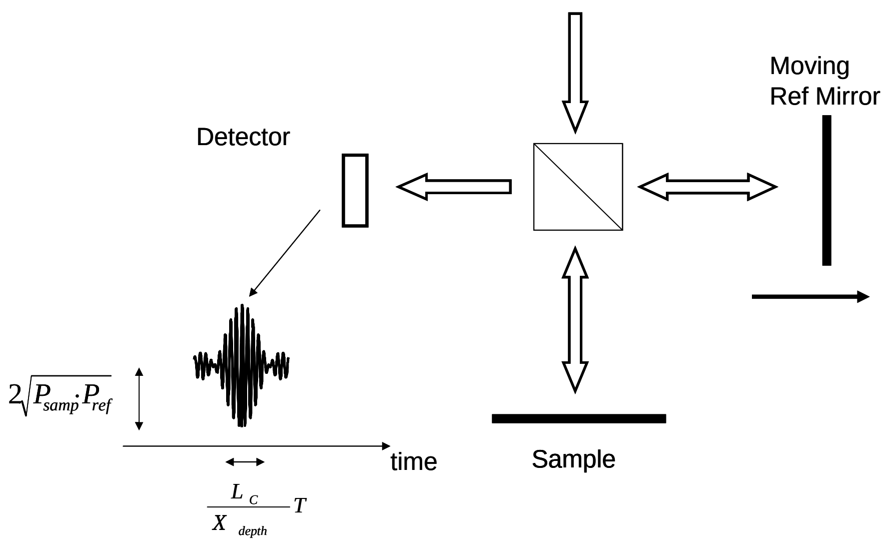
Time domain Optical Coherence Tomography (TDOCT) versus Fourier domain Optical Coherence Tomography (FDOCT):

The year of 2003 was an exciting year for researchers working in the research area of Optical Coherence Tomography (OCT). During that year, researchers came to the realization that there is a much more efficient way to do OCT imaging. This more efficient way of performing OCT imaging brings such a vast improvement to the equivalent signal to noise ratio (> 20dB), that the present approaches for OCT imaging will very likely be phased out in the very near future.

We begin the discussion of the efficiency improvement by pointing out the commonality of the old approach (time domain) and the new approach (Fourier domain). Then, we will take a look at the system design simplification (or complication, depending on where you are coming from) that the new approach entails. Finally, we will look at two different ways of looking at the signal to noise analysis for the two approaches and try to understand where the SNR improvement comes from.

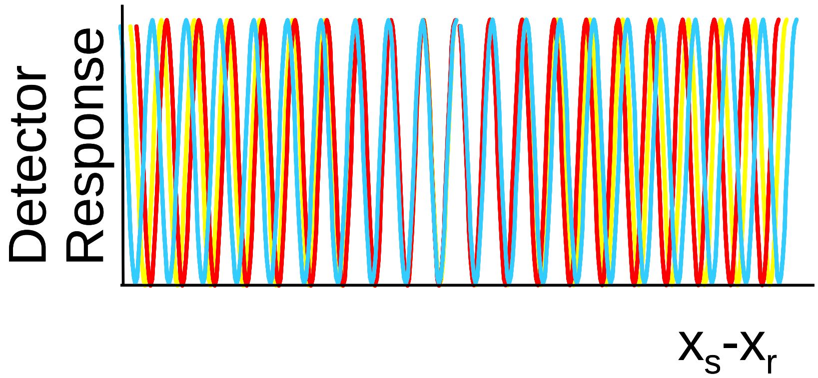
The best way to understand an OCT system is by first considering the behavior of a single wavelength component within an interferometer.



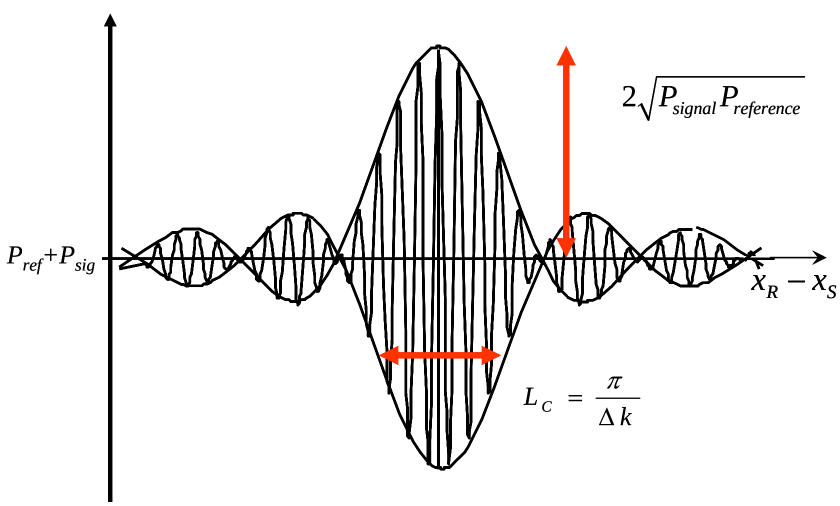
Consider the scheme shown in above. Now assume a monochromatic light source. If you scan the reference mirror, you will see a continuous sinusoidal interference signal. The expression for the interference signal is given by:

$$P_{combined}(k) = P_{reference}(k) + P_{signal}(k) + 2\sqrt{P_{signal}(k)P_{reference}(k)}\cos(k2(x_r - x_s)) \quad (1)$$

There is a very important thing to note here. Notice that at the point where $x_r = x_s$, all light components will constructively interfere ($\cos(k2(x_r - x_s)) = 1$), no matter what the wavelength is. Physically, this makes a lot of sense, when $x_r = x_s$, the light is going to take the exact same amount of time to travel either path, so the 2 different light trajectories have to end up with the same phase no matter what the wavelength is. Now, if the light source has a spectral span, we can treat each wavelength component in the same manner.



In the TDOCT case, we are collecting the light in a single detector. The resulting signal we see at the end is simply the sum of the individual trace. Since each sinusoidal component will have a slightly different wavelength, their period will be slightly different. Upon summing, we will get a resulting interference signal that drops off rapidly beyond the point where the two reflectors are matched in pathlengths.

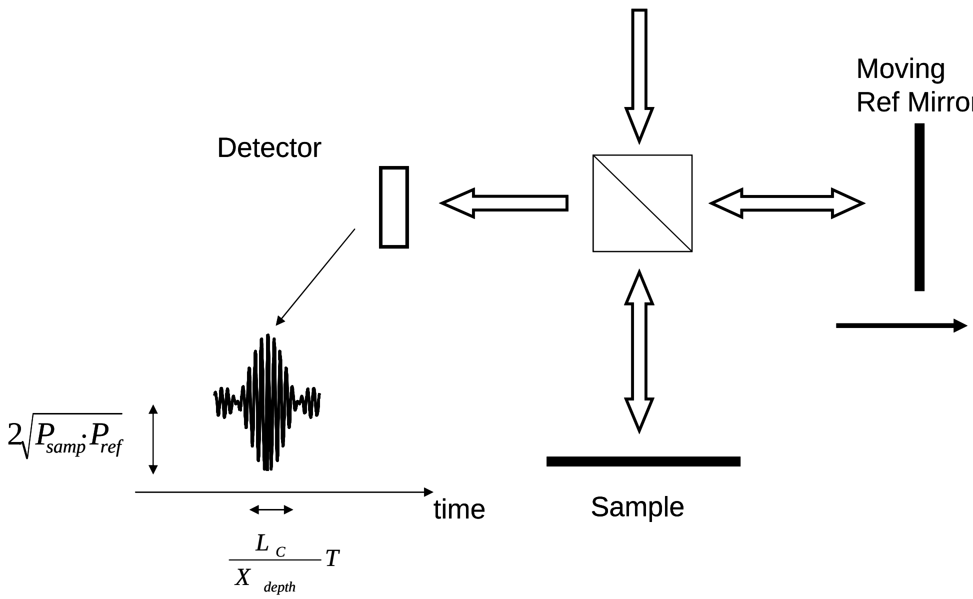


Mathematically, we can express the detected signal as:

$$\begin{aligned}
P_{combined} &= \int P_{combined}(k)dk \\
&= \int P_{reference}(k) + P_{signal}(k) + 2\sqrt{P_{signal}(k)P_{reference}(k)}\cos(k2(x_r - x_s))dk \qquad (2) \\
&= P_{reference} + P_{signal} + 2\sqrt{P_{signal}P_{reference}}\,\mathrm{sinc}\left(\Delta k(x_r - x_s)\right)\cos(k_o 2(x_r - x_s))
\end{aligned}$$

where $l_c = \frac{\pi}{\Delta k}$ is the coherence length (for a sinc function), and Δk is the spectral bandwidth of the light source. I am assuming a square spectral profile in this calculation and the rest of the essay.

And just so we are clear from the start, I am also going to assume that $P_{reference} >> P_{signal}$, which is the typical situation encountered in OCT.



Let's calculate the 'price' of SNR for a single resolvable point (and let's give a definition to 'resolvable point') in TDOCT. A resolvable point in OCT is given by the extent of the OCT coherence length. Suppose, we perform depth scanning at scan cycles of time period T and the scan depth is x_{depth}. We can calculate the time the system dwells on each resolvable point as equal to $\frac{l_c}{x_{depth}} T$. The effective signal strength collected for that point is basically equal to the number of useful detected signal photon within the acquisition time of each resolvable point. That is given by the modulation power producted by the time duration for acquisition of that single pixel. We write it in terms of detected photon count as:

$$Signal = 2\sqrt{P_{signal} P_{reference}} \frac{l_c T}{x_{depth}} \frac{\varepsilon}{h\nu}, \qquad (3)$$

where ε is the detection efficiency of the detector (the ratio of detected photon count to the actual photon number), and $h\nu$ is the energy quantum of the photons.

The heterodyne frequency shift is no relevance to the above consideration. The logic is the same as the consideration for an AC power line. The power delivery is dependent only on the voltage and current, the AC frequency is irrelevant.

The noise strength collected is equal to the number of noise photon accumulated within the pixel collection time. Assuming that the shot noise is white (which it is), the noise strength is given by:

$$Noise = \sqrt{\left(P_{signal} + P_{reference}\right)\frac{l_c T}{x_{depth}}\frac{\varepsilon}{h\nu}} \approx \sqrt{P_{reference}\frac{l_c T}{x_{depth}}\frac{\varepsilon}{h\nu}}\,. \tag{4}$$

The computed SNR for a given pixel is then:

$$SNR_{time\ domain} = \left(\frac{signal}{noise}\right)^2 = \left(2\sqrt{P_{signal}\frac{l_c T}{x_{depth}}\frac{\varepsilon}{h\nu}}\right) = 4P_{signal}\frac{l_c T}{x_{depth}}\frac{\varepsilon}{h\nu}\,. \tag{5}$$

SNR: What IS the convention?

If you have been working in the field of optics, you will very quickly come to realize that the definition of SNR can be very confusing.

In the normal engineering convention,

$$SNR = \frac{signal\ amplitude}{noise\ amplitude}$$

$$SNR_{dB} = 20\log_{10}\left(SNR\right)$$

In most optics related publications, (for example for OCT applications)

$$SNR_{OCT} = \left(\frac{signal\ amplitude}{noise\ amplitude}\right)^2$$

$$SNR_{OCT,dB} = 10\log_{10}\left(SNR_{OCT}\right)$$

At the end of the day, if we are discussing SNR in dB's, both conventions gets you to the same answer. But, if you are simply discussing SNR, the results are actually off by a square factor.

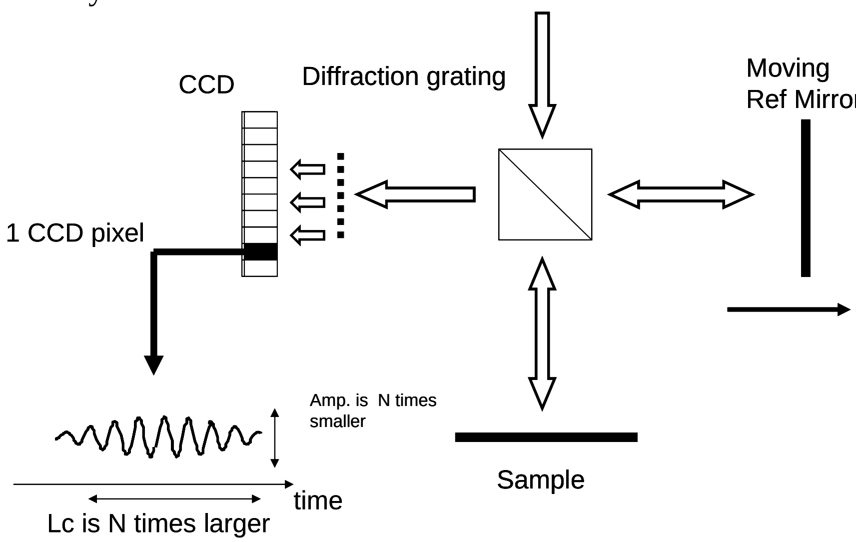
I am going to stick with the optics convention for the rest of this essay and subsequent essays. Just keep in mind the difference!

There are two ways to see that FDOCT has an SNR advantage over TDOCT.

FDOCT (deBoer's explanation)

The way proposed by J. de Boer and group in their OL paper (Optics Letters, Vol. 28 Issue 21 Page 2067 (November 2003)) is very intuitive. I really like it for its directness.

The experiment is very simple. Perform the experiment as you normally would (translating the reference mirror, etc). This time round, spectrally disperse the interference output onto N distinct detection channels. We are going to assume that the spectral bandwidth of the light source is equally divided into N parts and each component is detected by individual channels.



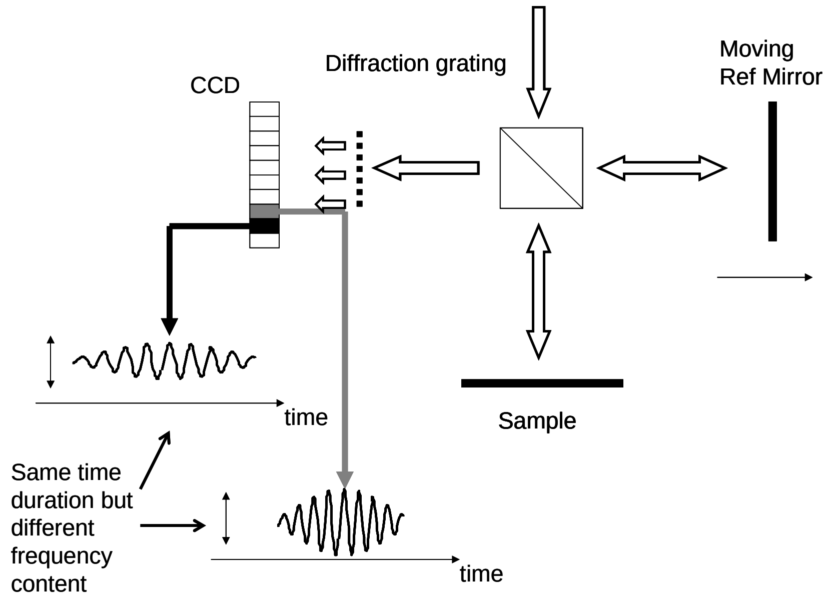
Now, each channel will have a lower incident power (by a factor of 1/N) when compared with the original time time domain experiment. But, the interference signal associated with each reflection site within the sample will last longer by a factor of N. Remember that the coherence length is proportional to the spectral bandwidth ($l_c = \frac{\pi}{\Delta k}$), so an N times reduction in bandwidth leads to N times increase in coherence length. This means that the effective signal strength from each channel in terms of detected photon count is given by:

$$Signal_{per\ channel} = 2\sqrt{\frac{P_{signal}P_{reference}}{N^2}}\frac{(l_c N)T}{x_{depth}}\frac{\varepsilon}{h\nu} = 2\sqrt{P_{signal}P_{reference}}\frac{l_c T}{x_{depth}}\frac{\varepsilon}{h\nu}. \qquad (6)$$

The signal strength is as good for every given channel as it is for the one channel situation for TDOCT. This is not too surprising. The power per channel is down but we are collecting for the signal for a longer time because the coherence length is lengthened $(l_c N)$. We do pay a price in this situation, and that is that we have to contend with an effectively longer coherence length $(l_c N)$.

The noise term associated with each channel can be likewise analyzed and calculated to be: $Noise_{per\ channel} = \sqrt{\frac{P_{reference}}{N} \frac{(l_c N)T}{x_{depth}} \frac{\varepsilon}{h\nu}} = \sqrt{P_{reference} \frac{l_c T}{x_{depth}} \frac{\varepsilon}{h\nu}}$. (7)

Again, a power reduction is compensated by a longer collection time. We further note that this noise term is the shot noise component that falls within the heterodyne frequency acceptance bandwidth of the light component falling on the channel. We note here that because each channel reads a different spectral bandwidth of the interferometer's output, the associated heterodyne frequency is different for each channel.



When we combine the processed data from the channels to constitute the OCT signal with the complete spectral bandwidth of the source, the signal contribution from each channel can simply be summed together.

$$Signal_{fourier-domain,deBoer} = \sum_N Signal_{per\ channel} = N2\sqrt{P_{signal}P_{reference}} \frac{l_c T}{x_{depth}} \frac{\varepsilon}{h\nu} \qquad (8)$$

This unsurprisingly results in the same signal maximum amplitude as you would get in a TDOCT scenario. But, this time round, you gain access to information far beyond the coherence envelop (N times beyond); even though at points beyond the coherence envelop, the different components simply cancel out. So yes, the information appears useless beyond the coherence envelop but it does have an advantage – it allows the extension of the base time frame for signal and noise collection. The noise terms has to be summed in a root mean squared approach as the noise content of the channel do not share the same frequency components.

$$Noise_{fourier-domain,deBoer} = \sqrt{\sum_N \left(Noise_{per\ channel}\right)^2} = \sqrt{NP_{reference} \frac{l_c T}{x_{depth}} \frac{\varepsilon}{h\nu}} \qquad (9)$$

So the computed SNR in this case is:

$$SNR_{fourier-domain,deBoer} = \left(2\sqrt{NP_{signal}\frac{l_c T}{x_{depth}}\frac{\varepsilon}{h\nu}}\right)^2 = 4NP_{signal}\frac{l_c T}{x_{depth}}\frac{\varepsilon}{h\nu}, \quad (10)$$

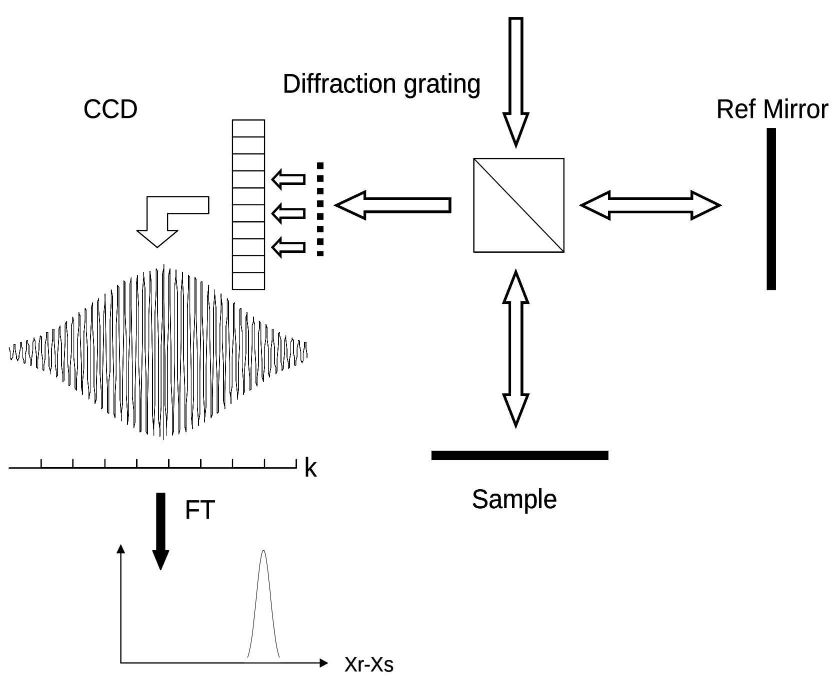
which is an improvement of $\sqrt{N}$ times over the time domain case. (Or a N times increase if you are using the OCT-speak SNR definition. Or a dB increase of $20\log_{10}\sqrt{N} = \log_{10} N$, which is the same in normal or OCT-speak.)

Another way to look at this is to consider a 2 channels case. We split the spectral bandwidth of the interferometer output in half and put each half on separate channels. In the 1st channel, the heterodyne signal spans from say fa to fb. And in the 2nd channel, it spans from fb to fc. In the 1st channel (2nd channel), the shot noise term that would have to be accepted along with the signal will be from the fa to fb region (fb to fc region). In comparison, in a single channel scenario, the contributions that would otherwise go to separate channels will be mixed, the shot noise to be accepted need necessarily span from fa to fc. This results in more total shot noise collected in the case than in the 2 channels case.

FDOCT (Choma's explanation)

The second way to appreciate the FDOCT advantage is developed by Mike Choma. This approach is a little less intuitive than de Boer's way, but it is a generalized formulation that applies directly to all FDOCT cases.

A generalized FDOCT system is based on a static interferometer in which the output is spectrally dispersed onto N channels. An array of data compiled from the measurements from the channel set can then be Fourier transformed to yield an equivalent A-scan data set.

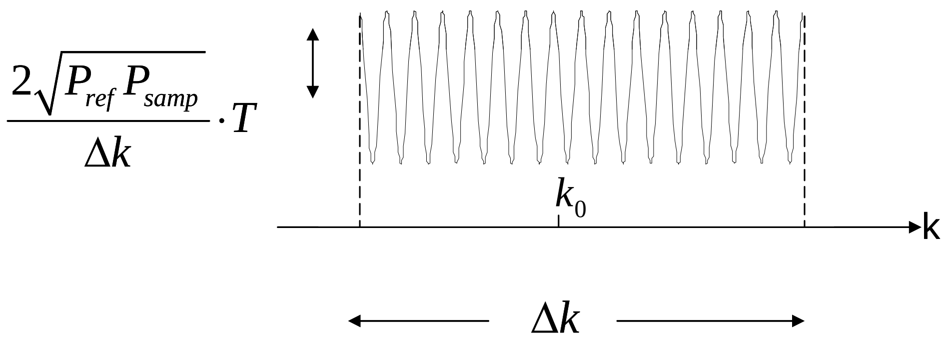


Each ‘frequency’ component within the collected channel data set can thus be associated with a particular depth within our sample. I have placed the word frequency in parenthesis because this frequency has to do with a periodic function expressed in k axis rather than time. The signal to noise issue in this case consists of finding the signal strength of a particular ‘frequency’ component and the associated noise at this ‘frequency’ component.

Assuming a simplifying situation where the spectrum is a square function. We can see that the interference signal resulting from a reflector at location x_s (assuming that the reference mirror is parked at x_r) and which has been accumulated over a time T is given by the following profile. The spectral function that we can expect the detector array to see will be (in energy units):

$$E_{combined}(k) = P_{reference}(k)T + P_{signal}(k)T + 2\sqrt{P_{signal}(k)P_{reference}(k)}T\cos(k2(x_r - x_s))$$
$$= \frac{P_{reference}T}{\Delta k} + \frac{P_{signal}T}{\Delta k} + \frac{2\sqrt{P_{signal}P_{reference}}T}{\Delta k}\cos(k2(x_r - x_s)) \quad , \quad (11)$$

where the last equality is arrived by using the square function spectrum assumption.



Much in the same manner as we derived the relevant signal photon count in the time domain case(Eq. 3), we simply product the modulation amplitude with the x-axis extent to get at the relevant signal photon count here. In this case, the x-axis extent is the spectral bandwidth in wavenumber. In both the time domain and Fourier domain case, we are really after the same thing – the number of signal photon that arises from the interference of the signal and reference beam. Whether we count them in k space or t space isn't important.

$$Signal_{fourier-domain,choma} = \frac{2\sqrt{P_{signal}P_{reference}}T}{\Delta k}\Delta k\frac{\varepsilon}{h\nu} = 2\sqrt{P_{signal}P_{reference}}T\frac{\varepsilon}{h\nu} \quad (12)$$

Now we come to dealing with the noise term. This is again easy to deal with. The k axis is analogous to the time axis. Since noise is white, the noise term from t1 is uncorrelated to the noise term from t2. The same can now be said of noise term from k1 and k2. So we can calculate the noise term in much the same way as we calculated it for the time domain situation. It is simply given by the square root of the total photons collected within the measurement process:

$$Noise_{fourier-domain,choma} = \sqrt{P_{reference} T \frac{\varepsilon}{h\nu}} . \quad (13)$$

The SNR of this case becomes:

$$SNR_{fourier-domain,choma} = \left(2\sqrt{P_{signal} T \frac{\varepsilon}{h\nu}} \right)^2 = 4 P_{signal} T \frac{\varepsilon}{h\nu} . \quad (14)$$

But, we are not quite done yet.

In order to do a fair comparison with TDOCT, we need one more piece of information - the depth scan of this particular FDOCT arrangement. The depth scan is limited by the channel number. This has to do with the fact that the highest 'frequency' component that can be sampled is limited to half the sampling 'frequency'. This means that the highest number of oscillation we can discern from our data set is N/2 oscillation where N is the number of channels in the spectrometer. From Eq. (11), we can see that this correspond to a maximum depth, $(x_r - x_s)_{\max} = x_{depth}$, by the following calculation:

$$\begin{aligned} &\Delta k 2 (x_r - x_s)_{\max} = \left(\frac{N}{2} \right) 2\pi \\ &\Rightarrow x_{depth} = \frac{N\pi}{2\Delta k} \end{aligned} \quad (15)$$

As $\frac{\pi}{\Delta k} = l_c$, the maximum depth is therefore also expressible as N/2 times the coherence length. The lowest frequency that can be measured in this system is of course zero, which works out to be associated with a depth of zero. From these calculations, we can see that the depth range of the system is given by $x_{depth} = \left(\frac{N}{2} \right) l_c$.

When we compare this to the formulation for the time domain case, we see that to compare a similar TDOCT to FDOCT setup, we need to have at least N pixel in the FDOCT setup so that it can cover the same depth range as a time domain system, where:

$$N / 2 = x_{depth} / l_c . \quad (16)$$

Upon, satisfying that condition, we can see that the SNR of equivalent TDOCT and FDOCT systems are:

$$SNR_{fourier-domain,choma} = 4P_{signal}T\,\frac{\varepsilon}{h\nu}, \qquad (17)$$

$$SNR_{time\ domain} = P_{signal}T\,\frac{l_c}{x_{depth}}\frac{\varepsilon}{h\nu} = P_{signal}T\,\frac{2}{N}\frac{\varepsilon}{h\nu}. \qquad (18)$$

FDOCT systems are clearly more efficient than TDOCT systems. In the case, we showed it to be N/2 times better.

If you have followed this analysis so far, you would have noticed one slight inconsistency. In the de Boer explanation, the calculated improvement was N while in the Choma scenario, the improvement is N/2. Why is there a difference? The answer is pretty subtle.

Remember that in de Boer's case, the reference arm is still scanned during the data acquisition process while in Choma's case, the reference arm is stationary. That movement breaks the symmetry in terms of +/- value in arm length mismatch $(x_r - x_s)$ in the interferometer and subsequently allows for access to the full depth range.

Okay, the above sentence is a little opaque, but I think it should make more sense if we take a closer look at Choma's case. We have previously said that the range of readable $(x_r - x_s)$ is from zero to $x_{depth} = \frac{N\pi}{2\Delta k}$. But, that is not strictly correct. $(x_r - x_s)$ from the the values of $-\frac{N\pi}{2\Delta k}$ to zero will show up in the measurement too. The only problem is that contribution from negative $(x_r - x_s)$ value is indiscernible from contribution from positive $(x_r - x_s)$ value. (See Eq. (11), notice that $(x_r - x_s) = -\Delta x$ and $(x_r - x_s) = +\Delta x$ gives you the same spectral profile.)

In this situation, in order to avoid the mix-up of signal from these two locations: $(x_r - x_s) = -\Delta x$ and $(x_r - x_s) = +\Delta x$, we choose to set the reference arm in such a way that we never get into a situation where there will be signal contribution from $(x_r - x_s) < 0$. This means that we are truncating by half the actual depth range that a N channel FDOCT system can actually handle in order to avoid the abovementioned problem. This subsequently leads to a decrease in calculated FDOCT efficiency of a half in Choma's scenario.

Of cos, there are ways to access the entire depth range that a N channel FDOCT system can provide. One approach is to do a phase step with the reference mirror. This phase step allows us to distinguish the $(x_r - x_s) = -\Delta x$ and $(x_r - x_s) = +\Delta x$ components from each other.

The way this works is the following. Suppose we have a signal that looks like this from a measurement. This can be contributed from a reflector at $x_s = x_r - \Delta x$ or $x_s = x_r + \Delta x$. But, if we displace the reference reflector by δx, if the reflector was at $x_s = x_r - \Delta x$, we would expect to see a tighter fringe pattern (because $x_r - x_s = \Delta x + \delta x$ now) (the fringes squeezes in around the zero point in k-axis based on Eq. (11)). The opposite is true for an opposite displacement (the fringes expands out from the zero point in k-axis based on Eq. (11)). Here's where it gets a little tricky. In general, we measure the spectral signal at a k location that is pretty well removed from zero. This means that there is an even easier way to separate the two types of contribution. The expansion/contraction is still there, but the relative shift of the spectral oscillation to one or the other direction is even more prominent. If we choose the reference reflector displacement to be +ve, we can expect the spectrum to shift in one direction or the other depending on whether the reflector is at $x_s = x_r - \Delta x$ or $x_s = x_r + \Delta x$. By the direction of the shift, we can then distinguish the two components.

As an exercise, try proving the following statement. As per the above described setup, show that given the same total acquisition time, but equally splitting the time between two measurement before and after a phase step, we can improve the effectively double the total usable depth range while retaining the same SNR as before. This doubling of the range and thus the SNR efficiency puts Choma's analysis on the same footing as de Boer's calculation.