequent cavity alignment, resulting in good compatibility with the troublesome Bell-state generation and connection to QM. At smaller finesse, it has been reported that the intensity of side (or additional) clusters of the main frequency comb increases [28]; however, because of their spectral separation, the side clusters can be removed using a spectral filter or can be used to increase the number of modes without filtering.

- When a short comparison with the bandwidth of a Pr quantum memory is carried out, it is not understood whether the source is compatible with existing memories. In detail: 1. is the overall bandwidth sufficiently compact for an AFC protocol to address it? 2. is the separation between the teeth compatible with that of an AFC? 3. if so, what would the rephasing time be? If not, what more is needed?

The short answers, in our consideration, are

1. Yes, the overall bandwidth of two photon is almost same order as that of an atomic frequency comb (AFC).
2. No, the teeth of two-photon spectrum are corresponding to the AFC position, not compatible with teeth of an AFC. (please see Fig.1)
3. Further requirement is wider-range wavelength converter, for instance.

Details are below:

Let us explain the outline of AFC quantum memory scheme using Pr^{3+}:YSO first of all. Pr^{3+} is the dopant of YSO crystal and the concentration corresponds to replacement percentage of Y atom with Pr^{3+} ion. Then individual Pr^{3+} ion possesses specific absorption line around ~ 494.7 THz or 606 nm because of surrounding environment or crystal field. As a whole ensemble, this crystal looks having wide-broadened absorption line, as named, inhomogeneous broadening (see Fig.2 right). Inhomogeneous broadening has a width of the order of from GHz to hundred GHz depending on dopant concentration. AFC structure is tailored into inhomogeneous broadening within a range of < 4 MHz, the width of < 100 kHz, and the comb separation of ~ 100 kHz (this separation is directly related to rephasing time or photon echo time). This comb structure can be established thanks to hyperfine structure (see Fig.2 left) and the long lifetime of that. Because AFC structure is some-order narrower than inhomogeneous width, we can tailor many AFC structure with separation of ~ 120 MHz, which is corresponding to the FSR of two-photon spectrum (Fig.1). Our wavelength converter has bandwidth of 25 GHz and it just fits to inhomogeneous width of 1-100 GHz. On the other hand, our (telecom) TPC has ~1-THz overall bandwidth and if it becomes necessary to use more than 25 GHz for more frequency multiplexing, wider bandwidth wavelength converter is required.

We bothered whether all of the above information complicate the article or not, and decided to add the following sentence:

> [new] line 214: In the case of Pr^{3+}:YSO, the achievable atomic-frequency-comb range is less than 4.6 MHz, and the inhomogeneous bandwidth is on the order of several to tens of GHz (when the concentration of the dopant Pr^{3+} is 1%); however, we could not measure the largest bandwidth accurately. Such a wide inhomogeneous broadening can show good connectivity with the overall comb range of WC-TPC limited by the bandwidth of the wavelength converter. Further, assuming

that the profile of a photon linewidth of ~1 MHz is pure Lorentzian and that one of the absorption range of 4.6 MHz is squarish, the coupling efficiency is greater than 90%.

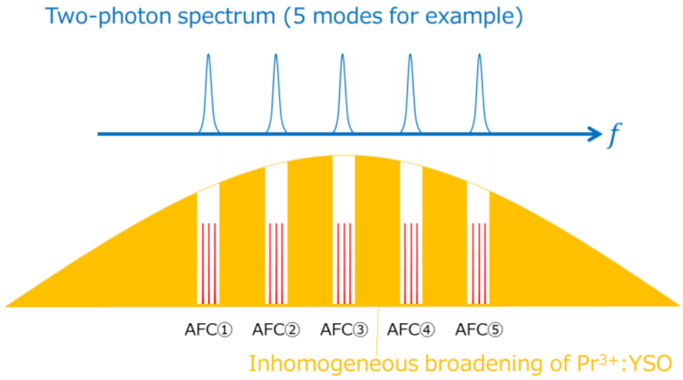

Fig 1. AFC structure for frequency multimode.

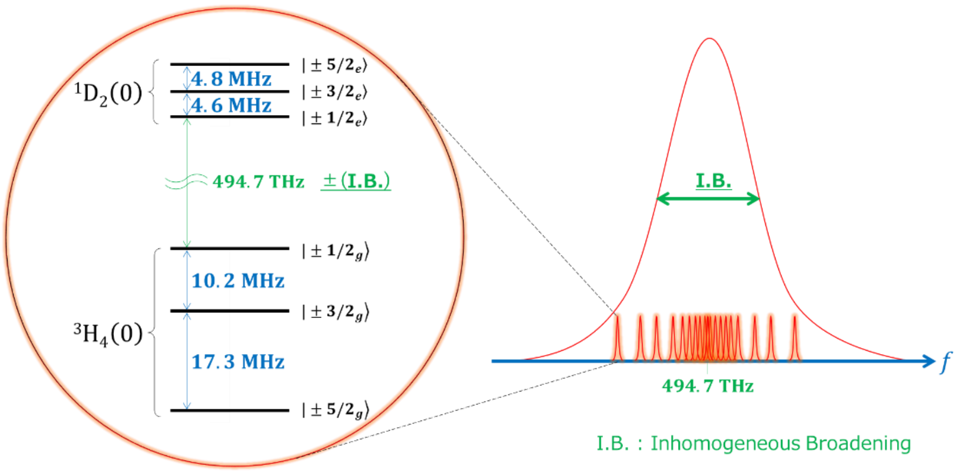

Fig 2. Energy diagram (left) and inhomogeneous broadening (right) of Pr^{3+}:YSO.

- The discussion on the correlation filter perhaps needs some revision and extensions.

We thank reviewer #1 for the thoughtful suggestion. According to this, we revised the corresponding paragraph and added the aspect of frequency/time domain, it now reads,

[old] line 196: The increase in $g^{(2)}_{s,i}(0)$ after WC was attributed to the different sources of noise before and after WC, and also indicates a post-selective effect in which an inconsecutive photon of a pair can be picked up by a “correlation filter” so to speak. The degenerate-SPDC has a gain curve whose center of degenerate point is the highest SNR per unit frequency. In this case, the wavelength converted range is positioned around degenerate point and it is possible to pick up

the highly correlated two photons. In terms of noise, the main sources of noise before WC are the generation of new pairs of photons within the coherence time of another pair that has not yet completed its round trip in the cavity.

[new] line 226: The increase in $g_{s,i}^{(2)}(0)$ after the WC was attributed to the different sources of noise before and after the WC, and it also indicates a post-selective effect in which an inconsecutive photon of a pair can be chosen by a so-called "correlation filter." In the process of WC, because SFG acts only as a spectral shift from telecom to visible wavelengths, it cannot exceed the original SNR with only shifting. Let us consider this effect from two perspectives: frequency and time domains. In the frequency domain, the degenerate-SPDC has a gain curve whose center of degenerate point shows the highest SNR per unit frequency. In this case, the wavelength converted range is positioned around the degenerate point, and it is possible to pick up the highly correlated two photons. In the time domain, in terms of noise, the main sources of noise before WC are the generation of new pairs of photons within the coherence time of another pair that has not yet completed its round trip in the cavity.

We do not feel able to provide a recommendation before these clarifications are take into account.

We appreciate very valuable comments and believe that this article has been improved dramatically.

Reply to the comments of reviewer #2 :

We are very grateful to reviewer #2 for the appropriate and constructive comments. We have taken all the comments into account in the revised version of our paper. Here we describe our response to each of them.

The manuscript presents an entanglement source targeted at fiber-based quantum internet. This work, while timely, requires substantial revision. Publication in current form is premature.

Given that the introductory text/context is all about converting entangled photons to the correct wavelength for quantum memories, one would expect that there would be some measure of the fidelity of these converted photons, but instead, all they present is Fig 4c where the comb sort of looks like the comb before conversion. They should have done tomography on the converted photons, not just the photons from the two-photon comb.

We thank reviewer #2 for the kind suggestion.

As he/she pointed out, we have not done tomography on the converted photons due to the following reasons:

① Aiming at long-distance quantum communication, polarization states could be converted to time-bin states. We are now developing this conversion system and still have some issues such as a visibility of interferometer or an accuracy of tomographic measurement. Our next step is to realize precise treatment of time-bin states, and the loss of this system will be added to the WC-TPC result of the main text. It enables us to estimate the whole quantum-communication system performance.

② Although tomography after both long-fiber transmission and WC is required for quantum communication, there have been no demonstration about wavelength-converted two-photon correlation in the previous literature so far. This is due to the difficulty of constructing a working system including photon source, long fiber and WC. Development of a telecom photon source involved is so hard that there are very few reports even about telecom bands. To obtain detectable signal after WC, system improvements that assume polarization rotation in a fiber are required including WC efficiency and photon detection rate. We have succeeded in constructing the system which is a crucial step into quantum repeater and therefore can observe comb structure of the converted photon after the long fiber.

③ As a latter part of discussion, we need to develop a polarization-insensitive wavelength converter, which requires a Sagnac-type interferometer, almost the same coupling rate both for horizontal and vertical polarization, and of course high conversion efficiency with low noise rate for fair tomography. This technology is still very challenging because we have to convert both two photons. Reference [Ikuta, R. *et al*. Nat. Commun. **9**, 1997 (2018)] demonstrated, indeed, wavelength converted state tomography. But this experiment was for only one-photon wavelength conversion and it took 83 hours for the state tomography. (As a result, they indicated that spin–photon entanglement is retained after one-photon wavelength conversion.)

Although this reported time includes loading time of atoms to a magneto-optical trap, we assume that it will take the same or more order of time when two photons are converted and we will obtain the same result with them: two-photon entanglement is retained after wavelength conversion.

Based on the above-mentioned discussion, we significantly modified the introduction (mainly, latter part) to show both what have not been done in the previous researches about quantum communication and the points of our research. Also, we added the sentences below into the last part of Discussion.

[new]

- line 263: Our single-polarization type WC can be directly applied to real quantum communication by using time-bin states because of the advantages of relative simplicity and higher efficiency: the polarization of time-bin photons will become unpolarized along with long-fiber transmission, and the condition will become the same as in this experiment, except for the quantum state.
- line 275: Our future plan is to convert polarization basis to time-bin basis and to add the loss of basis-conversion system to the WC-TPC result, which will enable us to estimate the whole quantum-communication system performance.

Let's say the intro were re-written to make it clearer that this is a very technical paper in experimental optics and not necessarily something very useful to the quantum information community. Then their major claims are that 1) they developed a way of creating a two-photon comb that is more robust to external conditions, setup, etc. and 2) they couple it to a wavelength converter.

We thank the reviewer #2 for the advice to improve the introduction. According to the advice, we revised the introduction to insist the technicality and added the sentence below to make it clearer.

[new] (last part of introduction) line 71: This very technical development can be applied not only to quantum information science but also to experimental optics.

Regarding 1), they claim their method gives linewidths an order of magnitude shorter than the versatile entanglement source for QI applications they cite in Ref. 19, but then if you look at the table in SI Table 1 (in note 6) comparing their results to other photon sources that were not explicitly developed as versatile entanglement sources, it seems that their line widths are not substantially smaller, so not sure if this work is significant for experimental optics community either.

We apologize that SI note 6 does not include all of the many related references. And this supplementary note aims only to show the comparison between our result and the top-class data. We would like to emphasize that there are only a few reports about telecom narrow-linewidth photon source due to the difficulty of constructing a low loss system at the wavelength regime. Moreover, this is the first demonstration of telecom narrow-linewidth entanglement source. To evaluate novelty, let us compare with narrow-linewidth photon sources in SI Table 1.

[8] J. Fekete (2013)

Their telecom wavelength of 1436 nm has a more loss in fiber than our wavelength of 1514 nm. They cannot go under 1 MHz and there is no entanglement.

[9] Z. -Y. Zhou (2014)

Their wavelength of 1560 nm is almost centered in telecom band but the linewidth of 8 MHz is wider than our demonstration. And there is no entanglement.

[10] K. Liao (2014)

In this table, only their work does not use optical cavity; Four wave mixing of trapped ^{85}Rb. They cannot create telecom band with reasonable rate.

[11] L. Tian (2016)

They use PPLN coated of mirror and cannot demonstrate sub MHz despite very high finesse. The wavelength is not telecom.

[12] M. Rambach (2016)

They use “flip trick” technique which effectively doubles the cavity length and can realize 0.67 MHz width. But this technique cannot be applied to entanglement source. Further, the wavelength is not telecom.

[13] J. Wang (2018)

Their wavelength has neither telecom nor short linewidth.

1st Author (Year)	Characteristics	λ (nm)	$\Delta\nu$ (MHz)	Fidelity (%)	Mode
[8] J. Fekete (2013)	One wavelength aims at direct absorption of Pr^{3+}:YSO	1436 606	1.7, 2.9	-	1
[9] Z. -Y. Zhou (2014)	Telecom	1560	8	-	Multi
[10] K. Liao (2014)	Four wave mixing using ^{85}Rb with magneto-optical trap	780 795	0.8	95.2	1
[11] L. Tian (2016)	Single Mode (Degenerate SPDC)	795	15	95.2±0.8	1
[12] M. Rambach (2016)	Flip-trick technique to realize very narrow linewidth	795	0.67*	-	Multi

[13] J. Wang (2018)	Alignment using optical wedges	935 880	9, 9.5	89.6 (~80)	1 (Multi)
This work	**Telecom narrow linewidth and entanglement with frequency multimode**	**1514**	**0.95**	**96.1**	**Multi**

From the above comparison, our result shows an excellent novelty in cavity-enhanced photon sources. In the other aspect, our technique is specialized for practical use: high finesse is not needed to realize a quantum-memory compatibility.

If necessary, please see "Table 2" in the review below to compare the linewidths of many photon sources. We will find that 1 MHz is one of the smallest values.

[[20] *Slattery, O., Ma, L., Zong, K. and Tang, X. Background and Review of Cavity-Enhanced Spontaneous Parametric Down-Conversion. Journal of Research of the National Institute of Standards and Technology 124, 124019 (2019).*]

According to his/her kind comment, we revised the sentence below.

> [new] (Supplementary note 6) line 184: We would like to emphasize that there is no demonstration of telecom narrow-linewidth entanglement source.

Regarding 2), the authors don't provide much context on how hard it is to convert photons or whether they did any innovation to increase efficiency/fidelity (not even in SI notes 4 where they talk about experimental details of wavelength conversion), so it's difficult for me to say whether this is a big deal.

We thank reviewer #2 for pointing out this issue. We divide the corresponding explanation into I.) conversion part and II.) fidelity part.

I. Let us explain the innovation to convert photons and increase efficiency.

Even only wavelength conversion of both two photons is reported for the first time, to our knowledge. Furthermore, only our system can easily achieve 10-km fiber transmission because loss of visible wavelength (the major one in the previous experiments so far) has approximately 20-30 dB. The failure reason was the low count rate or bad signal to noise ratio after conversion. Then we innovated so as to get high rate, for example, [A] by increasing telecom two-photon counts without losing fidelity, narrow linewidth and frequency multimode and [B] by improving total conversion efficiency with reduced noise and loss.

For [A], our main approach was to develop an SPDC source consisting of a long cavity. After all, count rate of photons strongly depends on the optical path in a nonlinear crystal and cavity. In the case of usual size of a cavity, it is difficult to realize both high rate and narrow linewidth because of complexity of axis alignment. Based on these backgrounds, long-cavity technique can achieve both high rate and narrow linewidth for quantum memory, furthermore, more multimode.

For [B], we examined a lot of experimental conditions like changes in the combination of lenses, careful observation of waveguides by using camera, and optimization of noise filters. We found the best setup and realized the wavelength conversion of both two-photons.

II. In this part, we refer to improving fidelity.

The most important factor to determine the fidelity of an entangled state is the relative phase θ in $|HH\rangle + e^{i\theta}|VV\rangle$. When there are some effects depending on polarization inside the cavity like birefringence, θ becomes θ_i, which means that the phase changes based on the number of cavity-round trips. Then the cavity-output state will be a mixed state, and the fidelity will decrease. Our main effort is to reduce this effect as much as possible. For example, a cavity mirror has a little polarization-dependent phase shift at some reflection angle, so that we construct a bow-tie cavity with a small reflection angle (approximately 2-3 degrees) to decrease senkrecht and parallel polarization dependence. Also, a PPLN crystal has birefringence, and we adjust the temperature of two PPLNs with a few milli-kelvin stabilities to reduce total birefringence. Furthermore, the mirrors after the cavity for coupling to the PMF are all removed, and two photons straightly couple into the PMF. These are only a few examples to realize high fidelity: dephasing mechanism is very complex in reality.

We added the sentences as below:

(about the fidelity)

[new] (introduction) line 56: It is still challenging to obtain all of high photon-count rate, narrow linewidth and high entanglement fidelity in the telecom regime.

(Supplementary note 1) line 23: Our main effort aiming at high fidelity is to reduce decoherence effect as much as possible. We construct a bow-tie cavity with a small reflection angle (approximately 2-3 degrees) to decrease senkrecht and parallel polarization dependence. This angle is limited by finite width of PPLN crystal; the minimum magnitude is proportional to arctangent of the half of the short arm of bow-tie configuration and the width of PPLN. Long cavity can make reflection angle smaller than ordinal-size cavity, resulting in decoherence. Also, a PPLN crystal has birefringence, and we adjust the temperature of two PPLNs with a few milli-kelvin stabilities to reduce total birefringence. Furthermore, the mirrors after the cavity for coupling to the PMF are all removed, and two photons straightly couple into the PMF. These are only a few examples to realize high fidelity: dephasing mechanism is very complex in reality.

(about the efficiency of WC)

(introduction) line 63: WC still has difficulty about conversion efficiency and noise photon to realize photonic interface.

(Method) line 318: To obtain higher efficiency, we examined a lot of experimental conditions like changes in the combination of lenses, careful observation of waveguides by using camera, and optimization of noise filters. We found the best setup and realized the wavelength conversion of both two-photons.

Paper writing clarity/style could be substantially improved. While this is not a significant factor in deciding whether a paper should move forward, I would like to suggest changes to the presentation; the current version was not an easy read.

We thank reviewer #2 for pointing out this issue. We handed in both main text and supplementary to English proofreading to revise grammars, phrases and terms. Furthermore, the introduction was changed to clarify the point of our research. We hope this manuscript now becomes easier to read.

Reply to the comments of reviewer #3 :

We are very grateful to reviewer #3 for the appropriate and constructive comments. We have taken all the comments into account in the revised version of our paper. Here we describe our response to each of them.

The manuscript I have been asked to review presents a spontaneous parametric down conversion based versatile entanglement source (VES) that operates at telecom wavelengths and is capable of generating photon pairs in a narrow linewidth. While telecom regime operation is attractive from a fiber-optic communication point of view, narrow linewidths are a requirement for being able to interface the photons with the narrow atomic transitions of quantum memories (QMs). The first part of the experiment describes the operation and characterization of the VES, e.g., via Hanbury-Brown-Twiss (HBT) measurements for two-photon correlations and Bell state tomography for entanglement fidelity, etc. Effects such as dispersion arising from transmitting the entangled pairs through an optical fiber channel are also analysed. In the second part, the authors present a wavelength conversion (WC) stage: as typical QMs operate in the visible regime, the generated photon pairs need to have their wavelength converted to be compatible to the QM. This WC is realized through the process of sum frequency generation (SFG) after which, they compare the results obtained in the telecom regime with the visible regime intensively, e.g., through normalized second-order signal-idler correlation coefficients. Overall, this is a beautiful and fairly complex/challenging experiment and the obtained results – from the first part at least – are definitely nice.

We are very pleasure to hear your positive comments.

However, I have a rather serious concern about the second part. In line 126, the authors claim, "Two photons in the telecom regime will be regenerated as two QM-compatible photons after passing through the PPLN waveguide without losing their original quantumness". Such a lossless quantum frequency conversion[1] (QFC) however happens only if a certain condition – depending on the interaction length of the nonlinear medium and the pump power along some other factors – is met. This is indeed also discussed in the Supplementary material: see the equation after line 101. From Supplementary Fig. 4a, I would infer that QFC happens at a SFG pump power well above 200 mW, whereas, the main results (presented in Fig. 4 in the manuscript) seem to be at a much lower power of 50 mW.

In light of that, I am not entirely sure what to make of claims such as "...clearly showed that the low SNR in the signal from the two photons in the telecom range can be increased by passing it

through the wavelength converter." in line 140. Could this be an instance of the (severe) limitations of the photon counting approach[2]? Or is this some other effect? I think the authors should clarify this further. As far as I understand, the reason that they do not present the final results at a pump power requisite for QFC is most likely because of the very high noise-photon creation rate; see Fig. 6c in Supplementary. But perhaps the quantumness of the wavelength-converted photon pair is not so critical for the actual QM operation anyway? It would be great if the authors could elaborate this.

We appreciate his/her kind focus and suggestion.

Yes, this is due to high noise rate. In the aspect of only efficiency, as you said, more than 200-mW pump power is appropriate for unity internal efficiency. But the increment of efficiency (that is, differential of supplementary Fig. 4 (a)) becomes less with increasing pump power. On the other hand, the noise becomes quadratically stronger, and there must be an optimal power for this system. To estimate the optimal value, we adopted the quotient of efficiency divided by noise. This function shows that approximately 10-mW pump power is suitable despite low conversion efficiency. Because it takes relatively longer time to obtain two-photon coincidences, we used 50-mW pump power. If more power are used, there was a concern that photon counting system might stop due to saturation within measuring time. Not only that, as in Ref. [[2]*Page 302 of Lvovsky, A. I., & Raymer, M. G. Continuous-variable optical quantum state tomography. Rev. of Mod. Phys.* ***81****, 1–33. (2009)*], multipartite photons generated by mainly pump laser would cause other bad effects perhaps like false estimation of two-photon coincidences.

We are now developing a special spectral filter as in the discussion part, and SNR improvement after conversion (claims such as "...clearly showed that the low SNR in the signal from the two photons in the telecom range can be increased by passing it through the wavelength converter") will become dramatic when our filter will be realized. After that, we can experimentally and perfectly demonstrate that "Two photons in the telecom regime will be regenerated as two QM-compatible photons after passing through the PPLN waveguide without losing their original quantumness", or unity conversion efficiency. We are very sorry that we should have written these in detail around supplementary Fig. 5, and so on. We thank him/her for great feedback.

> [new] line 141: Ideally, two photons in the telecom regime will be regenerated as two QM-compatible photons after passing through the PPLN waveguide without losing their original quantumness.
>
> (Supplementary note 4) line 139: Although this pump power (7.97 mW) was the best to obtain high SNR, we chose a higher power of 50 mW to obtain shorter experimental time for accurate comparison of $g_{s,i}^{(2)}(0)$ (**Supplementary Figure 2** (b)).

Apart from that, I have some other minor concerns.

1. Overall, this work is hard to read (Text in the Supplementary material is especially quite hard to understand). Apart from grammatical errors, strange phrases and sentence structures, the authors also use new technical terms such as frequency mode-locked multimode that I had never heard of before (and could also not find with a web search). Perhaps they could try to find a native speaker to improve the overall level of English? And try to use well-established technical terms, wherever possible?

We thank reviewer #3 for pointing out this issue.

We handed in both main text and supplementary to proofreading to revise whole English. And we try to adopt guidance sentence for more.

[old]

(abstract part) line 11: with mode-locked multimode

(introduction part) line 51: We utilize the two-photon comb (TPC) technique, which realizes a frequency mode-locked multimode and two correlated photons with a narrow sub-megahertz linewidth.

[new]

(abstract part) line 13: even with frequency multimode

(introduction part) line 66: We utilize the two-photon comb (TPC) technique, which realizes a large number of frequency multimodes and entangled photon pairs with a narrow sub-MHz linewidth even in a telecom band.

(result part) line 98: Cavity locking was achieved using the Pound–Drever–Hall technique to target the absorption line of QM and stabilize the relative phase of entanglement.

2. Is the transmission fiber length 10 km or 20 km? For instance, they write 20 km in the abstract, while in Fig. 1 it is 10 km. Maybe when they say 20 km, they mean the collective distance traveled by the photon pairs? In any case, would be nice to clarify this a bit.

We thank reviewer #3 for the kind attention, and we noticed that 20-km distribution and 10-km transmission were mixed. The correct is that transmission length is 10 km, while the collective length of two photons is 20 km, as he/she told.

[old] line 12: “transmission”, line 229: “transmission”

[new] line 13: “distribution”, line 268: “distribution”

3. The name ‘single-photon counting module’ (SPCM) is used in the same sense as a single photon detector, which is technically incorrect and leads to confusion. I assumed they used a Si based avalanche photodiode for photon detection at visible wavelengths and superconducting detector at telecom wavelengths; and obviously, both these detectors were interfaced to some

kind of counting apparatus. It would be good to differentiate this properly throughout the paper.

We thank reviewer #3 for pointing out this issue. And we recognized these abbreviations made the text hard to read. We chose the words of silicon avalanche photodiode (SiAPD) and superconducting single photon detector (SSPD). Also, we specified time-correlated single-photon counting (TCSPC) module of HydraHarp 400 in the main text.

[old] line 87: single-photon counting module (SPCM)

[new] line 103: silicon avalanche photodiode (SiAPD)

(methods part) line 303: We utilized HydraHarp 400 as a TCSPC module, whose resolution is 32 ps for that shown in **Fig. 2 (a)**, and 16 ps and for the others. The 32 ps resolution was used because it could record longer interval times, which was required for measuring the coincidences with an adequate margin.

4. Line 96: "When dispersion less than the time interval is occurred, the width of peak after long-fiber transmission reflects the distribution of frequency". This is somewhat arcane, are they alluding to dispersion in the 10 km channel?

We thank reviewer #3 for the great suggestion. Yes, we would like to express the dispersion in the 10-km fiber. We changed sentence structure and tried to make it easier to understand.

[new] line 108: In general, the ratio between the time interval and the pulse width of each peak gives the number of frequency modes or comb range: if some fiber dispersion exists, the number of modes can be evaluated because the pulse width becomes substantially larger than the timing jitter and thus the timing jitter becomes ignorable. Then, we performed 10-km fiber transmission and evaluated the width of the frequency range by measuring the dispersion. We observed ~2 ns broadening per one channel of TCSPC, and we estimated that TPC had a comb range of approximately 1–2 THz from the fiber dispersion of 15 ps·nm^{-1}·km^{-1} (please see Supplementary Information).

[1]**Kumar, P. (1990). Quantum frequency conversion. Optics Letters, 15(24), 1476**

[2]**Page 302 of Lvovsky, A. I., & Raymer, M. G. (2009). Continuous-variable optical quantum state tomography. Reviews of Modern Physics, 81, 1–33.**

We appreciate very careful comments and believe that the article has been improved dramatically.

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

The amendments operated in reply to the observations of the referees are, in my opinion, sufficient to grant a recommendation for acceptance.

As a minor comment, I am convinced that the discussion on the compatibility of the upconverted emission and the memory the authors included in my response should find a place in the supplementary material. Indeed, I found it clear and to the point.

Reviewer #2 (Remarks to the Author):

I have read the revised article and supplementary - most of my comments on the previous version have been sufficiently addressed. There are still some residual readability issues, mostly in the SI, but it's much improved on that front.

It is unclear from the manuscript whether this article belongs in this journal (I leave this up to the editor), but I find their results and conclusions to be solid.

Reviewer #3 (Remarks to the Author):

The authors have tried to take into account the concerns I raised. Here are my final comments on the revised manuscript and the rebuttal:

ˆ Ideally, two photons in the telecom regime will be regenerated as two QM-compatible photons after passing through the PPLN waveguide without losing their original quantumness

I would strongly urge the authors to emphasize that in practice, their experiment is far away from preserving the ‘original quantumness’ of the entangled pair during the SFG process. Simply starting the sentence with Ideally, as they have done in the revised draft, is not good enough in my opinion.

They should elaborate in the main paper (and not just supplementary) under what conditions a proper quantum frequency conversion would have been possible and where they actually are.

ˆ To estimate the optimal value, we adopted the quotient of efficiency divided by noise.

If this is a measure that has been used in previous publications, then please cite them. Otherwise, use already known measures such as NEP (see Langrock et al., Optics Letters, 30(13), 1725 (2005)). I believe this would give them a different value of the optimal pump power.

ˆ Finally, I do see better sentence structures / less grammatical mistakes now, nevertheless, I feel there is yet some scope for further improvement.

The authors have tried to take into account the concerns I raised. Here are my final comments on the revised manuscript and the rebuttal:

- Ideally, two photons in the telecom regime will be regenerated as two QM-compatible photons after passing through the PPLN waveguide without losing their original quantumness

  I would strongly urge the authors to emphasize that in practice, their experiment is far away from preserving the 'original quantumness' of the entangled pair during the SFG process. Simply starting the sentence with *Ideally*, as they have done in the revised draft, is not good enough in my opinion. They should elaborate in the main paper (and not just supplementary) under what conditions a proper quantum frequency conversion would have been possible and where they actually are.

- To estimate the optimal value, we adopted the quotient of efficiency divided by noise.

  If this is a measure that has been used in previous publications, then please cite them. Otherwise, use already known measures such as NEP (see Langrock et al., Optics Letters, 30(13), 1725 (2005)). I believe this would give them a different value of the optimal pump power.

- Finally, I do see better sentence structures / less grammatical mistakes now, nevertheless, I feel there is yet some scope for further improvement.

We thank all the three reviewers for careful reading of our manuscript, and we are greatly helped by their thoughtful comments. After rethinking and improving the manuscript based on the comments, we believe that our revised manuscript satisfies the requirements for the publication in *Communications Physics*. Following is our reply to the reviewers.

We made this document in the following manner:

The comments from reviewers: **Black bold**

Our reply: Blue

The previous manuscript: indicated by [old] (indented)

Tea revised manuscript: indicated by [new] (indented)

Reply to the comments of reviewer #1 :

We are very grateful to reviewer #1 for the appropriate and constructive comments. We have taken all the comments into account in the revised version of our paper. Here we describe our response to each of them.

The amendments operated in reply to the observations of the referees are, in my opinion, sufficient to grant a recommendation for acceptance.

As a minor comment, I am convinced that the discussion on the compatibility of the upconverted emission and the memory the authors included in my response should find a place in the supplementary material. Indeed, I found it clear and to the point.

We thank reviewer #1 for very positive recommendation and comment.
According to his/her suggestion, we added the discussion about quantum memory and compatibility as new supplementary note with number of 7. The content is almost same with previous response letter.

Reply to the comments of reviewer #2 :

We are very grateful to reviewer #2 for the appropriate and constructive comments. We have taken all the comments into account in the revised version of our paper. Here we describe our response to each of them.

I have read the revised article and supplementary - most of my comments on the previous version have been sufficiently addressed. There are still some residual readability issues, mostly in the SI, but it's much improved on that front.

It is unclear from the manuscript whether this article belongs in this journal (I leave this up to the editor), but I find their results and conclusions to be solid.

We thank reviewer #2 for very positive and kind comment.

Reply to the comments of reviewer #3 :

We are very grateful to reviewer #3 for the appropriate and constructive comments. We have taken all the comments into account in the revised version of our paper. Here we describe our response to each of them.

The authors have tried to take into account the concerns I raised. Here are my final comments on the revised manuscript and the rebuttal:

'Ideally, two photons in the telecom regime will be regenerated as two QM-compatible photons after passing through the PPLN waveguide without losing their original quantumness'

I would strongly urge the authors to emphasize that in practice, their experiment is far away from preserving the 'original quantumness' of the entangled pair during the SFG process. Simply starting the sentence with Ideally, as they have done in the revised draft, is not good enough in my opinion. They should elaborate in the main paper (and not just supplementary) under what conditions a proper quantum frequency conversion would have been possible and where they actually are.

We are very sorry for our inappropriate expression and insufficient revision. Thanks to his/her advice, we found the word 'original quantumness' very ambiguous and misleading. We erased this word and introduced the information about quantum frequency conversion.

> [old]
> Ideally, two photons in the telecom regime will be regenerated as two QM-compatible photons after passing through the PPLN waveguide without losing their original quantumness.
> [new]
> Our setup aimed at conversion of time-bin state which is suitable for long-distance fiber transmission and would achieve proper quantum frequency conversion because the temporal/linewidth profile of input photons would be preserved

Additionally, the related discussion exists in the paragraph from line 260.

"After developing a special filter, WC can be customized as a polarization-insensitive type [39], which is difficult to achieve with the same efficiency as that of a single-polarization type. Polarization-insensitive WC is beneficial in the case of randomly rotating polarization when the conversion efficiency is more than half that of single-polarization WC. Irrespective of the type, the flying qubits should be in a time-bin state that is resistant to polarization changes along the transmission path [52]. Our single-polarization type WC can be directly applied to real quantum communication by using time-bin states because of the advantages of relative simplicity and higher efficiency: the polarization of time-bin photons will become unpolarized along with long-fiber transmission, and the condition will become the same as in this experiment, except for the quantum state."

'To estimate the optimal value, we adopted the quotient of e□ciency divided by noise.'

If this is a measure that has been used in previous publications, then please cite them. Otherwise, use already known measures such as NEP (see Langrock et al., Optics Letters, 30(13), 1725 (2005)). I believe this would give them a di□erent value of the optimal pump power.

We thank very valuable comment. In supplementary, we added two references [[8] Yu, Y. et al. Entanglement of two quantum memories via fibres over dozens of kilometres. Nature **578**, 240–245 (2020).], [[9]Langrock, C. et al., Highly efficient single-photon detection at communication wavelengths by use of upconversion in reverse-proton-exchanged periodically poled LiNbO3 waveguides. Opt. Lett., **30**, 1725-1727 (2005).] In the former reference, the term 'signal to noise ratio (SNR)' is used as same meaning with us but there are few explanations about calculation. Additionally, we calculated noise equivalent power (NEP) of our system in Figure 1 below. It shows the minimum value is 806 $Hz^{1/2}$ at 46.8 mW pump power.

[new]

(Supplementary note 4)

This calculation is similar to reference [8].

As one of the indices, noise equivalent power (NEP) is often adopted [9], and the optimum pump power of that shows 46.8 mW.

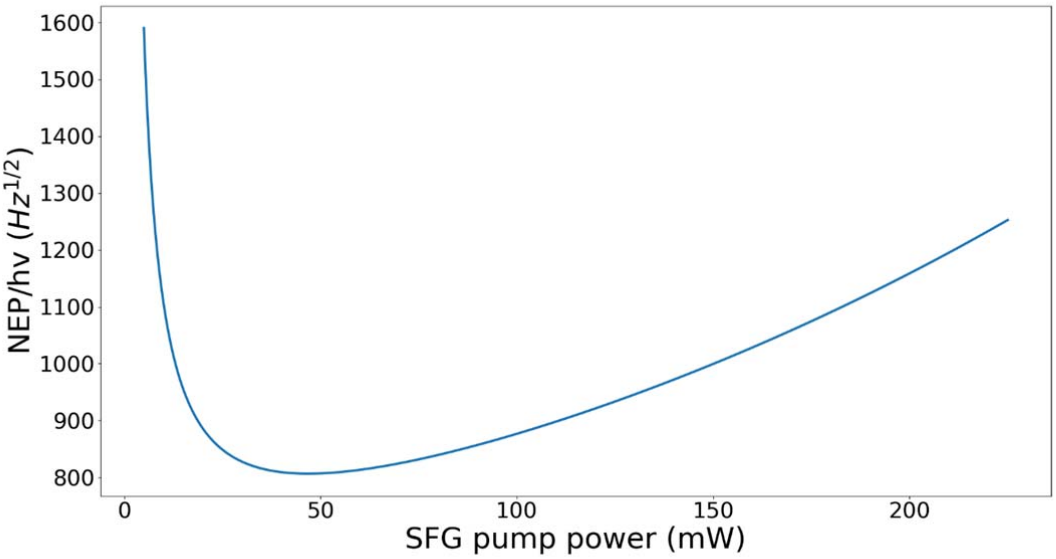

Figure 1. Noise equivalent power (NEP). The minimum value is 806 $Hz^{1/2}$ at 46.8 mW pump power.

- Finally, I do see better sentence structures / less grammatical mistakes now, nevertheless, I feel there is yet some scope for further improvement.

We thank reviewer #3 for very constructive and kind comments.