General statement:
This discussion manuscript is highly suitable for publication in AMT and of value to the general aerosol community. It presents the results of an intensive inter-comparison of an instrument that is widely used for particle size distribution analysis in the laboratory and ambient monitoring situations for particles in the 0.5μm and larger diameter range. The extensive data set is well presented and the statistical results are a valuable benchmark that quantify the precision and uncertainties of data from these instruments. The data set ultimately begs further analysis to determine the causes of the observed variability and, as the authors clearly point out, there is a need for a better reference standard for calibration and testing .

General questions and comments for consideration:

Given that the counting and sizing by the APS are so sensitive to flow rates in the aerosol and sheath flows a reference to the procedure used during the laboratory study for adjusting these flows would be valuable.

Can unit-to-unit variability, especially counting efficiency for particles less that 0.9μm be explained to some extent by analysis of the pulse pair types?

Page 11521 line 19
Furthermore, although no TOF-recalibration has been performed, the deviations in sizing were corrected roughly in a post-processing step.

Describe this correction in more detail.

Line 26 and following paragraphs
The large unit-to-unit variability in the sub-micron range certainly results from individual differences in unit counting efficiencies.

What parameter do you mean with respect to "variability"? Number-size distribution?
The term "detector error" needs explanation since the detector is an integrated unit of the APS consisting of optical sensor, laser source, optical components. The counting errors may be associated with low pulse height (detector sensitivity, flow alignment, optical alignment and

cleanliness or laser beam focusing) or the pulse processing algorithms.

Page 11524, line 9
Was there any significant difference in counting efficiency (variability) and lower acceptable useable size range between units for PSL vs. ammonium sulfate vs. ambient aerosols that you could evaluate?

Suggested changes to text:

Page 11515 line 5
In a laboratory study within the framework of ACTRIS (Aerosols, Clouds, and Trace gases Research Infrastructure Network), aerodynamic particle size spectrometers (APS model 3321, TSI Inc., St. Paul, MN, USA) were compared with a focus on flow rate~~s accuracy~~, particle sizing, and unit-to-unit variability of the particle number size distribution.

page 11519 line 7
No re-calibrations of the TOFs were performed. In the following sections only the results after the flow re-adjustments are analyzed.

Page 11519 line 18
This ~~fact is based on the~~may be a result of its flow re-adjustment, while the TOF calibration was untouched.
Was the sense of the change in sizing consistent with the change in acceleration flow rate? (It was in my quick analysis.) Were the flow adjustments great enough in magnitude to affect the sizing deviations by the amount observed?

Page 11523, line 24
With respect to the nominal PSL diameters, the mean deviation for the size accuracy was generally ±10 % with a systematic trend toward a negative deviation (0 to -10%) for larger particles.

11524, line 1
variability up to 60 %. This variability is most likely a result of individual counting efficiencies based on detector ~~errors~~sensitivity.

line 9
Only for the size range larger than 0.9 µm is the variability within the range of 10% for the majority of devices. Thus, without further, device-specific calibration and testing ~~further corrections~~ the size range below 0.9 µm should be rejected.