Dear colleagues,

My co-authors and I thank Daniella Goldfarb as editor and the reviewers for their efforts in re-evaluating our manuscript. We are pleased that the reviewer's comments are generally positive, and we appreciate their useful suggestions for improving the manuscript. In response, we edited the manuscript as detailed below. The reviewer's suggestions/remarks are indented while our replies not indented.

RC1

General comment

> The authors have very significantly improved the manuscript. In particular all my suggestions/remarks (maybe except one, see below) have been addressed properly and clearly.

We thank the reviewer for their kind words about our work and the revised manuscript.

Major comment

> Table 1: I still do not understand how it is possible to obtain an improvement in SNR of a factor 68 between the CW and RS experiments for a sample with T1=T2=100 ns and B1=25 uT (i.e., in almost optimal conditions for the CW experiment), and with a CW modulation frequency of 100 kHz. Since, in these conditions, the precessing magnetization in RS can be maximum a factor of 2 larger than in CW, such additional factor 34 (or more) should come from the noise. But the CW experiment is performed with a modulation at 100 kHz, so I guess the noise spectral density in the effective bandwidth of the RS experiment is probably not much smaller. In order to clarify this point, I would suggest to the authors to show the noise spectral density and, eventually, the signal voltages both "scaled" at the detection point (i.e., at the point Vx) for the two experiments.

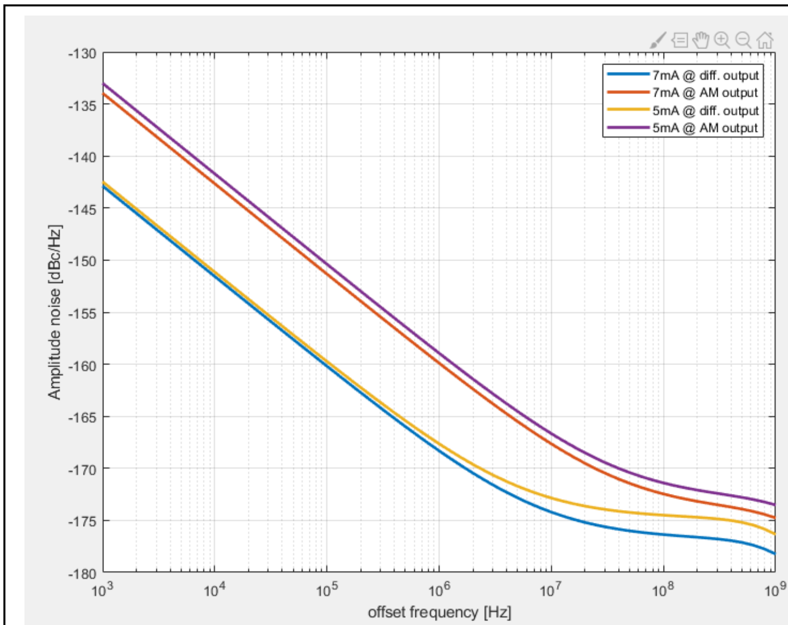

Figure 1 Simulated detector noise

We agree with the reviewer that the SNR improvement can only originate from an improvement in signal amplitude (a factor of 2x due to reduced saturation) and an improved noise floor. As it turns out, the low-frequency noise behavior of the intrinsic AM demodulator that has been used in this paper is relatively poor, as can be seen from Fig. 1 above, which shows the simulated AM noise of the VCO in its output signal compared to the noise inside the demodulated signal that we use in the paper (AM output for CW, *purple*, and RS, *red,* diff output is the noise floor in the differential oscillator output, AM output is the noise in the demodulated output). In the CW measurement, the lock-in amplifier extracts the noise around the modulation frequency of 100 kHz with the very small equivalent noise bandwidth. By contrast, the noise floor in the transient RS data is the integrated noise from the repetition rate of the RS frequency sweep to the full bandwidth of the detection system (200 kHz to 128 MHz). Due to the decreasing noise vs. frequency characteristic of Fig. 1, the RS noise floor is improved compared to a CW experiment. Here, although the integrated RS noise is much larger than the CW noise in a single scan, in the RS experiments, the noise floor can be reduced by significant averaging (~150,000 averages). For a (somewhat) fair noise comparison, one could average the same amount of time that is needed to acquire a single CW point.

We have calculated the anticipated SNR difference between CW and RS, given the data in the above figure, taking into account the different signal processing schemes for CW and RS mentioned above (narrowband filtering in CW, wide bandwidth detection and averaging in RS). These calculations indicate an improvement of a factor of 3.8x in the noise floor of the RS experiment compared to the CW measurement for the same measurement time. This corresponds to an agreement between the measured (not-normalized) SNR improved of 12.5 and the simulated improvement of 6.3 within a factor of two. At this point, we would like to thank the reviewer for their question, which has brought us to a deeper understanding of the different mechanisms that establish the noise floor in CW vs RS EPR experiments with the oscillator-based EPRoC detectors.

> In the answer, the authors mention: "At this point, however, we do experience considerable line broadening in the spectrum, which complicates quantitative analyses". I don't understand this answer. I would not expect a "considerable" line broadening for B1=25 uT with T1=T2=100 ns. If there is "considerable line broadening", I would suggest to discuss this in the manuscript (is it due to oscillator frequency noise due to the unfiltered VCO control voltage ?).

We apologize for the confusion. The sentence mentioned by the reviewer "At this point, however, we do experience considerable line broadening in the spectrum, which complicates quantitative analyses," refers to our statement: "In this view, the spin system is completely saturated at (1/2)*Mo in CW." It does not refer to the experimental/sample parameters of the measurements shown in the manuscript for which virtually no sample saturation is present.

> Line 343: I do not understand the meaning of " ….from an increased signal amplitude due to a later onset of sample saturation in the RS regime". It seems to me that the saturation arrives effectively at a larger B1 in RS but overall the signal amplitude improvement in RS, as discussed also above, should be factor of 2 or so with respect to CW operated at the optimal B1 and for T1=T2.

Thank you for bringing this point to our attention. The sentence “The improvement in SNR arises from an increased signal amplitude due to a later onset of sample saturation in the RS regime and an improved noise floor at higher frequencies of the EPRoC detector." Was included to explain the SNR improvement as partially coming from an increased signal (factor of 2x) and partially from an improved noise floor which is obtained via signal averaging. We have revised this sentence to read as follows,

"The improvement in SNR arises from a combination of an increased signal amplitude due to a later onset of sample saturation (a factor of approximately 2x) in the RS regime and an improved noise floor due to the significant signal averaging employed in the RS measurements."

to avoid this confusion in the final version of the manuscript.

Minor comments:

> Line 26: Reading this phrase, it seems that a larger sweep range would affect the "sensitivity". I would suggest the authors to rephrase it to make it more clear. I guess that the "spin sensitivity" will not be affected, but I agree that a larger sweep range would extend the applicability of the proposed method to a larger class of samples.

A larger sweep width will be beneficial if the B1 magnitude can be increased in the same manner, and a sample with longer relaxation times is used, such that an increase in the scan rate and an improvement of the spin sensitivity may be seen. Without the additional increase in B1; however, the reviewer is correct in saying that no increase in spin sensitivity is expected for the same repetition rate. We have elected not to alter the sentence structure in the manuscript because the maximum B1 of the current EPRoC was not utilized due to this sweep width limitation. The reviewer is also correct in stating that a larger sweep width indeed improves the applicability to a broader class of samples in the absence of spin sensitivity improvements.

> Line 129: Typo: "and an" instead of "andan"

This typo was corrected.

> Line 140: The fact that the amplitude and frequency mode should in principle provides the same sensitivity was discussed in details also in Matheoud et al. (2018) (Matheoud et al., Journal of Magnetic Resonance 294 (2018) 59–70). I would suggest to add this reference here (a reference which is already cited in the manuscript).

The citation has been added to the corresponding sentence.

> Line 348: Typo: "by" instead of "bu".

This typo was corrected.