Return-to-zero shift register DAC low-level distortion

Marcel van de Gevel, Version 3, 8 April 2024

1. Introduction

DIY audio member bohrok2610 found distortion at low signal levels, such as -60 dBFS, 1 kHz. Measuring with PCM2DSD at a DSD256 rate and zooming in, the distortion products turned out not to occur just at exact integer multiples of the signal frequency, but there were actually clusters of several peaks close to integer multiples of the signal frequency. Adding a deliberate offset changed the frequencies of the peaks.

Suspecting it might be related to intermodulating out-of-band tones, the measurement was repeated with two quasi-multibit modulators, namely my PWM8 modulator and HQPlayer's AMSDM7 modulator. There was no low-level distortion to be seen with these modulators.

2. FM hypothesis

You can regard an ordinary single-bit sigma-delta modulator as a kind of time-quantized frequency modulator: the average frequency of the ones coming out of the sigma-delta modulator is $f_s/2$ plus something that is proportional to the input signal. As the modulator's output signal is a discrete-time signal, the locations where ones can occur is defined by the clock, hence that "time-quantized".

At 0 dB DSD, the momentary frequency of the ones is modulated all the way from $f_s/4$ to $3f_s/4$, at -60 dB DSD, from 0.49975 f_s to 0.50025 f_s. That is, the FM carrier frequency is $f_s/2$ and the peak frequency deviation is 0.00025 f_s, which is 1411.2 Hz at DSD128 and 2822.4 Hz at DSD256 (which is the rate bohrok2610 used for most of his measurements). With a 1 kHz audio frequency, the frequency deviation is larger than the audio frequency, which means that there will be several sideband peaks around $f_s/2$.

If those peaks intermodulate with each other, they can produce distortion products in the audio band. If there is a small DC offset somewhere (like you can get in digital when you truncate a 32 bit number to 24 bit, for example) that will shift the carrier and the FM sideband peaks slightly in frequency. Because the modulator is a discrete-time circuit, there will be aliases that shift in the opposite direction. Intermodulation with the aliases might then produce the slightly frequency-shifted audio signal peaks, see section 3. An experiment where a small offset was deliberately subtracted from the digital audio signal shifted the positions of the audio tones (of the tones that were between 2995 Hz and 3005 Hz anyway).

Properly dithered quasi-multibit modulators should not produce an FM-like spectrum, but just shaped noise. When there are no out-of-band tones, they also can't intermodulate to produce distortion products in the audio band. That would explain why bohrok2610 didn't see the distortion with my PWM8 algorithm, nor with HQPlayer's AMSDM7 modulator (not to be confused with ASDM7).

The question is where the intermodulation between the out-of-band tones takes place.

If it is plain old intermodulation between out-of-band signals in the output filter, more passive filtering before the first op-amps should help. Using op-amps with a larger input stage "linear"

range, such as FET op-amps with a large ratio of their slew rate to their gain-bandwidth product, might also help, but as ThorstenL pointed out, the filter capacitors couple any spikes straight into the op-amp output stages, so it could just as well be output stage intermodulation distortion (see Figure 1). Forcing a larger DC current into the op-amp output stages might help in two ways: linearize the output stage and reduce open-loop output impedance, thereby reducing the spikes at the input stages as well.
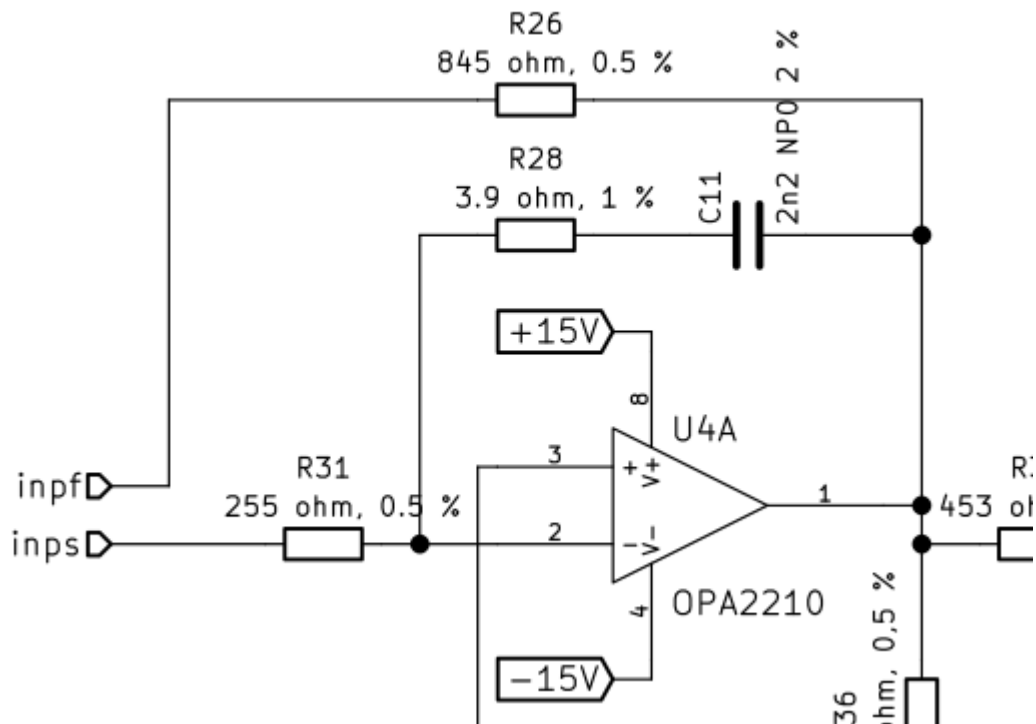

Figure 1: Filter input stage. Any peak that occurs at the input (despite the FIR filter and 8.2 nF *capacitor on the DAC board) couples into the op-amp output through* R_{26} *and* R_{31}-R_{28}-C_{11}*. Forcing an extra DC current into the output of the op-amp (on top of the* 1 mA *that already flows through* R_{26}*) may or may not reduce its open-loop output impedance. If it does, it also reduces the voltage peak at the negative op-amp input.*

As tones around half the sample rate often cause trouble in systems based on single-bit sigma-delta modulation, I already took two measures to reduce their effect: the notch at half the sample rate of the FIRDAC and the 8.2 nF capacitors to ground straight at the DAC outputs. Apparently what bohrok2610 has measured is what is left despite these measures. The levels he measured are small, the peaks are of the order of -130 dBFS.

It could also be intermodulation due to disturbances on the reference supply or the clock, for example crosstalk from the data signal to the reference or clock. However, bohrok2610 did many experiments with different reference supply decoupling and even with a different reference regulator and saw no difference in the low-level distortion, while he did see an effect when trying different output filters.

3. Attempt to figure out what intermodulates with what

3.1. Measurements of mid March 2024

Bohrok2610 sent me some most interesting plots on 16 and 18 March 2024, see Figure 2 ... Figure
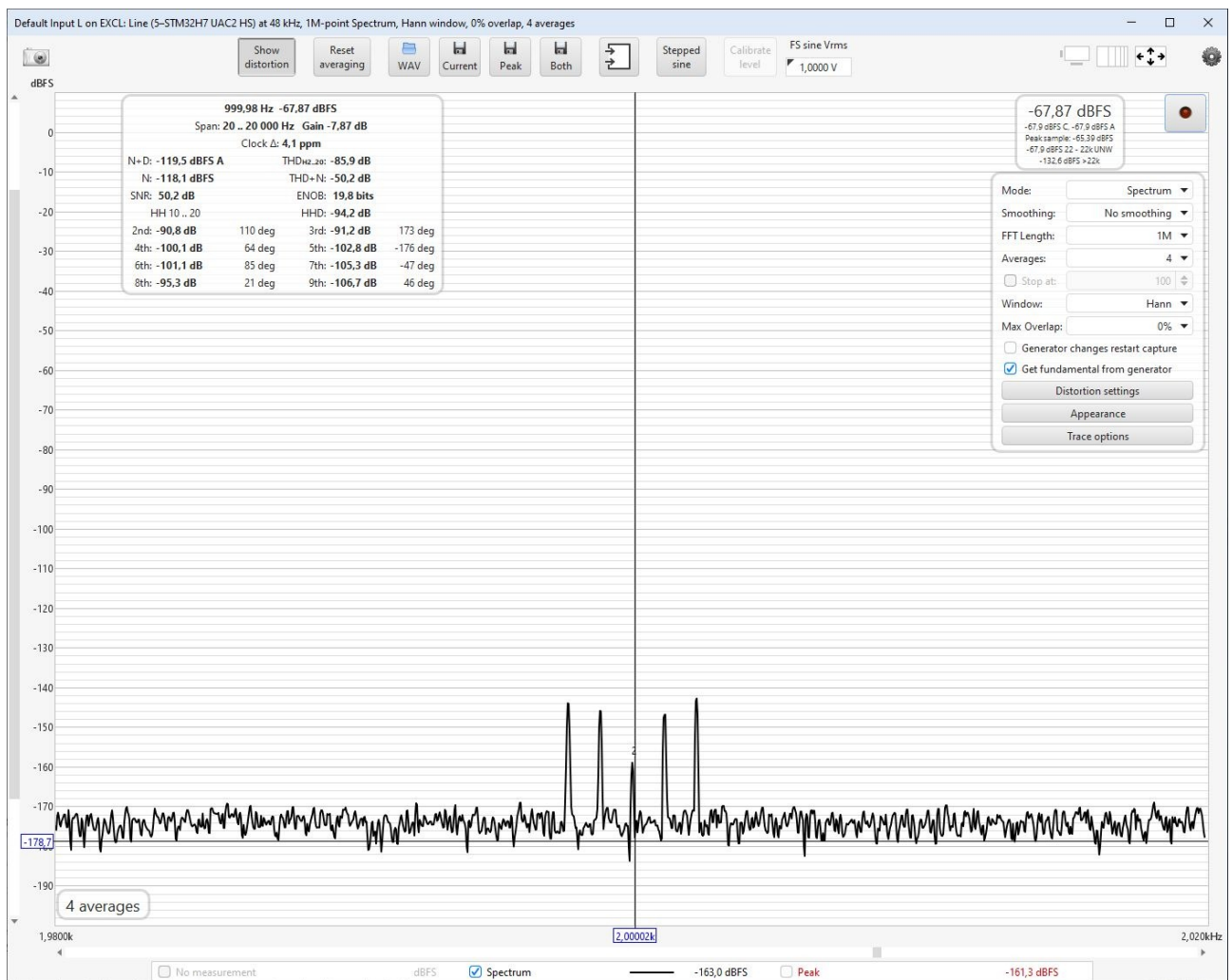
4.

Figure 2: Zoom around 2 kHz *when playing a* 1 kHz, -60 dB *sine wave using PCM2DSD, DAC and ADC not synchronized*
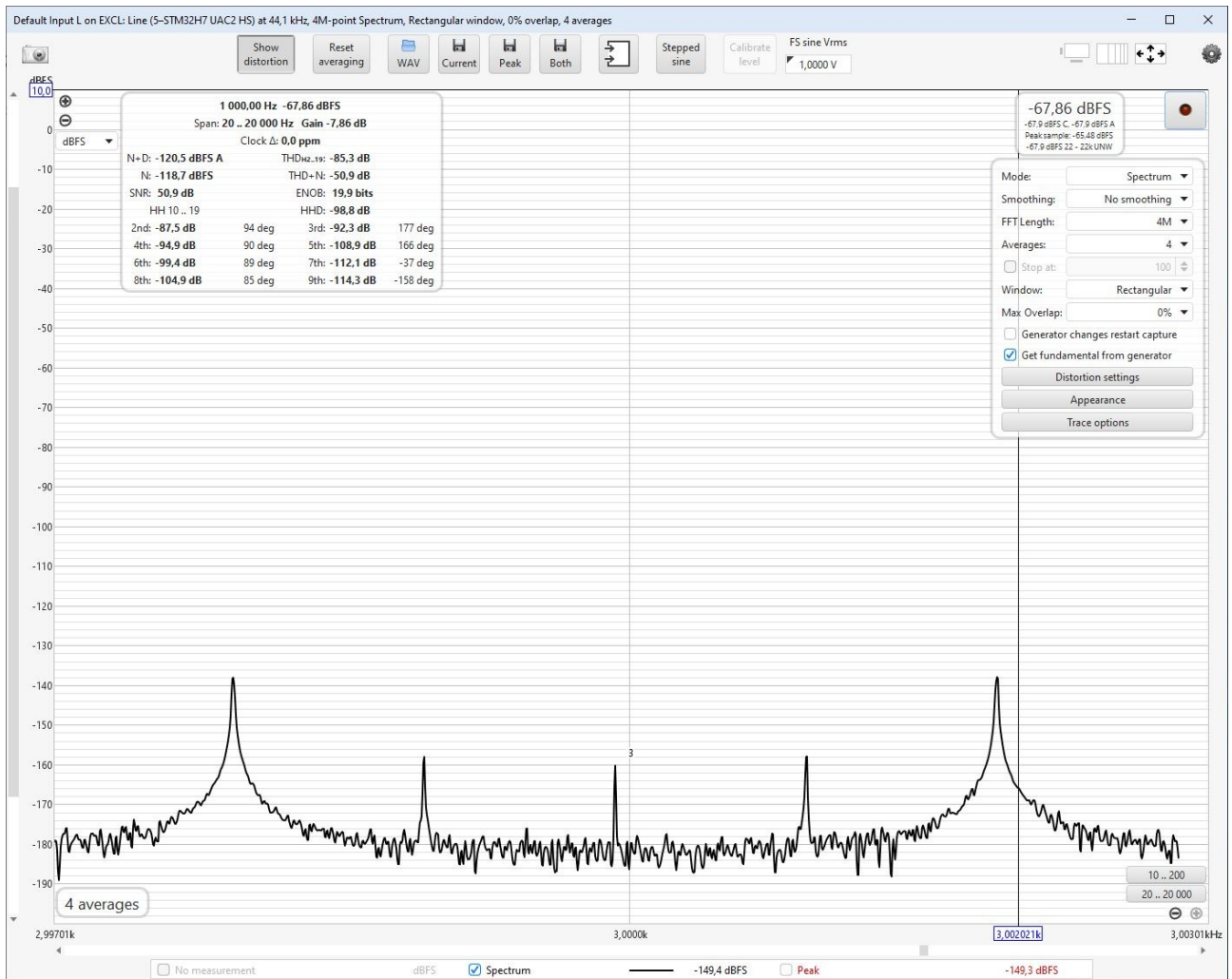
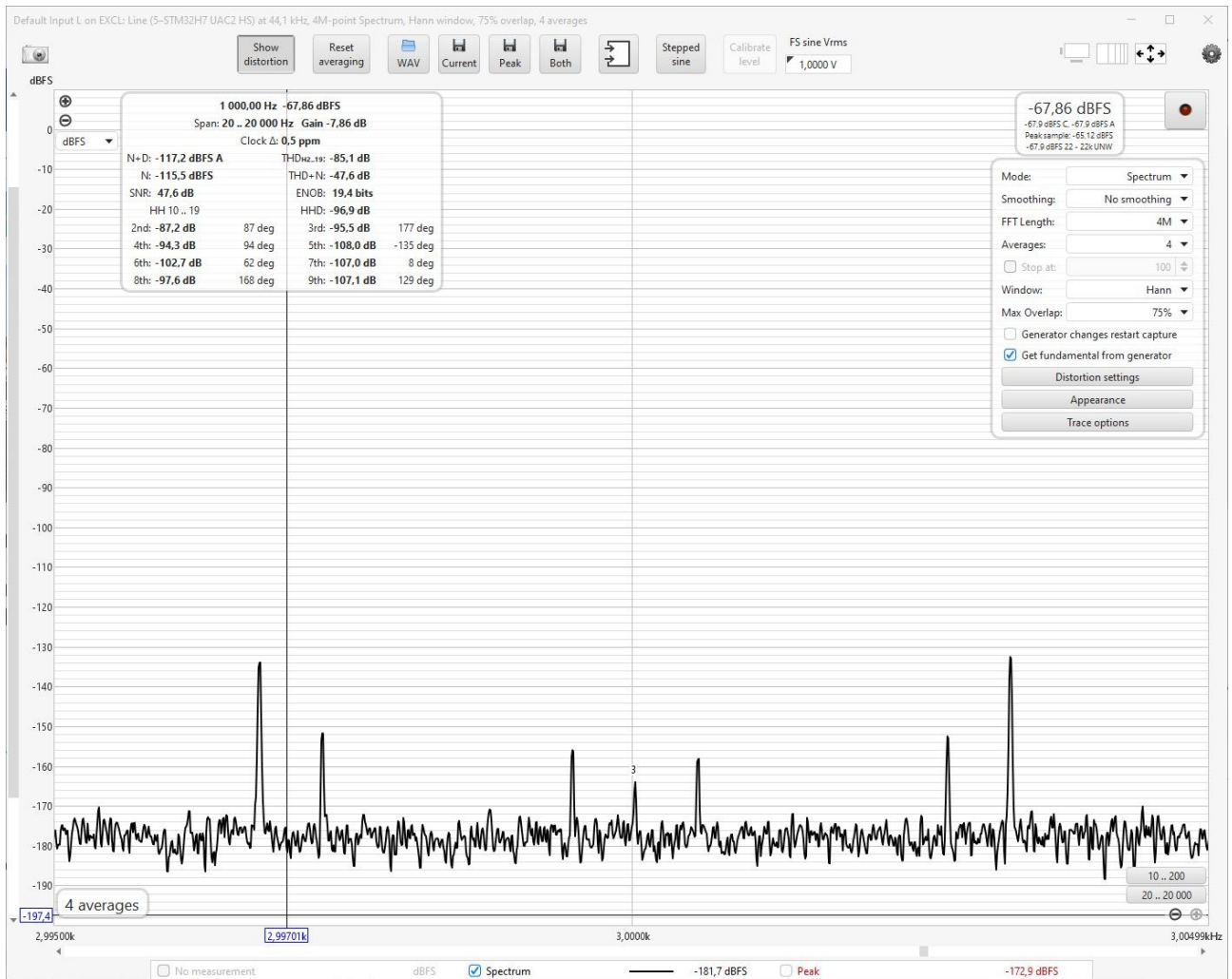*Figure 3: Zoom around* 3 kHz *when playing a* 1 kHz, -60 dB *sine wave using PCM2DSD, DAC and ADC synchronized*

Figure 4: Zoom around 3 kHz *when playing a* 1 kHz, -60 dB *sine wave with an offset of* 512 *subtracted from a* 32 *bit signal, again using PCM2DSD, DAC and ADC synchronized*

I've displayed the plots as large as possible on my computer screen and used a set square with ruler markings to measure the locations of the peaks. The range of the horizontal axes varies between 6 Hz and 40 Hz and the width on my computer screen was 22.6 cm. Assuming I read off the distances within ±0.25 mm, the resulting frequency error must have been within ±44.3 mHz for the coarsest scale and within ±6.64 mHz for the finest scale.

The resulting frequencies of the peaks in Hz were as shown in 1.

	1997.787611	1998.902655	2000	2001.132743	2002.247788	
	2997.939204	2998.929469	2999.925044	3000.912655	3001.90292	
2996.768142	2997.311845	2999.473398	3000.021522	3000.565226	3002.726779	3003.274903
Distances to centre peak:						
	-2.21238938	-1.09734513	0	1.132743363	2.247787611	
	-1.98584071	-0.99557522	0	0.987610619	1.977876106	
-3.25338053	-2.70967699	-0.54812389	0	0.54370354	2.705256637	3.253380531

Table 1: Frequencies determined with a set square and LibreOffice Calc. Top row: around 2 kHz, second row: around 3 kHz, bottom row: around 3 kHz with the deliberate offset of -512.

The first odd thing is that without the deliberate offset, the frequency distances between the peaks measured around 2 kHz are a bit larger than those around 3 kHz: on average about 1.115 Hz and about 0.991 Hz, respectively. When you carefully read the notes at the top of the plots, you see that Figure 2 was actually measured with a 48 kHz rather than a 44.1 kHz PCM sample rate. Presumably the DSD sample rate was proportional to the PCM sample rate, and 48/44.1 times 0.991 Hz is approximately 1.079 Hz. This is a reasonable match, considering the relatively coarse scale of Figure 2.

The case with the offset of -512 is particularly interesting. If the clusters of peaks around the harmonics were caused by a simple offset at the PCM2DSD modulator input, adding a small extra offset should change the frequency distance between the peaks, but one would expect them to remain equidistant and one would expect the number of peaks to stay the same. This doesn't happen, which would indicate that there is something else in the PCM2DSD causing the split peaks that you get without deliberate offset. The follow-up experiment of section 3.5 leads to a different conclusion, though.

3.2. Spectrum if the modulator were an ideal frequency modulator

The modulator is running at a DSD256 rate, that is, 11.2896 MHz sample rate. Without offset, the equivalent FM carrier frequency would be half the sample rate, 5.6448 MHz. At 0 dB DSD, the momentary frequency would vary ±2.8224 MHz around this value, but as the signal is only -60 dB DSD, it actually gets modulated over ±2.8224 kHz.

At a modulating frequency of 1 kHz, ideal frequency modulation with a 2.8224 kHz peak frequency deviation leads to fairly strong sideband peaks up to 4 kHz from the carrier:

0 Hz: -14.2381 dB with respect to the unmodulated carrier
±1 kHz: -7.9106 dB
±2 kHz: -6.3905 dB
±3 kHz: -11.1552 dB
±4 kHz: -19.2212 dB
±5 kHz: -29.5802 dB
±6 kHz: -41.713 dB

This can be calculated using Bessel functions of the first kind, see Wikipedia: https://en.wikipedia.org/wiki/Frequency_modulation LibreOffice Calc and Excel can calculate those. Mind you, this is treating the sigma-delta modulator as an ideal frequency modulator - it will only be approximately true for a real sigma-delta modulator.

According to bohrok2610, his PCM2DSD is scaled such that full-scale PCM is converted to -1 dB ± 0.05 dB DSD, see https://www.diyaudio.com/community/threads/simple-dsd-modulator-for-dsc2.370177/post-7642489
An offset of 512 subtracted from a 32 bit signed binary PCM signal corresponds to an offset of $-512/2^{31}$ times full scale. $-512/2^{31}$ times $10^{-1/20}$ times 2.8224 MHz is -0.5997339839 Hz. The whole FM spectrum will therefore be shifted by -0.5997339839 Hz, centred at 0.5997339839 Hz below half the sample rate, that is. Taking the estimated ±0.05 dB inaccuracy into account, it can actually be anything between -0.5962915552 Hz and -0.603196286 Hz. For simplicity, I will refer to this as about -0.6 Hz.

3.3. Aliases and components around multiples of f_s

As we are really talking about a discrete-time system, there will also be an alias at the same frequency distance on the other side of half the sample rate, so about 0.6 Hz above half the sample rate. There is, of course, also the small desired signal at 1 kHz (which wouldn't be generated by a real frequency modulator), a very small DC offset at 0 Hz and a lot of quantization noise. All of it will repeat around multiples of the sample rate.

That is, looking only at the spectral peaks (not the noise), calling the sample rate f_s, the modulating (audio) frequency f_m and the frequency offset f_{off}, we have:

A very small DC offset and its images at nf_s with integer n. To keep the equations simple, I will use double-sided spectra (saves lots of ± signs), so n can be negative, zero or positive.

The desired signal and its images at $nf_s \pm f_m$ with integer n.

The FM-like peaks and their aliases and images at $(n + \frac{1}{2}) f_s \pm f_{off} + k f_m$ with integer k, the strongest components occurring for $k \in \{-4, -3, -2, -1, 0, 1, 2, 3, 4\}$, at least for the given frequency deviation and modulating frequency.

The FM-like peaks are normally considerably larger than the -60 dB desired signal, but they are suppressed by the notches of the FIRDAC, while the desired signal is not. The offset is very small indeed and can probably neglected.

3.4. Resulting intermodulation products

The resulting intermodulation products end up at sums and differences of integer multiples of the frequencies of the FM-like peaks or their aliases/images and (possibly) of the desired signal or its images. As f_s is very much greater than f_m and f_{off}, the only way to end up at audio frequencies of a few times f_m is to have all f_s and ½ f_s terms cancel.

Only the FM-like peaks and their aliases and images have ½ f_s terms. Their orders have to be such that these terms end up at 0 or at some integer multiple of f_s. The intermodulation product then either ends up at a low frequency, or at a low frequency after it also intermodulates with a suitable image of the desired signal.

That is, the intermodulation product of p of these FM-like peaks ends up at

$$N_1\left(\left(n_1+\frac{1}{2}\right)f_s \pm f_{off} + k_1 f_m\right) + N_2\left(\left(n_2+\frac{1}{2}\right)f_s \pm f_{off} + k_2 f_m\right) + ... + N_p\left(\left(n_p+\frac{1}{2}\right)f_s \pm f_{off} + k_p f_m\right)$$

where the N_i are all integers that can be positive, negative or zero. The f_s terms sum up to

$$N_1\left(n_1+\frac{1}{2}\right)f_s + N_2\left(n_2+\frac{1}{2}\right)f_s + ... + N_p\left(n_p+\frac{1}{2}\right)f_s = \left(N_1 n_1 + N_2 n_2 + ... + N_p n_p + \frac{1}{2}\sum_{i=1}^{p} N_i\right)f_s$$

Therefore, $\sum_{i=1}^{p} N_i$ must be even.

If the f_{off} terms were all positive, the total f_{off}-related part would be

$$N_1 f_{off} + N_2 f_{off} \ldots + N_p f_{off} = \sum_{i=1}^{p} N_i f_{off}$$

which would be an even multiple of f_{off}.

There is actually a ± sign in front of the f_{off} terms, but the sum remains an even multiple of f_{off} anyway: when you change any of the terms from $+f_{off}$ to $-f_{off}$, you subtract an even multiple of f_{off}.

The offset-related peaks around the harmonics therefore have to be spaced at multiples of $2f_{off}$, which is about 1.2 Hz. This doesn't match the measurements at all: ±0.55 Hz, ±2.7 Hz and ±3.25 Hz are no multiples of 1.2 Hz.

If the PCM2DSD had been scaled such that 0 dB PCM is converted to 0 dB DSD, that 1.2 Hz would have been 1.3458251954 Hz. At least we would have an explanation for the peaks at ±2.7 Hz then, the other peaks would still be unexplained, as would the absence of peaks at ±1.35 Hz.

3.5. Experiment with various offset values

On 30 March 2024, bohrok2610 sent results from some new measurements where various different DC offsets were added (not subtracted) to the 32 bit signed integer PCM signal going into the PCM2DSD. These measurements were done with a 96 kHz PCM signal, resulting in a 12.288 MHz output sample rate (the PCM2DSD reduces its interpolation factor at high input rates). Bohrok2610 sent nine plots, see Figure 5 to Figure 13.
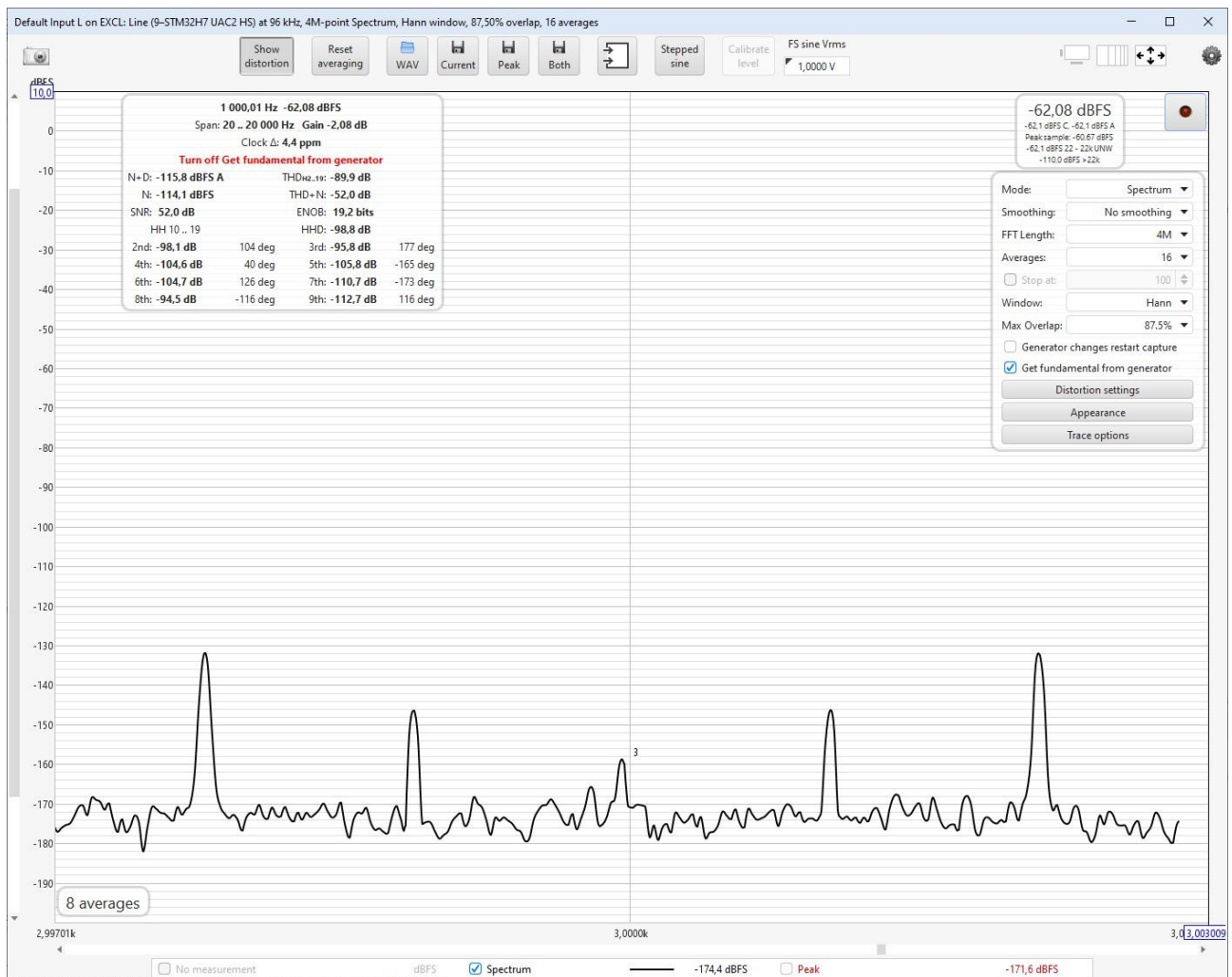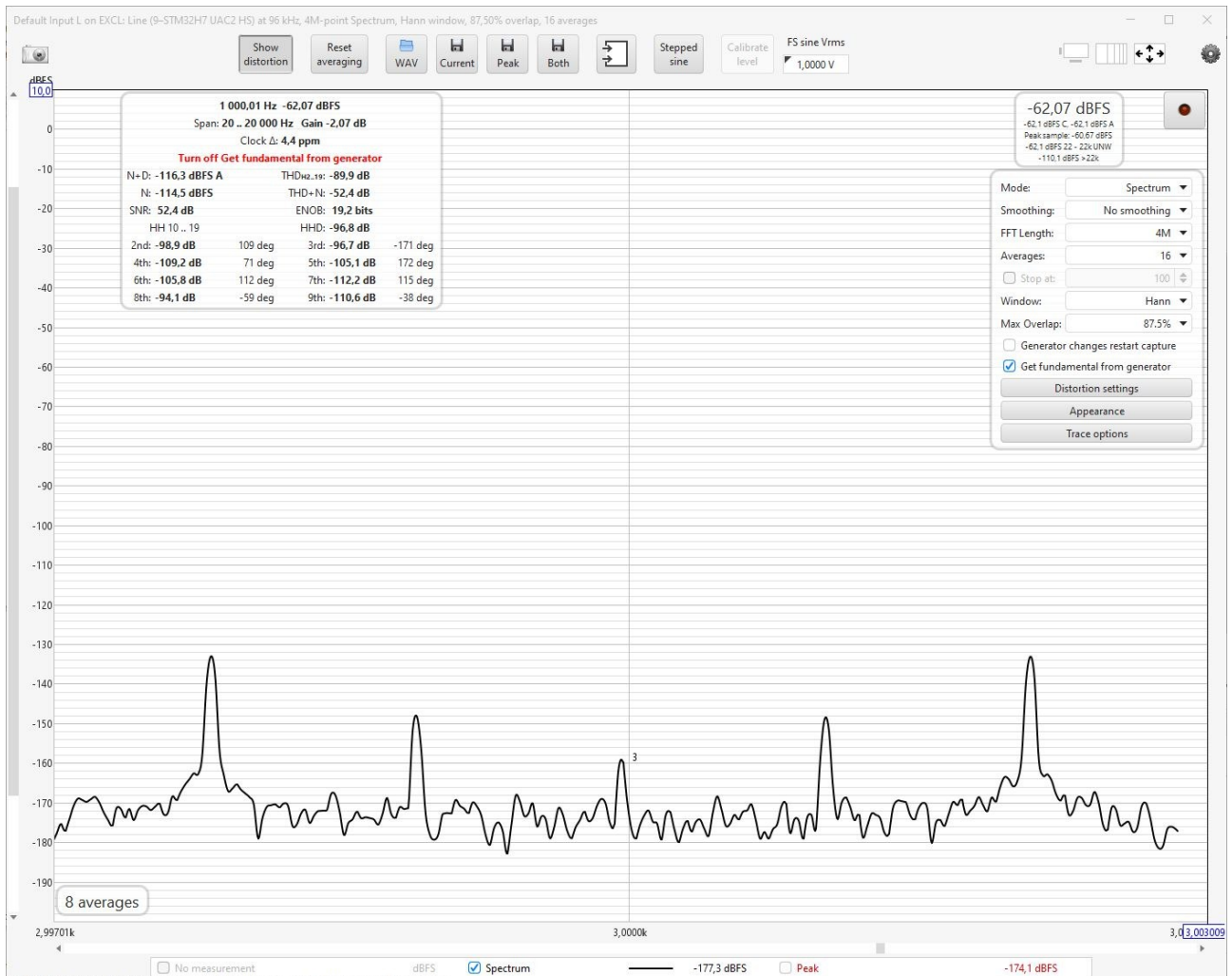
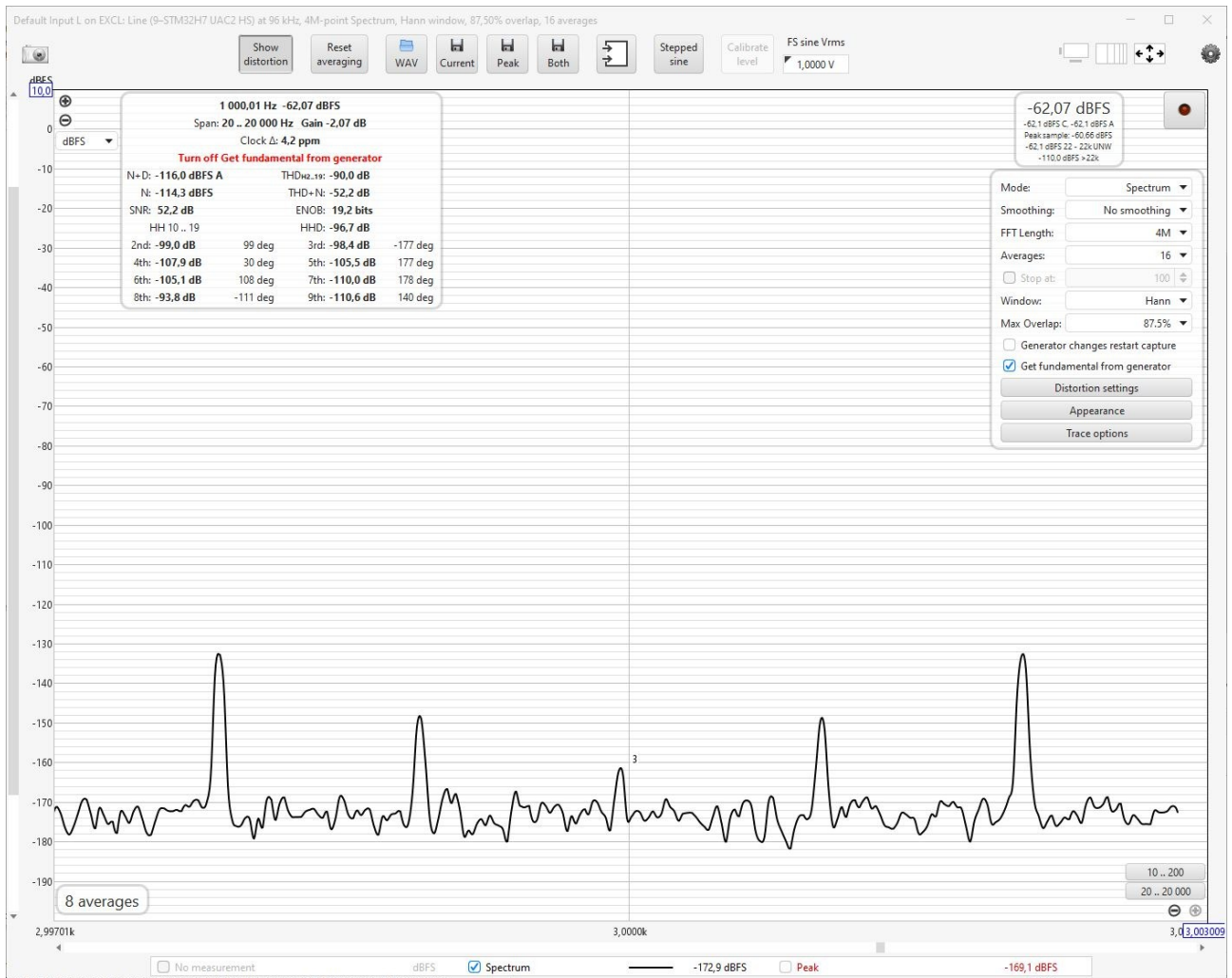*Figure 5: Offset 0, 96 kHz*

Figure 6: Offset 16, 96 kHz
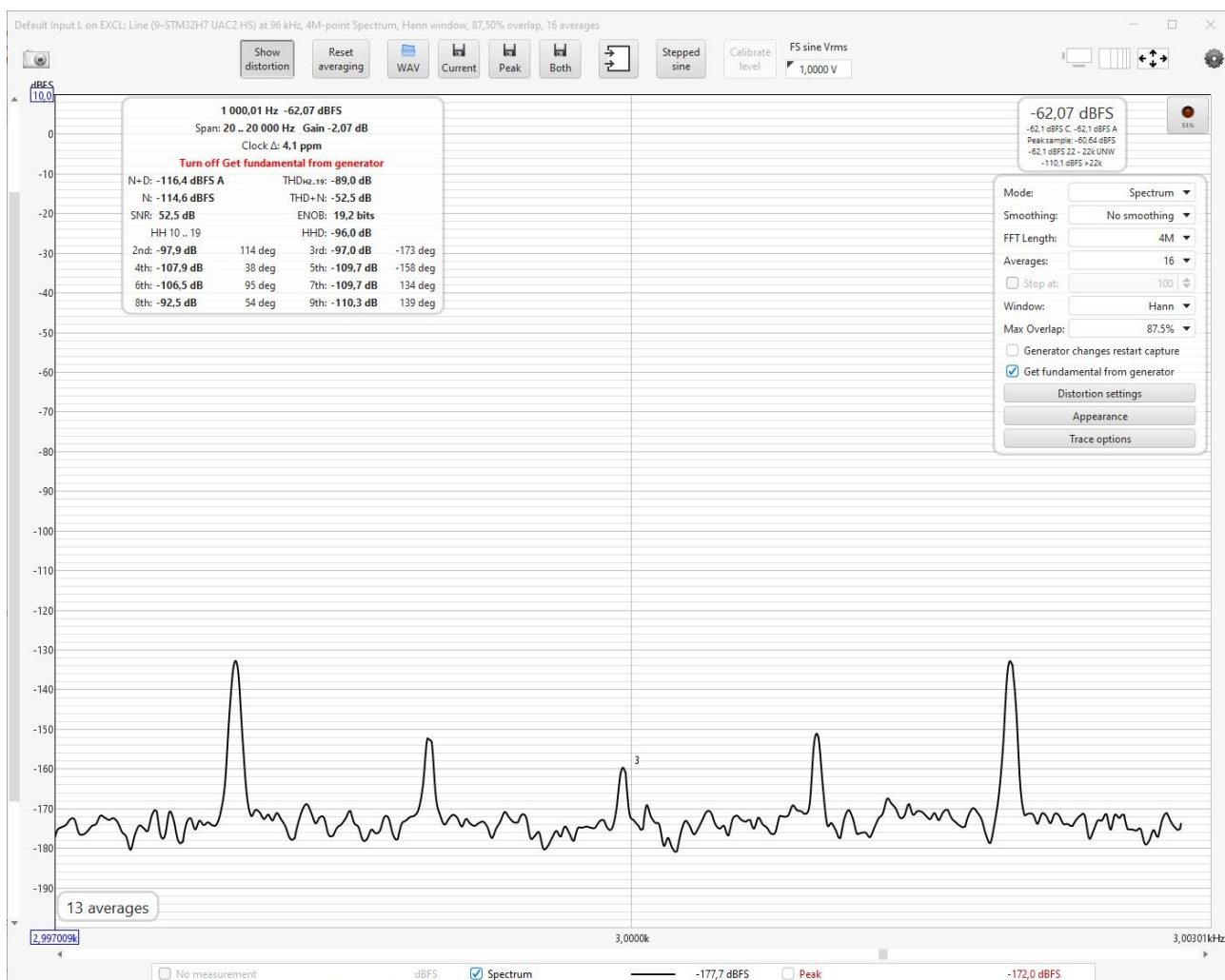
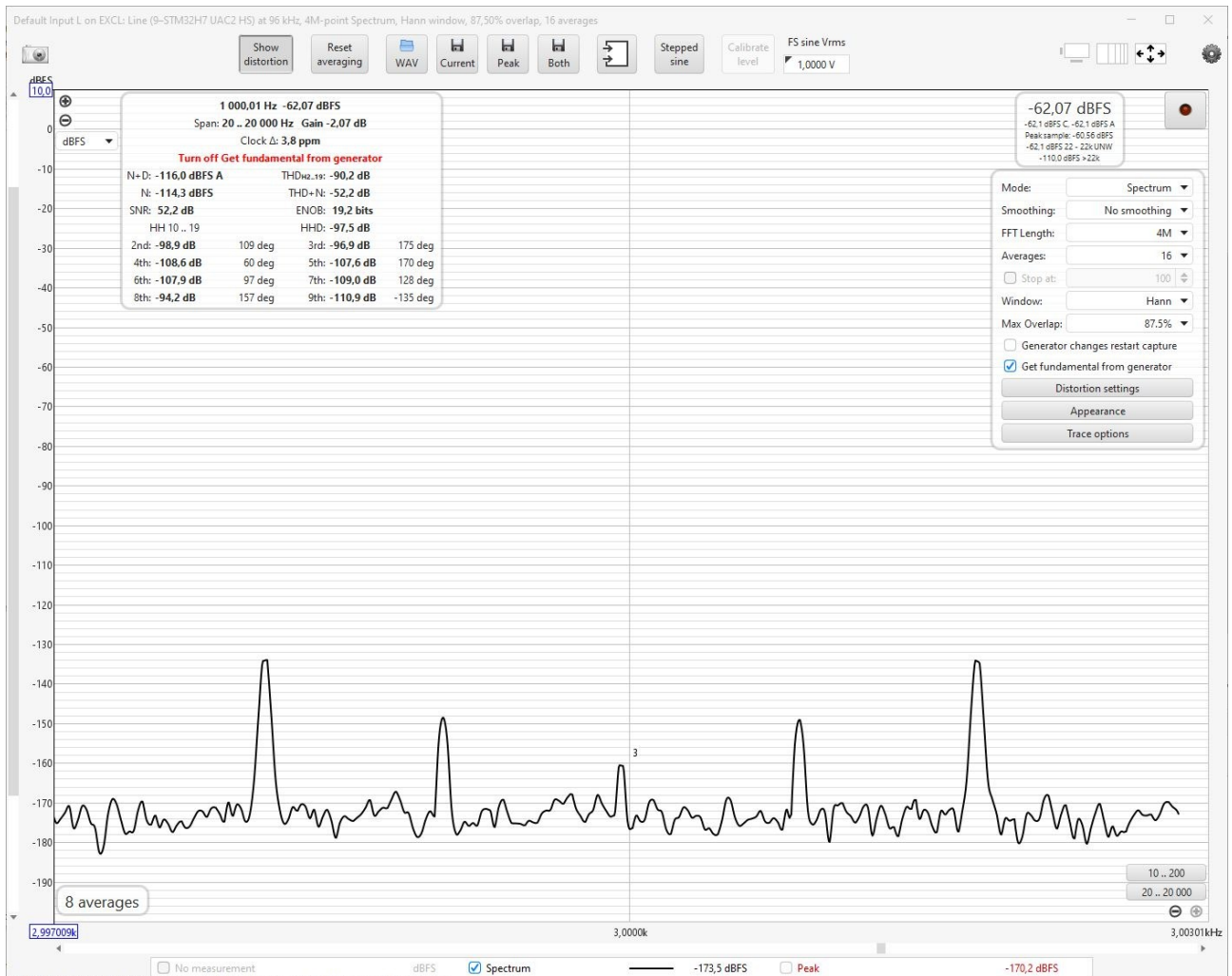*Figure 7: Offset 32, 96 kHz*

Figure 8: Offset 64, 96 kHz
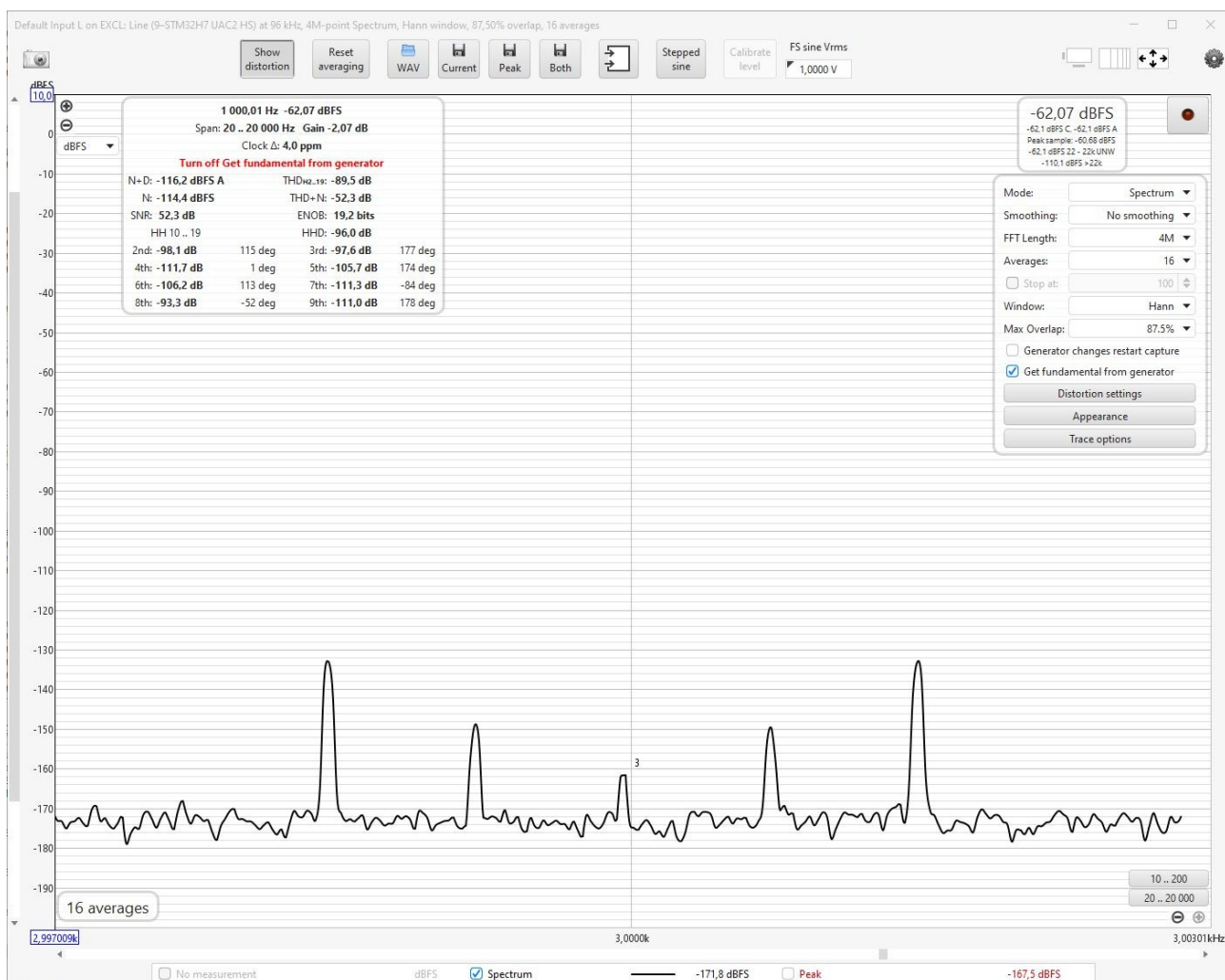
Figure 9: Offset 128, 96 kHz
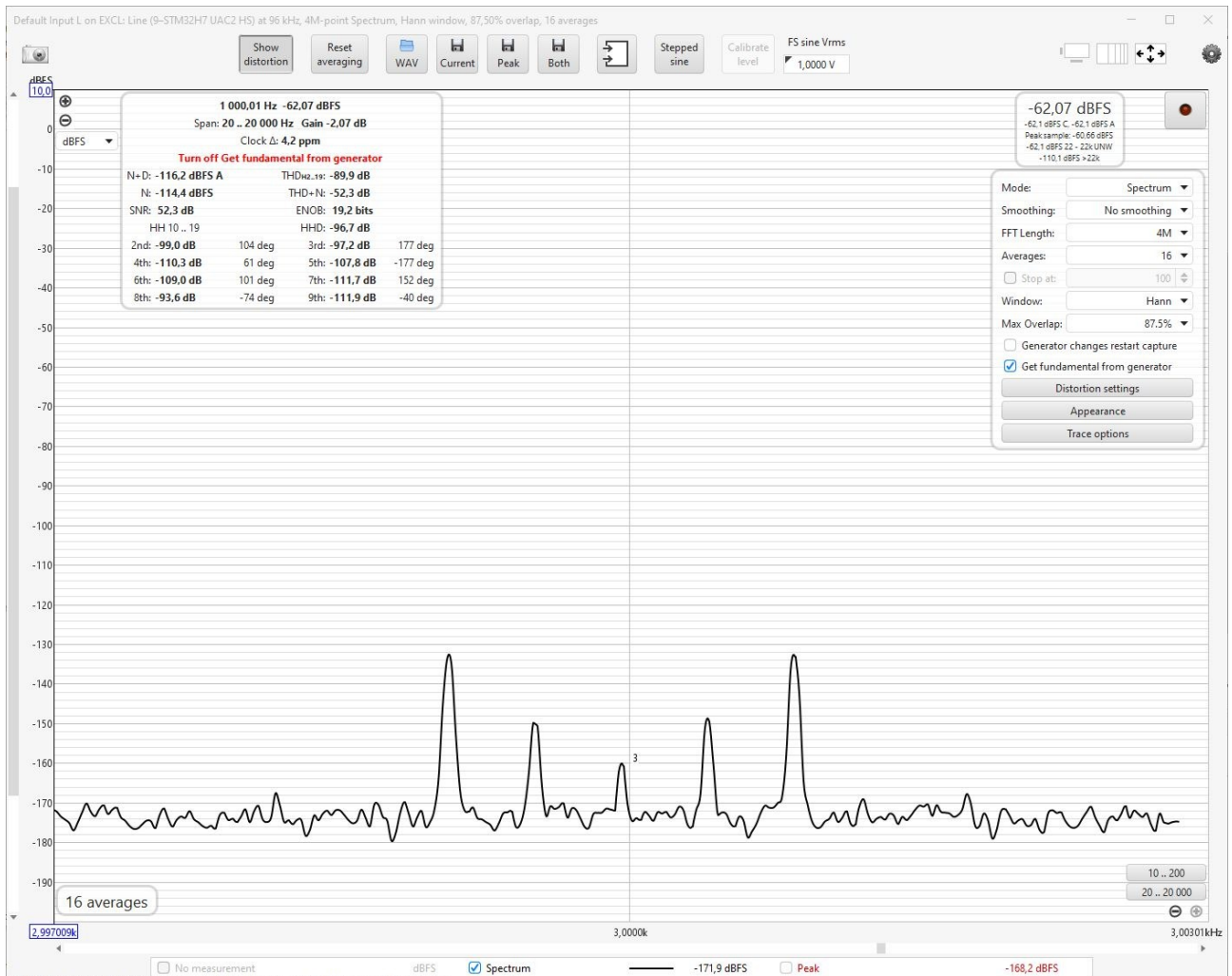
Figure 10: Offset 256, 96 kHz
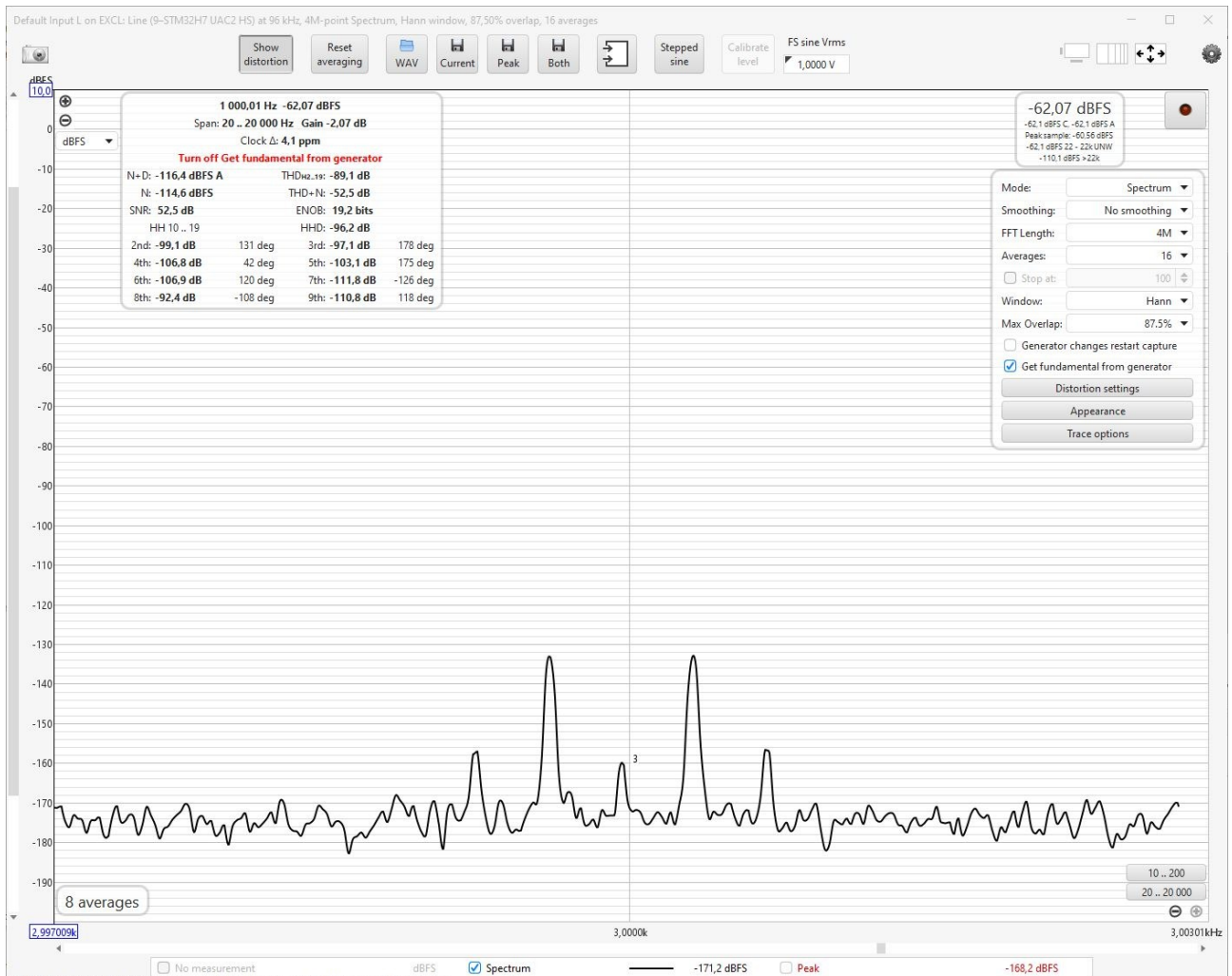
Figure 11: Offset 512, 96 kHz
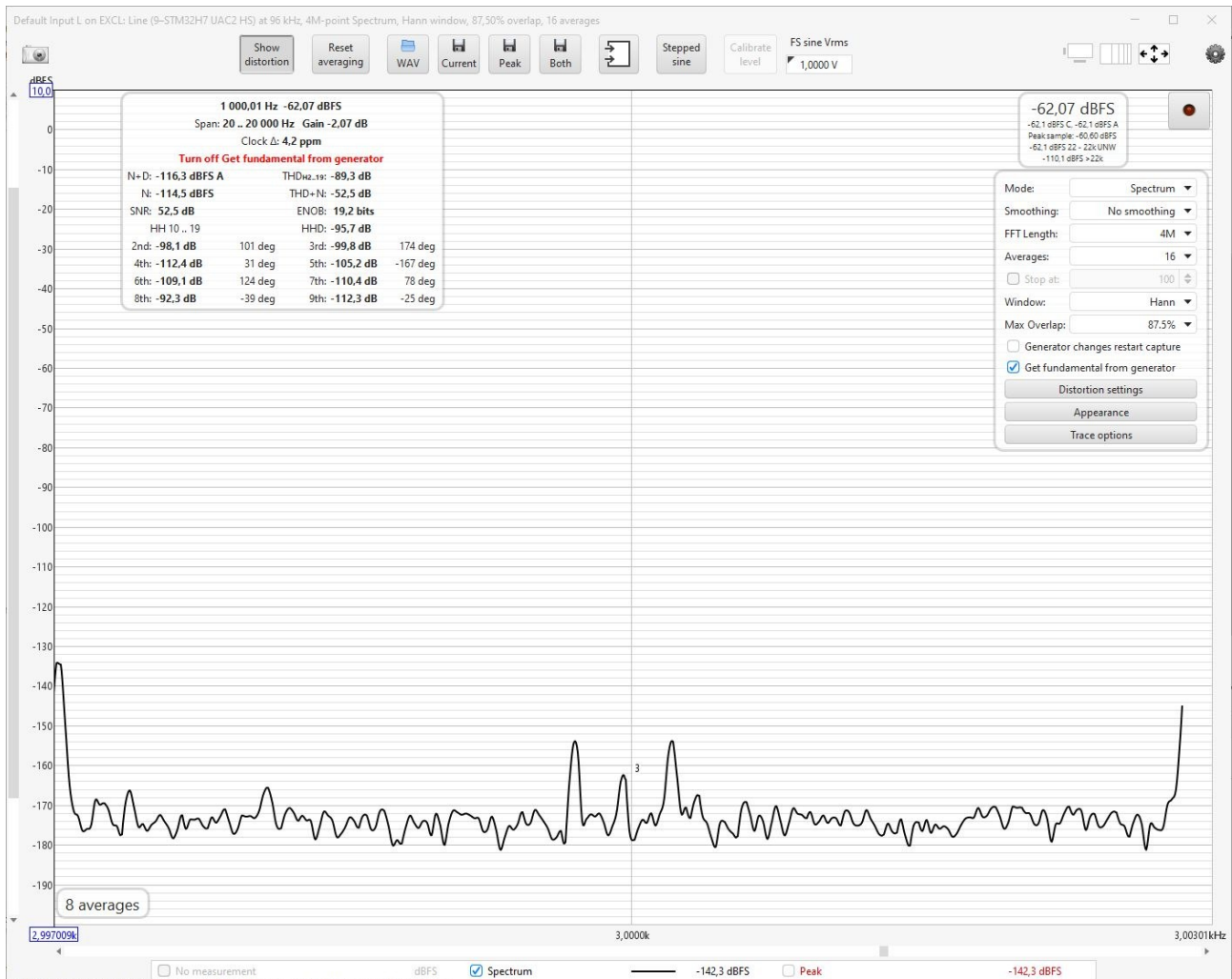
Figure 12: Offset 1024, 96 kHz

Figure 13: Offset 2048, 96 kHz

Comparing the plots, there is always except at an offset of 1024 the pattern large peak, small peak, smaller peak, small peak, large peak. At an offset of 1024, the pattern is small peak, large peak, smaller peak, large peak, small peak. The smaller peak is very close to 3 kHz and doesn't move, so this is presumably the third harmonic frequency.

Ignoring the small peaks and only looking at the large ones, the pattern is quite consistent with theory if the PCM2DSD has an equivalent input offset of about -870. At an added offset of 1024, the peaks have then just crossed the centre. See 2 and Figure 14.

offset	peak 1	peak 2	peak 3	peak 4	peak 5
0	2997.78775	2998.8681	2999.94845	3001.03411	3002.11711
16	2997.82491	2998.88933	2999.96437	3001.02614	3002.09322
32	2997.86472	2998.90791	2999.96437	3001.00756	3002.04013
64	2997.94898	2998.94472	2999.95108	3000.96541	3001.96115
128	2998.1083	2999.02172	2999.94843	3000.88575	3001.80714
256	2998.42959	2999.18901	2999.95639	3000.72643	3001.48585
512	2999.0589	2999.50499	2999.9617	3000.40249	3000.85654
1024	2999.19698	2999.59262	2999.96702	3000.32814	3000.71581
2048	2997.0329	2999.70679	2999.95108	3000.21396	3002.90379

Table 2: Frequencies (read off with a set square) as a function of offset, colours indicate how the big peaks cross the frequency of the centre peak
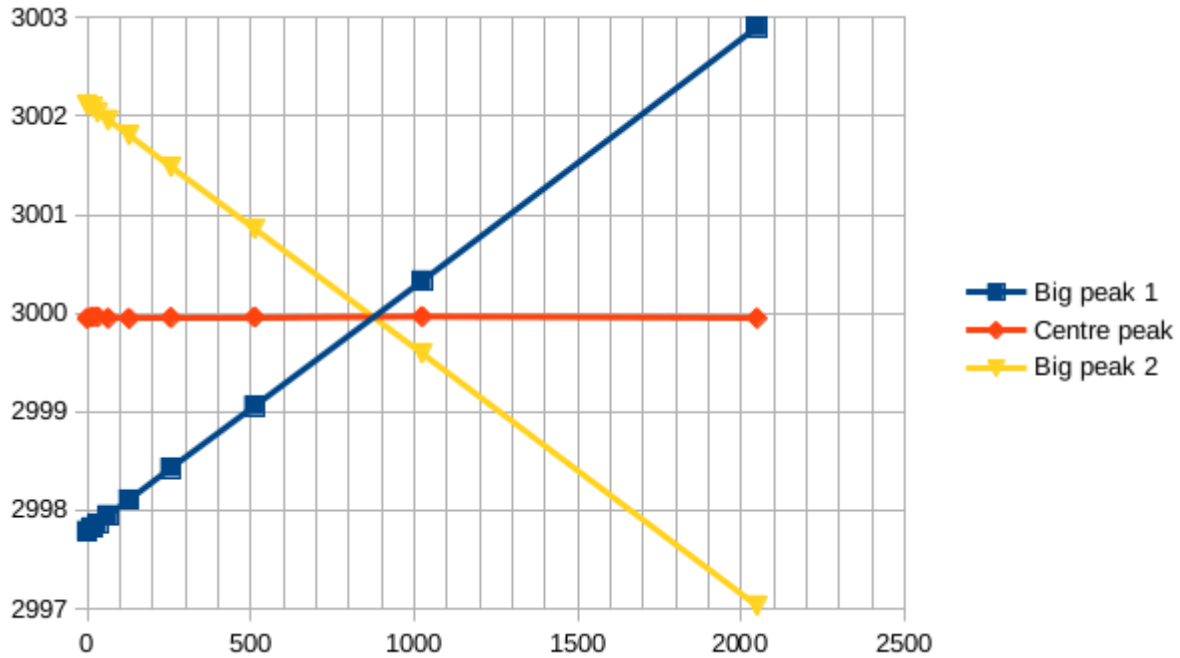

Figure 14: Frequencies of the big peaks and the centre peak as a function of the added offset

In absolute value, the slope of the frequencies of the big peaks is on average about 2.490297 mHz/LSB, corresponding to 5347872 Hz for 2^{31} LSB.

As explained in section 3.4, when the FM-like components around $f_s/2$ are shifted by a frequency offset f_{off}, low-frequency intermodulation products between these components and their aliases will shift by some even multiple of f_{off}. In the simplest case, second-order intermodulation distortion between a component that shifts f_{off} to the right and a component that shifts f_{off} to the left will result in a difference frequency that shifts by $2f_{off}$.

Hence, f_{off} changes by (at most and most probably) one half of 2.490297 mHz/LSB. This corresponds to a frequency change of 2673936 Hz for 2^{31} LSB (PCM full-scale).

A sigma-delta modulator running at 12.288 MHz and scaled to produce 0 dB DSD at 0 dB PCM in,

should change the frequency of its ones by 3.072 MHz when driven by full-scale PCM. 2673936 Hz is 1.2054 dB less. Hence, 0 dB PCM produces -1.2054 dB DSD, which is close to the -1 dB mentioned by bohrok2610.

All in all:

-The large peaks around 3 kHz behave according to theory, the small ones do not

-The PCM2DSD is scaled for -1.2054 dB DSD out at 0 dB PCM in

-It has an input offset of about -870 LSB

-The distortion products can be simple second-order intermodulation products between the FM-like components around $f_s/2$ and their aliases. For example, the frequency difference between

$\frac{1}{2}f_s + f_{off} + f_m$ and $\frac{1}{2}f_s + f_{off} - 2f_m$ is precisely 3 f_m, that is, three times the audio frequency.

Intermodulation between signals in the megahertz range can therefore contribute to harmonic distortion. The frequency difference between $\frac{1}{2}f_s + f_{off} + f_m$ and $\frac{1}{2}f_s - f_{off} - 2f_m$ is 3 f_m + 2 f_{off}, resulting in one of the weird peaks around the third harmonic.

3.6. Time varying amplitudes

When f_{off} is very small and you measure with a resolution bandwidth much greater than $2f_{off}$ and a sweep time much smaller than $1/(2f_{off})$, you see peaks of which the amplitude goes up and down, rather than split peaks. This is easy to explain: for small f_{off}, an FM component at ½ f_s - f_{off} - kf_m is quite close to the alias of ½ f_s - f_{off} + kf_m that lies at ½ f_s + f_{off} - kf_m. In fact, each FM component is at a frequency distance of $2f_{off}$ from an alias of another FM component.

You can look at such a pair of frequency components as a double-sideband modulate with modulating frequency f_{off}, which has an envelope that passes through zero twice per cycle of f_{off}. In other words, there are moments when they are in phase and enhance each other, and moments when they are in antiphase and cancel each other. When the envelope of a pair of frequency components passes through zero, the same happens with any audio intermodulation products related to them.

4. Conclusion

The fact that the distortion products are not exact multiples of the signal frequency and that adding a small offset shifts them, is in my opinion pretty good evidence that the distortion products are intermodulation products of FM-like out-of-band peaks. The same holds for their disappearance when quasi-multibit modulators are used.

For the original experiment with one deliberate offset of -512, I have been unable to explain what exactly intermodulates with what. For the follow-up experiment with various deliberately added offsets, as long as you ignore the small peaks and concentrate on the big ones, everything seems to match theory if the PCM2DSD has an offset of about -870. The intermodulation can then be simple second-order intermodulation between the FM-like components around $f_s/2$ and their aliases.

The experiments bohrok2610 did showed a clear sensitivity of the distortion products to the output filter, so it appears that the intermodulation occurs in the output filter. Adding more passive filtering

between the DAC core and the active part of the reconstruction filter would be the logical way to deal with this, although switching to other op-amps or forcing more DC current through their output stages could also help.