Referee #2:

I congratulate the authors for the nice manuscript, and I would like to start by offering my sincere apologies for the delay in providing this review. The authors compared a standard ultrafine CPC with a modified Airmodus A20 CPC to investigate the effect of poor counting statistics on the calculation of growth and formation rates. Uncertainties on these quantities are often neglected and this work offers new and interesting results. The manuscript is well-written and fits the scope of the journal. The analysis is sound, and the results are presented clearly and concisely. However, some minor comments need to be addressed, and the manuscript's clarity can also be improved in a few places.

We thank the reviewer for their insightful suggestions (original comment in black) and please find our responses below (in blue) and changes to the revised manuscript in red.

Minor comments:

Lines 42-44: This sentence about technological development can probably be used for every measurement system. Consider removing or rephrasing it.

We agree with the reviewer and adjusted the sentence to:

"However, there are still significant discrepancies between different particle size-distribution data sets, especially for the sub-10 nm size-range (Kangasluoma et al., 2020)"

Line 44: "a large fraction" can you provide more quantitative information? Losses will clearly depend on the instrumental setup but please mention what the typical range is.

We thank the reviewer for asking to be more quantitative here and put (typically >95%) in the text to give the order of magnitude, based on e.g. Stolzenburg et al. (2017), AMT.

Line 48: The "PI" parameter is not used in the rest of the manuscript, and I do not see the reason for mentioning it here in the introduction. The only relevant information is that a system with lower losses, higher flow rate and sampling time will have better counting statistics. You already explained this in lines 47-48, so you can remove this part on the PI parameter.

We agree with the referee and removed that paragraph.

Line 51: "the PI parameter … describes the instrument sensitivity towards low number concentrations" instrument sensitivity represents the smallest absolute amount of change that can be detected. So, sensitivity should be the same at low and high number concentrations. I think speaking of the signal-to-noise ratio is more meaningful in this context.

Paragraph about the PI-parameter is removed, so no need to change the wording here.

Line 54: Is there any drawback in using a CPC with a higher aerosol flow? I guess there must be an optimum range; otherwise, why not use a 10lpm flow rate or higher? I could think of coincidence (probably not a big issue when the CPC is used behind a DMA), a higher DMA sheath flow is required to keep the same resolution, and probably some other technical issues with the CPC construction itself. Adding a short sentence on the cons/problems of having a larger flow would be useful.

Please see our response to referee #1 concerning this point.

Section 2.2:

- did you also intercompare the counting efficiency of the two CPCs? From Figure S1, it seems there is a ~5% difference; if not corrected, can this affect your results?

The counting efficiency curves as shown in Fig. S1 are included in the inversion and hence the 5% difference is always accounted for when inverted data are used. For growth rate calculations inverted data is not needed as only the relative rise of the signal is considered, which we clarified again in line with the comment of referee #1.

- What about the response time of the two CPCs? This is a relevant parameter for your analysis, so it would be important to report it here (CPC3776 response time is well characterized, but I don't know if it is the same for the modified A20).

  We agree with the reviewer that the response time could be an issue (~0.2 s for the TSI2778 and ~1s for the A20). However, in our entire analysis we compare the counts during a measuring period at a fixed DMPS voltage (i.e. one size), which is always longer than 3.5 seconds for each DMPS step. To get to the total number of counts in that period we took the average concentration and multiplied it by the flow rate and time interval. Over these intervals the different response times should not play any role, especially as the DMPS has a short stabilizing time interval (1 second) before every new measurement, such that we can assume that during the measurement at one size, the number of counts at both CPCs should not be affected by differences in response time. We agree with the referee that this should be clarified and added the following explanation to the manuscript.

  In Section 2.2 we added:

  "Apart from their differences in activation efficiency and effective detector flow rate, the two CPCs have different response times to a change in aerosol concentration, which are ~0.1 s for the TSI 3776 and ~1 s for the (unmodified) Airmodus A20 (Enroth et al., 2018). However, as we will see below that small difference does not affect our approach in comparing the counting statistics of the two CPCs."

  In the newly drafted Section 2.3 we added:

  "In our DMPS, the voltage is stepped from 3 to 1000 V in 17 steps (corresponding to selected mobility diameters of 2.07 to 40 nm assuming singly charged particles), with a settling time of 1 second at the beginning of each voltage step (which should remove any bias from different response times of CPCs, if they are ≤1s). The measured particle number concentration C (in cm^{-3}) for size is determined by the number of particles N counted in the time interval τ where the voltage is kept constant (which varies between 3.5 seconds for the largest size and 64 seconds for the smallest size) a specific measurement volume and the number concentration C can be calculated by using the volumetric flow rate through the optics Q_{opt}: (…)"

Line 110: here, you define 'time t' but in equation (1) it is defined as tau; please use the same definition.

Thanks for noting. We now consistently use τ throughout the manuscript.

Lines 118-125: this part is confusing because you provide the Poissonian distribution before saying what a Poissonian process is. Additionally, you are mixing the general description of a Poissonian process with its applicability to an optical counting system. For example, coincidence applies to certain types of measurement systems but not to a general Poissonian process. My suggestion is first to define a general Poissonian process, then describe the Poissonian distribution and conclude with the applicability of Poissonian statistics to a counting system like a CPC.

We agree with the reviewer that this part is not well-written and reformulated it:

"A random variable N has a Poisson distribution with the parameter $\mu\tau > 0$, where τ is the measurement time, and μ is the intensity (rate) of the process, if the random variable can obtain discrete values (0,1,2,3,…) within the time interval τ. If the process is characterized by the following properties: 1) For $\tau = 0$ we have $N(0) = 0$, 2) in separate time intervals, the numbers of detected events are independent of each other and 3) the number of events in any interval of length τ obey the Poisson distribution:

$$P(N(\tau) = N) = \frac{e^{-\mu\tau}(\mu\tau)^N}{N!} \quad (1)$$

A Poisson distribution can be shown to have the following properties: the expected value $\mathrm{E}[N]$ of the distribution can be calculated as $\mathrm{E}[N] = \mu\tau$, and the standard deviation (σ) can be calculated as $\sigma = \sqrt{VAR[N]} = \sqrt{\mu\tau} = \sqrt{\mathrm{E}[N]}$.

In a CPC, the particles are counted in the optical unit of the CPC, where a nozzle directs the particle stream to cross a laser beam perpendicularly. Light is scattered from the laser beam as the particles cross it, and the scattered light is collected by a photodiode. In typical optics with ~1 lpm aerosol flow, the probability of coincidence in the counting process is negligible with moderate number concentrations (< 30 000 cm^{-3}), which are typically measured downstream of a DMPS system. In our DMPS, the voltage is stepped from 3 to 1000 V in 17 steps (corresponding to selected mobility diameters of 2.07 to 40 nm assuming singly charged particles), with a settling time of 1 second at the beginning of each voltage step (which should remove any bias from different response times of CPCs, if they are $\leq$ 1s). The measured particle number concentration C (in cm^{-3}) for size is determined by the number of particles N counted in the time interval τ where the voltage is kept constant (which varies between 3.5 seconds for the largest size and 64 seconds for the smallest size) by using the volumetric flow rate through the optics Q_{opt}:

$$C = \frac{N}{Q_{opt} \cdot \tau} \quad (2)$$

If we assume that the number concentration remains constant during the voltage scan of the DMPS (which is anyways also a requirement for any inversion procedure which considers multiply charged aerosols), the counting process in the DMPS can be considered a Poisson process.

In our setup, we can neglect the total penetration of the system since the compared CPCs measure in parallel in the same DMPS system and the total penetration is the same for both. This allows us to compare the raw data from the CPCs without an inversion and the uncertainties related to it (Stolzenburg et al., 2022). As our DMPS outputs the average concentration during each voltage step, we need to rearrange Eq. (2) for the counted particles N. This also shows that we can predict that a factor 50 increase of Q_{opt} (effective undiluted optics flow of 0.05 lpm in the TSI 3776 versus 2.5 lpm in the modified Airmodus A20) should lead to a factor 50 increase of N:

$$N = C \cdot \tau \cdot Q_{opt} \quad (3)”$$

Lines 126-129: when reading this part, I was confused about the applicability of Poissonian statistics to your problem because N during an NPF event is a function of time and is not Poissonian. It is easy to see this if you think that for Poisson P(t1) = P(t2) for any t1 and t2 but for NPF P(t1)<P(t2) if t2>t1 (the particle number increases with time during NPF). After reading the manuscript, I understood that this is probably not a concern because you are working with narrow concentration intervals where the time dynamics likely do not play a role. However, I think it is necessary to comment on this and on the general applicability of Poissonian statistics to describe NPF events.

We thank the reviewer for thinking about the applicability of the Poisson process in that respect. We think that it is the assumption of a constant size-distribution and hence CPC inlet concentration during one voltage step of the DMPS (hence size) which justifies the applicability of the Poisson distribution. Within that measurement interval, two separate time intervals would be independent of each other.

It does not matter if during a day the measured concentration changes, as each measurement by itself is a Poisson process. We clarified this in the newly drafted Section 2.3:

"If we assume that the number concentration remains constant during the voltage scan of the DMPS (which is anyways also a requirement for any inversion procedure which considers multiply charged aerosols), the counting process in the DMPS can be considered a Poisson process."

Line 132: Any observed system is characterized by random fluctuations leading to some sort of inherent variability, but I would not classify this as 'random uncertainty'. I attribute random uncertainty to fluctuations in the measuring system (e.g., small changes in the flow rate, laser current,…).

Lines 142-145, a few comments regarding this approach:

- As mentioned before, what is the effect of the instrument response time? I would expect that if the response time is substantially different, then, with this approach, you would amplify the error. However, this is not a real error because of the different instrument transfer functions. Ideally, you should account for it before performing this analysis (especially considering that you are working with narrow time intervals).
  We thank the reviewer for his careful comment and please see our response above. The response time is no issue here, as we compare counts within time intervals of constant voltage which are significantly larger than the CPC response times of both instruments.
- Instead of considering a narrow interval, why didn't you select a single count value (e.g. pick 300 counts per unit time in the A20 and compare it with the corresponding distribution in the 3776 CPC)? This would remove the uncertainty related to the finite interval selection in the A20.
  We needed to choose a narrow interval as otherwise there would not be sufficient events at a single count value in order to obtain reasonable fits. As also requested by referee #1, we clarified this in the text, please see our response there.
- Mention explicitly that a gaussian distribution is a good approximation of a Poissonian when mu*tau is sufficiently large (>~10), which is why you are using a gaussian PDF to fit the data.
  Agreed. We added that to the text.
  "(which is a good approximation to a Poisson distribution when $E[N] > 10$)"

Line 188: Did you use the square root of N as underlying uncertainty? If so, please mention it explicitly.

We added that.

"(assuming a $\sqrt{N}$ uncertainty)"

Line 192: Replace the second "using" with a different verb and the correct form (e.g. considers).

Thanks. Corrected to "considers".

Lines 219-222: You could make the same scatter plot for periods with no NPF to exclude the sizing effect. It should be an easy check with your dataset and would probably resolve this open question (Fig. S1 already shows a difference in the counting efficiency).

This is an interesting idea. We checked it and could not find a significant improvement for the agreement when the NPF days are removed from the dataset. Therefore, we removed our statement about the cutoff uncertainty from the text.

Figure 5: to what extent can the Poissonian statistics explain the observed discrepancies in GR and J? I guess that for a quantitative answer, you would have to run the MC simulation for all events, which is not what I am asking for, but a comment on this aspect would be useful for the paper.

We thank the reviewer for that interesting question. We now indicate the counting error of the GR3-6 and J3 measurements in Fig. 5 for all three example days. Within these errorbars the obtained GR and J values all fall indeed onto the 1:1 line, which gives some indication that the observed scatter can be explained by the uncertainty from the counting error. We added the following text:

"Altogether, the counting uncertainties derived for all three days analyzed by the Monte Carlo approach can explain the observed scatter between the values derived by the two instruments (see errorbars on the three selected events in Fig. 5), which implies that the counting uncertainty is a major issue when GR and J values are compared between different instruments."

Figure 6 and Figure 7: is the GR and J distribution centered around the "real" value? It would be useful to report the value measured for the real event (or mention that is the same as the distribution mean if this is the case).

We agree with the reviewer that this information should be added to the manuscript. For the strong NPF day, the distributions are indeed well centered around the original result obtained from the actual measurement data, which we added to the text:

"(…) demonstrating the observed variations shown in Fig. 5 and with the mean of the distributions roughly centered around the original result."

For the weaker NPF days this changes, especially for the 5th of May, where the distribution is significantly offset from the result obtained from the original data. We thus added:

"At very low J3 (5th May 2017, Fig. 7), the Monte Carlo distributions for the TSI 3776 data get skewed (with the mean of the distribution also deviating significantly from the original result) and (…)"

Line 263: is the statistical uncertainty defined as one standard deviation? Please report which type of statistical uncertainty was used.

Agreed. We added "(defined as 1σ standard deviation of the Monte Carlo derived distribution divided by the initial GR_{3-6} result obtained from the actual measurement data)".

Line 281: "with" instead of "which".

Changed. Thanks.

Line 295: "the" instead of "that"

Changed. Thanks.

Section 5.3: This part is interesting because it shows that counting statistics is the main source of uncertainty for GR and J determination. You show that other CPC measurement errors can be neglected even with your upper-limit approach (you are essentially attributing all A20 measurement errors to the TSI CPC). However, this message is not very explicit, my suggestion is to restructure this section to clarify this point. I think this is an important conclusion and should be highlighted better.

We agree with the reviewer that this could be highlighted better and added the following text:

"For the events at reduced J_3 (Fig. S3 in the Supplement and Table 1) the influence of the measurement error on the size distribution-derived quantities GR_{3-6} and J_3 becomes almost negligible compared to the even higher counting uncertainties as almost no further broadening of the result distributions are observed. Altogether, this clearly demonstrates that the counting uncertainty is the dominant source of error for nucleation and growth rate determination when a TSI 3776 ultrafine CPC is used."

Line 338: I would remove 'sub-10 nm range', your findings regarding Poissonian statistics apply to any size range, and absolute counts are often lower for larger particle sizes.

Agreed and removed.

Lines 349-351: I would rephrase this because, for most practical applications (the majority of NPF studies are performed with UCPC having a small flow rate), this additional source of error seems negligible, as you have shown in the previous section. So, I would say that the additional measurement error becomes important only when using a system with high counting statistics, as in the case of the A20 CPC.

We agree with the reviewer and adjusted these sentences to:

"However, we showed that the counting uncertainty is main source of error for the size distribution-derived quantities J and GR for the widely used TSI 3776. The additional sources of uncertainty might only become important in the derivation of the nucleation and growth rates when the counting uncertainties are reduced as in the case of the modified Airmodus A20."