Measuring black carbon mixing states using a tandem centrifugal particle mass analyzer and single particle soot photometer system

by

Arash Naseri

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Mechanical Engineering
University of Alberta

Abstract

This work aimed to improve the evaluation of atmospheric black carbon (BC) mixing states by using a tandem Centrifugal Particle Mass Analyzer (CPMA), and Single Particle Soot Photometer (SP2) system. To derive a BC mixing state using the CPMA SP2 system, the measurements should be deconvolved with a proper inversion scheme that accounts for multiple charging of particles and physical constraints. Four objectives were followed to improve the accuracy of the method: (*i*) developing a novel deconvolution scheme; (*ii*) determining optimum measurement configurations for the CPMA SP2 system response collection; (*iii*) experimentally characterizing the CPMA transfer function; and (*iv*) incorporating CPMA-SP2 system into real-world applications to demonstrate its performance.

The first objective was achieved by developing a novel inversion scheme for determining the two-dimensional number distribution of non-refractory materials on black carbon (*i.e.,* m_{p}— m_{BC}), taking into account both physical constraints and a lack of knowledge about black carbon particle mobility. The novel approach was tested with several well-established regularization methods to determine which method outperforms the other. The results revealed that the *exponential distance method* outperformed other regularization methods tested. Also, to derive the optimum sampling settings of the CPMA SP2 responses, Monte Carlo sampling was utilized to examine

the sensitivity of reconstruction error to experimental inputs. Based on the aerosol concentration and measuring time, general recommendations for typical aged atmospheric aerosols are: (*i*) large numbers of SP2 bins per decade ($n_{\mathrm{s}}^{'} = 64$), (*ii*) a preference for higher SP2 sampling counts per CPMA setpoint ($N_{\mathrm{s}} > 10^4$), (*iii*) a moderate number of CPMA setpoints per decade ($n_{\mathrm{c}}^{'} = 3$ to 8), (*iv*) CPMA resolution around 1.0, and (*v*) a high CPMA flow rate (1.5 L/min).

For improving the evaluation of the mathematical model of the CPMA-SP2 system, which directly impacts the reconstruction accuracy of m_{p}— m_{BC} distribution, the transfer function of CPMA was experimentally evaluated and approximated with a triangular transfer function. Deviation of the idealized transfer function from the actual case was measured, and it was found that the actual CPMA transfer function is generally a little bit narrower than expected, while the losses are much higher than expected.

Finally, the improved CPMA-SP2 system was demonstrated for the measurement of wildfire BC mixing states and compared to other measurement techniques established in the literature. The comprehensive comparison between methods indicates that while the CPMA-SP2 measurement may suffer from poor temporal resolution, it enjoys higher counting statistics to retrieve information, thus the reconstructed results are less uncertain, and also information is retrieved on a larger validated detection size range.

Preface

This thesis was written as part of a research collaboration with Prof. Jason Olfert being the lead collaborator at the University of Alberta. The inversion scheme used for recovering true particle mass — rBC mass distribution using the CPMA-SP2 measurements was developed in collaboration with Dr. Timothy Sipkens, a postdoctoral fellow at the University of British Columbia at the time of the collaboration. I ran all the experiments and analyzed the data presented in this thesis myself, except Chapter 5 for which Dr. Joel Corbin helped me analyze the SP2-LEO data. Also, the tandem CPMA measurements described in Chapter 4 were facilitated by the use of software developed by Dr. Tyler Johnson that synchronized the instruments. Chapters 2 and 3 of this thesis have been published as

- Naseri, A., Sipkens, T. A., Rogak, S. N., & Olfert, J. S. (2021). An improved inversion method for determining two-dimensional mass distributions of non-refractory materials on refractory black carbon. Aerosol Science and Technology, 55(1), 104-118.

- Naseri, A., Sipkens, T. A., Rogak, S. N., & Olfert, J. S. (2022). Optimized instrument configurations for tandem particle mass analyzer and single particle-soot photometer experiments. Journal of Aerosol Science, 160, 105897.

This dissertation is dedicated to my homeland, Iran, and the free people throughout history who have bravely fought for "Women, Life, Freedom" and have sacrificed their lives in this way

Acknowledgements

Jason, my best-ever supervisor, deserves special recognition for his invaluable support and scientific leadership during this research. I have been fortunate to have a supportive supervisor who has greatly contributed to improving this study with his broad knowledge of aerosol science. My years at the University of Alberta were enriched by a supervisor I couldn't imagine a better.

I would also like to express my gratitude to Dr. Timothy Sipkens for his support, enthusiasm, and brilliant ideas relating to my research. His extensive knowledge of inversion and coding always taught me something new. My thanks to Dr. Joel Corbin for teaching me how to interpret SP2 data, and for his insightful comments during our collaboration. I would like to thank the National Research Council Canada and Dr. Allan Bertram of the University of British Columbia for providing us with their instruments and facilities for this study. My gratitude goes to the funding agencies which generously supported my research through numerous scholarships, particularly Alberta Innovates and NSERC (FlareNet Research Network).

Last but not least, I would like to give my deepest thanks to my family for all the encouragement and support they have given me throughout my entire life. I am especially grateful for my mom and dad, Maryam and Behzad, and my beloved sister, Reyhaneh, whose love and support is instrumental to my success.

Contents

List of Tables

List of Figures

Nomenclature

LIST OF SYMBOLS

$p(\mathbf{b})$	The evidence in Bayesian framework
α	Regularization parameter
$\arg\min_{\mathbf{x}}\{\ldots\}$	a minimization with respect to a quantity of interest ($\mathbf{x}$)
β	The inverse of the CPMA resolution
$\boldsymbol{\mu}$	mean of the refractory black carbon mass and particle mass in logarithmic reconstruction space
$\boldsymbol{\Sigma}$	covariance of the refractory black carbon mass and particle mass in logarithmic reconstruction space
$\frac{\mathrm{d}N}{\mathrm{dlog}m_{\mathrm{i\,AAC}}}$	The number distribution of aerosols classified by the AAC
$\frac{\mathrm{d}N}{\mathrm{dlog}\,m_{\mathrm{p}}}$	the marginal number concentration distribution of the particle mass
$\frac{\mathrm{d}N}{\mathrm{dlog}\,m_{\mathrm{rBC}}}$)	the marginal number concentration distribution of the rBC mass (*i.e.,* number distribution of rBC mass measured directly by SP2 while the CPMA is bypassed)
$\frac{\partial^2 N}{\partial \log m_p \partial \log m_{\mathrm{rBC}}}$	two-dimensional number concentration of particle-rBC mass distribution
$\Delta\tau$	the time difference between the peak intensity of the time-resolved SP2 scattering and incandescence signals
$\delta(\cdot)$	the δ-Dirac function

η	The counting efficiency of the CPC at a CPMA mass setpoint
Γ^{D}	a correlation matrix describing the correlation between points in the reconstruction
λ	Loss factor of the CPMA trafer function
$\mathbf{A}$	discretized CPMA-SP2 kernel function
$\mathbf{b}$	discretized CPMA-SP2 response
$\mathbf{D}_{i,j}$	the Mahalanobis distance between pairwise (i and j) points in the reconstruction domain
$\mathbf{L}^{\mathrm{pr}}$	the Cholesky factorization of the inverse covariance matrix of prior
$\mathbf{L}^{\mathrm{TK}}$	a discretized differential operator (*i.e.,* Tikhonov matrix)
$\mathbf{r}$	any point in the reconstructed particle mass-rBC mass distribution
$\mathbf{x}$	descretized quantity of interests (two-dimensional number concentration of particle-rBC mass distribution)
$\mathbf{x}^{\mathrm{ex}}$	The exact solution
$\mathbf{x}^{\mathrm{ph}}$	The ground truth
$\mathbf{y}_{\mathrm{t}}$	Measured data
μ	Width factor of the CPMA trafer function
$\odot$	the elementwise product
Ω	Analytical CPMA transfer function
ω	Rotational speed of CPMA Classifier
ω_{c}	Rotational speed at the center of the gap between the cylinders of the CPMA classified
$\det(\cdot)$	denotes the determinant
$\overline{\mathbf{y}_{\mathrm{t}}}$	The series mean of measured data

Φ	Integer charge state of particle
ρ_{rBC}	rBC material density
ρ_{coating}	Material density of coating
Σ^{pr}	the prior covariance
σ_{p}	geometric standard deviation of particle mass
σ_{rBS}	geometric standard deviation of rBC mass
$\tilde{m}$	Non-dimensional particle mass in tandem CPMA measurments
ε_i	CPMA-SP2 measurement error
$\widehat{\mathbf{y}}_{\mathrm{t}}$	Predicted values by the fitting model
${}_{adj}R^2$	Adjusted R-square of the fitting model
$B_{\max}$	The mobility of aerosols with the maximum mass that transverses the CPMA classification region at a given mass setpoint
d_{opt}	optical diameter of particle
d_{rBC}	mass equivalent diameter of the rBC portion of the particle
d_m	Mobility diameter
E	Electric field strength between two concentric cylinders
e	The elementary charge of the electron
f	the charge fraction
f_{coating}	the mass fraction of non-rBC (coating) in the particles
f_r	relaxation factor in Towomey-Markowski regularization scheme
h	The height of the CPMA transfer function
$H(\cdot)$	the Heaviside function
K_{i}	CPMA-SP2 kernel for the i^{th} CPMA setpoint
L	The length of the CPMA classification region

l_{P}	the correlation lengths for particle mass
l_{rBC}	the correlation lengths for rBC mass
m_{12}^{*}	The mass-offset factor between the two CPMA setpoints in the tandem CPMA measurement
m_{1}^{*}	Upstream CPMA mass setpoint in tandem CPMA measurments
m_{2}^{*}	Downstream CPMA mass setpoint in tandem CPMA measurments
$m_{\mathrm{p,g}}$	the geometric mean of particle mass
m_{P}	Total particle mass
m_{p}	the mass of a particle subjected to the CPMA-SP2 system
$m_{\mathrm{rBC,g}}$	the geometric mean of rBC mass
m_{max}	The maximum mass that transverses the CPMA classification region at a given mass setpoint
m_{rBC}	Refractory black carbon mass
$n_{\mathrm{c}}^{\prime}$	The number of CPMA setpoints per decade
$n_{\mathrm{s}}^{\prime})$	the number of SP2 bins per decade
N_{2}	Number concentration of the twice-classified aerosols in tandem CPMA measurement
N_{1}	Number concentration of aerosols passing through the upstream CPMA in tandem CPMA measurement
$N_{\mathrm{P},i}$	the number concentration of refractory black carbon particles at the i^{th} CPMA setpoint
N_{s}	Number of SP2 counts per CPMA setpoint
$N_{\mathrm{s.avg}}$	Average number of particles collected over all of the CPMA setpoints
N_{tot}	total number concentration of rBC-containing particles

p_{li}	the likelihood in Bayesian framework
p_{po}	the posterior distribution in Bayesian framework
p_{pr}	The prior in Bayesian framework
Q	The aerosol sample flow rate through the CPMA
r	Radii of the CPMA classifier
R^2	R-square of the fitting model
r_{c}	Radii at the center of the gap between the cylinders of the CPMA classified
R_{12}	the correlation between the logarithm of particle and rBC
R_{s}	The size resolution parameter of the AAC
R_m	CPMA resolution
S	Effective SP2 kernel
$S(m_{\mathrm{rBC}})$	an effective SP2 kernel
S_f	Markowski smoothing factor in MART and Towomey-Markowski regularization schemes
t_{coating}	coating thickness
$t_{\mathrm{max,incand}}$	time of the peak intensity of SP2 incandescence signal
$t_{\mathrm{max,scat}}$	time of the peak intensity of SP2 scattering signal
t_{so}	The SP2 sampling time when the CPMA is bypassed in the CPMA-SP2 measurement
V	Electric potential difference across the cylinders,

ACRONYMS

AAC	Aerodynamic Aerosol Classifier
AMS	Aerosol Mass Spectrometer
ANOVA	Analysis of Variance
APD	Avalanche Photo Diode
APM	Aerosol Particle Mass Analyzer
BBHG	Broad Band High Gain
BBLG	Broad Band Low Gain
BC	Black Carbon
CCN	Cloud Condensation Nuclei
CDF	Cumulative Distribution Function
CI	Confidence Intervals
CMD	Count Median Diameter
CPC	Condensation Particle Counter
CPMA	Centrifugal Particle Mass Analyser
DMA	Differential Mobility Analyser
EC	Elemental Carbon
FWHM	Full Width Half Maximum
GSD	Geometric Standard Deviation
LEO	Leading Edge Only
LII	Laser Induced Incandescence
MAE	Mean Absolute Error
MART	Multiplicative Algebraic Reconstruction Technique
ME	Mean Error
MMD	Mass Median Diameter

Nd:YAG	Neodymium doped Yttrium Aluminium Garnet Nd: Y_3Al5O_{12}
ND	Normalized Derivative
PM2.5	Particulate Matter less than 2.5 microns in diameter
PSL	Polystyrene latex
rBC	refractory Black Carbon
RMSE	Root Mean Square Error
SASS	Scanning Aerodynamic Aerosol Sizer
SMPS	Scanning Mobility Particle Sizer
SOA	Secondary Organic Aerosol
SP-AMS	Soot Particle Aerosol Mass Spectrometer
SP2	Single Particle Soot Photometer
TEM	Transmission Electron Microscopy
VOC	Volatile Organic Compound

Chapter 1

Introduction, Problem Statement, and Thesis Objectives

1.1 Atmospheric aerosols

A population of liquid or solid particulate matter suspended in the atmosphere is called *atmospheric aerosol*. Atmospheric aerosols, particularly originating from the burning of fossil fuels, contribute considerably to both human morbidity/mortality and climate change (Pachauri et al., 2014). Air pollution currently contributes to a global human mortality rate of 5.5 million premature deaths per year (Brauer et al., 2016). Atmospheric aerosols with a diameter less than 2.5 μm, also referred to PM2.5, have been singled out as the most hazardous portion of atmospheric aerosol populations (Robinson, 2020). Atmospheric aerosols are closely linked with a diverse range of severe and long-lasting illnesses, more importantly, cardiovascular and respiratory diseases (Burnett et al., 2014), and to less extent, mental illnesses (Oudin et al., 2016), skin-related diseases (Drakaki, Dessinioti, and Antoniou, 2014), and sexual health (Sram, 1999). Atmospheric aerosols also have a substantial impact on many critical processes in atmospheric chemistry and physics. Aerosols extensively

influence the dynamics of the climate in either direct and indirect ways. Aerosols scatter and absorb solar radiation, and are regarded as *direct contributors* to climate change (Bellouin et al., 2005; Ravishankara, Rudich, and Wuebbles, 2015), *e.g.,* black carbon (BC) particles heat the atmosphere by directly absorbing sunlight (Bond and Bergstrom, 2006; Jacobson, 2001). Also, aerosols indirectly influence large-scale atmospheric dynamics (Penner, Dong, and Chen, 2004) via providing nucleation sites for water droplets to form, resulting in cloud formation (Andreae and Rosenfeld, 2008; DeMott et al., 2010, 2011; Farmer, Cappa, and Kreidenweis, 2015; Jacobson, 2006; Qi et al., 2017), causing *indirect effects* on climate by enhancing the global albedo which lessens the amount of solar radiation striking the Earth surface; and as sites for heterogeneous reactions to occur (Lee et al., 2017), changing not only aerosol properties but also gaseous concentrations (Abbatt, Lee, and Thornton, 2012). Also, BC particles settle on snow/ice, cause more heating by decreasing snow/ice albedo (He et al., 2014; Lee et al., 2017; Liou et al., 2014), causing faster snow/ice melting resulting in less snow-covered land reflecting sunlight back to space.

1.2 Refractive black carbon and its impact on the climate change

Produced primarily by the incomplete combustion of fossil fuel or biomass (*e.g.,* wildfires), BC is estimated by some to be the second-largest human-caused contributor to global warming after carbon dioxide (Bond et al., 2013; Ramanathan and Carmichael, 2008). Black carbon is aggregates of spheroidal particles within the range of 10 and 50 nm in diameter (Bellouin et al., 2005; Ravishankara, Rudich, and Wuebbles, 2015), which are strongly light-absorbing and insoluble in water and

atmospheric organic solvents.

To determine the degree to which worldwide climate is sensitive to individual atmospheric constitutions, *e.g.,* BC-containing aerosols, the Intergovernmental Panel on Climate Change (Houghton et al., 1995; Houghton, 1992; Houghton, Jenkins, and Ephraums, 1990) defined a metric known as *radiative forcing*. Commonly defined for the Earth surface-troposphere system, the radiative forcing [W/m^2] is referred to the change in the net irradiance at the tropopause after letting stratospheric temperatures reach radiative equilibrium, while both Earth surface and tropospheric temperatures are kept fixed at their initial unperturbed values. However, estimation of the radiative forcing for some atmospheric aerosols like BC-containing particles is not always easy, and is subject to a high degree of uncertainty (Prather, Hatch, and Grassian, 2008; Ramanathan and Carmichael, 2008). One primary source of uncertainties is the diversity and variation of physicochemical properties and morphology of particles across the aerosol population during their evolution in the atmosphere. Discriminating this diversity of particle physical characteristics and chemical compounds within the atmospheric aerosol population is highly challenging and is further complicated as the physical and chemical properties of particles are continuously changing within their atmospheric lifespan. Despite these immense challenges for both experimental measurements and modeling, it is essential to evaluate the impacts of aerosols on the atmosphere and climate. For instance, some atmospheric aerosols are composed of complex mixtures of various chemical substances (*e.g.,* a single 100-nm particle could comprise of thousands of distinct chemical substances, Junge, 1952; Murphy and Thomson, 1997; Prather, Hatch, and Grassian, 2008) and could be composed of both primary and secondary species. Primary species are those

species that are produced during emission and reside in the condensed phase, secondary species are formed later in the atmosphere through chemical reactions. Two common examples of primary particles in the atmosphere are combustion aerosols such as BC particles and metals from unburnt lubricating oils (Toner et al., 2008). Secondary species in the atmosphere are typically nitrates, sulfates, and secondary organic aerosol (SOA), produced respectively from the oxidation of NO_x, SO_2, and volatile organic compounds (VOCs) (Hallquist et al., 2009; Winkler, 1973). SOAs are formed through the photo-oxidation of VOC's and radicals in the atmosphere.

Winkler (1973) introduced the concept of *aerosol mixing state* to define the diversity in chemical composition across an atmospheric aerosol. He considered two extremes; (*i*) *internal mixing* and (*ii*) *external mixing*. The former refers to a mixing state where the population of particles within an aerosol are all composed of the same chemical species with the same fractional composition. The latter represents states that all particles within an aerosol are made up of pure chemical substances and have separate compositions. As BC-containing particles undergo constant evolution during their lifespan in the atmosphere due to coagulation processes and condensation of secondary aerosols, ambient BC mixing states are found neither fully internally nor externally mixed. Primary BC particles are often found in external mixing states, and they are primarily water-insoluble and hydrophobic (Khalizov et al., 2013; Miljevic et al., 2012), and they cannot act as cloud condensation nuclei and form cloud droplets. As shown in Figure 1.1, after being emitted from a combustion source, BC aggregates grow coatings of secondary aerosol and become considerably restructured, bringing about significant changes to BC particle optical properties (He et al., 2016; Moteki, Kondo, and Adachi, 2014), morphology (Cappa et al., 2012; Weingartner,

Burtscher, and Baltensperger, 1997), particle physical and chemical properties (*e.g.,* particle size, and surface chemical properties, Khalizov et al., 2013). The structural change happens when BC particles are internally mixed with (*i.e.,* BC surfaces are coated by) other atmospheric aerosol components via condensation or chemical oxidation processes (He et al., 2016; Khalizov, Cruz-Quinones, and Zhang, 2010; Moteki, Kondo, and Adachi, 2014). When BC particles are internally mixed with hygroscopic compounds (*i.e.,* the materials attracting water in either vapor or liquid phase from their environment), they act as cloud condensation nuclei (CCN), which directly and indirectly influence the radiative forcing of BC particles by forming cloud droplets and altering cloud lifetimes. Nevertheless, practically, it is not common to find exclusively internal or external aerosol mixing states in the atmosphere (Bondy et al., 2018; Healy et al., 2014; O'Brien et al., 2015).

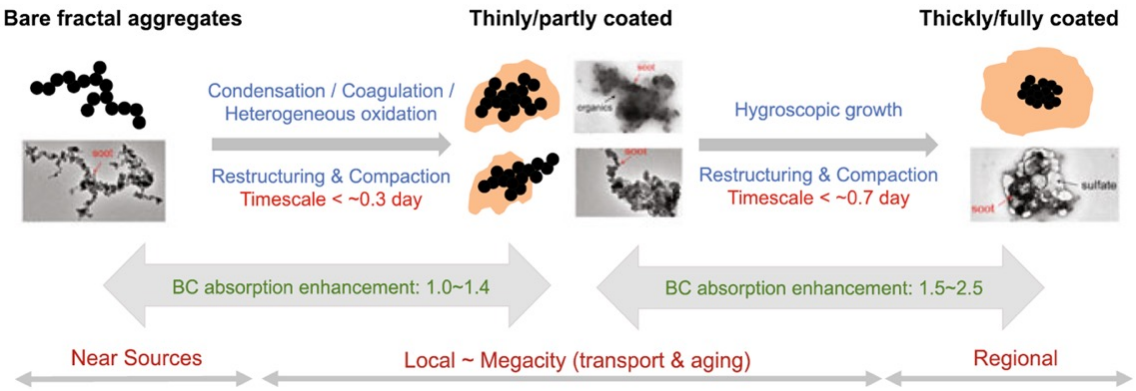

Figure 1.1: A summary of BC structure evolution during its aging process in the atmosphere (adapted from Kokhanovsky (2019))

BC mixing states in the atmosphere and, consequently, their radiative forcing are not fully characterized as a result of BC spatiotemporal alterations (due to various degrees of emissions in different regions), and its intricate microphysical properties (*e.g.,*

particle morphology and size distribution, Lee et al., 2015; Stephens, Turner, and Sandberg, 2003) that are constantly changing during the aging process. Although considerable attention has been paid to establish more accurate climate models to calculate BC optical properties and its radiative forcing (Andreae and Gelencsér, 2006; Schwarz et al., 2010), climate models are still unable to explain the extensive variations in BC radiative forcing. Consequently, BC mixing states need to be measured to lessen the uncertainties in climate forcing due to BC. The single particle soot photometer (SP2) is often used to measure atmospheric BC concentrations and mixing states of BC-containing particles (Stephens, Turner, and Sandberg, 2003). How the SP2 operates and how it measures BC mixing states will be explained in detail in the next section.

1.2.1 The single particle soot photometer measurements

The SP2 measures the mass of refractory black carbon (rBC) in individual particles, where refractory black carbon is defined as carbonaceous material which is insoluble and vaporizes only at temperatures near 4000 K (Petzold et al., 2013). In the operation of the SP2, as shown in Figure 1.2, rBC-containing particles are directed toward a laser cavity and are hit by an intracavity laser beam (*i.e.,* Nd: YAG, Neodymium-doped Yttrium Aluminum Garnet), causing the rBC portion to scatter and absorb the incident light and heat up, whereas the non-rBC portion only scatters the incident light. Increasing the particle temperature results in simultaneously inducing vaporization of the non-rBC portion and the emission of laser-induced incandescence (LII), *i.e.,* blackbody radiation (Zhang et al., 2018b).

Compound lens assemblies direct scattered and incandescent light to four distinct

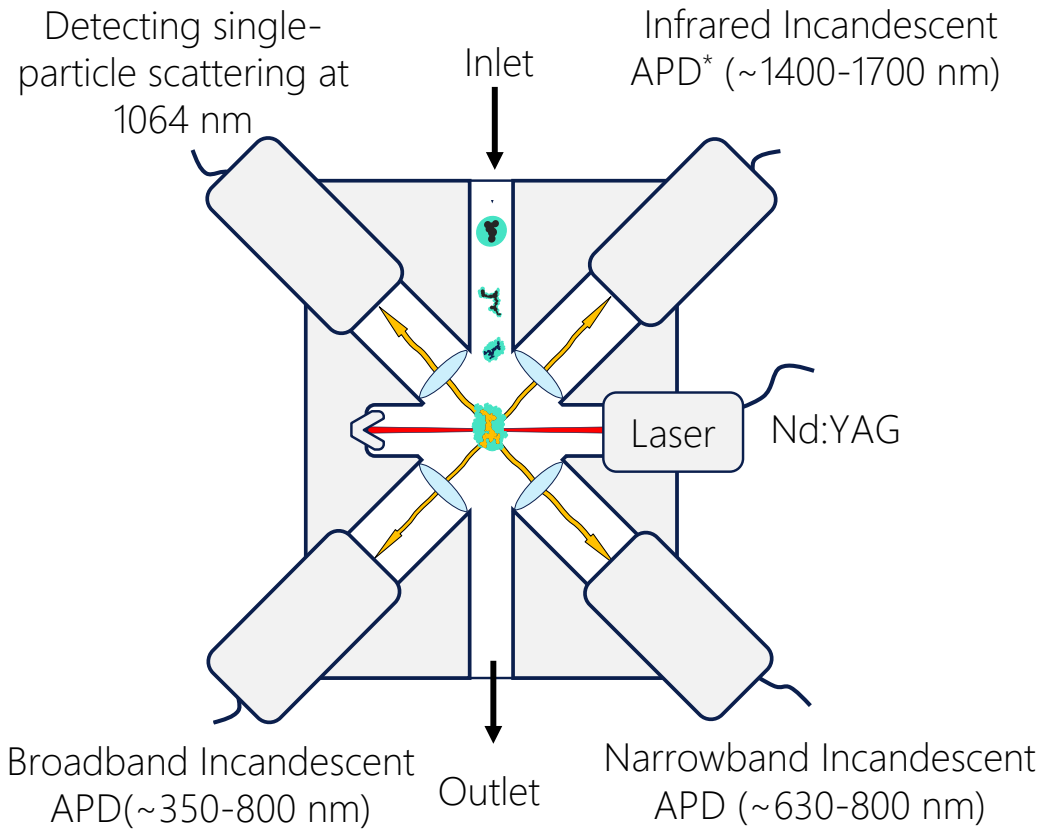

Figure 1.2: A simple schematic representative the operating principles of the SP2

avalanche photodiodes (APDs) which convert the light to electricity using the photoelectric effect. The SP2 is equipped with four APDs: one for measuring scattered light, two for the detection of incandescence light in the visible light range (with a wavelength $\sim$ 350–800 nm and 630–800 nm), and one for incandescence light within the infrared range ($\sim$ wavelength of 1400–1700 nm). The peak LII intensity (*i.e.,* the blackbody radiation signal) in the SP2 occurs just before particle vaporization at $\sim$ 4000 K, and is correlated to the BC mass of a particle using information derived from the calibration of LII amplitude to particles of a known BC mass (Stephens,

Turner, and Sandberg, 2003). The SP2 also monitors particle light scattering from its 1064 nm laser. For non-rBC particles, the maximum light-scattering signal can be used to measure an optical-equivalent particle size. However, for large non-rBC particles (*i.e.,* $>\sim$ 400 nm), the detector saturates and direct correlation of the peak height and particle size is not possible. Additionally, this size measurement is also confounded for rBC-containing particles because the non-rBC material evaporates due to the heating of the rBC during the passage of the particle through the laser beam. Therefore, the scattering signal decreases as the non-rBC material evaporates, and after evaporation, the scattering signal from the rBC core will vary depending on its location in the laser beam. Furthermore, the scattering signal of the commercial SP2 has a narrow detection range ($\sim$ 200–400 nm in optical-equivalent diameter, which depends on the laser power, optical alignment, and the sensitivity of the detectors (Schill et al. (2018), Sharma et al. (2017), Wu et al. (2021), and Zhang et al. (2018c)) and predicting the amount of non-rBC material requires assumptions for morphology (*e.g.,* core-shell morphology), material density and refractive indices which are highly uncertain (Cappa et al., 2012; Taylor et al., 2015; Zhang et al., 2016). Nevertheless, SP2 users attempt to use the scattering signal to estimate the amount of non-rBC material in the rBC-containing particles and large non-rBC particles. Three common methods are: (*i*) leading-edge only (LEO; Gao et al. (2007)), (*ii*) normalized derivative (ND; Moteki and Kondo (2008)), and (*iii*) incandescence lag-time (Moteki and Kondo, 2007). Fundamentals of these measurement methods are briefly described in the next sections, and more detail can be found in studies of Moteki and Kondo (2007, 2008).

1.2.1.1 Leading-edge only

In the literature, the amount of coating material in the rBC-containing particle is typically represented as a coating thickness (t_{coating}), which is calculated as

$$t_{\text{coating}} = \frac{d_{\text{opt}}(\text{RI}) - d_{\text{rBC}}}{2}, \tag{1.1}$$

where, $d_{\text{opt}}(\text{RI})$ is the optical diameter of the particle (a function of refractive index, RI), and d_{rBC} is the mass equivalent diameter of the rBC portion of the particle. According to the core-shell type morphology assumption and assuming a core material density (*i.e.,* $\rho_{\text{rBC}} = 1800\ \text{kg/m}^3$), the amount of m_{rBC} for each particle, which is directly derived from the incandescence signal, is converted to d_{rBC};

$$d_{\text{rBC}} = \left(\frac{6m_{\text{rBC}}}{\pi\rho_{\text{rBC}}}\right)^{1/3}. \tag{1.2}$$

To estimate d_{opt}, the scattering cross-section (Leighton, 1994), is needed to be determined as an input to the Mie theory. The scattering cross-section of the particles, can be estimated from the peak intensity of the Gaussian scattering signal, which was calibrated with measurements of PSL of known size. There is, however, a confounding effect for scattering size measurements of rBC-containing particles, such that as coated rBC particles pass through the laser beam, the rBCs heat up, and the coating material evaporates, resulting in a lower peak intensity of the scattering signal. Accordingly, the peak intensity of the scattering signal cannot be directly derived from the originally detected signal and requires the reconstruction of the undistorted scattering signal. Developed by Gao et al., 2007, the leading-edge-only (LEO) approach reconstructs the expected scattering signals of rBC-containing particles by fitting a Gaussian function to the part of the scattering signal that precedes the evaporation

of the coating material, *i.e.,* the leading edge of the scattering signal. We defined the leading edge of the scattering signal as 3% of the maximum laser intensity based on scatter plots of LEO and standard analysis for non-absorbing particles, similar to previous studies (Gao et al., 2007; Taylor et al., 2015).

Apart from the reconstructed scattering signal, several assumptions, like the refractive indices of the coating and rBC core, are required as inputs to Mie theory to obtain the optical diameter (scattering-equivalent size) of rBC-containing particles. In reality, the exact value of rBC's refractive index is unknown and may vary from one rBC material to another; nonetheless, it has been empirically shown that $\kappa \approx n - 1$, (Bond and Bergstrom, 2006; Moteki, Kondo, and Nakamura, 2010). Taking this constraint into account, the SP2 Toolkit uses a range of lookup tables containing scattering cross-sections at the corresponding SP2 laser wavelength ($\lambda = 1064$ nm) for diverse core and coating refractive indices. The user then selects the most appropriate rBC refractive index by ensuring that the rBC volume predicted from scattering by uncoated rBC particles is similar to the rBC volume predicted from the incandescence signal. Uncoated rBC particles are defined from the light-scattering signal just prior to incandescence onset, when coatings are assumed to have evaporated. In our study, this resulted in $n_{\mathrm{rBC}} = 2.26 + i\ 1.26$. It should be noted that this approach must not be mistakenly interpreted as quantifying the rBC refractive index, which requires more supplementary measurements, but as a mechanism to ensure internal consistency in the LEO results.

The LEO approach is commonly employed to derive the size or mixing state of rBC-containing aerosols (Raatikainen et al., 2017; Sedlacek III et al., 2012; Zhang et al., 2018a,b). However, the scattering detection range for pure scattering particles

(with no rBC portion) is narrow (*i.e.,* $\sim$ 200–500 nm). The size range for LEO measurements of rBC-containing aerosols depends not only on the size of the particle, but also on the amount of rBC in the particle which affects the refractive index and the evaporation rate of the coating portion.

To measure coating thickness/mass of rBC-containing particles using the LEO method, a spherical core-shell morphology with several core/shell property assumptions are required, namely, (*i*) a spherical void-free rBC core and values for: (*ii*) rBC material density (ρ_{rBC}), (*iii*) coating material density (ρ_{coating}), (*iv*) core refractive index (n_{rBC}), and (*v*) coating refractive index (n_{coating}). The following values are often used in the literature: ρ_{rBC}= 1800 kg/m^3 (Bond and Bergstrom, 2006; Corbin et al., 2018; Liu et al., 2020; Moteki, Kondo, and Nakamura, 2010; Ouf et al., 2019; Park et al., 2004), ρ_{coating} = 1000 kg/m^3 (Ditas et al., 2018; Liu et al., 2019), $n_{\mathrm{rBC}} = 2.26 + i\ 1.26$ (Dahlkötter et al., 2014; Laborde et al., 2013; Moteki, Kondo, and Nakamura, 2010; Taylor et al., 2015; Zanatta et al., 2018), and $n_{\mathrm{coating}} = 1.5 + i\ 0$ (Laborde et al., 2013; Liu et al., 2015; Nakayama et al., 2010; Yuan et al., 2021). It is well-known that the core-shell assumption is physically inaccurate (Sedlacek III et al., 2012) and the accuracy of LEO in determining coating thickness will depend on the four property values used in the calculation (Taylor et al., 2015; Zhang et al., 2016).

1.2.1.2 Normalized Derivative

The normalized derivative approach (Moteki and Kondo, 2008) in evaluating rBC mixing states is similar to the LEO approach except a different methodology is used to estimate the undisturbed particle diameter. While a split-detector signal

is used in the LEO to derive the particle positional information, which is essential to recovering the undisturbed scattering signals, the *normalized derivative* of the scattering signal is used in the ND approach to obtain analogous information. Thus, the ND approach is similar to the LEO approach except that the detection limits and noise in the optical sizing of the coated particle may be slightly different.

1.2.1.3 Lag-time analysis

Lag-time analysis is a categorical approach in determining a general picture of the mixing state of BC. SP2 scattering signals can have two peaks, with the second one occurring almost at the peak of the incandescence. While the second peak of the scattering signal is typically higher than the first peak in thinly to moderately coated rBC particles, the opposite is true for thickly coated rBC particles. For the latter, the global maximum of scattering signals occurs well before the peak intensity of the incandescence signals. Thus, the time difference between the peak intensity of the time-resolved SP2 scattering and incandescence signals ($\Delta\tau = t_{\mathrm{max,scat}} - t_{\mathrm{max,incand}}$) is used to categorize rBC-containing particles into two groups: (*i*) rBC particles with *thick* coating, or (*ii*) *thin-to-moderate* coating (Corbin et al., 2018; Moteki and Kondo, 2007; Schwarz et al., 2006; Subramanian et al., 2010). Generally, particles are categorized as rBC with thick coating if the time lag between the peak intensities of the scattering and incandescence signals is more than ~ 2 μs, indicating considerable loss of coating material due to the heat absorbed by the particles as they transverse the laser beam. For rBC particles with thin-to-moderate coating, the peak scattering and incandescence signals occur nearly coincidentally (*e.g.,* time lag < 2 μs).

One limitation of this method is that it cannot categorize rBC-containing particles

whose rBC core and coating material are fragmented by laser light (Moteki and Kondo, 2007; Sedlacek III et al., 2012). In such a case, the rBC portion reaches its boiling point and evaporates, while the unevaporated non-rBC portion is fragmented into smaller particles. A key to recognizing such a phenomenon is that the scattering signal does not entirely vanish after the incandescence signal goes to zero. In addition to the mentioned difficulty, the lag-time method cannot distinguish between uncoated and thinly-to-moderately coated rBC particles. For more complex categorizations, advanced methods should be implemented, *e.g.,* supervised machine learning method (Lamb, 2019).

Depending on the rBC-containing particles measured and the SP2 used, the time-lag threshold varies slightly. For instance, Liu et al. (2022) and Zhang et al. (2018c) used the threshold of $\Delta\tau = 1.8$ μs and 1.6 μs, respectively, while $\Delta\tau = 2.0$ μs was used by Wang et al. (2014b) and Wu et al. (2016). Nevertheless, this variation, which may be caused by the different coating compositions in different regions, is almost negligible, *e.g.,* less than $\sim 0.4\,\mu$s (Moteki and Kondo, 2007), regardless of the atmospheric conditions and pollution level (Laborde et al., 2012; Wu et al., 2016, 2017).

1.2.2 Soot particle aerosol mass spectrometer

The soot particle aerosol mass spectrometer (SP-AMS), which is analogous to having an Aerosol Mass Spectrometer (AMS) and an SP2 in tandem, is designed to measure the mass distribution of rBC and organic components in terms of vacuum aerodynamic diameter, as well as determine the chemical characteristics of sources (Onasch et al., 2012). A schematic of SP-AMS is shown in Figure 1.3. Initially,

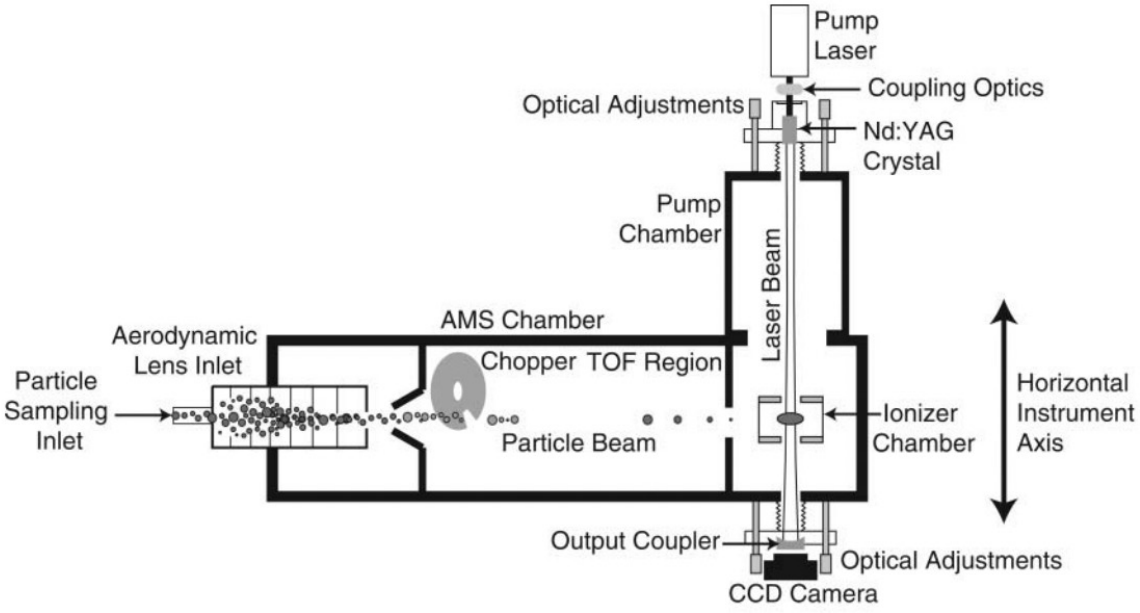

Figure 1.3: A schematics of SP-AMS instrument (Onasch et al., 2012).

aerosols are directed to the vacuum chamber through the inlet aerodynamic focusing lens system, which comprises several orifice lenses that focus the particles into a very narrow aerosol beam. The vacuum chamber (*i.e.,* time-of-flight tube) is essentially the particle sizing compartment, where the particle aerodynamic diameter is determined by measuring the velocity of particles passing through the chamber. The aerodynamic diameter of aerosols measured under the vacuum condition (*i.e.,* in the free-molecular regime) is referred to as the vacuum aerodynamic diameter (d_{va}) (DeCarlo et al., 2004). Afterward, particles enter the aerosol composition detection chamber, where particles are passed through a tungsten plate heated to approximately 600°C and Nd:YAG laser beam to evaporate non-refractory (*e.g.,* organic compounds) and refractory components (*e.g.,* rBC), respectively. In addition to

measuring the scattering and incandescence signals of refractory particles, similar to the approach explained for SP2-only measurements in Section 1.2.1, the vaporized materials are then ionized and collected for determining chemical properties by mass spectroscopy. The fundamental SP-AMS measurements return an average aerosol mass spectrum which are used to derive the average mass distribution of particles in terms of vacuum aerodynamic diameter (*i.e.,* $\frac{\mathrm{d}M}{\mathrm{dlog}d_{\mathrm{va}}}$)(Jayne et al., 2000; Onasch et al., 2012).

1.2.3 tandem CPMA-SP2 measurements

The aforementioned measurement techniques, however, are either highly uncertain (*i.e.,* indirect measurements based on assumptions Moteki and Kondo, 2007), or have a poor range (*e.g.,* SP-AMS with a detection limit of 25 fg of rBC per particle, which is far larger than the median particle mass for most atmospheric rBC containing aerosols Lee et al., 2015). The purpose of this thesis is to develop another method of measuring BC mixing state that overcomes the limitations of the prescribed approaches, such that it requires no optical model nor assumptions to obtain the total particle mass, and can measure the entire population of rBC-containing particles which can be measured by the SP2. Both of these advantages can be achieved when the measurement method does not require the SP2 scattering signal to determine the total particle mass. Alternately, a centrifugal particle mass analyzer (CPMA) can be placed in front of the SP2 to classify particles by their mass-to-charge ratio prior to SP2 measurements. Consequently, the SP2 determines the mass of rBC in each particle whose total mass (BC and non-rBC mass) is known from CPMA classification (Liu et al., 2017).

1.2.3.1 Operating principles of the centrifugal particle mass analyzer

A summary of how the CPMA operates is briefly described here, and more detail can be found in Olfert and Collings (2005). The CPMA is a particle mass classifier made of two coaxial rotating cylindrical electrodes with a voltage difference applied between them (the CPMA schematics are shown in Figure 1.4). Coaxial electrodes,

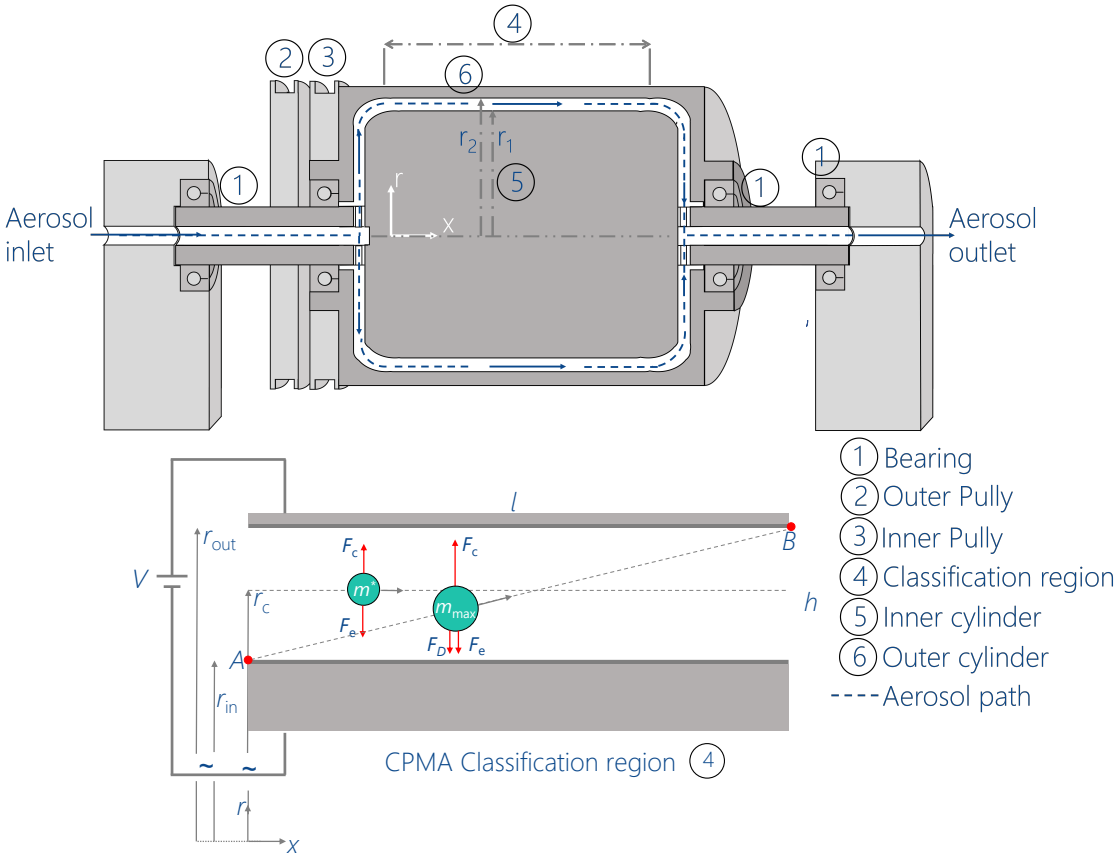

Figure 1.4: Simple schematics of CPMA operating principles. F_{D}, F_{C}, and F_{E} represent drag force, centrifugal force and electric force, respectively, that may act on the particles passing through the classification regions.

which rotate with different rotational speeds, impose centrifugal and electrical forces

on charged particles simultaneously. Particles, based on their mass and electric charge, experience different centrifugal and electrical forces, in which the centrifugal force acts toward the outer cylinder, whereas the electrical force acts toward the inner electrode. If the electrical and centrifugal force are nearly balanced, the particle will exit the classification region. There are three fundamental concepts to consider in CPMA analysis and data interpretation:

i. Aerosols with a mass-to-charge ratio marginally higher and lower than the selected mass transverse the classifier. The probability of an aerosol of a chosen mass-to-charge ratio passing through the CPMA classification region is described by a distribution, known as a transfer function.

ii. The CPMA sorts aerosols by *mass-to-charge ratio*, and the classified aerosols are not strictly monodisperse in mass. The reason being is, in addition to singly charged particles of the desired mass, doubly charged particles of twice the desired mass (and particles up to the maximum number of charges in the aerosol) will also pass through the classification region.

iii. When the CPMA operates at low rotational speeds, for instance, when it classifies fairly large aerosols, small uncharge particles may not be removed and can successfully transverse the classifier.

1.2.3.2 CPMA-SP2 tandem measurements

As a morphology-independent measurement, a tandem CPMA-SP2 system can be used to determine the mixing state of BC-containing particles, such that the mass of rBC (m_{rBC}) and the total mass of each particle (m_{p}) can be measured, regardless

of the particle morphology and coatings (Liu et al., 2017). A schematic of this measurement system is shown in Figure 1.5.

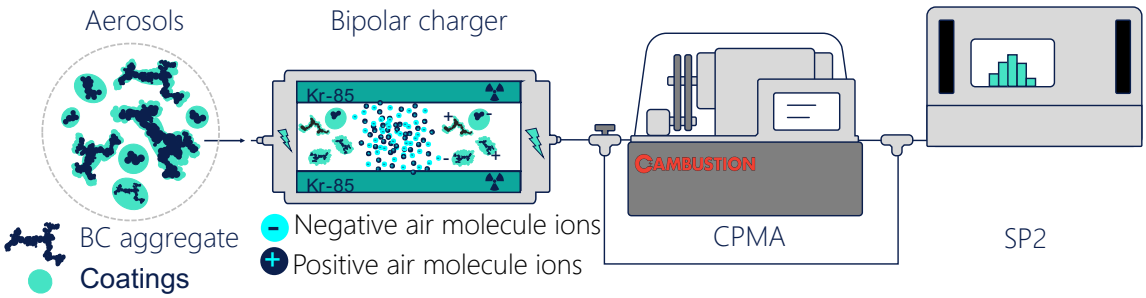

Figure 1.5: A simple schematics of the tandem CPMA-SP2 measurements.

In the system, aerosols are passed through a neutralizer to be electrically charged and are then directed toward the CPMA inlet. The CPMA is successively stepped through a desirable mass range, and the SP2 measures the mass of rBC in each CPMA-classified particle. Figure 1.6*a* and *b* show the normalized SP2 responses (concentrations) at different CPMA setpoints (m_{p}^{*}) for sample smoke particulates of BC with no coating and BC with heavy coating, respectively. In order to take the advantage of such morphology-independent measurements to determine the mixing states of rBC particles, some studies (Hu et al., 2021b; Liu et al., 2017; Liu et al., 2022), calculated the amount of non-BC material contained in the ensemble of particles by

$$m_{\mathrm{non-rBC}} = m_{\mathrm{p}}^{*} - m_{\mathrm{rBC}}, \tag{1.3}$$

where $m_{\mathrm{non-rBC}}$ is the mass of non-BC material in a BC-containing particle. Although this measurement system eliminates scattering signal-associated uncertainties and substantially improves a particle detection range of SP2-only methods, the ap-

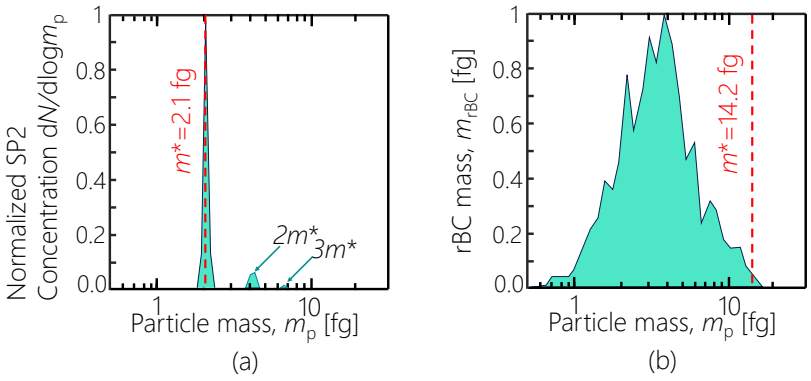

Figure 1.6: Normalized SP2 responses (concentrations) at different CPMA setpoints (m_{p}^{*}) for a sample (a) uncoated and (b) heavily coated distribution. Multiple modes in panel (a) correspond to the multiple integer charge states. The red dotted-dash lines represent the location where m_{rBC} equals m_{p}.

proach taken by these studies has two major issues. Neither the width of the CPMA transfer function nor the concept of multiply charged particles (as discussed in Section 1.2.3.1) are taken into account in this approach. Using this method, the results would be valid only if the CPMA transfer function was extremely narrow and almost all multiply-charged particles could be either removed by inspection or their portion are insignificant. However, the atmospheric rBC population could have a wide range of coating mass (*i.e.,* some rBC contain little or no coating while some may contain significant amounts) and fractions of the multiply-charged particle are not negligible. Figure 1.6*b* illustrates an example of atmospheric rBC population with a wide range of coating masses, making it virtually impossible to distinguish the multiply-charged particles by inspection. As a result, the CPMA-SP2 measurements cannot directly be used to determine the amount of non-rBC content.

Accordingly, in order to address these difficulties in using the CPMA-SP2 responses and rather derive average values for rBC mixing states, Broda et al. (2018) used tandem CPMA-SP2 measurements and set up a system of equations in such a way as to derive a two-dimensional map of total particle mass (m_{p}) and rBC mass (m_{rBC}), *i.e.,* $\partial^2 N/\partial \log m_{\mathrm{p}} \partial \log m_{\mathrm{rBC}}$ by linking the CPMA-SP2 responses with the mathematical model of the measurement system. This link can be achieved by using Fredholm integral equations (Kandlikar and Ramachandran, 1999). Applying Fredholm theory to the CPMA-SP2 tandem system, the system of equations becomes:

$$N_{\mathrm{R},i} = \int_0^\infty \int_0^\infty K_i\left(m_{\mathrm{p}}, m_{\mathrm{rBC}}\right) \left(\frac{\partial^2 N}{\partial \log m_{\mathrm{p}} \partial \log m_{\mathrm{rBC}}}\right) \mathrm{d}\log m_{\mathrm{rBC}} \mathrm{d}\log m_{\mathrm{p}} + \varepsilon_{i.}, \quad i = 1, 2, \ldots, n \tag{1.4}$$

where $N_{\mathrm{R},i}$ is the number concentration of refractory black carbon particles at the i^{th} CPMA setpoint; $\partial^2 N/\partial \log m_{\mathrm{p}} \partial \log m_{\mathrm{rBC}}$ is the two-dimensional particle-rBC mass distribution, which is the quantity-of-interest; $K_i\left(m_{\mathrm{p}}, m_{\mathrm{rBC}}\right)$ is the kernel of system, which describes the relationship between the CPMA-SP2 system responses and a mathematical model of the measurement system, for the i^{th} CPMA setpoint; and ε_i is a measurement error. The kernel of the system can be described as the product of two functions

$$K_i\left(m_{\mathrm{p}}, m_{\mathrm{rBC}}\right) = S\left(m_{\mathrm{rBC}}\right) \Gamma_i\left(m_{\mathrm{p}}\right), \quad i = 1, 2, \ldots, n \tag{1.5}$$

where $S\left(m_{\mathrm{rBC}}\right)$ is an effective SP2 kernel that acts to bridge between the true aerosol distribution and the binned measurements, and $\Gamma_i\left(m_{\mathrm{p}}\right)$ is the CPMA transfer function and includes the charging fraction.

In this approach, aerosols are passed through a neutralizer to be electrically

charged and are then directed toward the CPMA inlet. The CPMA is successively stepped through a desirable rBC-containing aerosol mass range, and the SP2 measures the mass of rBC portion of CPMA-classified particles using solely incandescence light (*i.e.,* LII signals). A time series of rBC masses of CPMA-classified particles is then discretized into bins, and placed in a two-dimensional m_{p}— m_{rBC} to compute BC mixing states. However, as can be seen from Equation(1.4), the true distribution $\partial^2 N/\partial \log m_{\mathrm{p}} \partial \log m_{\mathrm{rBC}}$ measured by the CPMA-SP2 system is masked by a convolution with the corresponding kernel functions (Kuwata, 2015; Sipkens, Olfert, and Rogak, 2020a). As such, performing the inverse to find $\partial^2 N/\partial \log m_{\mathrm{p}} \partial \log m_{\mathrm{rBC}}$ from the tandem CPMA-SP2 response (*i.e.,* $N_{\mathrm{R},i}$), includes a deconvolution, which comes with some mathematical challenges. First, the system of equations of these problems are regularly rank deficient, *i.e.,* the number of unknowns is more than the number of equations, as a result of estimating the rBC mass distribution at more positions than where CPMA-SP2 measurements are taken. Besides, even if the problem was full rank, the convolution presented in Equation 1.4 masks the ground truth distribution with various instrument functions (*e.g.,* including the effect of multiple charging of the particles, and blurs the data. This issue is considered an *ill-posed problem* in which (*i*) several solutions can equally represent the measured data and (*ii*) the inversion procedure tends to magnify the error in the data. Therefore, suitable attention must be paid to retrieve the particle mass-rBC mass distribution, even if the problem is full-rank. Several mathematical techniques have been developed and implemented for one-dimensional particle size distribution inversion (see the review paper of Voutilainen, Kolehmainen, and Kaipio, 2001). Nevertheless, as the derivation of two-dimensional distributions of aerosol properties using tandem ar-

rangements of aerosol instruments is relatively new in aerosol science, the extension of such techniques to two-dimensional distributions of aerosol properties, especially for the CPMA-SP2 system, is an evolving field that needs careful consideration. Broda et al. (2018) used Twomey-Markowski method (*i.e.,* Twomey with routine smoothing steps) as an inversion method and applied it to CPMA-SP2 measurements to obtain corresponding m_{p}— m_{rBC} distributions. There are however limitations to this method in terms of computational cost and measurement accuracy, which need to be enhanced.

Hypothetical two-dimensional distributions of rBC mass (m_{p}) over total particle mass (m_{p}) which could be derived from a tandem CPMA-SP2 measurements for a rBC population with little to no non-rBC material and rBC population with large amounts non-rBC material, are shown in Figure 1.7a and b. In these distributions, $\partial^2 N$ represents the number concentration of particles, which have the total mass between m_{p} and $m_{\mathrm{p}} + \mathrm{d}m_{\mathrm{p}}$ and rBC mass between m_{rBC} and $m_{\mathrm{rBC}} + \mathrm{d}m_{\mathrm{rBC}}$. In two-dimensional m_{p}— m_{rBC} space, the straight vertical lines correspond to a fixed particle mass located on the x-axis (see Figure 1.7c). It is not physically possible to have particles where the total mass is less than the rBC mass. In these plots, lines of constant mass fraction of rBC, *i.e.,* $m_{\mathrm{rBC}}/m_{\mathrm{p}}$, would correspond to straight lines through the domain at an angle that is parallel to the main diagonal where $m_{\mathrm{rBC}}/m_{\mathrm{p}} = 1$. Lines of constant non-rBC mass (shown as red dotted lines), which we use as a surrogate for the mass of non-rBC materials in each particle and denote as $m_{\mathrm{non-rBC}}$, will follow $\log(m_{\mathrm{rBC}}) = \log(m_{\mathrm{p}} - m_{\mathrm{non-rBC}})$, and are indicated by dash lines in Figure 1.7c. Intuitively, these lines approach $-\infty$ as $m_{\mathrm{non-rBC}}$ approaches m_{p} (where the particle would be all non-refractory). Moving vertically and horizontally

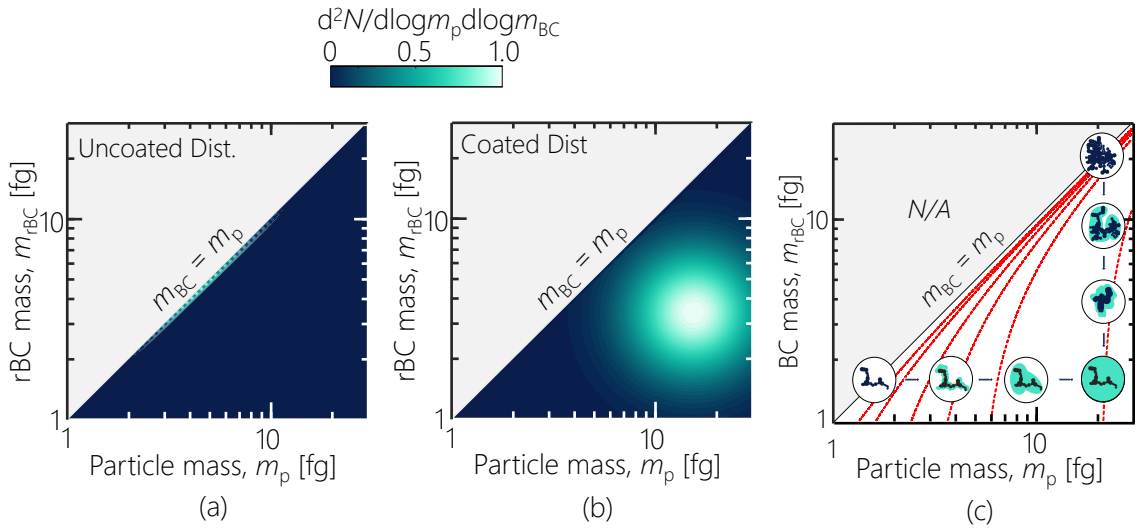

Figure 1.7: (a,b) Normalized two-dimensional m_{rBC} — m_{p} distribution for a sample (a) rBC population with no or little coating, and (b) rBC population with heavy coating. (c) A representation of two-dimensional m_{rBC} — m_{p}.

away from the diagonal results in a decrease of rBC mass fraction (*i.e.,* increasing the coating mass fraction), where the latter one represents an increasing particle mass while the rBC mass remains constant. Additionally, the m_{p}— m_{rBC} distribution can also be mapped out into different spaces to derive explicit information about other quantity of interests. Figure 1.8 illustrates the transformation of m_{p}— m_{rBC} distribution of Figure 1.7b to other spaces. Figure 1.8a shows the normalized number concentration of non-rBC component of rBC-containing particles, which explicitly demonstrates how non-rBC material is distributed on the total mass of rBC-containing particles. Likewise, Figure 1.8b and c show $\partial^2 N/\partial \log m_{\mathrm{rBC}} \partial \log m_{\mathrm{non-rBC}}$, and $\partial^2 N/\partial \log m_{\mathrm{p}} \partial \log f_{\mathrm{non-rBC}}$, which characterize how the the non-rBC material and non-rBC fraction ($f_{\mathrm{non-rBC}} = m_{\mathrm{non-rBC}}/m_{\mathrm{p}}$) are distributed on the rBC mass and total particle mass, respectively. It is evident that detailed information regarding

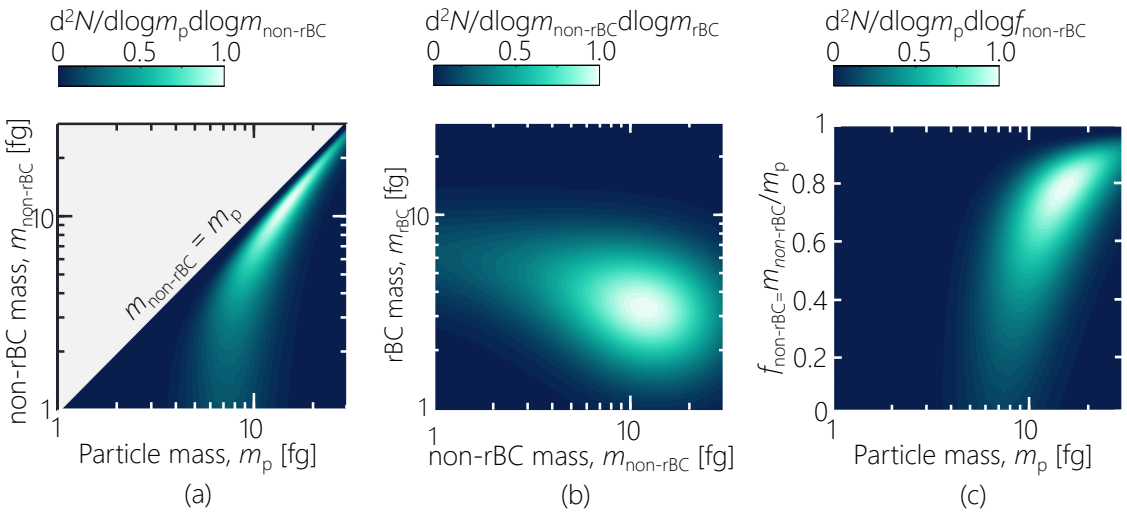

Figure 1.8: (a,b) Normalized two-dimensional m_{rBC} — m_{p} distribution for a sample (a) rBC population with no or little coating, and (b) rBC population with heavy coating. (c) A representation of two-dimensional m_{rBC} — m_{p}.

non-rBC material in rBC-containing particles, which cannot be directly measured with the CPMA-SP2 system, can be derived with this approach. Moreover, the m_{p}— m_{rBC} distribution can also be integrated to obtain one-dimensional distributions; for instance, the particle mass concentration as a function of particle mass can be derived using,

$$\frac{\partial^2 M}{\partial \log m_{\mathrm{p}}} = \int_0^\infty \frac{\partial^2 N}{\partial \log m_{\mathrm{p}} \partial \log m_{\mathrm{rBC}}} m_{\mathrm{p}} \mathrm{d} \log m_{\mathrm{rBC}}, \tag{1.6}$$

or likewise, the total mass concentration, M, can be derived through double integration of the distribution. It is worth noting that while N represents the number concentration of BC-contacting particles, the mass concentration distribution derived in Equation (1.6) refers to the total mass concentration of BC-containing particles.

1.3 Problem Statement and Research Objectives

The fundamental goal of the present study is to improve the accuracy of retrieving m_{p}— m_{rBC} distributions from CPMA-SP2 measurements. This is accomplished by either optimizing how to collect and prepare data prior to the inversion or improving the corresponding inversion approach. Four key objectives are:

i. The initial inversion scheme (*i.e.,* Twomey-Markowski) introduced by Broda et al. (2018) for retrieving the m_{p} — m_{rBC} distributions has limitations (Naseri et al., 2021a). First, the reconstruction accuracy of Twomey-Markowski is inaccurate for narrow distributions, and the approach incurs a high computational cost. Further, the CPMA transfer function and particle losses in the CPMA are poorly quantified, and the inversion of Broda et al. (2018) was a crude correction scheme. Therefore, the first objective is to build on the efforts of Broda et al. (2018) by improving the accuracy of the inversion scheme by *i*) setting up a system of equations to allow for higher resolution reconstructions, *ii*) reducing the computational effort and reconstruction accuracy by truncating the solution domain to remove unphysical particles (*i.e.,* the mass of rBC in a particle cannot exceed the mass of the particle), and *iii*) introducing a marginal distribution constraint to account for uncertainties in the CPMA transfer function and losses in the CPMA.

ii. The accuracy of CPMA-SP2 measurements depends on operating conditions like (*i*) the CPMA resolution, (*ii*) the number of SP2 counted at each CPMA setpoint, (*iii*) the number of CPMA setpoints, and (*iv*) the number of SP2 bins to discretize the SP2 measurements. Hence, the next objective is to determine

the optimum operating conditions for a range of aerosol populations that might be encounted in the atmosphere.

iii. Another objective is to experimentally determine the transfer function of the CPMA to improve the accuracy of the CPMA-SP2 kernel function, $K(m_{\mathrm{p}}, m_{\mathrm{rBC}})$, resulting in a more accurate deconvolution of CPMA-SP2 measurements when the inversion is performed to derive a m_{p}—m_{rBC} distribution.

iv. The last objective is to compare the performance of CPMA-SP2 method with SP2-only methods in terms of: detection size range, temporal resolution, counting statistics, and errors associated with the measurements, in real-world measurements.

Chapter 2

Developing an improved inversion scheme for retrieving two-dimensional distributions using CPMA-SP2 measurements

2.1 Introduction

This Chapter describes an improved inversion scheme for determining the mass distribution of non-refractory materials on refractory black carbon using the CPMA-SP2 system. As described in section 1.2.3.2, the distribution of non-rBC mass on rBC-containing particles can be obtained by deriving two-dimensional number distributions of total particle mass and particle rBC mass. Accordingly, Broda et al. (2018) used tandem CPMA-SP2 measurements and applied a novel Twomey-Markowski (*i.e.,* Twomey with routine smoothing steps) inversion scheme to calculate the distribution of non-rBC mass in rBC-containing particles and a two-dimensional map of total particle mass (m_{p}) and rBC mass (m_{rBC}), *i.e.,* $\frac{\partial^2 N}{\partial \log m_{\mathrm{p}} \partial \log m_{\mathrm{rBC}}}$. Yu et al. (2020) later employed this approach to measure the distribution of non-refractory coatings

on BC in Beijing and found that the distribution of coatings varies greatly with the pollution level.

However, the Twomey-Markowski approach is not without its limitations (Sipkens, Olfert, and Rogak, 2020a). Reconstruction accuracy is limited, particularly for narrow distributions, and the approach incurs a high computational cost. Further, though the width of the CPMA transfer function is a function of mobility, the mobility of the particles is unknown. Also, the particle losses in the CPMA are poorly quantified. Broda et al. (2018) tried to account for this by multiplying the reconstruction by a scalar such that the total number concentration of particles would agree with that of measured directly by the SP2 when the CPMA was bypassed. However, this approach neglects any size-dependent losses and broadening/narrowing due to the transfer function.

In this chapter, the aim is to build on efforts of Broda et al. (2018) by (*i*) improving the accuracy of the inversion scheme, (*ii*) setting up a system of equations to allow for higher resolution reconstructions, (*iii*) reducing the computational effort and reconstruction accuracy by truncating the solution domain to remove unphysical particles (*i.e.,* the mass of rBC in a particle cannot exceed the mass of the particle), and (*iv*) introducing a marginal distribution constraint to account for particle mobility and losses in the CPMA. This work also builds on Sipkens, Olfert, and Rogak (2020a,b), who demonstrated how more advanced inversion schemes can be implemented for aerosol two-dimensional mass-mobility inversions. Following from that work, we consider the performance of unregularized least-squares, Twomey-Markowski, maximum entropy regularization, and Tikhonov regularization methods on synthetic data *phantoms* representing a range of aerosols that could exist in the

atmosphere, including uncoated rBC, moderately-coated rBC, heavily-coated rBC, and a mixture of uncoated and coated rBC particles.

2.2 Problem definition

The concept of deriving two-dimensional distributions from tandem aerosol measurements was first introduced by Rawat et al. (2016) for mass-mobility distributions (using a differential mobility) and improved upon by Buckley et al. (2017) and Sipkens, Olfert, and Rogak (2020a,b). Broda et al. (2018) extended this concept to the problem of inverting CPMA-SP2 data. The problem involves a deconvolution of the following integral equation,

$$N_{\mathrm{P},i} = \int_0^\infty \int_0^\infty \eta\left(m_\mathrm{p}, d_\mathrm{m}\right) K_i\left(m_\mathrm{p}, m_\mathrm{rBC}, d_\mathrm{m}, \Phi\right) \left(\frac{\partial^2 N}{\partial \log m_\mathrm{p} \partial \log m_\mathrm{rBC}}\right) \cdot \mathrm{dlog}m_\mathrm{rBC}\mathrm{dlog}m_\mathrm{p} + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{2.1}$$

where $N_{\mathrm{P},i}$ is the tandem CPMA-SP2 system response, *i.e.,* the number concentration of refractory black carbon particles at the i^{th} CPMA setpoint; $\partial^2 N / \partial \log \ m_\mathrm{p} \partial \log m_\mathrm{rBC}$ is the two-dimensional particle-rBC mass distribution, which is the quantity-of-interest and is discussed in more detail at the beginning of Section 2.3; ε_i is a measurement error; and $K_\mathrm{i}(m_\mathrm{p}, m_\mathrm{rBC}, d_\mathrm{m}, \Phi)$ is the kernel for the i^{th} CPMA setpoint. η is the particle transmission efficiency, and it accounts for particle losses in the CPMA-SP2 system (*e.g.,* diffusion and impaction) not included in the instrument transfer functions. Generally, diffusional losses are a function of the particle mobility (d_m), while impaction losses are a function of mobility and mass (m_p).

The kernel is the product of two functions

$$K_i\left(m_{\mathrm{p}}, m_{\mathrm{rBC}}, d_{\mathrm{m}}, \Phi\right) = S\left(m_{\mathrm{rBC}}\right)\Gamma_i\left(m_{\mathrm{p}}, d_{\mathrm{m}}, \Phi\right), \quad i = 1, 2, \ldots, n \tag{2.2}$$

where $S(m_{\mathrm{rBC}})$ is an effective SP2 kernel that acts to bridge between the true aerosol distribution and the binned measurements, and $\Gamma_i(m_{\mathrm{p}}, d_{\mathrm{m}}, \Phi)$ is the CPMA transfer function and includes the charging fraction.

2.2.1 The CPMA kernel

The theoretical instrument response of the CPMA, Γ_i, is defined for the i^{th} setpoint, characterized by a single-charged, target particle mass ($m_{\mathrm{p}}{}^*$) and the number of charges on a particle (Φ), as

$$\Gamma_i = \sum_{\Phi=1}^{\Phi\mathrm{max}} f\left(d_{\mathrm{m}}, \Phi\right)\Omega\left(m_{\mathrm{p},i}{}^*, m_{\mathrm{p}}, d_{\mathrm{m}}, \Phi\right), i = 1,\ 2,\ \ldots, n \tag{2.3}$$

where m_{p} denotes the mass of a particle subjected to the CPMA; $\Omega(m_{\mathrm{p},i}{}^*, m_{\mathrm{p}}, d_{\mathrm{m}}, \Phi)$ is the CPMA transfer function evaluated using the analytical expressions of Sipkens, Olfert, and Rogak (2020c) (using Case 1C in that work, which has computational and other advantages over finite-difference simulations); and $f\left(d_{\mathrm{m}},\ \Phi\right)$ is the charge fraction, *i.e.,* the fraction of the particles at an integer charge state of Φ, taken from Wiedensohler (1988) with a maximum integer charge state of $\Phi_{\mathrm{max}} = 3$. Implicit in both the CPMA transfer function and the charge fraction is some assumption about the mobility diameter, which is not explicitly specified in this procedure. It is common to use an empirical mass-mobility equation (Kelly and McMurry, 1992) to estimate the mobility diameter from the effective density. Olfert and Rogak (2019)

found that

$$\rho_{\text{eff}} \equiv \frac{m_{\text{p}}}{\frac{\pi d_{\text{m}}^3}{6}} \approx \left(510\frac{\text{kg}}{\text{m}^3}\right)\left(\frac{d_{\text{m}}}{100\text{ nm}}\right)^{-0.52} \tag{2.4}$$

describes a large range of combustion-generated particles and is used to convert the particle mass to a mobility diameter whenever the CPMA transfer function is evaluated. This assumption will have impacts on the width (and in the case of diffusion, the amplitude) of the CPMA transfer and charging functions that propagate through to influence the width and amplitude of reconstructed distributions. Broda et al. (2018) partially correct for this by scaling two-dimensional reconstructions by a factor F such that the total number concentration of black carbon particles agrees with SP2-only measurements, that is $F = N_{\text{SP2}}/N$. However, such an approach neglects size-dependent loss and broadening/narrowing due to the CPMA transfer function. We instead address this issue by placing a new constraint into the inversion, as is discussed in connection with Equation (2.31) later in this work.

2.2.2 The effective SP2 kernel

In contrast with the work of Broda et al. (2018), we present the measurements as the product of a double convolution in Equation (2.1), incorporating an *effective* kernel for the SP2, $S\left(m_{\text{rBC}}\right)$, into the procedure. We say *effective* in that the function is user-imposed, representing the binning of the discrete SP2 data, rather than from the physical spreading of the distribution during measurement, as is the case for a CPMA or differential mobility analyzer. More specifically, SP2 data is fundamentally given as a series of u discrete points representing m_{rBC} for each particle detected (*i.e.* the SP2 is a *single-particle instrument*). In this respect, we

can consider the SP2 data as a vector containing the u measurements of m_{rBC}, $\mathbf{m}^{\mathrm{SP2}} = \left[m_1^{\mathrm{SP2}}, m_2^{\mathrm{SP2}}, m_3^{\mathrm{SP2}}, \ldots, m_u{}^{\mathrm{SP2}}\right]^{\mathrm{T}}$. A list of this form is not immediately amenable to being combined with CPMA data or for use in Equation (2.1), which requires count-based data about a series of setpoints. In both the present work and in Broda et al. (2018), this is overcome by binning the SP2 data to generate a set of responses that act as an input to the inversion (in fact, data binning is one option for data output in new models of the SP2). This binned response can be represented mathematically by considering two observations:

(1) We note that $\mathbf{m}^{\mathrm{SP2}}$ contains a set of masses that, in the absence of measurement errors, are expected to be drawn from the distribution $\mathrm{d}N/\mathrm{dlog}m_{\mathrm{rBC}}$.

In this respect, the cumulative distribution function (CDF) for m_{rBC}, scaled by the number of particles, can be represented as

$$\int_0^{m_i} \frac{\mathrm{d}N}{\mathrm{dlog}m_{\mathrm{rBC}}} \mathrm{dlog}m_{\mathrm{rBC}} \approx \sum_{l=1}^{u} H\left(m_i - m_l^{\mathrm{SP2}}\right), \tag{2.5}$$

where $H(\cdot)$ is the Heaviside function. This expression can be intuitively understood to represent the CDF in that the function on the right side steps up everytime a measurement, m_l^{SP2}, is encountered. Thus, when many m_l^{SP2} are clustered together, the function increases more quickly. The two terms would be equal when $u \to \infty$, with an increasing error as the number of observations, u, decreases. Differentiating, the m_{rBC} distribution can now be represented as

$$\left.\frac{\mathrm{d}N}{\mathrm{dlog}m_{\mathrm{rBC}}}\right|_{m_i} \approx \sum_{l=1}^{u} \delta\left(m_i - m_l^{\mathrm{SP2}}\right), \tag{2.6}$$

where $\delta(\cdot)$ is the δ-Dirac function. In other words, if we assume that the SP2 output is exact, we can represent the rBC mass distribution as an infinite sum of δ-Dirac

functions centered on the observed m_l^{SP2}. Though outside of the scope of this work, one could also replace the δ-Dirac function in this expression with a function having a finite width that represents the uncertainties in each measured m_l^{SP2}.

(2) More importantly, for a continuous m_{rBC} distribution, the theoretical binned response from the SP2 (*i.e.,* the counts in the i^{th} histogram bin) can be represented by the convolution

$$N_{R,i}^{\mathrm{SP2}} = \int_0^\infty \frac{\mathrm{d}N}{\mathrm{dlog}m_{\mathrm{rBC}}} \left[H\left(m_{\mathrm{rBC}} - m_i^{\mathrm{low}}\right) - H\left(m_{\mathrm{rBC}} - m_i^{\mathrm{high}}\right) \right] \mathrm{dlog}m_{\mathrm{rBC}} \tag{2.7}$$

where the combination of Heaviside functions collectively results in a boxcar function and m_i^{low} and m_i^{high} are the lower and upper limit of the i^{th} bin. Here, only particles between m_i^{low} and m_i^{high} counts towards the i^{th} measurement (*i.e.,* the data is binned based on the single-particle output of the SP2). Noting similarity with Equation (2.1), one can directly retrieve an *effective* kernel from this integral equation as

$$S_i\left(m_{\mathrm{rBC}}\right) = H\left(m_{\mathrm{rBC}} - m_i^{\mathrm{low}}\right) - H\left(m_{\mathrm{rBC}} - m_i^{\mathrm{high}}\right). \tag{2.8}$$

Combining these two principles, the equivalent binned response from the discrete data can be evaluated as

$$\begin{aligned} N_{R,i} &= \int_0^\infty \sum_{l=1}^{u} \eta\delta\left(m_i - m_l^{\mathrm{SP2}}\right) \left[H\left(m_{\mathrm{rBC}} - m_i^{\mathrm{low}}\right) - H\left(m_{\mathrm{rBC}} - m_i^{\mathrm{high}}\right) \right] \mathrm{dlog}m_{\mathrm{rBC}} \\ &= \sum_{l=1}^{u} \left[H\left(m_l^{\mathrm{SP2}} - m_i^{\mathrm{low}}\right) - H\left(m_l^{\mathrm{SP2}} - m_i^{\mathrm{high}}\right) \right] \end{aligned} \tag{2.9}$$

Conveniently, the tandem response for a continuous distribution, as is relevant for Equation (2.1), follows from substituting Equation (2.8) into Equation (2.1),

$$
\begin{aligned}
N_{\mathrm{R},i} =& \eta \int_0^\infty \int_0^\infty \Gamma_i\left(m_\mathrm{p}\right) \left[H\left(m_\mathrm{rBC} - m_i^\mathrm{low}\right) - H\left(m_\mathrm{rBC} - m_i^\mathrm{high}\right)\right] \\
& \times \left(\frac{\partial^2 N}{\partial \log m_\mathrm{p} \partial \log m_\mathrm{rBC}}\right) \mathrm{d}\log m_\mathrm{rBC} \mathrm{dlog} m_\mathrm{p} + \varepsilon_i.
\end{aligned}
\tag{2.10}
$$

2.2.3 Discretization

Practically, the problem must be discretized to allow for computation. To this end, we discretize $\partial^2 N / \partial \log m_\mathrm{p} \partial \log m_\mathrm{rBC}$ into n elements in which the quantity takes on a uniform value (taking on the form of a *reconstruction grid*), such that

$$
\begin{aligned}
b_i = N_{R,i} =& \sum_{j=1}^{j=n} \int_{\left(m_\mathrm{p}^\mathrm{low}\right)_j}^{\left(m_\mathrm{p}^\mathrm{high}\right)_j} \int_{\left(m_\mathrm{rBC}^\mathrm{low}\right)_j}^{\left(m_\mathrm{rBC}^\mathrm{high}\right)_j} \eta \Gamma_i\left(m_\mathrm{p}\right) \left[H\left(m_\mathrm{rBC} - m_i^\mathrm{low}\right) - \right. \\
& \left. H\left(m_\mathrm{rBC} - m_i^\mathrm{high}\right)\right] \cdot \left(\left.\frac{\partial^2 N}{\partial \log \mathrm{m}_p \partial \log m_\mathrm{rBC}}\right|_j\right) \mathrm{dlog} m_\mathrm{rBC} \mathrm{dlog} m_\mathrm{p} + \varepsilon_i.
\end{aligned}
\tag{2.11}
$$

We explicitly note that the m_rBC bin edges used in the reconstruction do not need to be the same as those used to bin the data (*i.e.,* $m_\mathrm{i}^\mathrm{high}$ does not necessarily equal $m_\mathrm{rBC}^\mathrm{high}$. Since $\partial^2 N / \partial \log m_\mathrm{p} \partial \log m_\mathrm{rBC}|_j$ is constant within the integral bounds, we can move the term outside of the integrand

$$
\begin{aligned}
b_i = N_{\mathrm{P},i} =& \sum_{j=1}^{n} \eta \left.\frac{\partial^2 N}{\partial \log m_\mathrm{p} \partial \log m_\mathrm{rBC}}\right|_j \int_{\left(m_\mathrm{p}^\mathrm{low}\right)_j}^{\left(m_\mathrm{p}^\mathrm{high}\right)_j} \int_{\left(m_\mathrm{rBC}^\mathrm{low}\right)_j}^{\left(m_\mathrm{rBC}^\mathrm{high}\right)_j} \Gamma_i\left(m_\mathrm{p}\right) \\
& \times \left[H\left(m_\mathrm{rBC} - m_i^\mathrm{low}\right) - H\left(m_\mathrm{rBC} - m_i^\mathrm{high}\right)\right] \mathrm{d}\log m_\mathrm{rBC} \mathrm{d}\log m_\mathrm{p} + \varepsilon_i.
\end{aligned}
\tag{2.12}
$$

This expression now forms the basis for a set of linear equations, where the linear operator is defined as

$$A_{ij} = \int_{\left(m_{\mathrm{p}}^{\mathrm{low}}\right)_j}^{\left(m_{\mathrm{p}}^{\mathrm{high}}\right)_j} \int_{\left(m_{\mathrm{rBC}}^{\mathrm{low}}\right)_j}^{\left(m_{\mathrm{rBC}}^{\mathrm{high}}\right)_j} \Gamma_i\left(m_{\mathrm{p}}\right) \times \left[H\left(m_{\mathrm{rBC}} - m_i^{\mathrm{low}}\right) - H\left(m_{\mathrm{rBC}} - m_i^{\mathrm{high}}\right)\right] \times \mathrm{dlog}\, m_{\mathrm{rBC}} \mathrm{dlog} m_{\mathrm{p}} \tag{2.13}$$

and

$$x_j = \left.\frac{\partial^2 N}{\partial \log m_{\mathrm{p}} \partial \log m_{\mathrm{rBC}}}\right|_j, \tag{2.14}$$

which results in the system $\mathbf{b} = \mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon}$, where $\mathbf{b} \in \mathbb{R}^{ni \times 1}$, $\mathbf{A} \in \mathbb{R}^{ni \times nj}$, $\mathbf{x} \in \mathbb{R}^{nj \times 1}$, $\boldsymbol{\varepsilon} \in \mathbb{R}^{ni \times 1}$, and n_{i} and n_{j} are the number of data and reconstruction points, respectively. Broda et al. (2018) used a simplified version of the problem in which $\left(m_{\mathrm{rBC}}^{\mathrm{high}}\right)_j = m_l^{\mathrm{high}}$, and $\left(m_{\mathrm{rBC}}^{\mathrm{low}}\right)_j = m_l^{\mathrm{low}}$, such that one reconstructs $\partial^2 N / \partial \log m_{\mathrm{p}} \partial \log m_{\mathrm{rBC}}$ using an identical set of bins or grid points as was used to bin the original SP2 data. In this case, the boxcar function in A_{ij} is also uniform over each element and can be removed from the expression, allowing for simplification of Equation (2.13) to

$$A_{\mathrm{ij}} = \int_{\left(m_{\mathrm{p}}^{\mathrm{low}}\right)_j}^{\left(m_{\mathrm{p}}^{\mathrm{high}}\right)_j} \int_{\left(m_{\mathrm{rBC}}^{\mathrm{low}}\right)_j}^{\left(m_{\mathrm{rBC}}^{\mathrm{high}}\right)_j} \Gamma_i\left(m_{\mathrm{p}}\right) \mathrm{dlog} m_{\mathrm{rBC}} \mathrm{dlog} m_{\mathrm{p}}. \tag{2.15}$$

The remaining integrand in this latter form is not a function of m_{rBC} and can be evaluated to yield

$$A_{ij} = \left(\Delta \log m_{\mathrm{rBC}}\right)_j \int_{\left(m_p{}^{\mathrm{low}}\right)_j}^{\left(m_p^{\mathrm{high}}\right)_j} \Gamma_i\left(m_{\mathrm{p}}\right) \mathrm{d} \log m_{\mathrm{p}} \tag{2.16}$$

The remaining integrand in this latter form is not a function of m_{sBC} and can be evaluated to yield

$$A_{ij} = \left(\Delta \log m_{\mathrm{rBC}}\right)_j \int_{\left(m_{\mathrm{p}}\ \mathrm{low}\ \right)_j}^{\left(m_{\mathrm{p}}^{\mathrm{high}}\right)_j} \Gamma_i\left(m_{\mathrm{p}}\right) \mathrm{d} \log m_{\mathrm{p}} \tag{2.17}$$

Equation (11) in Broda et al. (2018), noting that the entire problem is reduced by a factor of $(\Delta \log m_{\mathrm{pBC}})_j$. The generalization to Equation (2.13) offers a couple advantageous over the treatment of Broda et al. (2018):

1. There are several occasions in which the measured data may be binned at a coarse resolution, such as when there are limited counts or when the data is pre-binned during collection. In such scenarios, the approach of Broda et al. (2018) dramatically limits one's ability to produce high-resolution reconstructions of $\partial^2 N / \partial \log m_{\mathrm{p}} \partial \log m_{rBC}$.

2. For simulated cases, such as the current work, the approach of Broda et al. (2018) requires that the same discretization scheme to be used for both generating and interpreting the m_{rBC} data. This commits an inverse crime that can otherwise be avoided if this simplification is not made.

The number of CPMA setpoints and SP2 data bins should be chosen by the experimentalist based on the duration of measurements and the accuracy of results (*i.e.,* more CPMA setpoints will result in longer experiments but higher accuracy).[1]

The problem now resembles that from Buckley et al. (2017) and Sipkens, Olfert, and Rogak (2020a). However, unlike mass-mobility measurements, the CPMA-SP2 problem presents a unique constraint in that m_{rBC} cannot exceed the total particle mass. Broda et al. (2018) included these points in the reconstruction and made them zero at the beginning of each smoothing step. However, this may result in

[1]This topic itself is the subject of discussion later in Chapter 3, which demonstrates the overall procedure for making CPMA-SP2 measurements using an optimal sampling configuration which is defined as the conditions that yield the lowest Euclidean error relative to the input m_{p}—m_{rBC} distribution.

unnecessary smoothing over the diagonal where $m_\mathrm{p} = m_\mathrm{rBC}$, when priors are applied during inversion, and unnecessarily increases the size of the problem. Here, we instead accommodate the elements, the half of the diagonal elements is weighted in the computations to perfectly follow the no-negative coating constraint. This results in the reduced system,

$$\mathbf{b} = \mathbf{A}^1\mathbf{x}^1 + \boldsymbol{\varepsilon} \tag{2.18}$$

where the invalid entries of $\mathbf{x}$ and columns of $\mathbf{A}$ (the adjusted matrix will hereafter be denoted as $\mathbf{A}^1$ and the size of $\mathbf{x}$ is determined therefrom). For the unique scenario in which the reconstruction grid has the same discretization in the m_p and m_rBC directions, with $n_{\mathrm{j},0}$ points in each direction, this involves removing the upper left triangle of the grid, as noted in Figure 2.1, and reduces the number of unknowns from $n_{\mathrm{j},0}^2$ to $n_{\mathrm{j},0}\left(1+n_{\mathrm{j},0}\right)/2$, such that $\mathbf{A}^1 \in \mathbb{R}^{n_i \times [n_{j,0}(1+n_{j,0})/2]}$ and $\mathbf{x} \in \mathbb{R}^{[n_{j,0}(1+n_{j,0})/2]\times 1}$. Consistent with Sipkens, Olfert, and Rogak (2020a), a linear constraint is used in conjunction with a linear least-squares solver to retrieve non-negative distributions.

2.3 Particle-rBC mass phantoms and data fabrication

Consider now hypothetical two-dimensional $m_\mathrm{p} - m_\mathrm{rBC}$ distributions or phantoms. To provide context, first let us briefly describe $m_\mathrm{p} - m_\mathrm{rBC}$ space. As noted in the preceding section, the space is naturally limited by the $m_\mathrm{p} = m_\mathrm{rBC}$ diagonal or 1:1 line, as the mass of rBC cannot exceed the total particle mass. For log-log plots, lines parallel to this main diagonal will correspond to isolines of constant mass fraction rBC, *i.e.,* $\log m_\mathrm{rBC} = \log\left(m_\mathrm{rBC}\frac{m_\mathrm{p}}{m_\mathrm{p}}\right) = \log\left(m_\mathrm{p} f_\mathrm{rBC}\right) = \log m_\mathrm{p} + \log f_\mathrm{rBC}$,

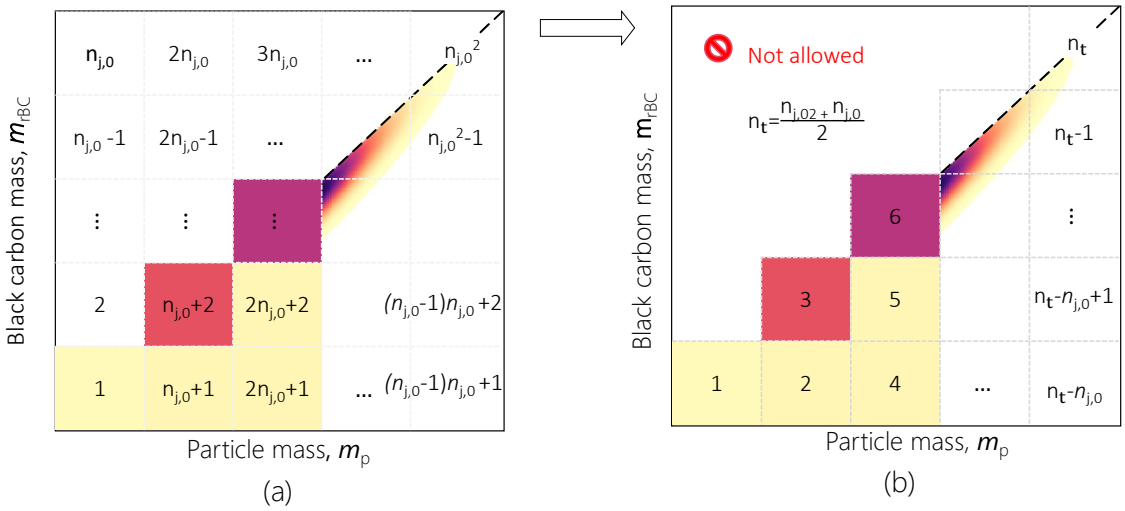

Figure 2.1: Vector representation of x (a) with and (b) without considering the upper triangle elements to implement non-negative mass of non-refractory component of particles and accommodating the sharp edge where m_{rBC} equals mp.

where $f_{\mathrm{rBC}} = m_{\mathrm{rBC}}/m_{\mathrm{p}}$ is the mass fraction of rBC in the particles. In this regard, movement toward the lower, right portion of these plots corresponds to an increase in the fraction of the particles that is not rBC. Horizontal lines, which correspond to m_{rBC}-selected or conditional m_{p} distributions, correspond to particles that contain the same amount of rBC. One interpretation, then, would be that these conditional distributions represent the range in the thickness of non-rBC coating that is found on the same rBC core. However, we need not assume a core-shell structure, which allows for a generalization of this statement to the amount of non-rBC material in otherwise similar rBC particles.

2.3.1 Bivariate lognormal phantoms

Similar to the work of Sipkens et al. (2020 a, 2020b), we now use bivariate lognormal distributions to construct phantoms. For any point in reconstruction space, the bivariate lognormal distribution can be defined as

$$\begin{aligned}\left.\frac{\partial^2 N}{\partial \log m_\mathrm{p} \partial \log m_\mathrm{rBC}}\right|_{\log_{10}\mathbf{r}} &= p\left(\left[m_\mathrm{rBC}, m_\mathrm{p}\right]^\mathrm{T}\right) = N_\mathrm{tot}\det(2\pi\boldsymbol{\Sigma})^{-\frac{1}{2}} \\ &\times \exp\left\{-\frac{1}{2}\left(\log_{10}\mathbf{r}-\boldsymbol{\mu}\right)^\mathrm{T}\left(\boldsymbol{\Sigma}\right)^{-1}\left(\log_{10}\mathbf{r}-\boldsymbol{\mu}\right)\right\},\end{aligned} \tag{2.19}$$

where $\partial N / \partial \log_\mathrm{rBC}|_{\log_{10}\mathbf{r}}$ is the joint $\mathrm{m_p} - \mathrm{m_{rBC}}$ probability density function, $\mathbf{r} = [m_\mathrm{rBC}, m_\mathrm{p}]^\mathrm{T}$ is any point in the reconstructed $m_\mathrm{p} - m_{rBC}$ distribution, N_tot scales the probability density function by the total number concentration in the aerosol, $\det(\cdot)$ denotes the determinant, and $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ denote the mean and covariance of the refractory black carbon mass and particle mass in logarithmic reconstruction space, respectively. The latter two quantities are defined as

$$\boldsymbol{\mu} = \left[\log_{10} m_\mathrm{rBC,g}, \log_{10} m_\mathrm{p,g}\right]^\mathrm{T}, \tag{2.20}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \left(\log_{10}\sigma_\mathrm{rBC}\right)^2 & R_{12}\log_{10}\sigma_\mathrm{p}\log_{10}\sigma_\mathrm{rBC} \\ R_{12}\log_{10}\sigma_\mathrm{p}\log_{10}\sigma_\mathrm{rBC} & \left(\log_{10}\sigma_\mathrm{p}\right)^2 \end{bmatrix} \tag{2.21}$$

where $m_\mathrm{p,g}$ and $m_\mathrm{rBC,g}$ are the geometric mean of particle and rBC mass, respectively; σ_rBC and σ_p are the geometric standard deviation of particle and rBC mass; and R_{12} is the correlation between the logarithm of particle and rBC mass. The latter quantity can be understood to control the slenderness of the distributions (*e.g.,* Phantom 1 in Figure 2.2 corresponds to a large degree of correlation). For more information on the bivariate lognormal size distribution, we refer the reader to Appendix B of Sipkens, Olfert, and Rogak (2020b). Five hypothetical phantoms are considered in an attempt

to span a range of applications (*i.e.,* atmospheric measurements, smog chamber studies, and emission source characterization). In all cases, $\sigma_{\mathrm{rBC}} \sim \sigma_{\mathrm{p}}$ such that the mass fraction of rBC in the particles is largely constant across the full range of particle sizes and the distributions are approximately parallel to the 1 : 1 line. Further, $\mathrm{m}_{\mathrm{rBC,g}} = \mathrm{m}_{\mathrm{p,g}}$ will correspond to mostly bare soot, with some width allowing for the addition of small amounts of non-rBC material, and $\mathrm{m}_{\mathrm{rBC,g}} < \mathrm{m}_{\mathrm{p,g}}$ will correspond to distributions where a majority of the particles contain non-rBC material. The distribution parameters are given in Table.2.1, with the corresponding phantoms shown in Figure 2.2. Figure 2.2 also shows the marginal number concentration distribution of the particle mass ($\mathrm{d}N/\mathrm{dlog}m_{\mathrm{p}}$) and refractory black carbon mass ($\mathrm{d}N/\mathrm{dlog}m_{\mathrm{rBC}}$), obtained by integrating the two-dimensional mass distribution over m_{p} and m_{rBC}, *,e.g.,* the marginal m_{rBC} distribution would be

$$\frac{\mathrm{d}N}{\mathrm{dlog}m_{\mathrm{rBC}}} = \int_{-\infty}^{\infty} \frac{\partial^2 N}{\partial \log m_{\mathrm{p}} \partial \log m_{\mathrm{rBC}}} \mathrm{dlog}m_{\mathrm{p}} \tag{2.22}$$

where, in practice, the bounds are adjusted to span the considered reconstruction domain.

Table 2.1: Parameters used in the generation of hypothetical two-dimensional refractory black carbon mass-total particle mass distributions. For table entries with multiple comma-separated values, each entry refers to a different distribution mode, in a consistent order.

Phantom	$\boldsymbol{m}_{rBC,g}$(fg)	$\boldsymbol{m}_{p,g}$(fg)	$\boldsymbol{\sigma}_{rBC}$	$\boldsymbol{\sigma}_{p}$	$\boldsymbol{R}_{12}$	$\boldsymbol{m}_{p}$[fg] range	$\boldsymbol{m}_{rBC}$[fg] range
1) Uncoated soot particles	2.0	2.0	0.204	0.204	1	0.1-20	0.1-20
2) Uncoated soot mixed with slightly coated soot particles	2.69	3.0	0.24	0.23	0.95	0.1-20	0.1-20
3) Moderately coated soot particles	2.30	1.22	0.24	0.23	0.95	0.1-20	0.1-20
4) Heavily coated soot particles	3.38	15	0.23	0.23	0.01	0.1-20	0.1-20
5) Bimodal uncoated and coated soot particles	0.87, 2.56	0.90,8.00	0.15,0.24	0.15,0.23	0.96,0.97	0.1-20	0.1-20

2.3.2 Synthetic data generation

The phantoms from the preceding section are discretized on a grid of 256×256 elements to give the ground truth of each phantom, $\mathbf{x}^{\text{ph}}$, with the value in each element of the grid given by Equation (2.19) evaluated at the center of the element. Synthetic data, $\mathbf{b}$, is generated by evaluating the forward model, Equation (2.17), with this $\mathbf{x}^{\text{ph}}$ and corrupting the result with Poisson-Gaussian noise, following Foi et al. (2008) and Sipkens et al. (2017). The high-resolution grid and noise were included to avoid inverse crime (Wirgin, 2004). The level of the Gaussian component was set to 5% of the max counts (*i.e.,* 5% of 100,000). Data was generated to correspond to 8 setpoints per decade for the CPMA and 29 bins per decade for the SP2 data, resulting in $\mathbf{b} \in \mathbb{R}^{(12 \cdot 64) \times 1}$ over the particle mass range defined in Table.2.1. The reconstruction grid is composed of 64×64 elements with centers defined identically in the $m_{\text{p}BC}$

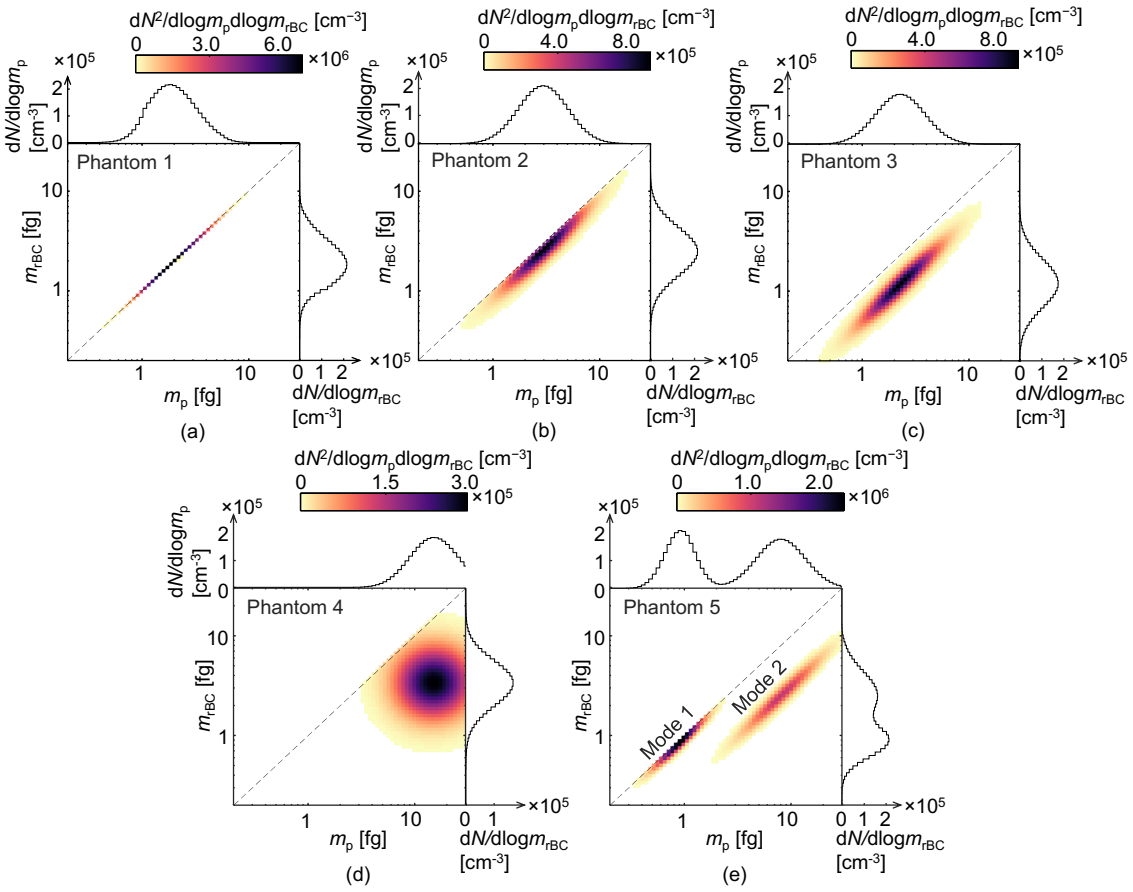

Figure 2.2: Discretized mp-mrBC distributions considered in the present study for all five phantoms.

and $m_{\mathfrak{R}}$ directions (*i.e.*, $n_{\mathrm{j},0} = 64$), such that $\mathbf{x}^1 \in \mathbb{R}^{2080\times 1}$ and $\mathbf{A}^1 \in \mathbb{R}^{768\times 2080}$. Note that since, $2080 > 768$, the problem is rank deficient, having more points in the reconstruction than in the data. The CPMA parameters used to evaluate $\mathbf{A}$ are given in Table.2.1. The particle transmission efficiency is set to a constant $\eta = 0.8$ for all particle masses. This assumption is made since it is neither easy to quantify the size/mass-dependent losses in the system or non-idealilities in the CPMA transfer function. In the present study, the marginal distribution prior is introduced, to

properly address this issue without knowing the size/mass-dependent particle losses in the CPMA-SP2 system.

Table 2.2: CPMA parameters used in the current work.

Parameter	Value
Inner radii (r_1 [mm])	60
Outer radii (r_2 [mm])	61
Flow rate (Q [m^3/s])	$5x10^{-6}$
Classifier length (L [m])	0.2
The speed ratio ($\omega=\omega_1/\omega_2$)	0.9697
Temperature (T [K])	293
Pressure (P [atm])	1
CPMA Resolution (R_m)	10

2.4 Inversion schemes

Several commonly used methods (Carfora, Esposito, and Serio, 1998; Kandlikar and Ramachandran, 1999; Sipkens, Olfert, and Rogak, 2020a; Voutilainen, Kolehmainen, and Kaipio, 2001) were implemented to invert Equation (2.17), including; (*i*) least squares, (*ii*) Twomey, (*iii*) Twomey-Markowski, (*iv*) the multiplicative algebraic reconstruction technique (MART), (*v*) derivative-based Tikhonov regularization, and (*vi*) exponential distance regularization. The reconstruction quality is assessed using the relative Euclidean error,

$$\epsilon = \frac{\|\mathbf{x} - \mathbf{x}^{\text{ex}}\|_2^2}{\|\mathbf{x}^{\text{ex}}\|_2^2}, \tag{2.23}$$

where $\|\cdot\|_2$ denotes the l^2-norm and $\mathbf{x}^{\text{ex}}$ is the exact solution. The implementation of the various inversion methods is described in Appendix.A.

Briefly, the unregularized, least-squares solution is given by

$$\mathbf{x}^{\text{LSQ}} = \arg\min_x \left\{ \left\| \mathbf{A}^1\mathbf{x} - \mathbf{b} \right\|_2^2 \right\} = \arg\min_x \left\{ \sum_{i=1}^{N_i} \left[\left(\sum_{j=1}^{N_j} A_{ij}^1 x_j \right) - b_i \right]^2 \right\}, \quad (2.24)$$

where $\arg\min_{\mathbf{x}}\{\ldots\}$ denotes a minimization with respect to $\mathbf{x}$. The Twomey approach (Twomey, 1975) and and MART (Gordon, Bender, and Herman, 1970) are iterative approaches that use updating rules of

$$x_j^{k+1} = \left[1 + \left(\frac{b_{\text{i}}}{\mathbf{a}_i\mathbf{x}^k} - 1 \right) A_{ij} \right] x_j^k, \quad (2.25)$$

and

$$x_j^{k+1} = \exp\left(\sum_i f_r A_{ij} d_i^k \right) x_j^k = \left(\frac{b_i}{\mathbf{a}_i\mathbf{x}^k} \right)^{f_r A_{ij}} x_j^k, \quad (2.26)$$

respectively. Here, $\mathbf{x}^{k+1} \in \mathbb{R}^{[nj,0(1+nj,0)/2]\times 1}$ is the set of updated values obtained from multiplying the two-dimensional distribution (the initial guess or the one calculated in the previous iteration) to the kernel of the system, $\mathbf{a}_i$ is the i^{th} row of kernel matrix $\mathbf{A}$, and f_r is a relaxation factor. Twomey-Markoski (Markowski, 1987) improves on Twomey's approach by introducing intermittent smoothing steps in the main updating rule, with the two-dimensional adaptation described in Broda et al. (2018), Buckley et al. (2017), and Sipkens, Olfert, and Rogak (2020a). An analogous MART equivalent is also applied, using the same intermittent smoothing approach proposed for the two-dimensional case proposed by Buckley et al. (2017). (We refer the reader to the Appendix.A for additional information on how these schemes are applied to the present application and Sipkens, Olfert, and Rogak (2020a) for a

comparison of these approaches to Tikhonov regularization in the context of mass-mobility inversion.) Tikhonov regularization adds a penalty term to the least-squares scheme, most commonly aimed at minimizing the 0th, 1st, or 2nd order derivative of the solution to dampen noise or encourage smoothness (Bashurova et al., 1991; Crump and Seinfeld, 1982; Leighton, 1994; Seinfeld and Pandis, 1998; Talukdar and Swihart, 2003; Wang et al., 2006; Wolfenbarger and Seinfeld, 1990). Regardless of the derivative being minimized,

$$\mathbf{x}_{\alpha}^{\mathrm{Tk}} = \arg\min_{x} \left\{ \left\| \mathbf{A}^{1}\mathbf{x} - \mathbf{b} \right\|_{2}^{2} + \alpha \left\| \mathbf{L}^{\mathrm{Tk}}\mathbf{x} \right\|_{2}^{2} \right\}, \tag{2.27}$$

where $\mathbf{L}^{\mathrm{TK}}$ is a discretized differential operator called the Tikhonov matrix, and $\alpha > 0$ is called the regularization parameter. The properties of the Tikhonov matrix are described in the Appendix.A(Section A.2.1).

Exponential distance regularization (Sipkens, Olfert, and Rogak, 2020b) follows from the general Tikhonov framework, Equation (2.27), but replaces $\mathbf{L}^{\mathrm{Tk}}$ with an exponential function of the distance between points in the reconstruction, such that adjacent elements in the reconstruction should take on similar number concentrations. Mathematically,

$$\mathbf{L}^{\mathrm{Tk}} = \mathrm{chol}\left[[\exp(\mathbf{D})]^{-1} \right], \tag{2.28}$$

where chol(...) denotes the Cholesky factorization and $\mathbf{D}$ is a symmetric matrix containing the Mahalanobis distance between pairwise (i and j) points in the reconstruction domain. Accordingly, the entires of $\mathbf{D}$ are

$$D_{i,j} = \left[\left(\log_{10}\mathbf{r}_i - \log_{10}\mathbf{r}_j \right)^{\mathrm{T}} \left(\mathbf{\Gamma}^{\mathrm{D}} \right)^{-1} \left(\log_{10}\mathbf{r}_i - \log_{10}\mathbf{r}_j \right) \right]^{1/2}, \tag{2.29}$$

where $\mathbf{\Gamma}^{\mathrm{D}}$ is a correlation matrix describing the correlation between points in the reconstruction,

$$\mathbf{\Gamma}^{D} = \begin{bmatrix} {l_{\mathrm{p}}}^2 & {R_{12}}^{\mathrm{D}} l_{\mathrm{p}} l_{\mathrm{rBC}} \\ {R_{12}}^{\mathrm{D}} l_{\mathrm{p}} l_{\mathrm{rBC}} & {l_{\mathrm{rBC}}}^2 \end{bmatrix}; \tag{2.30}$$

l_{p} and l_{rBC} are the correlation lengths for m_{p} and m_{rBC}, respectively; and R_{12}^{D} is the correlation in the distances (preferentially allowing smoothing along certain directions). For the CPMA-SP2 problem, these correlation lengths are expected to be identical (i.e., $l_{\mathrm{p}} = l_{\mathrm{rBC}}$), such that smoothing preferentially occurs along the lines of constant mass fraction rBC. For cases of very narrow distributions, *e.g.,* when there is little-to-no non-refractory material in the particles, the off-diagonals in $\mathbf{\Gamma}^{\mathrm{D}}$ can be set such that smoothing is only applied along the length of the distributions. More information on the exponential distance prior for the aerosol problem is given in Sipkens, Olfert, and Rogak (2020b) and the Appendix.A, including discussion of the choice of parameters for multimodal phantoms. Though not considered in previous works, the value of $\mathrm{d}N/\mathrm{dlog}m_{\mathrm{rBC}}$ can also be measured by the SP2 directly by bypassing the CPMA. Logically, $\mathrm{d}N/\mathrm{dlog}m_{\mathrm{rBC}}$ can be used to provide prior information. Adding this term to the Tikhonov functional,

$${\mathbf{x}_{\alpha,\lambda}}^{\mathrm{TK}} = \arg\min_{x} \left\{ \lambda \left\| \mathbf{A}^1 \mathbf{x}^{\mathrm{TK}} - \mathbf{b} \right\|_2^2 + \left\| \mathbf{C}\mathbf{x}^{\mathrm{TK}} - \mathbf{d} \right\|_2^2 + \alpha \left\| \mathbf{L}^{\mathrm{TK}} \mathbf{x}^{\mathrm{TK}} \right\|_2^2 \right\}, \tag{2.31}$$

where λ is a second regularization parameter to account for the CPMA losses that would cause a discrepancy between $\mathbf{b}$ and $\mathbf{d}$; $\mathbf{d}$ is the discrete measurements of the m_{rBC} marginal distribution; and $\mathbf{C}$ is a composite kernel composed of the discrete, effective SP2 transfer function, $\mathbf{S}\left(m_{\mathrm{rBC}}\right)$, and an operator $\mathbf{C}^1$ summing up the 2D distribution (i.e., $\mathbf{x}$) in the corresponding direction, such that $\mathbf{C}\mathbf{x} = \mathbf{S}\mathbf{C}^1\mathbf{x} = \mathbf{d}$. We

hereafter refer to this prior as the *marginal distribution prior*. (Note that since $\mathbf{S}$ only incorporates the effective SP2 transfer function - (*i.e.,* only incorporates binning of the measurements) and that this portion of the kernel is independent of the particle mass and mobility diameter. The $\mathbf{A}^1$ term retains the same form as before, incorporating the particle mass and mobility diameter, with that latter by way of the effective density). Solutions obtained by Tikhonov methods are highly dependent on the regularization parameters (*e.g.,* α and λ), leading to broad variation in converged solutions. Here, the ground truth is known, such that the regularization parameters can be obtained by calculating some form of error metric, *e.g.,* the relative Euclidean error. However, as in practice the ground truth is not known, we also consider two heuristics for choosing the regularization parameter: (i) the well-established L-curve approach, implementing the automated methods of Calvetti, Golub, and Reichel (1999) , and Cultrera and Callegaro (2016), and (ii) the new Bayes factor method introduced by Sipkens, Olfert, and Rogak (2020b), based on the work of Thompson and Kay (1993) (see the Appendix.A for more information). The authors note that while the L-curve framework can be generalized for determination of multiple regularization parameters (*e.g.,* Belge, Kilmer, and Miller, 1998, 2002), such as simultaneously estimating the α and λ, the automated L-curve approach is limited to estimating α since the method commonly underestimates the regularization parameter in this work, even in this simpler case. This is also true of the exponential distance prior parameters, which are not simple regularization parameters that scale terms in the Tikhonov functional. Accordingly, whenever multiple prior parameters must be chosen, the Bayes factor scheme is the only approach employed.

2.5 Results and discussion

2.5.1 Comparison of inversion methods without considering the marginal distribution prior

Figure 2.3 shows example solutions for Phantom 3 for the Twomey-Markowski, 1st order Tikhonov, and the exponential distance methods, without the marginal distribution prior. Solutions for the other inversion methods and phantoms are shown in the Appendix.A. As expected, the least-square method is poor (relative error of

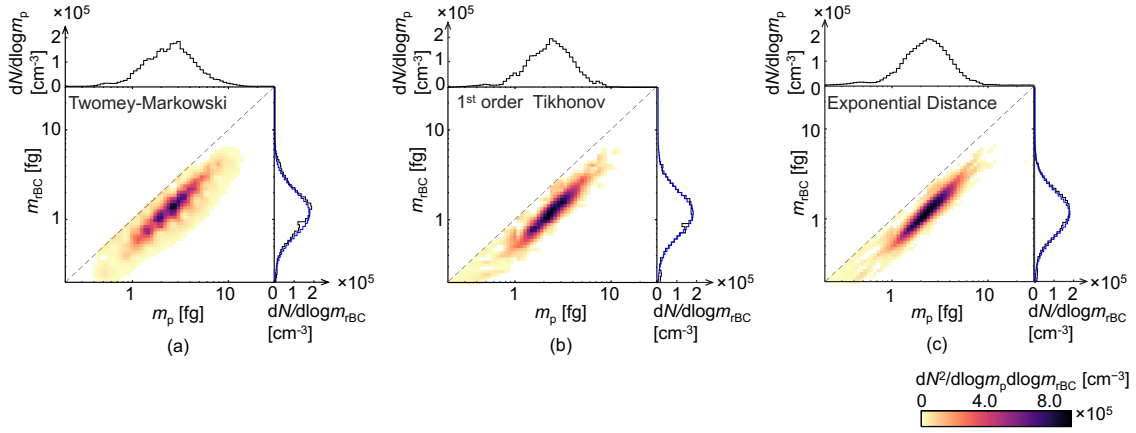

Figure 2.3: Distributions retrieved by the Twomey–Markowski, first-order Tikhonov, and exponential distance methods for synthetic data fabricated from Phantom 3. Marginal distributions are shown on either axis, where blue lines are the actual marginal refractory black carbon ($\mathrm{d}N/\mathrm{dlog}m_{\mathrm{rBC}}$ of the original phantom.

$\sim 150\%$), resulting from the rank deficiency of the problem and indicates the need for some kind of regularization.

The Twomey-Markowski method results in a substantial improvement to the retrieved distributions (reducing the relative error to 30%) but is prone to hotspots, consistent with previous work (Sipkens, Olfert, and Rogak, 2020a). The 1st order Tikhonov and the exponential distance reconstructions give a substantial further

enhancement in reconstruction accuracy (relative errors of $\sim 18\%$ and 15%, respectively), again consistent with previous work (Sipkens, Olfert, and Rogak, 2020a). Even a small degree of under-regularization was noted to result in artificial peaks to the right of the main peaks, such that sensitivity to the regularization parameter close to the minimum should be checked. Other, minor artifacts remain for the Tikhonov and exponential distance priors at the lower limits of the distributions, *i.e.,* for $m_\mathrm{p} < 1\mathrm{fg}$, which stem from the presence of noise in the data.

Figure 2.4 shows the relative Euclidean error across the broader range of inversions methods for all phantoms. Reconstructions generally improve moving from least-squares to Twomey-type to Tikhonov-type inversion schemes. Twomey regu-

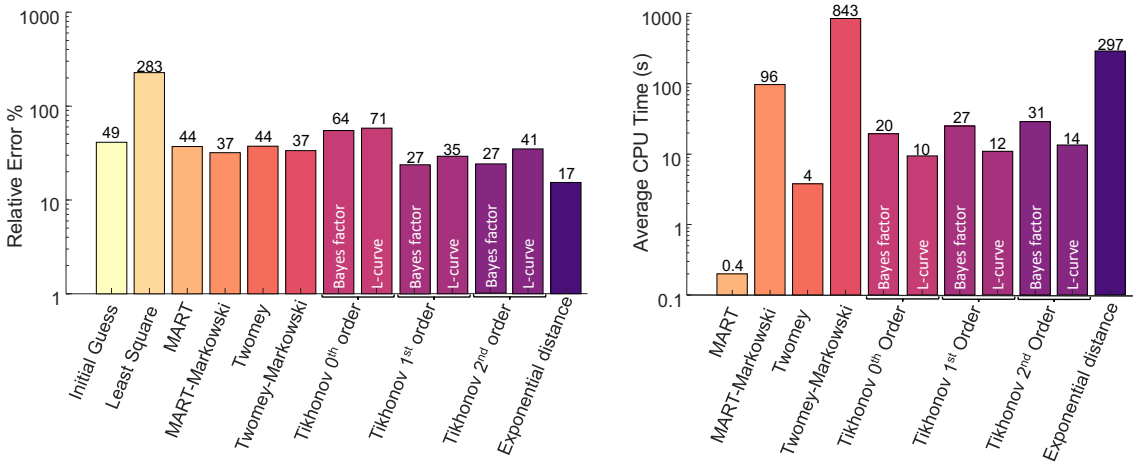

Figure 2.4: (Left) Relative Euclidean error and (right) average CPU time for convergence for each reconstruction technique averaged over all phantoms. The CPU times presented here are to be considered in comparative terms only as exact the CPU times vary depending on the processor configuration and reconstruction resolution. Both error and CPU times are presented on a logarithmic scale.

larization can only predict the rough phantom shape and magnitude of the original phantoms, with the Twomey-Markowski method improving reconstructions at a

higher computational cost (the Twomey-Markowski scheme is the slowest reconstruction method), again consistent with previous work. MART and MART with intermittent smoothing retrieve solutions similar to the original Twomey and Twomey-Markowski algorithms, respectively, though with convergence rates much faster than the Twomey-type methods (nearly 23 times faster). Sipkens, Olfert, and Rogak (2020a) also noted this speed enhancement comes with a small risk of erratic or divergent behavior by the algorithm.

For Tikhonov-type reconstructions of all phantoms, Figure 2.4 indicates the relative Euclidean error is generally lower following the Bayes factor approach to regularization parameter optimization, though at the cost of a doubling of convergence time. Zeroth-order Tikhonov, which does not impose smoothness on the solution, is only able to control the largest fluctuations and is generally considered insufficient. The first and second-order Tikhonov schemes result in nearly 10% reductions in the relative error, when compared to Twomey–Markowski. The exponential distance approach marks a further improvement in terms of reconstruction accuracy.

2.5.1.1 Effect of the marginal distribution prior

Results presented up to this point are obtained without considering the marginal distribution prior. Figure 2.5 shows the first-order Tikhonov and exponential distance solutions for phantoms of 1, 2, and 3, adding the marginal distribution prior. Figure 2.5 reveals that the marginal distribution prior decreases the reconstruction accuracy of the Tikhonov methods as it produces some artifacts in the reconstruction area, especially near the boundaries of the domain. Notably, for phantom 1, Figure2.5 shows a slender artifact near the upper limit of m_{p} space. The shape, elongated in

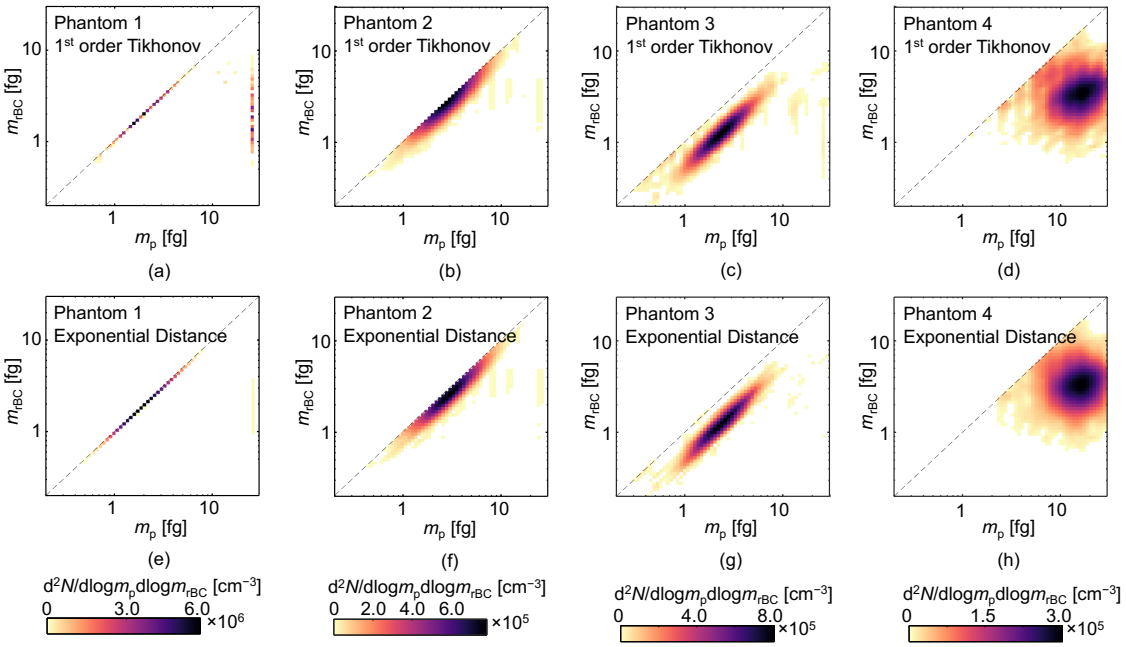

Figure 2.5: Distributions retrieved by considering both Tikhonov-type priors and a marginal distribution as prior information for (top) the first-order Tikhonov and (bottom) exponential distance methods. Blue lines are the actual marginal refractory black carbon ($\mathrm{d}N/\mathrm{dlog}m_{\mathrm{rBC}}$) of the original phantom.

m_{rBC} space, can be understood to stem from attempting to increase the height of the marginal distribution by adding particles in a region that least affects the solution norm. Similar statements can be made for artifacts located in similar regions for the other phantoms.

Figure 2.6, which compares the performance of different Tikhonov and exponential distance methods with and without the marginal distribution prior, demonstrates that while considering the m_{rBC} marginal distribution prior improves the reconstruction accuracy of the exponential distance method, it would not boost the accuracy of Tikhonov schemes. The authors note that by considering the rBC marginal distribution, the difference in reconstruction accuracy increases very significantly in moving

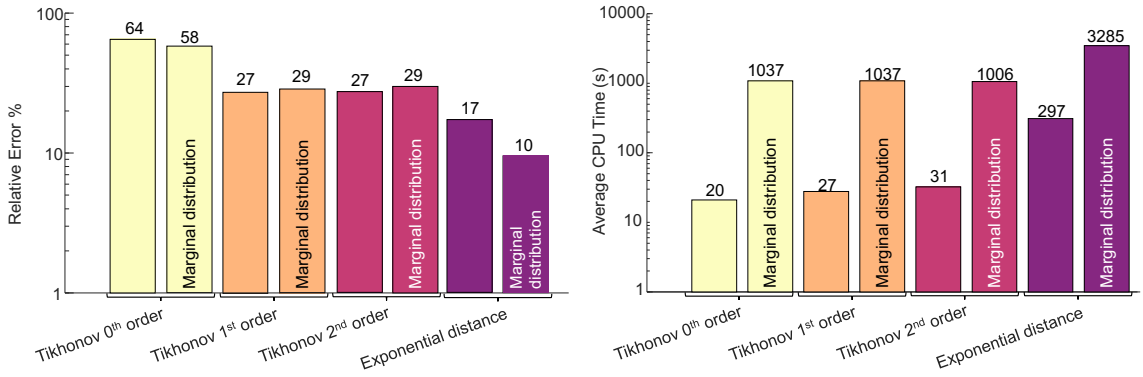

Figure 2.6: The average (left) Euclidean error and (right) convergence time, CPU time of reconstructions regularized solely with and without the marginal distribution priors across all phantoms.

toward narrower phantoms, marking an improvement relative to first-order Tikhonov regularization, from 8% across phantoms 2, 3, 4, and 5 to 65% for phantom 1, see Figure S11 (Appendix.A). This implies that while the exponential distance reconstructions of phantom 1 are highly accurate, with a relative Euclidian error of only 3%, neither of the Tikhonov regularization methods, whether the marginal distribution prior is considered or not, can reconstruct it properly (see Figures S8 and S11 of the Appendix.A), as the structure of the Tikhonov matrix correlates distant points in reconstructions. Generally, Figure 2.6 shows the average relative error for the exponential distance method is preferentially improved, hovering around one-third of the error for the 1st-order Tikhonov reconstructions.

Despite the accuracy afforded by the combination of exponential distance and marginal distribution priors, this method does require the practitioner to define $\mathbf{\Gamma}^{\mathrm{D}}$, which can be challenging when multiple mixing states of black carbon are present, such as for Phantom 5 (see Appendix.AFigure A.9). If l_{p} and l_{rBC} are defined in

a similar manner to the unimodal distributions, the relative Euclidian error of the reconstructions will not reduce enough to obtain the optimal solution, derived by the exponential distance method and depicted in Figure 2.7. Even averaging values over

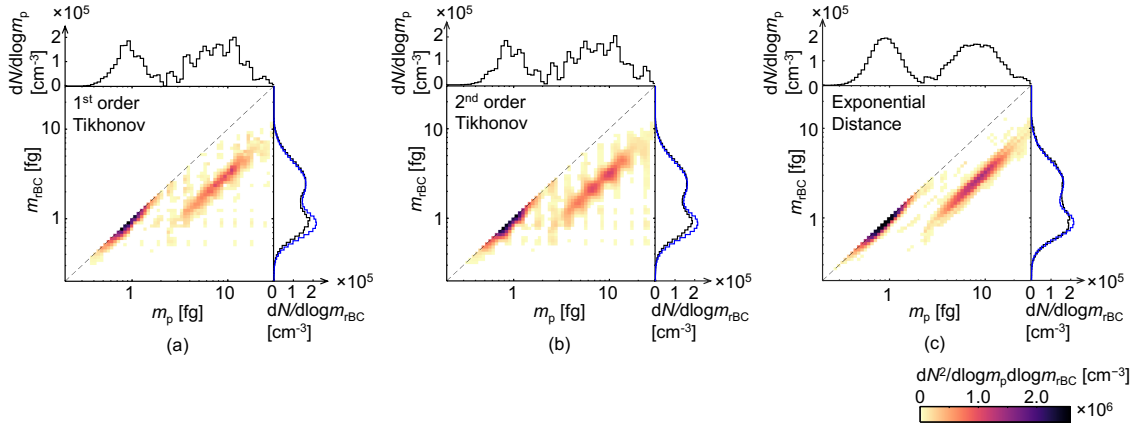

Figure 2.7: The reconstructed distribution Phantom 4 regularized by considering both smoothing and actual marginal distribution as prior information for the first and second-order Tikhonov and Exponential distance methods. Blue lines are the actual marginal refractory black carbon ($\mathrm{d}N/\mathrm{dlog}m_{\mathrm{rBC}}$) of the original phantom.

multiple modes may over-regularize narrow distributions, making them wider, and under-regularized broad distributions, resulting in excessive noise. In the present article, the elements of optimal $\mathbf{\Gamma}^{\mathrm{D}}$ are derived by doubling the arithmetic average of correlation matrices of modes 1 and 2 of Phantom 4, which is similar to the result of Sipkens, Olfert, and Rogak (2020b) suggested for the unimodal distributions. This approach makes the elements of $\mathbf{\Gamma}^{\mathrm{D}}$ lie within $\sim$ 55% of the correlation lengths and off-diagonals that would be expected for the phantom parameters from Table 1. Further consideration of multimodal m_{p}—m_{rBC} distributions and their effect on the optimal value of $\mathbf{\Gamma}^{\mathrm{D}}$ and reconstruction accuracy will be a topic of future work.

2.6 Conclusions

In the present study, we present a new approach to obtain the m_{p}—m_{rBC} distribution from tandem CPMA-SP2 data. The new approach formally removes (non-physical) elements where $m_{\mathrm{rBC}} > m_{\mathrm{p}}$, which allows for a faster convergence rate; better regularization in the vicinity of $m_{\mathrm{rBC}} = m_{\mathrm{p}}$ by preventing smoothing over the constraint; and generally improves reconstruction accuracy. A range of priors, including Twomey, Twomey-Markowski, MART (including the option for adding smoothing function), and various Tikhonov-type inversion schemes, are considered. MART with a smoothing function was shown to have a higher convergence rate, nearly eight times faster than the current state-of-the-art Twomey-Markowski method. First- and second-order Tikhonov methods further improve reconstruction accuracy, reducing the relative error in the reconstructions by 10%. However, neither of the derivative-based Tikhonov approaches can retrieve the pure uncoated state of black carbon, phantom 1 , as the structure of the prior correlates distant points in reconstructions. In response, we demonstrate that the exponential distance prior, which only correlates local points in reconstructions, largely outperforms the other regularization methods, with further improvements afforded by introducing a marginal distribution prior (*i.e.,* a prior that incorporates SP2-only measurements, where the CPMA is bypassed).

In summary, for narrow distributions, we recommend that the exponential distance prior approach be used. For more general, wider, lognormal, or Gaussian-like distributions, either first-order Tikhonov or the exponential distance prior are considered most appropriate, with the former being slightly faster and simpler and the

latter being more accurate.

Chapter 3

Optimum sampling configurations for the CPMA-SP2 measurement system

3.1 Introduction

This Chapter derives the overall procedure for making CPMA-SP2 measurements using an optimal sampling configuration (*i.e.,* optimal simulation), which is defined as the conditions that give the lowest Euclidean error relative to the input $m_{\mathrm{p}} - m_{\mathrm{rBC}}$ distribution. Accordingly, the number of CPMA setpoints per decade, the number of SP2 bins per decade, CPMA resolution, and the number of SP2 counts per CPMA setpoint are derived as a function of aerosol concentration and desired measurement time to accurately reconstruct $m_{\mathrm{p}} - m_{\mathrm{rBC}}$ distributions.

In a typical classifier-detector system, in which common aerosol classifiers (*e.g.,* DMA; Knutson and Whitby, 1975, and CPMA; Olfert and Collings, 2005) are used in tandem arrangement with a particle detector, such as a condensation particle counter (CPC; Quant et al., 1992), optical particle counter (OPC; Sorensen et al.,

2011), or SP2 (Stephens, Turner, and Sandberg, 2003), the classifier is stepped or scanned through a range of set points, with the output measured by the detector. Data inversion is then used to relate the detector data and classifier set points to an aerosol distribution (*e.g.,* count size or mass distribution). Common examples include the scanning mobility particle sizer (SMPS; Wang and Flagan, 1990) which is a DMA-CPC system, and the scanning aerodynamic aerosol sizer (SASS; Johnson et al., 2018) which is a AAC-CPC system. An important consideration in classifier-detector systems is choosing optimal operating conditions. For example, for an SMPS, scan time and classifier resolution (a parameter describing the width of the monodisperse classifier output) can be specified in such a way to derive accurate particle size distributions with appropriate counting statistics under a range of sampling conditions. In most atmospheric studies, where number concentrations are low, it is recommended to have longer scan times (*e.g.,* 5 minutes) while maintaining high resolutions (Birmili et al., 2007; Wallace, 2006; Weber et al., 1996; Zikova and Zdimal, 2012). In contrast, when dealing with a transient source such as combustion processes (Zhu et al., 2002) or some rapid particle nucleation processes (Kulmala et al., 2012; Riccobono et al., 2014), such operating conditions may be suboptimal. In such experimental situations, different recommendations are required to achieve accurate size distributions at different sampling conditions (*e.g.,* Shah and Cocker, 2005).

Tandem measurements, where two or more classifiers are placed in series with a detector, have the promise of providing more detailed measures of aerosol characteristics. However, tandem techniques also introduce an extra set of potentially interrelated experimental parameters that must be optimized. For instance, in stud-

ies with tandem arrangements of mass and mobility classifiers (*e.g.,* the arrangements in Buckley et al., 2017; Rawat et al., 2016; Sipkens, Olfert, and Rogak, 2020a,b), one must simultaneously optimize the parameters of a mobility spectrometer (*e.g.,* scan time, flow rates, resolution) and a particle mass analyzer (*e.g.,* device resolution), including a number of interdependent parameters (*e.g.,* the flow rates and charging affect both devices). As such, experimental design is more complicated for tandem measurements. Relative to the SMPS, optimizing the experimental properties for a system that contains a CPMA is more critical, as the device takes longer to stabilize. Thus, longer collection times are often preferred due to improvements in counting statistics. However, as with SMPS measurements, these collection times come at the expense of added work-hours and is inaccurate for non-stationary aerosols.

Similarly, this chapter aims to optimize the experimental parameters of the CPMA-SP2 system, using Monte Carlo simulations to account for the non-linear relationship between the experiment parameters and the reconstruction accuracy of m_p—m_rBC distributions. We start by considering a scenario in which the aerosol source is assumed to be stationary, such that measurement duration is unconstrained, and examine trends with different experimental parameters (*i.e.,* the number of CPMA setpoints per decade, the number of SP2 bins per decade, CPMA resolution, and the number of SP2 counts per CPMA setpoint). For each simulated experiment, we also estimate the measurement duration, allowing for a discussion of the optimal experiment settings for specific measurement durations as may be relevant if the aerosol is non-stationary or the experiments are expensive. Next, we discuss the effect of phantom (*i.e.,* a synthetic m_p—m_rBC distribution) characteristics, ranging from very narrow to very broad rBC distributions. We conclude by making prac-

tical recommendations for the operating parameters across a range of measurement durations and aerosol concentrations, describing a web app we have developed for general recommendations and making specific recommendations for a moderately broad distribution representative of an aged soot.

3.2 Problem definition

As discussed in Section 2.2.3, to derive $m_{\mathrm{B}} - m_{\mathrm{rBC}}$ distribution a linear system of equations of Equation (2.18), which is set up according to the using the CPMA-SP2 measurements, should be solved. In this study, the reconstruction grid is held constant with 24 elements per decade ($\mathbf{x} \in \mathbb{R}^{5184\times 1}$). By contrast, the dimension of system response ($\mathbf{b} \in \mathbb{R}^{[n_{\mathrm{c}}\times n_{\mathrm{s}}]\times 1}$) is variable, given as the product of the number of CPMA setpoints(n_{c}) and the number of SP2 bins (n_{s}). For the remainder of this work, s subscripts denote experimental parameters related to the SP2, and c subscripts parameters related to the CPMA.

3.2.1 Inverse procedure

As discussed in Chapter 2, the overall system is ill-posed, and the system must be supplemented with prior information, *i.e.,* regularized. We here invert the problem using 1$^{\mathrm{st}}$ order Tikhonov regularization (Kandlikar - Ramachandran, 1999; Voutilainen et al., 2001) following the procedure described in Section 2.4. While this requires an adaptation of the traditional scheme as it has historically applied to one-dimensional size distributions (*cf.* Sipkens, Olfert, and Rogak, 2020a,b), the overall effect is the same: to add prior information that encourages smoothness. The reconstruction quality is assessed using the relative Euclidean error, $\epsilon = \|\mathbf{x} - \mathbf{x}^{\mathrm{ex}}\|_2^2 \,/\, \|\mathbf{x}^{\mathrm{ex}}\|_2^2$, where

$\|\cdot\|_2$ denotes the L2 norm and $\mathbf{x}^{\text{ex}}$ is the exact solution. The optimal regularization parameter is taken as that which minimizes the Euclidian distance between the known ground truth and the reconstructed distribution.

3.2.2 Synthetic data, parameterization, and sampling procedure

The present work generates a series of synthetic data by sampling a range of parameters relevant to the experiment and analysis. Synthetic data is generating using Equation (2.18) and requires knowledge of three components: a synthetic m_{p}—m_{rBC} distribution or phantom to give $\mathbf{x}^1$; a set of experimental parameters that are used to generate a kernel ($\mathbf{A}^1$); and a noise vector ($\boldsymbol{\varepsilon}$) which is composed of Poisson-Gaussian noise (Foi et al., 2008; Sipkens et al., 2017). Phantoms are taken as bivariate log-normal, which can be described by Equation (2.19) to Equation (2.21), to represent atmospheric m_{p}—m_{rBC} distributions.

3.2.2.1 Phantom parameterization for representative atmospheric distributions

Parameters of Equation (2.19) to Equation (2.21), are chosen to represent a wide range of black carbon mixing states found in the atmosphere following a condensation process and based on several studies Gong et al. (2016), McMeeking et al. (2010), and Yu et al. (2020). McMeeking et al. (2010), for example, conducted aircraft measurements of rBC aerosol concentrations in the lower troposphere over Europe using the SP2, spanning remote continental regions to highly populated urban areas. They reported a mass median diameter (MMD) and geometric standard deviation (GSD) of rBC mass of $d_{\text{MMD}} = 183$ nm and $\sigma_{\text{rBC}} = 2.08$, respectively. Transforming

to the quantities of interest here follows from using the Hatch-Choate equation,

$$d_{\mathrm{CMD}} = d_{\mathrm{MMD}} \exp\left[-3\left(\ln \sigma_{\mathrm{rBC}}\right)\right], \tag{3.1}$$

to convert this mass median diameter of rBC to its equivalent count median diameter (CMD) and $m_{\mathrm{rBC,g}} = \rho_{\mathrm{eff}} \frac{\pi d_{\mathrm{CMD}}^3}{6}$ to convert to a median rBC mass using the same rBC density of $\rho_{\mathrm{eff}} = 1800$ kg/m^3 as used by McMeeking et al. For the McMeeking et al. parameters, $d_{\mathrm{CMD}} = 135.6$ nm and $m_{\mathrm{rBC},g} = 2.35$ fg. Phantoms were created which model a simple condensation process in the atmosphere, that is, the geometric mean rBC mass and geometric standard deviation of m_{rBC} were the same in all phantoms (*i.e.,* the rBC particles are unchanged through the condensation process and coagulation is neglected)

As noted above, the covariance in Equation (2.21) is parameterized using a correlation, R_{12}, between two types of particle mass, thereby specifying the breadth of the m_{p}—m_{rBC} distribution. Specifically, we assume that high rBC mass fractions are associated with very thin or non-existent coatings and thus very narrow distributions. Conversely, particles with a substantial non-refractory component are associated with thick coatings and broad distributions. Futhermore, it was assumed that the GSD of the particle mass is equal to the GSD of the rBC mass (*i.e.,* $\sigma_{\mathrm{p}} = \sigma_{\mathrm{rBC}} = 2.08$), which is an arbitirary constraint, but one that is approximately observed by Yu et al. (2020). This is implemented by stating R_{12} as a function of $m_{\mathrm{p,g}}$,

$$R_{12} = -\left(3.0 \times 10^{-4}\right) {m_{p,g}}^2 + \left(2.2 \times 10^{-3}\right) m_{p,g} + 9.9 \times 10^{-1}. \tag{3.2}$$

Given that $m_{\mathrm{rBC,g}} = 2.35$ fg is taken as a constant in the associated paper, this acts as a surrogate for the mass fraction of non-rBC in the particles for the sampling

procedure described subsequently. It must be emphasized that the form of this relation is purely empirical, derived qualitatively based on observations of uncoated aerosols and the coated urban aerosols observed by Yu et al. (2020), using two-dimensional particle mass-rBC mass distributions, and studies by Liou et al. (2014) and Gong et al. (2016), which demonstrate how the aging processes affects the coating thickness of rBC-containing particles.

Throughout this work, $m_{\mathrm{rBC,g}}, \sigma_{\mathrm{p}}$, and σ_{rBC} are fixed at the values reported in Table 3.1, while $m_{\mathrm{p,g}}$ is sampled to realize the range of phantoms shown in Figure 3.1.

Table 3.1: Parameters used in the generation of hypothetical two-dimensional particle mass-BC mass distributions.

$m_{\mathrm{rBC,g}}$ [fg]	$m_{\mathrm{p,g}}$ [fg]	σ_{rBC}	σ_{p}	m_{p} [fg]	m_{rBC} [fg]
2.35	Sampled	2.08	2.08	0.3-300	0.3-300

3.2.2.2 Parameterizing the kernel

The kernel, $\mathbf{A}^1$, is parameterized using three independent, experimental parameters: (i) the aforementioned CPMA resolution (R_{m}); (ii) the number of CPMA setpoints per decade (n'_{c}), noting the prime is used to denote per decade); and (iii) the number of SP2 bins per decade (n'_s). The latter two quantities define the size of the system that is being inverted, with larger values incurring more computational effort. The remaining kernel parameters, *i.e.,* the flow rates and instrument dimensions, are fixed. The CPMA flow rate is set to $Q_{\mathrm{c}} = 0.42$ L/min and is assumed not to be equal to the SP2 flowrate ($Q_{\mathrm{s}} = 0.12$ L/ min). (*i.e.,* generally, $Q_{\mathrm{s}} \neq Q_{\mathrm{c}}$ as

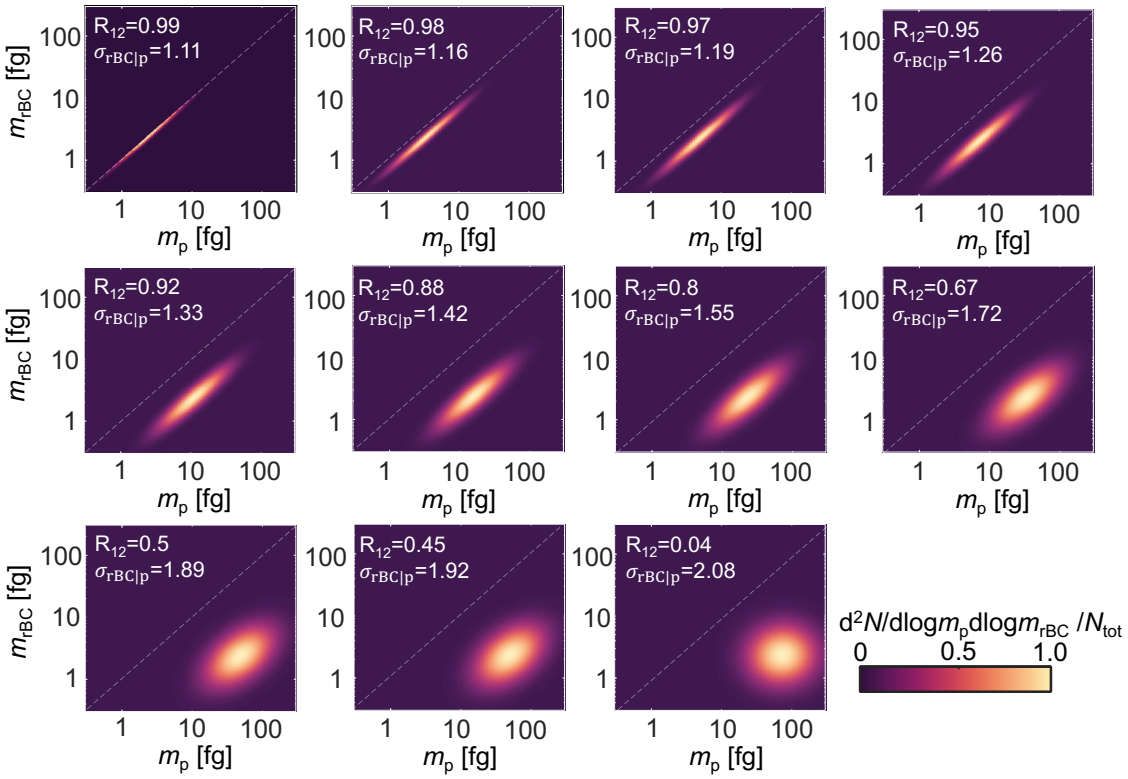

Figure 3.1: Different non-dimensionalized phantoms with diverse variability of non-refractory components between particles randomly generated for synthetic data fabrication.

some of the CPMA flow can be bypassed around the SP2). The dimensions of the commercial CPMA (Cambustion) are used for evaluation of transfer function, though most observations can be generalized to other particle mass analyzers (PMA).

3.2.3 A note on measurement duration

A key remaining quantity is the measurement duration. There are two strategies for conducting experiments. First, the experimentalist can allocate a *fixed time* for collecting particles at each CPMA setpoint, regardless of how many particles are

detected by the SP2. Such a strategy could result in poor counting statistics away from the distribution peak but does not greatly extend measurement time when there are few particles available to be counted. Alternatively, the experimentalist can allow the SP2 to continue measuring particles until a *fixed number* of SP2 counts, which we denote as N_{s}, is reached for each CPMA setpoint. Such a strategy ensures sufficient counting statistics at the expense of longer measurement durations for low particle concentrations. In this study, we generally opt for the fixed number strategies - such that N_{s} is included as part of the sampling procedure - to achieve improved counting statistics, with two caveats discussed subsequently in reference to the SP2 sampling time.

The measurement duration can then be estimated as a dependent variable made up of three components: (*i*) the total CPMA settling time, where t_{c} denotes the time it takes for the CPMA to change rotational speed and voltage for all of the CPMA setpoints; (*ii*) the sum of SP2 sampling times, where $(t_{\mathrm{s}})_{\mathrm{i}}$ is the time it takes for the SP2 to measure a certain number of particles at the i^{th} CPMA setpoint; and (*iii*) the SP2-only sampling times (t_{so}, the time it takes for the SP2 to measure a certain number of particles when the CPMA is bypassed). Then,

$$t = t_{\mathrm{c}} + \sum_{i=1}^{n_{\mathrm{c}}} (t_{\mathrm{s}})_i + t_{\mathrm{so}} \tag{3.3}$$

As shown in Figure 3.2, the CPMA sampling time is a weak function of resolution and flow rate if the CPMA flowrate is within (commonly used) limits ($Q_{\mathrm{c}} \sim 0.4 - 1.5$ L/min). Thus, the CPMA sampling time was derived empirically from a fit of

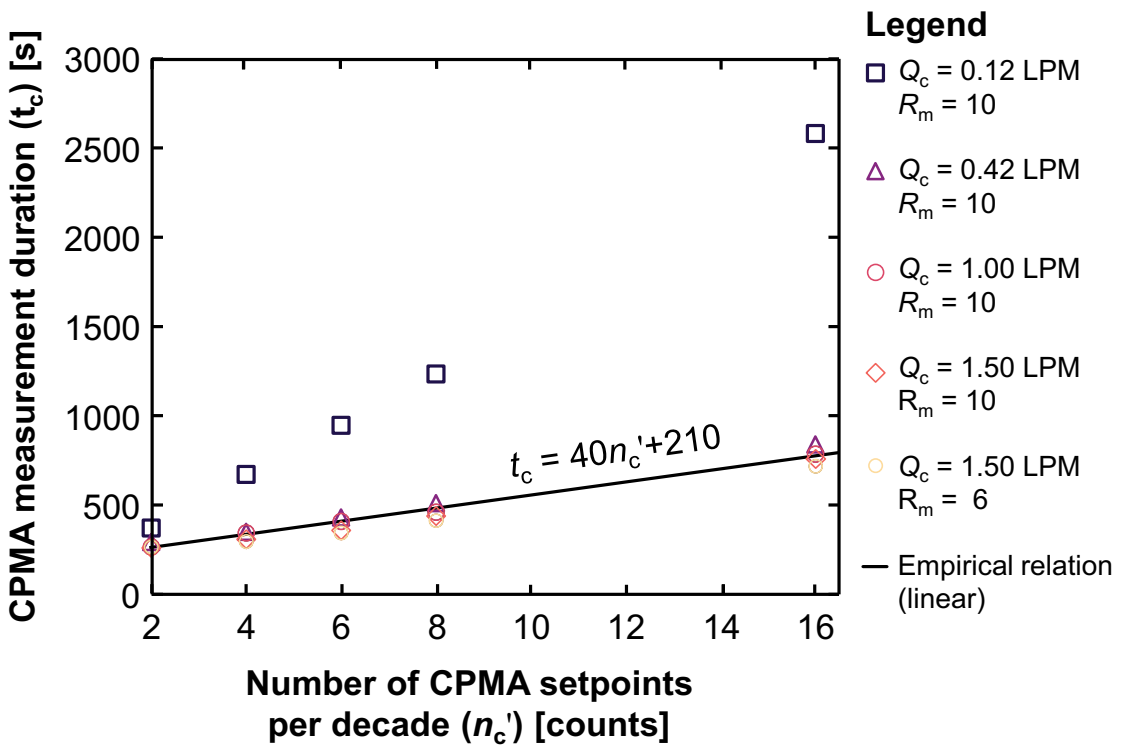

Figure 3.2: CPMA measurement durations, t_c, for stepping through 0.12 to 100 [fg] across a range of flow rates and resolutions. With the exception of the lowest flow rate, CPMA measurement durations are well-described by a single empirical function of the number of CPMA setpoints.

the data shown in Fig 2 and is

$$t_c = 40n_c' + 210. \tag{3.4}$$

Note that the n_c' intercept in Equation (3.4) is associated with the time to reach the initial rotational speed.[1]

In this scheme, the SP2 sampling time can be estimated as

[1]This is often chosen as the largest rotational speed of all of the setpoints as subsequent setpoints can be reached faster when the angular speed is ramping down.

$$(t_s)_i = \frac{N_\mathrm{s}}{Q_\mathrm{s}\,(N_c)_i}, \tag{3.5}$$

where $(N_\mathrm{c})_i$ is the number of particles output from the CPMA measurement stage at the i th setpoint, defined as

$$\begin{aligned}(N_c)_i &= \int_{-\infty}^{\infty} \Gamma_i\left(m_\mathrm{p}, d_\mathrm{m}, \Phi; {m_\mathrm{p,i}}^*, R_\mathrm{m}\right) \left.\frac{\partial N}{\partial \log m_\mathrm{p}}\right|_{m_\mathrm{p}} \mathrm{d}m_\mathrm{p} \\ &= N_\mathrm{tot} \int_{-\infty}^{\infty} \Gamma_i\left(m_\mathrm{p}, d_\mathrm{m}, \Phi; m_\mathrm{p,i}, R_\mathrm{m}\right) p\left(m_\mathrm{p}\right) \mathrm{d}m_\mathrm{p}\end{aligned}$$

where $\partial N/\partial \log m_\mathrm{p}$ it the (marginal) total particle mass distribution and $p(m_\mathrm{p})$ is the corresponding probability density function (integrating to unity). Practically, however, low number concentrations, high resolutions, and large N_s requirements can result in unrealistic measurement durations, particularly at the edge of the distributions, which can extend beyond a hundred hours. As such, we modify the above procedure using two additional caveats to generate more realistic measurement durations:

- *The 10 minute rule*: The measurement time at any given CPMA setpoint was limited to 10 minutes, which occasionally acts to truncate the counts at the edge of the distributions for the lower number concentrations, with the consequences noted in Section 3.3.

- *The 10 second rule*: Sampling is stopped if the SP2 measures less than 3 particles in 10 seconds, which implies $N_\mathrm{c} < 0.1$ cm^{-3} when the SP2 flowrate is 0.12 L/min. The 10 second rule implicitly incorporates the total number concentration (N_tot), CPMA resolution (R_m), and flowrate (Q_s).

When applied, these rules will reduce the observed number of particles measured

by the SP2 at the i^{th} setpoint, which we denote as $(N_{\text{s.col}})_i$. We then summarize using the average number concentration output by the CPMA,

$$N_{\text{s,avg}} = \frac{1}{n_{\text{c}}} \sum_{i=1}^{n_{\text{c}}} \left(N_{\text{s,col}}\right)_i . \tag{3.6}$$

As $N_{\text{s.avg}}$ is an average over all of the CPMA setpoints, it will have a single value for a given set of experimental parameters, acting as a summary quantity for a given set of experimental parameters that will recur in subsequent analysis.

The last term in Equation (3.3), *i.e.,* t_{so}, is the SP2 sampling time when the CPMA is bypassed and is calculated based on Equation (3.5) when $N_{\text{c}} = N_{\text{tot}}$ (because there is no CPMA to classify the particles). The same sampling rule as to what is described for calculation of $(t_{\text{s}})_i$ is considered here (*i.e.,* the 10 minute rule). It is worth mentioning that it is only important when collecting a high number of SP2 counts at low number concentration of particles, *e.g.,* the ten-minute rule is applicable when the goal is to collect $50,000$ SP2 counts at $N_{\text{tot}} = 40$ cm^{-3}.

3.2.4 Ranges considered for the experimental parameters

Overall, the preceding procedure results in a set of six parameters, listed in Table 3.2, which are sampled randomly in a Monte Carlo procedure. In general, parameters are sampled uniformly on a log-scale within a physical range, with modifications made to account for parameters that are necessarily discrete (*i.e.,* for $N_{\text{s}}, n_{\text{c}}$ and n'_{s}).

Table 3.2: The experimental parameters that are sampled in the Monte Carlo procedure. All parameters are sampled approximately from a log-uniform distribution, with special consideration for discrete variables. Secondary parameters, including measurement duration, the average number of SP2 counts per CPMA setpoint, and phantom correlation, were computed from these quantities. Note that the CPMA resolution is sampled to values lower than that typically used in experiments (and, on occasion, to values that cannot be realized for the given flow rates), to emphasize trends at the extreme lower range of the quantity.

Symbol	Sampled parameters	Distribution	Range	Unit
R_m	CPMA resolution	Log-uniform	[0.01, 10]	-
n_s'	Number of SP2 bins per decade	Log-uniform, rounded	$[2, 10^3]$	counts
n_c'	Number of CPMA setpoints per decade	Log-uniform, rounded	[2, 20]	counts
N_{tot}	Total number concentration of particles	Log-uniform	$[10, 10^6]$	cm^{-3}
N_s	Number of SP2 counts per CPMA setpoint	Log-uniform	$[100, 10^6]$	counts
$m_{p,g}$	The geometric mean of particle mass distribution	Log-uniform	[2.37, 70]	fg

The effective SP2 kernel described in Chapter 2 allows for an arbitrary number of bins. While a large n_s' is precise, it will result in noisier counts in each bin and in a large $\mathbf{A}^1$ matrix, with an associated increase in computational effort for the inversion. Accordingly, given that the number of bins does not affect the measurement duration, we chose a large interval for n_s': $[2, 10^3]$, with subsequent notes about the computational effort.[2]

Unlike n_s', n_c' is related to the measurement duration, which limits the reasonable range of the CPMA setpoints that were considered. Also, unlike n_s, n_c' does not have an effect on the noise, as the experiments are generally performed such that a specific number of particles is collected at each CPMA setpoint. Accordingly, we considered

[2]We explicitly note that the number of measurements (*i.e.,* the size of the $\mathbf{A}^1$ matrix and $\mathbf{b}$ vector) vary considerably with the sampled parameters, causing the problem to move from being underdefined to overdefined (the latter particularly occurs when n_s' is at the upper end of its range).

the smaller domain of $[2, 20]$, where the upper limit yields measurement durations that still occasionally extended above 10 hours for low number concentrations.

The CPMA resolution has consequences for the measurement duration, the precision with which particles are mass-selected, and the degree to which the CPMA transfer function spans between CPMA setpoints. The latter aspect relates back to n_c', where high R_m and low n_c' will result in a significant contingent of particles falling between CPMA transfer functions and and thus not having an impact on the instrument response.

For instance, the peak of the distribution could be missed if it is not aligned with a CPMA setpoint (*e.g.,* if the peak of the distribution is at $m_p = 3$ fg and the CPMA setpoints are such that measurements are taken at $m_p = 1$ and 5 fg with a high resolution, then $m_p = 3$ fg particles are not measured and the peak will be underrepresented). The inversion then has no knowledge of these particles, with consequences for reconstruction accuracy (*e.g.,* the peak of the distribution could be missed entirely, depending on the positioning of the CPMA setpoints). As such, the present work considers R_m below that typically used in (or even possible with) the commercially-available CPMA (Cambustion). These theoretical operating points are also modeled to better elucidate larger trends (showing where the inversions are poor). We discuss the feasibility of the recommended CPMA resolutions in Section 3.4. The total number concentration of particles (*i.e.,* $[10, 10^6]$) is selected to span the range of atmospheric number concentrations corresponding to clean and heavy air pollution episodes (Cheng et al., 2012; Reddington et al., 2013; Wang et al., 2014a) and the total number concentration of particles generated from combustion (Sipkens et al., 2021).

As noted above, a range of phantoms is realized by sampling the geometric mean of particle mass, $m_{\mathrm{p,g}}$, which is sampled randomly over a domain of $[2.35, 70]$ fg. When combined with the empirical relation for R_{12}, Equation (3.2), considered phantoms (*i.e.,* distributions) span from uncoated particles to the heavily coated particles that were observed by Yu et al. (2020) for an urban aerosol.

3.3 Results and discussion

Reconstructions are assessed using the relative Euclidean error, *i.e.,* the Euclidian distance between the reconstructed distribution and the known ground truth divided by the ℓ^2-norm of the ground truth. Note that because the reconstruction grid is identical for all of the reconstructions, the considered error metric is reasonable for all of the simulations.

3.3.1 Trends in reconstruction error with experimental parameters

Figure 3.3 depicts trends in the reconstruction error as a function of the input parameters. The relative Euclidean error appears on the vertical axis with favorable reconstructions with lower reconstruction errors towards the bottom and unfavorable reconstructions towards the top. Each panel depicts one of the six independent variables on the horizontal axis, except for the first panel, which indicates $N_{\mathrm{s,avg}}$, following Equation (3.6). A strong dependence (for a main effect) manifests as a discernable slope in these plots (*e.g.,* Figure 3.3*a*), whereas a flat scattering of points indicates less dependence (*e.g.,* Figure 3.3*d*). The significance of the input variables is also assessed using an analysis of variance (ANOVA), including two-factor

interactions, shown in Table 3.2., sorted by $p-$value.

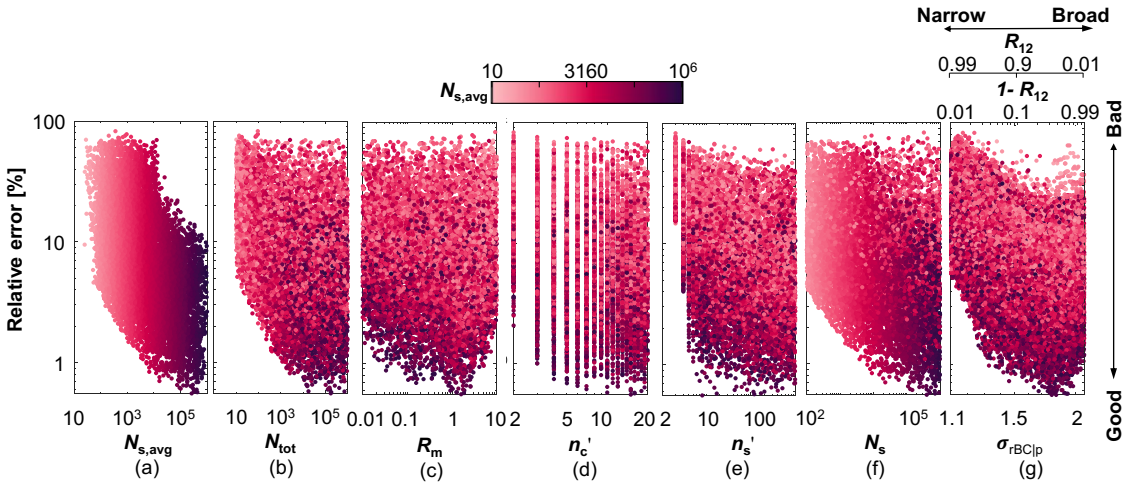

Figure 3.3: The sensitivity of reconstruction accuracy to (a) the average number of SP2 counts per CPMA setpoint, (b) the total number concentration of aerosol, (c) the CPMA resolution, (d) the number of CPMA setpoints per decade, (e) the number of SP2 bins per decade, (f) the number of SP2 counts per CPMA setpoint, and (g) the distribution width. All panels are log-log. The results give an indication of the main effects. The results are most sensitive to the average number of SP2 counts per CPMA setpoint ($N_{\mathrm{s.avg}}$), such that points are colored according to it throughout the panels. The phantom parameters are varied, with trends with respect to that parameter examined later in this work.

Table 3.3: Six-way ANOVA table identifying the significance of the main effects and two-factor interactions on the reconstruction error, sorted by ascending p-value. Interaction effects are denoted with two symbols separated by an asterisk. The quantity $1 - R_{12}$, which is the dependent variable plotted in Figure 3.3, is used in the place of $m_{\mathrm{p,g}}$. The analysis is run using the logarithm of both the stated quantities and reconstruction error and treats all of the parameters as continuous. The last three effects (*i.e.,* N_{s}) are considered negligible as its p-value exceeds 0.05.

Source	F	p-value
N_{s} * N_{tot}	1145.0	$< 10^{-100}$
R_{m}* n_{c}'	750.6	$< 10^{-100}$
n_{s}'	404.7	1.6×10^{-88}
R_{m}* N_{tot}	282.7	1.0×10^{-62}
n_{c}'* N_{tot}	265.6	4.6×10^{-59}
R_{m}	182.3	3.0×10^{-41}
N_{tot} * $(1 - R_{12})$	175.7	8.1×10^{-40}
R_{m}* N_{s}	147.1	1.2×10^{-33}
n_{s}' * n_{c}'	105.4	1.3×10^{-24}
n_{s}' * N_{s}	90.7	2.0×10^{-21}
n_{s}'* $(1 - R_{12})$	77.3	1.6×10^{-18}
N_{tot} * n_{s}'	74.5	6.7×10^{-18}
N_{tot}	74.3	7.6×10^{-18}
N_{s} * $(1 - R_{12})$	57.4	3.8×10^{-18}
R_{m} * $(1 - R_{12})$	49.5	2.1×10^{-14}
1-R_{12}	32.3	1.4×10^{-12}
n_{c}'* $(1 - R_{12})$	11.2	8.0×10^{-8}
R_{m} * n_{s}'	7.3	7.0×10^{-3}
n_{c}'*N_{s}	2.0	1.6×10^{-1}
n_{c}'	0.8	3.6×10^{-1}
N_{s}	0.3	6.1×10^{-1}

Figure 3.3 indicates that the strongest relationship is associated with the average number of SP2 counts per CPMA setpoint, $N_{\text{s.avg.}}$ Trends are as expected. Higher counts measured by the SP2 generally allow for better signal-to-noise ratio and thus result in lower errors, from an average of 40% at $N_{\text{s.avg}} \lesssim 50$ to 4% at $N_{\text{s.avg}} \gtrsim 50,000$. The data is scattered around this central trend due to the effects of the other input parameters (*e.g.,* the effect of a range of R_{m}) and different realizations of noise. Points are colored according to $N_{\text{s.avg}}$ throughout the other panels. The discussed observation is generally supported by the ANOVA, where the $N_{\text{s}} * N_{\text{tot}}$ interaction - which is largely determinant of $N_{\text{s avg}}$ is the most significant source of variance in the reconstruction error. This follows from noting that (i) $N_{\text{s}} = N_{\text{s,avg}}$ when measurement durations are reasonable and that (ii) N_{tot} is the primary independent variable determining if enough particles are present to reach that threshold measurement duration.

Structure in the $N_{\text{s}} * N_{\text{tot}}$ interaction is demonstrated in Figure 3.4*a*, by plotting trends with N_{s} for the extremes of N_{tot} . For $N_{\text{tot}} > 160,000$, the reconstruction error declines with N_{s}, consistent with the overall trend noted for $N_{\text{s,avg}}$ in Figure 3.3. In contrast, for $N_{\text{tot}} < 54$, the reconstruction error is nearly constant, as the 10 minute and 10 second rules are nearly always applied and thus N_{s} has no bearing on the measurements (and thus no effect on the error). These observations also drive the shape of the trends with N_{tot} and N_{s} in Figure 3.3*b* and *f*, where the larger spread of errors at the upper end of those panels derives from the effect of the other variable. Comparing Figure 3.3*a* and *f* indicates that the ten minute and second rules are invoked rather rarely in terms of all simulations, and it clearly shows where they are significantly invoked (*i.e.,* the points where the linear change in colour is disrupted).

As expected, these points are associated with low N_{tot} and high R_{m}, resulting in very low $N_{\mathrm{s,avg}}$, and consequently larger errors on average. This demonstrates that N_{s} is often an acceptable surrogate for N_{savg}, which is important given that N_{s} is parameter the experimentalist can control.

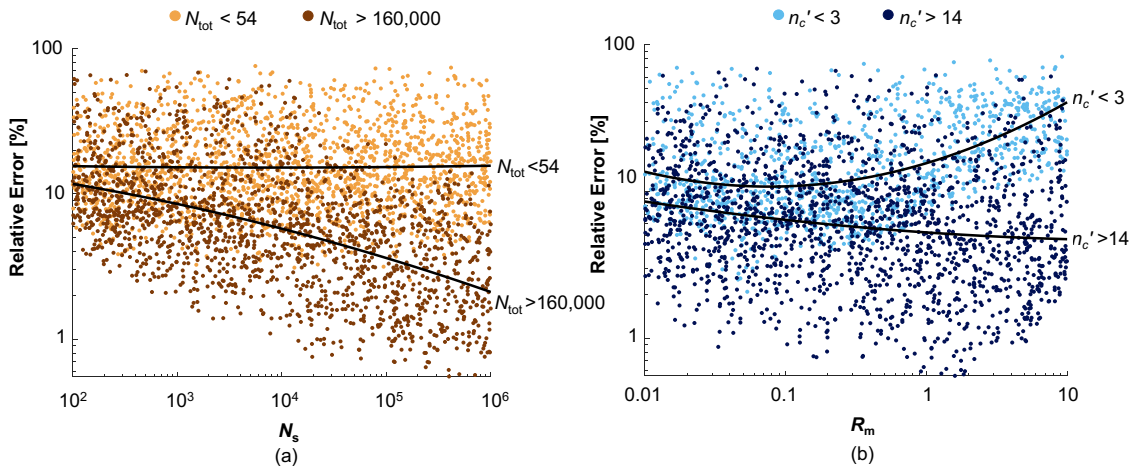

Figure 3.4: Trends in the reconstruction error (a) with N_{s} at the limits of N_{tot}, and (b) with R_{m} at the limits of n_c'. The two data clouds in panel a diverge for higher values of N_s, with an upward trend in error for $N_{\mathrm{tot}} < 54\,[\,\mathrm{cm}^{-3}]$ and a persistent downward trend for $N_{\mathrm{tot}} > 160,000\,[\,\mathrm{cm}^{-3}]$. The two data clouds in panel b diverge somewhat for higher values of R_{m}, with an upward trend in error for $n_{\mathrm{c}}' < 3$ and a persistent downward trend for $n_c' > 14$. Solid lines correspond to the moving average of the different data sets in log-log space, with the upper lines correspond to $N_{\mathrm{tot}} < 54\,[\,\mathrm{cm}^{-3}]$ and $n_{\mathrm{c}}' < 3$; and the lower lines corresponds to $N_{\mathrm{tot}} > 160,000\,[\,\mathrm{cm}^{-3}]$ and $n_{\mathrm{c}}' > 14$ in panel a and b, respectively. Trend lines are only provided to guide the eye.

The second most significant effect, $n_{\mathrm{c}}' * R_{\mathrm{m}}$, can be similarly visualized by plotting trends in the error at the limits of n_{c}'. In Figure 3.4*b*. For standard resolutions ($R_{\mathrm{m}} > 1$), opposite trends are observed in the error. Overall, for $n_{\mathrm{c}}' > 14$, the reconstruction error mostly trends downwards with increasing R_{m}. This is consistent with expectations, where these long experiments often benefit from the increased speci-

ficity of having narrow CPMA transfer functions. Inspection of the cloud of errors presents limits to the simple interpretation mentioned above, however, where the bottom of the cloud shows a clear inflection upward for $R_{\mathrm{m}} > 3$ even when $n'_{\mathrm{c}} > 14$. The upward inflection likely stems from the narrow transfer functions effectively reducing $N_{\mathrm{s,avg}}$, with the corresponding impact on the signal-to-noise ratio. Collecting considerably less $N_{\mathrm{s,avg}}$ via the narrow transfer functions would be particularly true when N_{tot} is low, such that the $N_{\mathrm{tot}} * R_{\mathrm{m}}$ interaction effect is also listed high in Table 3.3. As such, there appears to be little motivation to use $R_{\mathrm{m}} > 5$, a sentiment repeated in Section 3.4.1. In contrast, for $n'_{\mathrm{c}} < 3$, the reconstruction error trends upwards through the $R_{\mathrm{m}} > 3$ region. Here, large gaps between the CPMA setpoints yields two possibilities: (*i*) maintain a high degree of specificity with respect to m_{p} at the expense of any information about particles between the CPMA setpoints or (*ii*) use low resolutions so as to still measure the particles between the CPMA setpoints. The observations here indicate the latter scenario is preferred, that is, a limited number of CPMA setpoints (*e.g.,* because of limited measurement duration) should be accommodated by using lower resolutions. We note that similar principles (*i.e.,* measure the particles between the PMA setpoints via using higher PMA resolution) are likely to apply to particle mass analyzer-SMPS measurements. In fact, this effect receives some qualitative support in Sipkens, Olfert, and Rogak (2020b) for moderate CPMA resolutions used in that work. Interestingly, trends for the main R_{m} and n'_{c} effects (*cf.* Figure3.3*c* and *d*) are relatively weak and are better governed by rules derived from the $R_{\mathrm{m}} * n'_{\mathrm{c}}$ interaction effect.

The ANOVA analysis indicates that the most dominant main effect is n'_{s}. Increasing n'_{s} to hundreds of SP2 bins per decade improves reconstruction accuracy by

nearly an order of magnitude. The inclination toward using higher n'_s indicates a clear preference for resolution in terms of m_{rBC}, over signal-to-noise ratio improvements obtained by grouping the measurements into larger bins, and suggests that the practitioner should use as large a n'_s as possible. However, there are practical limitations to the above-mentioned recommendation. For instance, increasing n'_s results in a linear increase in the size of the system to be solved during inversion. As computational effort scales with the size of the system squared, such an increase can quickly become computational impractical.[3] At the same time, there are limits in terms of the precision of the particle mass output by the SP2. We recommend that n'_s be set such that the bin width approximates the uncertainties in the m_{rBC} measured by the SP2, beyond which the extra resolution is no longer useful and computationally expensive. We propose $n'_s = 64$, which implies an SP2 accuracy of 3.7%[4] . Relevant interaction effects involving n'_s include $n'_s * n'_c$ and $n'_s * N_s$. The latter interaction likely refers to the limitations for low N_s, where a high number of SP2 bins only acts to increase the number of empty bins rather than increasing precision.

[3]number of entries in $\mathbf{A}^1$, which ultimately needs to be inverted, scales as a multiple of n'_s. Specifically, the overall number of entries in $\mathbf{A}^1$ is given by $n_A = n'_s n'_c n_j(1 + n_j)/2$, where n_j is the number of elements per dimension of the reconstruction grid (assuming a square grid). For an appropriate reconstruction grid (*e.g.,* 100×100, that is $n_j = 100$) and moderate $n'_c = 10$, the upper end of $n'_s = 1000$ yields more than 454 million entries in $\mathbf{A}^1$ (*i.e.,* approx. 3.6 GB assuming double precisions floats).

[4]Logarithmically spaced points, by their nature, imply that adjacent bins are a constant multiple apart. For instance, $[1, 3.1622, 10]$ is logarithmically spaced data with a multiple of 3.1622. The precision of the input to the inversion algorithm is limited by the bin widths, with individual particle information grouped below this. Then, the minimum precision implied by the spacing above, normalizing by the lower bound (i.e., Precision $= 100\% \times (\text{UB} - \text{LB})/\text{LB} = 100\% \times (\text{UB}/\text{LB} - 1)$, where LB is the lower bound of the bin and UB is the upper bound of the bin). As UB/LB is a constant, the precision can be stated as a single percentage over the entire range. In such a case, for 64 logarithmically spaced points per decade, Precision $= 100\% \times (1.037 - 1) = 3.7\%$

3.3.2 Effect of phantom width

In general, Figure 3.3*g* indicates that wider phantoms (small R_{12}) are associated both with lower reconstruction errors (on average) and with a larger range of errors. The downward trend with increasing phantom width is partially rooted in the nature of the Tikhonov regularization scheme, which dampens noise at the expense of sharp edges, a feature that is limiting for very narrow m_{p}—m_{rBC} distributions (*cf.,* Chapter 2).

ANOVA analysis indicates that interaction effects dominate trends in the reconstruction error with distribution width. Figure 3.5 shows these effects by plotting trends in the reconstruction error for very broad ($R_{12} < 0.5$) and very narrow phantoms ($R_{12} > 0.98$). Figure 3.5 affirms observations from the ANOVA analysis, with distinct reconstruction error profiles for the broad and narrow phantoms. Of note, Figure 3.5*a* indicates a strong preference for higher CPMA resolutions for narrow distributions; while $R_{\mathrm{m}} \sim 1.5$ is optimal for broad distributions. Figure 3.5*b* also shows that while employing higher values of n_{s}' would improve the reconstruction accuracy of broad distributions, there is no benefit for thin distributions. The reconstruction errors in other panels trend in the same direction for the broad and narrow phantoms, with a simple shift to larger errors for the narrow phantoms.[5]

3.3.3 Trends with measurement duration

Figure 3.6 shows trends in the reconstruction error as a function of the estimated measurement duration, which ranges from 5 minutes to 10 hours. The various ex-

[5]There is a structural relationship between the optimal Tikhonov regularization parameter and the distribution width (*cf.,* Appendix C for more information.

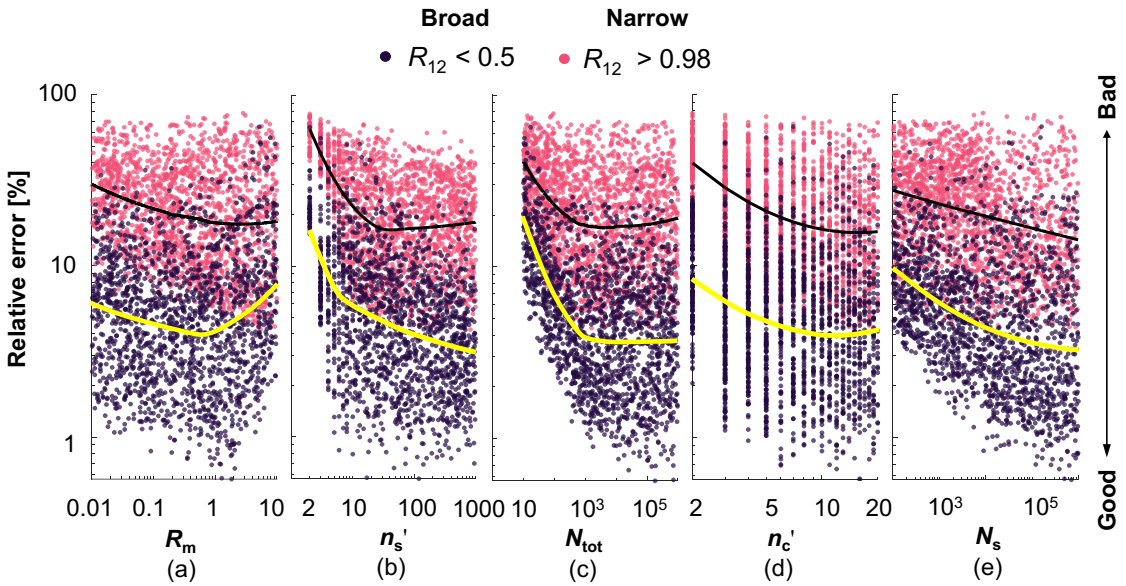

Figure 3.5: Trends in the reconstruction error with (a) R_m, (b) n_s', (c) N_{tot} , (d) n_c, and e) N_s for broad ($R_{12} < 0.5$) and narrow ($R_{12} > 0.98$) phantoms. Solid, black lines correspond to moving average of the different data sets in log-log space, with the upper black line correspond to $R_{12} > 0.98$ and the lower yellow line corresponds to $R_{12} < 0.5$, which are provided to guide the eye.

perimental parameters are used to color the data in each panel. Within these panels, vertical layering of colour (*e.g.,* progressing from light to dark from left to right, as is the case for N_{tot} N_s, and n_c') indicate a strong functional dependence between the measurement duration and a given parameter. By contrast, strong horizontal layering indicates a stronger dependence between a parameter and the error, without any dependence on measurement duration. The minimum measurement duration corresponds to the lowest values of n_c', when N_{tot} is high ($> 10^5$) and, when N_s is low (*e.g.,* 100). In these cases, the time required to stabilize the CPMA at each setpoint dominates measurement duration. The longest measurements durations are associated with cases where $n_c' = 20$ (the maximum), very low N_{tot} (*e.g.,* $N_{\mathrm{tot}} = 10 \sim \mathrm{cm}^{-3}$);

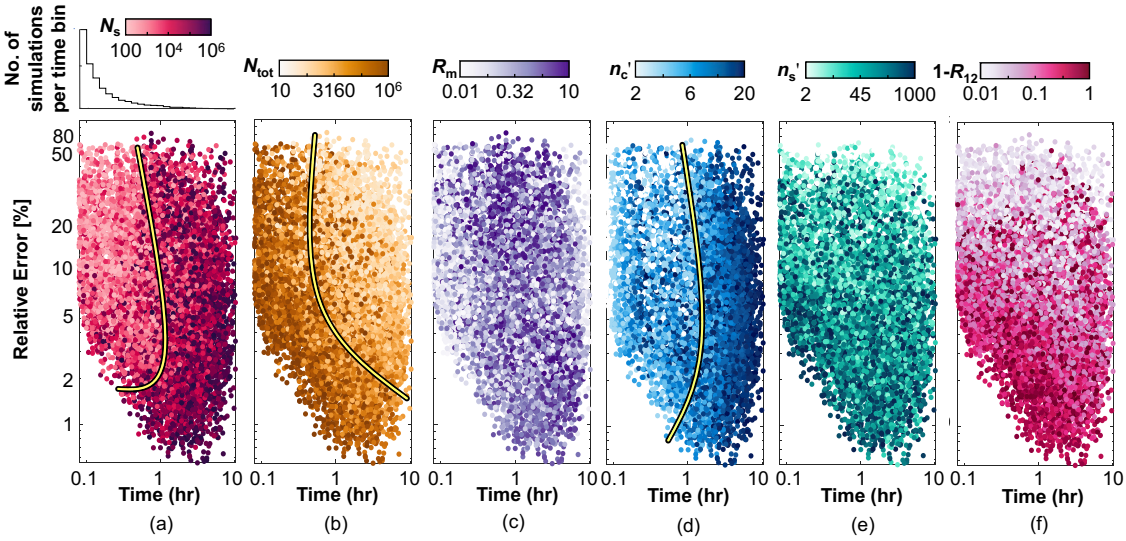

Figure 3.6: The sensitivity of reconstruction accuracy and experiment time to (a) the number of SP2 counts per CPMA setpoint, (b) the number of SP2 bins per decade, (c) the number of CPMA setpoints per decade, (d) the number concentration of aerosol, (e) the CPMA resolution, and (f) width of phantoms. The histogram in panel (a), indicates the number of simulations that fall into each time interval. Annotated lines are intended to guide the reader to distinguishing layering of colors, generally splitting darker and lighter points.

resembling the remote area condition), and large N_s (*e.g.*, $N_s > 10^5$), where the ten-minute rule is applicable for all CPMA setpoints and t_s dominates measurement duration. Horizontal layering of colour in panels Figure 3.6*e* and *f* indicates that phantom width and the number of SP2 bins do not affect the measurement duration, but that they still have significant effects on the overall reconstruction error, consistent with discussion in Section 3.3.1.

3.4 Recommendations and experimental procedure

Thus far, we have demonstrated the overall trends in the error as a function of experimental parameters. These results can be used to make recommendations by choosing the experimental parameters that give the lowest Euclidean error for a given set of experimental conditions. Figure 3.7 demonstrates the overall procedure for making CPMA-SP2 measurements using an optimal sampling configuration (*i.e.*, optimal simulation), which is defined as the conditions that give the lowest Euclidean error relative to the input $m_{\mathrm{p}} - m_{\mathrm{r}BC}$ distribution. Start by obtaining estimates of N_{tot} and R_{12}, using an SP2-only measurement (where the CPMA is bypassed) to estimate $N_{\mathrm{tot},}$, collecting around 50,000 particles, and using a single CPMA setpoint to estimate R_{12}.

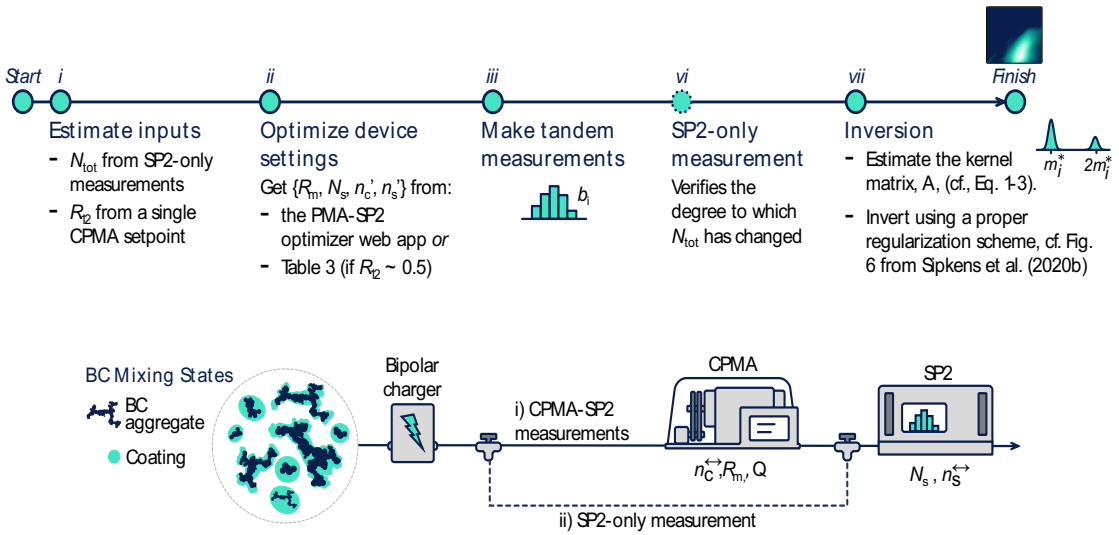

Figure 3.7: Schematic demonstrating the recommended procedure to conduct a proper CPMA-SP2 measurement using optimal sampling configurations recommended by PMA-SP2 Optimizer.

For the latter quantity, the distribution at a single CPMA setpoint can be used to estimate the conditional width, *i.e.,* $\sigma_{\mathrm{rBC|p}}$, which is the geometric standard deviation of the measurements at the single CPMA setpoint, and converted to R_{12} using (Bertsekas and Tsitsiklis, 2002)

$$R_{12} = \left[1 - \left(\frac{\log_{10}\sigma_{\mathrm{rBC|p}}}{\log_{10}\sigma_{\mathrm{rBC}}}\right)^2\right]^{1/2} \tag{3.7}$$

where σ_{rBC} is the width from the SP2-only measurements. We refer the reader to Appendix B for the derivation of this expression.

One can use the estimated R_{12}, $N_{\mathrm{tot},}$ with the desired measurement duration as inputs to determine the optimal sampling configuration. We have provided a web app (PMA-SP2 optimizer) as supplement to this thesis to accomplish this task generally. The web app bins the simulation and reports the experimental settings that gave the lowest Euclidean error within the specified bin. The number concentration range within the web app includes all simulations above and below the specified value, thus spanning half an order of magnitude. For the input time, the app considers any simulation with a time below that specified. For R_{12}, the range is shown explicitly in the drop down menu in the app. Finally, use the optimum sampling configurations to conduct the CPMA-SP2 measurements through an entire BC-containing population. Following the experiment, the practitioners may wish to perform a second SP2-only measurement, to determine whether the number concentration of rBC particles has changed over the measurement period. The average of these two SP2-only measurements can be used as a sort of prior information and incorporated into the inversion procedure as prior information distribution (*cf.*, Chapter 2).

Following measurement, the collected data is preprocessed and discretized based

on N_s (*i.e.,* computing **b**, *cf.,* Section 2.2.3). Finally, compute the kernel matrix, **A**, and invert the system using an appropriate regularization scheme to retrieve the m_p—m_{rBC} distribution.

3.4.1 Case study: Recommendations for moderately broad distributions

Rather than present detailed recommendations for all distribution shapes, for which it is challenging to define a simple set of rules, we present a case study for a moderately broad $m_{rBC} - m_p$ distribution ($R_{12} = 0.5$), consistent with an aged aerosol, for which we run a dedicated set of 8,000 Monte Carlo simulations. Figure 3.8 shows examples of the distributions retrieved by data fabricated by choosing different sampling configurations for CPMA-SP2 measurements. A sampling of the resultant

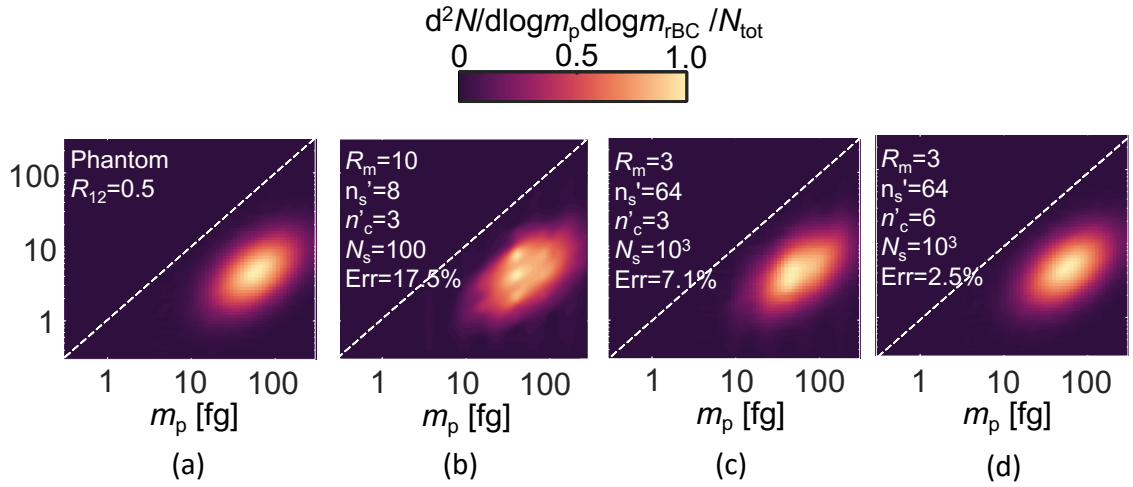

Figure 3.8: a) The synthetic distribution of rBC particles characterized with $R_{12} = 0.5$ and $N_{tot} = 1000$ cm^{-3} and b-d) examples of distributions retrieved from fabricated data derived by setting different CPMA-SP2 configurations. Reconstruction accuracy improves from panel b to d.

reconstructions shown in Figure 3.8*a* to *d* illustrates how the quality of the reconstruction depends on the CPMA-SP2 configurations relative to the phantom, shown in Figure 3.8*a*.

Figure 3.9 shows the optimal settings for the $R_{12} = 0.5$ case study, after binning the new set of simulations with respect to measurement duration and total number concentration, which allows for more precise predictions for such a specific case. Figure 3.9 shows that long experiments (large t) and high number concentrations (large N_{tot}) - corresponding to the upper, right corner of the panels - yield lower reconstruction errors. These experiments correspond to optimal values with larger n'_{c} and R_{m}. Consistent with Figure 3.9*b* and the surrounding discussion, n'_{c} and R_{m} often change together. This is particularly true for short experiments, where a reduction in n'_{c} (necessary to reduce measurement duration) is associated with a substantial drop in R_{m} that allows the instruments to still measure the particles between the CPMA setpoints. Generally, however, a value of $R_{\mathrm{m}} = 1$ to 2 covers all but a few scenarios (low N_{tot} and t).

The optimal number of SP2 bins, n'_s, is uniform relative to the input sampling range, with a preference towards $n'_{\mathrm{s}} > 100$, consistent with the discussion in Section 3.3.1. As noted in Section Section 3.3.1, there are limitations in terms of computational effort and precision, such that realizing these larger values is often impractical. As such, we recommend $n'_{\mathrm{s}} = 64$ across the entire domain, consistent with the previous discussion. However, in some exceptional circumstances, at other values of R_{12} when the particle concentration is low (e.g., $N_{\mathrm{tot}} < 100$), and the measurement duration is short (*e.g.,* less than 30 minutes), it is advantageous to discretize the SP2 counts on fewer SP2 bins (i.e., $n'_{\mathrm{s}} < 64$, refer to PMA-SP2 Optimizer for optimum

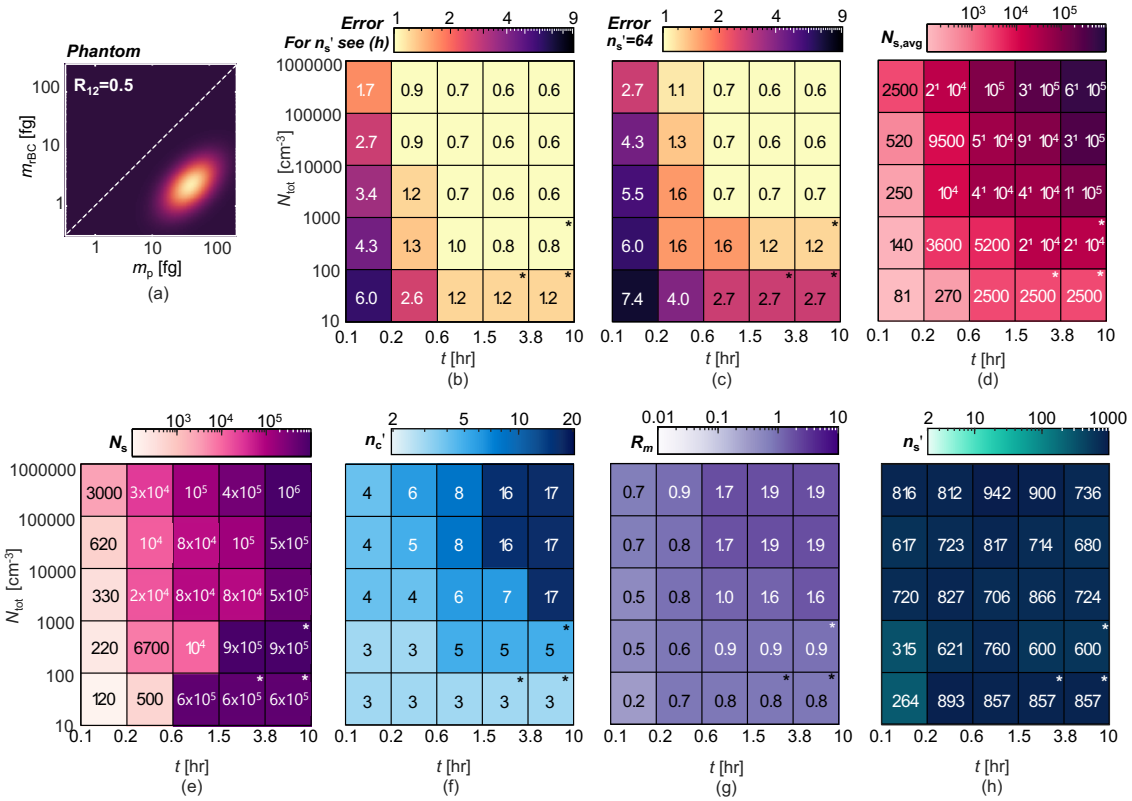

Figure 3.9: Optimum sets of experimental parameters as a function of number concentrations and measurement durations for the case study where $R_{12} = 0.5$. (a) Phantom with $R_{12} = 0.5$. (b) The reconstruction accuracy for (d) the average number of SP2 counts per CPMA setpoint, (e) the number of SP2 counts per CPMA setpoint, (f) the number of CPMA setpoints, (g) the CPMA resolution, and (h) the number of SP2 bins per decade. Results are binned over a 5×5 grid of discretized time-number concentration space spanning from 0.1 to 10 [hr] and 10 to 10^6 [cm^{-3}]. The plots are colored based on the minimum and maximum possible range of all variables considered in the present study. Panel (c) shows the reconstruction accuracy when the optimum simulations are repeated by replacing n_s' with 64 , while (b) indicates the unconstrained reconstruction accuracy corresponding to the n_s' in (h). Cells decorated by an asterisk indicate that the same set of experimental parameters is derived for each concentration limit (i.e., each row).

n'_s in such cases). Figure 3.9*c* shows that the reconstruction accuracy that results by using $n'_s = 64$ (a similar resolution as what is commonly used in SMPS inversion) is normally within 2% of that produced for higher values of n'_s.

Finally, a note on CPMA flow rate. Figure 3.10 indicates that the full range of recommended resolutions (generously, within $R_m = 0.6$ to 2) are achievable with the commercially-available CPMA (Cambustion) for the range of masses considered here (*i.e.,* 0.3-300 fg) with a flowrate of $Q_c = 1.5$ L/min.

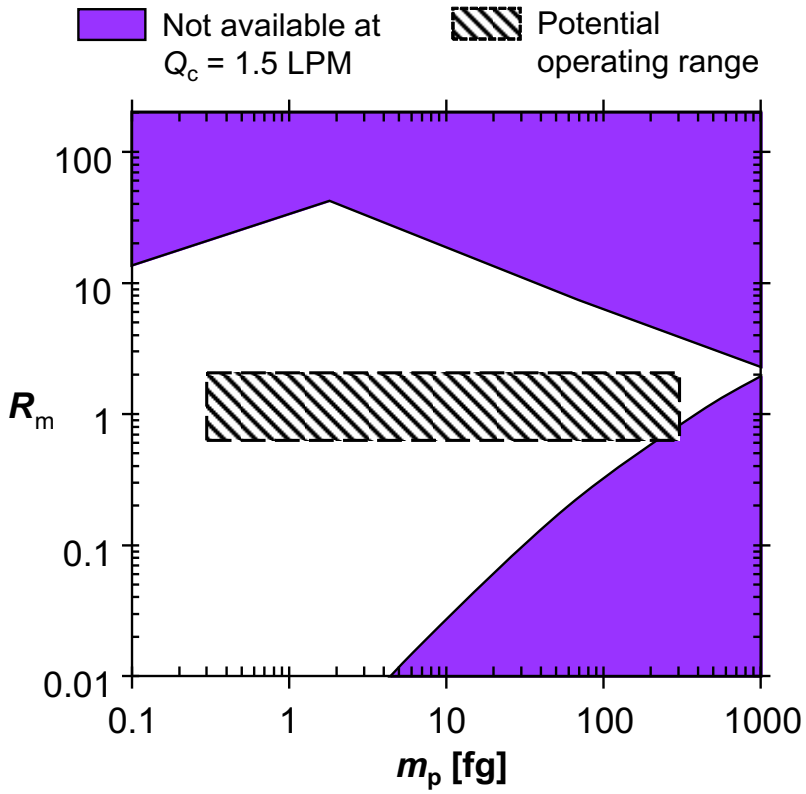

Figure 3.10: Limitations in Cambustion's CPMA operating conditions at a flowrate of $Q_c = 1.5$ L/min. The grey, shaded regions represents the optimum range of CPMA resolution recommended in the present study.

While $Q_c = 0.3$ L/min would also cover nearly 75% of the recommended operational range, flowrates of $Q_c < 1$ L/min are not recommended due to the increases in the particle residence time in the CPMA, which increases diffusional losses. On the other hand, while higher CPMA flowrates are theoretically possible, flowrates above 1.5 L/min are rarely used in the literature. It is also worth noting that the CPMA transfer function had only been experimentally validated only for moderate and high resolutions (Olfert et al., 2006) at the time of this study, but not at lower resolutions (*e.g.,* $R_m < 1$). In the next chapter, the CPMA transfer function will be experimentally evaluated for broad ranges of operating configurations (*i.e.,* $2 < R_m < 15$, 0.3 LPM $< Q <$ 8 LPM, and 0.05 fg$< m^* <$ 100 fg). Furthermore, running the CPMA with lower resolution at high mass setpoints results in low rotational speeds and allows a larger number of small uncharged particles to penetrate the CPMA.[6] The observation that these lower resolutions are preferred in cases of limited measurement duration suggests the utility in pursuing experimental validation of the transfer function in these regions, noting the increased need for inversion to properly interpret these measurements (wider transfer functions result in measurements that deviate more significantly from the shape of the underlying size distribution).

In summary, when limiting the measurement duration to an hour and assuming $R_{12} = 0.5$ (moderately broad distributions, such as those for atmospheric black carbon), practitioners should use the settings recommended in Table 3.4.

[6]The presence of uncharged particles can be addressed by either (*i*) including uncharged particles in the kernel function (i.e. including $\Phi = 0$ in the summation in Equation (2.2.1) or (*ii*) by preprocessing the collected data, *e.g.,* by comparing scattering and incandescent signals, as the small, rBC-containing uncharged particles have very small masses and would have tiny scattering signals relative to the charged particles.

Symbol	Instrument configuration	Recommended	Unit	Notes
R_m	CPMA resolution	1-2	-	Lower resolutions are recommended when n'_c is at the lower end of the range.
n_s'	Number of SP2 bins per decade	64	counts	Higher is better, but lower values can be used when testing.
n_c'	Number of CPMA setpoints per decade	6-8	counts	cf. note for R_m
N_s	Number of SP2 counts per CPMA setpoint	5000-50000	counts	This will depend on the total BC number concentration (e.g., 5000 at N_{tot} = 100, 50000 at N_{tot} = 10000)
Q_c	CPMA flow rate	1.5	LPM	Recommendation is not explicitly derived from sampling procedure.

Table 3.4: Recommended operating envelope for conducting the CPMA-SP2 measurements within an hour for the case study of a moderately broad phantom ($R_{12} = 0.5$).

3.5 Conclusions

The present study identifies the optimum sampling settings (*i.e.,* the number of CPMA setpoints per decade n'_c, the number of SP2 bins per decade (n'_s), CPMA resolution (R_m), and the number of SP2 counts per CPMA setpoint (N_s) to determine accurate two-dimensional m_p—m_{rBC} distributions. The non-linear relationship between the reconstruction error and experimental parameters is realized using a Monte Carlo analysis. Optimal experimental parameters are stated for a moderately broad m_p—m_{rBC} distribution. Generally, optimal results are achieved by collecting as many particles as possible and by using as many SP2 bins per decade (n'_s) as is computationally reasonable. The authors noted that reconstruction accuracy that results by conditioning on $n_s < 64$ is almost within 2% of that produced for higher values of n'_s. Thus, n'_s is recommended to be on the order of 50 to 100. Note, that n'_s can be adjusted in post-processing of the data and thus can be adjusted by the user based on a trade-off between computational time and solution convergence. For a

moderately broad $m_{\mathrm{rBC}}-m_{\mathrm{p}}$ distribution (*cf.,* Figure 8; $R_{12}=0.5$), and also valid for most of a given range of phantoms, reconstruction accuracies below 3% are achieved by using 3-8 CPMA setpoints over three decades (according to measurement time availability) with CPMA resolutions around 1. These operating conditions can be achieved using a flowrate of $Q_{\mathrm{c}}=1.5$ L/min for Cambustion's CPMA.

In this study we only explicitly show optimized results for moderately broad $m_{\mathrm{p}}-m_{\mathrm{rBC}}$ distributions ($R_{12}=0.5$). For the convenience of CPMA-SP2 users, we have posted a web-based application capable of making recommendations for broader range of experimental conditions. This simple web app is designed to estimate the optimal experimental parameters for tandem PMA-SP2 experiments. The app uses data from this study, and optimal experimental parameters are chosen as the minimum simulated error for a specific range about chosen for (*i*) measurement duration, (*ii*) total number concentration of BC particles and (*iii*) R_{12} as and input parameters to the app. For the number concentration, the range is up and down by a factor of $\sqrt{10}$, thus spanning an interval of an order of magnitude. For time, the app considers any simulation with a time below that specified. In other words, if the optimal simulation occurs in half the time, that simulation is shown in the output. For the correlation, specifying the width of the m_{p}—m_{rBC} distribution, the range is shown explicitly in the dropdown menu in the app (*cf.*, Figure B.1 of Appendix. B for further understanding).

Observations here are generally extendable to other particle mass analyzers and have some ramifications for tandem CPMA-SP2 measurement more generally (*i.e.,* not just to reconstructing 2D distributions). For example, overall trends in the errors with the experimental properties here will also generally apply to summary

parameters (*e.g.,* the distribution mode) inferred from CPMA-SP2 measurements (*e.g.,* errors will grow as the number of CPMA setpoints is reduced).

Chapter 4

Measurement of the centrifugal particle mass analyzer transfer function

4.1 Introduction

Developed by Olfert and Collings, 2005, the centrifugal particle mass analyzer (CPMA) is an instrument that classifies particles by their mass-to-charge ratio. The CPMA is made up of two coaxial rotating cylindrical electrodes with a voltage potential between them. The cylinders, which rotate at different rotational speeds, impose centrifugal and electrical forces on charged particles simultaneously. Aerosols, based on their mass and electric charge, experience different centrifugal and electrical forces, with the centrifugal force acting toward the outer cylinder and the electrical force acting toward the inner electrode. Consequently, large aerosols are subjected to a larger centrifugal force and pushed toward the outer cylinder, while small ones are attracted toward the inner electrode. Particle will exit the classifier if a balance between the electrical and centrifugal force occurs.

CPMAs have been increasingly used in various aerosol applications such as the measurement of particles emitted from combustion engines, (Graves et al., 2015; Graves, Koch, and Olfert, 2017; Momenimovahed and Olfert, 2015; Olfert, Symonds, and Collings, 2007; Sakai and Rothamer, 2017; Ubogu et al., 2018) and the measurement of particulates emitted from other sources, *e.g.,* cigarette smoke (Johnson et al., 2014, 2015a,b), gas flaring (Kazemimanesh et al., 2021; Sipkens et al., 2021), spark-discharge generators (Nilsson et al., 2015), plasma generator (Graves et al., 2020), powdered mineral dust (Marsden et al., 2018). Moreover, pairing CPMAs with other aerosol classifiers in a tandem arrangement has yielded extensive information about such particles. In these tandem measurements, the CPMA is operated in various *static* modes, *i.e.*, at a fixed mass set point, to allow aerosols of a particular mass-to-charge ratio to pass through the CPMA classification region and be measured by instruments downstream. Of particular note are measurements of the mass-mobility relationships of aerosols using a CPMA in tandem with a differential mobility analyzer (DMA; Knutson and Whitby, 1975; Liu and Pui, 1974) or with a condensation particle counter (CPC; Agarwal and Sem, 1980), also known as a scanning mobility particle sizer (SMPS), (Afroughi et al., 2019; Ghazi et al., 2013; Graves, Koch, and Olfert, 2017; Johnson et al., 2014; Olfert, Symonds, and Collings, 2007; Quiros et al., 2015; Sipkens, Olfert, and Rogak, 2020a,b; Sipkens et al., 2021). These measurements have typically been used to determine either summary parameters or the relationship between the distribution of desired parameters of measured aerosols, such as the effective density (Johnson et al., 2014; Olfert et al., 2017; Olfert, Symonds, and Collings, 2007), dynamic shape factor (Beranek, Imre, and Zelenyuk, 2012; Kuwata and Kondo, 2009), and mass-mobility exponent (Shapiro et al., 2012).

Using a similar methodology, some studies have paired the CPMA with other aerosol instruments, for instance, optical instruments (*e.g.,* single-particle soot photometer; Broda et al., 2018; Liu et al., 2017; Naseri et al., 2021a,b; Sedlacek III et al., 2018; Sipkens et al., 2021 and cavity attenuated phase shift spectrometers (CAPS) (Dastanpour et al., 2017), or aerosol mass spectrometers (Bell et al., 2017; Marsden et al., 2018), to measure the average coating mass and optical properties, or chemical compositions of measured aerosols, respectively.

While the CPMA classifies particles by mass-to-charge ratio, the classified particles are not strictly monodisperse in mass because, in addition to singly charged particles of the desired mass, doubly charged particles of twice the desired mass (and particles up to the maximum number of charges in the aerosol) also pass through the classification region. When the CPMA operates at low rotational speeds (*e.g.,* when it classifies relatively large aerosols), small uncharged particles may not be sorted out and may escape the classification region. Also, aerosols with a mass-to-charge ratio marginally higher and lower than the selected mass transverse the classifier. The performance of the CPMA in the classification of a chosen mass-to-charge ratio is described by a probability distribution known as a transfer function. The transfer function has profound implications for interpreting the mass-to-charge ratio distributions measured by the CPMA. Of note, in tandem measurement systems where the CPMA is used (*e.g.,* the CPMA-SP2 or CPMA-SMPS), the CPMA transfer function plays a crucial role in evaluating the kernel function (*i.e.,* the theoretical response) of the system, which is used for retrieving desired information via inversion (Broda et al., 2018; Buckley et al., 2017; Naseri et al., 2021a,b; Rawat et al., 2016; Sipkens, Olfert, and Rogak, 2020a; Sipkens et al., 2021). Accordingly, it is essential to

characterize the CPMA transfer function carefully.

The theoretical CPMA transfer function was initially derived by Olfert and Collings, 2005, using a finite difference scheme and mathematical methods. Sipkens, Olfert, and Rogak, 2020c revisited the evaluation of the CPMA transfer function, implementing novel numerical approaches to improve the computational efficiency of calculating the transfer function. Despite the theoretical ability to calculate the transfer function of the device, to date, little data have been collected regarding how the actual transfer function differs from the theoretical model (Olfert et al., 2006). Furthermore, no data are available for the commercially available CPMA (Cambustion Ltd., Cambridge, UK).

The transfer functions of the differential mobility analyzer (DMA; Knutson and Whitby, 1975) aerodynamic aerosol classifier (AAC; Tavakoli and Olfert, 2013) have been experimentally characterized using a *tandem measurement system* (Birmili et al., 1997; Collins et al., 2004; Fissan et al., 1996; Hummes et al., 1996; Johnson et al., 2018; Karlsson and Martinsson, 2003; Li, Li, and Chen, 2006; Martinsson, Karlsson, and Frank, 2001; Stratmann et al., 1997). In this system, two identical aerosol classifiers are employed in series (*i.e.,* a tandem arrangement). The first instrument is configured so that it operates at one set point and the total number concentration of particles exiting the classifier is counted. The second instrument, downstream of the first, is then stepped or scanned over the range of particle sizes exiting the first classifier, and the number concentration of particles exiting the second classifier is measured. Consequently, there is a distribution of particle counts measured at the exit of the second classifier as a result of the convolution of the transfer functions of the two classifiers. The actual transfer function or how it differs from an ideal case

is estimated using data deconvolution or fitting routines (Li, Li, and Chen, 2006), assuming that the two classifiers have identical transfer functions.

In this study, the tandem CPMA methodology was used to characterize the CPMA transfer function. Following previous work on the DMA and AAC, it is assumed that the ideal CPMA transfer function is triangular and experiments are used to determine factors which account for deviations from the ideal triangular transfer function namely the width factor (μ), which accounts for diffusional broadening or other factors which may increase or decrease the width of the transfer function, and the height factor (λ), which accounts for losses in the CPMA due to impaction or diffusion losses, etc. This study is of considerable significance as the experimentally evaluated CPMA transfer function can significantly enhance CPMA data analysis routines, *e.g.,* data inversion.

4.2 Theory and Methodology

4.2.1 CPMA Transfer Functions

Figure 4.1 shows a simple schematic of the Couette CPMA, in which aerosols flow between two coaxial cylinders rotating around their axis while a voltage is applied between them. Particles with different mass (m) and charge states (n_{q}) are exposed to centrifugal (F_{C}) and electrical (F_{E}) forces, such that a limited range of particles can pass through the classification region. The performance of the CPMA in classifying particles is characterized by a probability distribution known as a transfer function. A number of theoretical models have been developed for the Couette CPMA (Olfert and Collings, 2005; Sipkens, Olfert, and Rogak, 2020c) in order to numerically es-

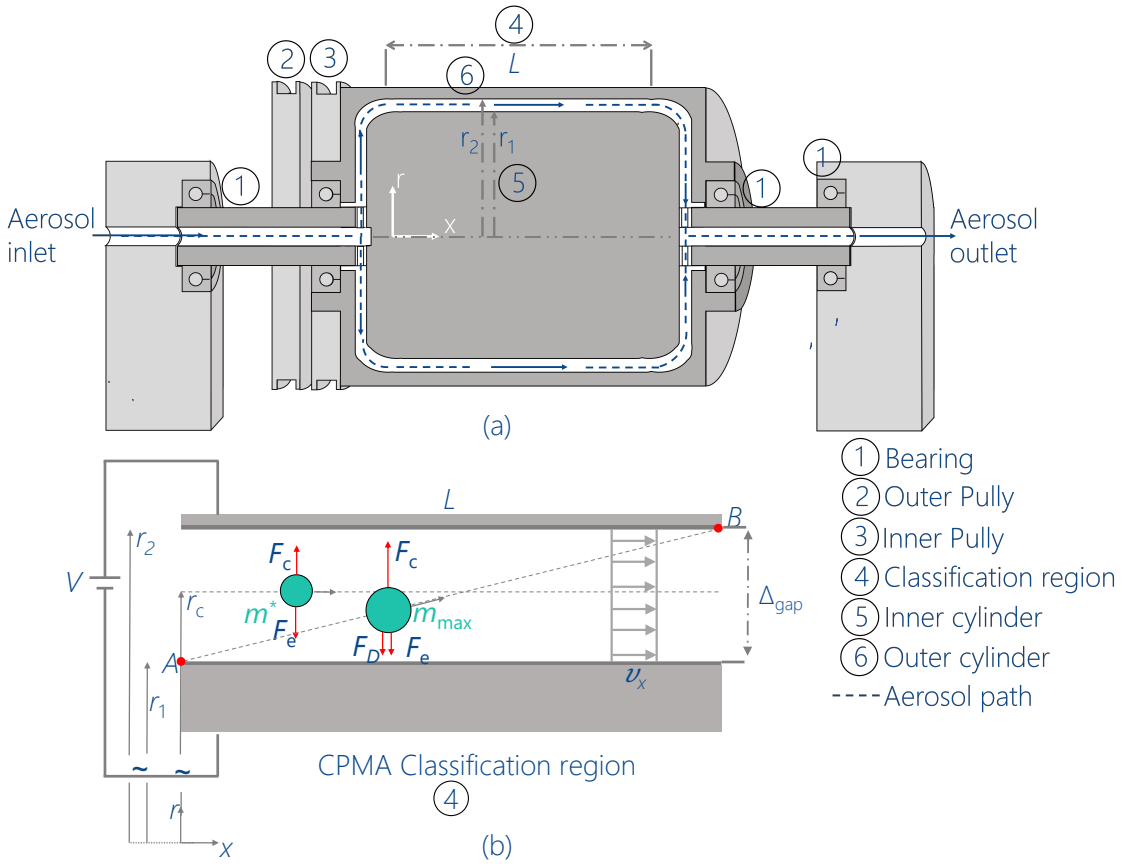

Figure 4.1: Schematics demonstrating (a) geometric features of the Couette CPMA, (b) operating principles in the CPMA classification region.

timate the CPMA transfer functions. In contrast to these numerical models which are computationally intensive for evaluating the CPMA transfer function, there is a simplified representation of the CPMA transfer function in the form of a triangular shape that has been implemented in the CPMA software (Cambustion Ltd., Cambridge, UK).

To derive the triangular transfer function, it is necessary to examine the motion of particles in the *classification region* (see Fig. 4.1) in order to determine a mass

range of particles which can pass through the classifier. Accordingly, the following assumptions will be made to simplify the equation of motion: (*i*) it is assumed that the channel is sufficiently narrow that the forces on the particles in the gap are constant, (*ii*) particle diffusion and image forces are negligible, (*iii*) particles acquire their terminal velocity instantly[1], (*iv*) the flow field is quasi-steady and the axial flow (v_x) in the classification region is laminar plug flow, and (v) the gap between the coaxial cylinders is much smaller than the CPMA radii ($\Delta_{\text{gap}} \ll r$). Consequently, the equation of motion of particles in the CPMA classification region in the radial (r) direction, can be expressed as

$$\sum F_r = F_{\text{C}} - F_{\text{E}} - F_{\text{D}} = \frac{m\text{d}v_r}{\text{d}t}, \tag{4.1}$$

where F_D is viscous drag force, and v_{r} is the particle velocity in the radial direction. Using Stokes' law, the viscous drag force is expressed as:

$$F_{\text{D}} = \frac{3\pi\mu_{\text{f}}d_{\text{m}}}{C_{\text{c}}(d_{\text{m}})}v_r = \frac{v_r}{B}, \tag{4.2}$$

where d_{m} is the mobility equivalent diameter, C_{c} is Cunningham slip correction factor, μ_{f} is the viscosity of the carrier gas, and B is the particle mechanical mobility. The electrical and centrifugal forces on an aerosol at a distance r from the rotational axis are

$$F_{\text{E}} = \frac{qV}{r\ln(r_2/r_1)}, \tag{4.3}$$

and

$$F_{\text{C}} = mr\omega^2, \tag{4.4}$$

[1]This assumption holds true since the relation time of the particles is much shorter than their residence time in the classification region.

respectively. Here, r_1 and r_2 are the inner and outer cylinder radius, respectively, V is the potential difference applied between the inner and outer cylinder, and q is the charge on the aerosol, and ω is rotational speed of the gas. Since the forces on the particles are assumed to be constant across the narrow gap between cylinders, the centrifugal and electric forces are calculated at the center of the gap ($\omega = \omega_c$ and $r = r_c$).

At a given operational condition of the CPMA (*i.e.,* a particular rotational speed and a voltage applied between cylinders), the centrifugal and electrical forces become equal for particles with a certain mass and charge state, such that they transit axially along the classifier. As a result, Equation 4.1 for equilibrium (*i.e.,* $F_D = 0$) in the radial direction can be written as $F_E = F_C$;

$$\frac{qV}{r_c \ln(r_2/r_1)} = m r_c \omega_c^2. \tag{4.5}$$

Rearranging Equation 4.5 in terms of the mass-to-charge ratio yields:

$$m^* = \left(\frac{m}{q}\right)_{eq} = \frac{V}{\omega_c^2 r_c^2 \ln(r_2/r_1)}. \tag{4.6}$$

The particle with the greatest mass that can pass through the CPMA is the particle that enters the classification region at the inner edge of the classifier (point A in Figure 4.1*b*) and reaches the outer edge of the classifier at the end of the classification region (point B in Figure 4.1*b*). Accordingly, Equation 4.1 for the maximum particle mass (m_{max}) that exits the classifier is,

$$m_{max} r_c \omega_c^2 - \frac{qV}{r_c \ln(r_2/r_1)} - \frac{v_r}{B_{max}} = \frac{m \mathrm{d}v_r}{\mathrm{d}t} = 0, \tag{4.7}$$

where B_{max} is the mobility of the particle with the mass of m_{max}. The term on the right-hand side is zero as the particles reach their terminal velocity instantly.

Due to the plug flow assumption, the radial particle velocity, v_r, is approximated by $Q\Delta_{\text{gap}}/A_\Delta L$, where Q is the volumetric aerosol flow rate, Δ_{gap} and L are the gap and the length of the classifier, respectively, and A_Δ is the flow cross-sectional area which can be estimated as $2\pi r\Delta_{\text{gap}}$. Substituting Equation 4.6 into Equation 4.7 gives,

$$m_{\text{max}} = m^* + \frac{Q}{2\pi r_{\text{c}}^2 \omega_{\text{c}}^2 B_{\text{max}} L}, \tag{4.8}$$

where the second term on the right-hand side is the full-width half max (FWHM) of the triangular transfer function in mass space. Consequently, the resolution of the triangular transfer function of the CPMA, which is the dimensionless form of FWHM (normalized by the equilibrium mass), is estimated by,

$$\frac{1}{R_m} = \frac{m_{\text{max}}}{m^*} - 1. \tag{4.9}$$

The voltage and rotational speed in Equation 4.5 are adjustable variables to obtain the user-selected mass set point (*i.e.,* the equilibrium mass, m^*) and resolution.

Alternative transfer functions have been derived. Olfert, 2005 and Sipkens, Olfert, and Rogak, 2020c derived an identical expression (albeit calculated differently) for the CPMA transfer function by using the same assumptions as the triangular transfer function, except that they assumed the axial flow to be parabolic and forces to vary radially. Due to the variations in forces and the assumption of parabolic flow, the CPMA transfer function has a *trapezoidal* shape in their model. Furthermore, they have presented an expression for the transfer function that incorporates particle diffusion, which here is referred to as the *diffusion* model.

Figure 4.2*a* illustrates the transfer function of the triangular model together with the trapezoidal and diffusion models of Sipkens, Olfert, and Rogak, 2020c (*i.e.,*

using Cases 1C and 1C-diff in that work) evaluated for the mass set point of $m^* = 0.05$ fg ($\sim$ 42 nm) with a constant flow rate of $Q = 0.3$ LPM and resolutions of $R_m = 2$ and $R_m = 15$, respectively. It is observed that while the triangular model is symmetric around the equilibrium mass-to-charge ratio, the trapezoidal model is positively skewed (*i.e.,* concentrated on higher particle masses) and has a narrower width at low resolution. The difference arises from the fact that while the forces are assumed to be constant in evaluation of the triangular model, they vary slightly with radius, such that the electrostatic and centrifugal forces are larger near the inner cylinder. As a result, smaller particles are removed more efficiently than larger ones, resulting in slight skewness towards higher masses. Moreover, it can be seen that the diffusion model is more skewed than the trapezoidal model when the sample flow rate and resolution are low. This is because diffusion is dominant for smaller particles resulting in a greater removal of smaller particles. A comparison between Figure 4.2*a* and *b* reveals that as the resolution increases from a low resolution of $R_m = 2$ to a high resolution of $R_m = 15$, both trapezoidal and diffusion models become nearly symmetrical, and the width of trapezoidal and diffusion models approach that of triangular transfer functions.

4.2.2 Determining Deviations from the Triangular Transfer Function

It will be shown in the Results section that none of the transfer function models accurately predict the actual transfer function of the CPMA. Therefore, an empirical model can be used to correct a theoretical transfer function model. Theoretically, this could be done by applying correction factors to any of the idealized transfer

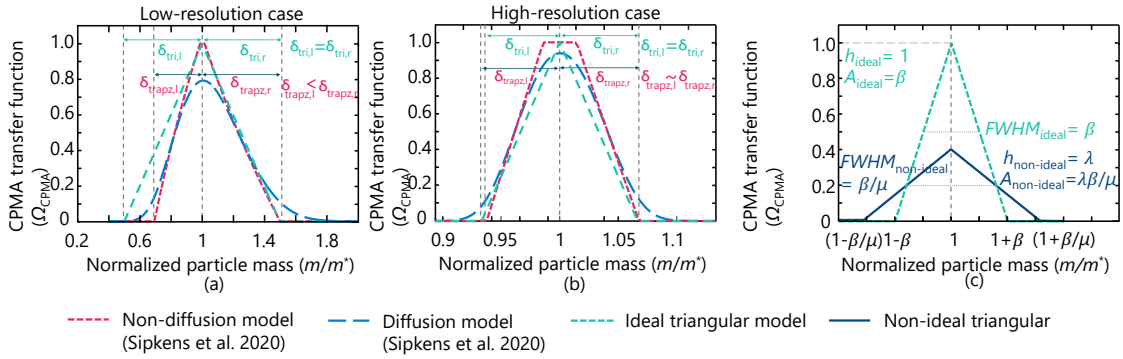

Figure 4.2: Comparison of three theoretical transfer functions of the CPMA; (*i*) triangular, (*ii*)trapezoidal, and (*iii*) diffusion model of Sipkens, Olfert, and Rogak, 2020c evaluated for a mass set point of $m^* = 0.05$ fg ($\sim$ 42 nm) and flow rate of Q=0.3 LPM at (a) low ($R_m = 2$) and (b) high ($R_m = 15$) resolutions. (c) Comparison of idealized (Equation 4.2.2) and non-idealized (Equation 4.2.2) triangular CPMA transfer functions.

functions described above. However, the diffusion and trapezoidal transfer functions are computationally expensive which prohibits their use on the hardware of the commercial CPMA, and they also make two-dimensional inversion very time-consuming. Thus, the work presented here will examine how the actual transfer function differs from the triangular transfer function and use correction factors to the triangular transfer function to create an empirical transfer function.

An ideal triangular (*i.e.,* non-diffusing) transfer function of the CPMA (Ω_{ideal}) in non-dimensional particle mass space, $\tilde{m} = m/m^*$, can simply be expressed as

$$\Omega_{\text{ideal}}\left(\tilde{m}, \beta\right) = \frac{1}{2\beta}\left[\left|\tilde{m} - (1+\beta)\right| + \left|\tilde{m} - (1-\beta)\right| - 2\left|\tilde{m} - 1\right|\right], \tag{4.10}$$

where β is the inverse of the CPMA resolution ($1/R_m$) or equivalently the full-width of the triangle at half-maximum (FWHM). The area of the prescribed transfer function, which determines the number of particles passing through the classifier, can

then be calculated by $A = hm^*/R_m$, where $h = 1$ is the height of the ideal triangular transfer function at the instrument set point. This simple expression is the CPMA transfer function derived in Equation 4.8 and 4.9, but expressed in non-dimensional mass space.

As a result of non-idealities in particle classification, the actual shape of the CPMA transfer function will differ from a triangle and may narrow or broaden the width or increase or decrease the height of the transfer function relative to the idealized triangle. Previous, researchers using the DMA or AAC (Birmili et al., 1997; Johnson et al., 2018; Martinsson, Karlsson, and Frank, 2001) have used various adjustable factors to correct Eq. 4.2.2 to account for non-idealities, although the definitions of the adjustable factors varies between studies. Following the work of Birmili et al., 1997, two adjustable factors are introduced to Equation 4.2.2 as

$$\Omega_{\text{non-ideal}}\left(\tilde{m}, \beta, \mu, \lambda\right) = \frac{\lambda\mu}{2\beta}\left[\left|\tilde{m} - \left(1 + \frac{\beta}{\mu}\right)\right| + \left|\tilde{m} - \left(1 - \frac{\beta}{\mu}\right)\right| - 2\left|\tilde{m} - 1\right|\right], \tag{4.11}$$

where λ and μ are *height factor* and *width factor*, respectively. While the height factor scales the transfer function's integrated area (A) by modifying the height of the transfer function (h) to account for particle losses (*e.g.,* losses occurring through diffusion and impaction), the width factor adjusts the width of the triangular transfer function ($FWHM_{\text{non-ideal}} = \beta/\mu$) to account for the widening of the transfer function due to particle diffusion or any other non-idealities in the CPMA (*e.g.,* flow effects). An illustration of the idealized transfer function (Equation 4.2.2) and how its width, height, and area are modified by the adjustable factors (*i.e.,* non-idealized transfer function, Equation 4.2.2) are shown in Figure 4.2*c*. In this example, the non-idealized

transfer deviates from the idealized transfer function with μ and $\lambda < 1$, demonstrating that along with losses, the actual transfer function may be broader than the idealized one.

The adjustable factors λ and μ are derived by conducting the tandem CPMA measurements as shown in Figure 4.3*a*. In the experiment, the number concentration of aerosols counted downstream (N_2) and upstream (N_1) of the second CPMA are counted with a CPC as the second CPMA is stepped through a range of set points. The theoretical number concentrations can also be calculated through the convolution of the incoming particle count distribution and the CPMA transfer functions,

$$\frac{N_2\left(m_2^*,m_1^*\right)}{N_1\left(m_1^*\right)} = \frac{\int \eta\left(m_2^*\right)\mathrm{d}N/\mathrm{dlog}m_\mathrm{i}\Omega_1\left(m_1,m_1^*,\lambda_1,\mu_1\right)\Omega_2\left(m_2,m_2^*,m_{12}^*,\lambda_2,\mu_2\right)dm_i}{\int \eta\left(m_1^*\right)\mathrm{d}N/\mathrm{dlog}m_\mathrm{i}\Omega_1\left(m_1,m_1^*,\lambda_1,\mu_1\right)dm_i}, \tag{4.12}$$

where $\mathrm{d}N/\mathrm{dlog}m_\mathrm{i}$ is the number distribution of aerosol entering the first CPMA, $\eta\left(m^*\right)$ is the counting efficiency of the CPC at a mass set point (m^*), Ω is the non-idealized triangular CPMA transfer function, in which the subscript 1 and 2 refer to the first and second CPMA, and m_{12}^* represents the mass set point agreement between the two CPMA set points (*i.e.*, $m_{12}^* = m_2^*/m_1^*$). This factor was included in Equation 4.2.2 to account for the offset between the actual and ideal N_2/N_1 ratios potentially resulting from any discrepancies in the mass set points of the two CPMAs that might have been caused by differences in voltage and rotational speed calibrations, etc. The equation is simplified by assuming that the counting efficiency of the CPC is constant over the narrow mass domain stepped through by the downstream CPMA, *i.e.*, $\eta\left(m_1^*\right) = \eta\left(m_2^*\right)$. It is also assumed that both CPMA transfer function width factors were identical ($\mu_1 = \mu_2$). (Note, the relative resolution of both CPMAs (R_m

or β) is the same at all set points in an experiment; however, the absolute width of the transfer function of the second CPMA (e.g., $m_{\max} - m^*$) changes at each set point). Additionally, as N_1 was directly measured, all losses within the upstream CPMA were already taken into account, and accordingly, $\lambda_1 = 1$ was used in the theoretical convolution.

Minimizing the difference between the theoretical and experimental values of the N_2/N_1 concentration ratio (Equation 4.2.2) using chi-squared minimization determined λ, μ and m_{12}^*. This analysis was based on the assumption that the likelihood of having multiply-charged aerosols at the entrance of the first CPMA was negligible so that their effects on the convolution of would likewise be negligible. It is shown in the Appendix that the fraction of multiply-charged aerosols was very small and does not have a significant affect on the results.

Unlike previous studies with DMAs and AACs, which assumed a uniform distribution of aerosols entering the tandem measurements, (*e.g.,* Martinsson, Karlsson, and Frank, 2001 and Johnson et al., 2018) the number distribution of aerosols passing through the AAC was measured by an SMPS and considered in the convolution. This helped to increase the accuracy of the fits.

4.3 Experimental setup and procedure

4.3.1 Tandem CPMA measurements

A schematic of the TCPMA experimental setup is shown in Figure 4.3*a*. Since the width of the CPMA transfer function and its calculation depends on particle mobility, it was necessary to measure an aerosol with known mobility. To this end, a $\sim 1\%$

solution of Santovac 5 in acetone (purity $\geq 99.5\%$) was used to generate spherical aerosols with a known density (1198 kg/m^3). The TSI atomizer (Model 3076) was employed to atomize the solution and generate the aerosols. In order to evaporate the acetone, the generated aerosol was diluted by a factor of ~ 25 with dilution air at a temperature of 60 °C using a two-stage diluter (Dekati® eDiluter Pro).

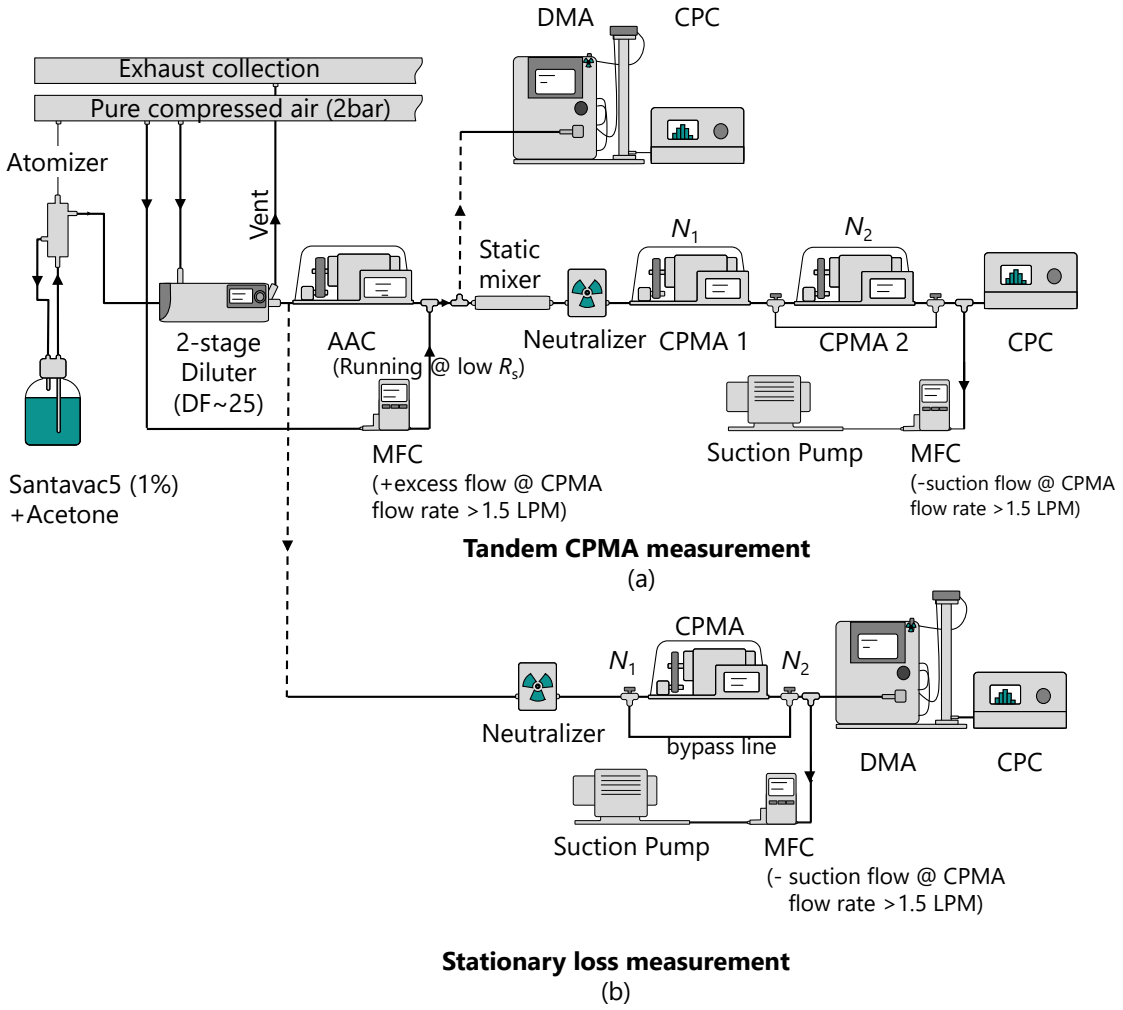

Figure 4.3: A schematic of a) the tandem CPMA measurement b) the stationary loss measurement setups.

TCPMA measurements can be complicated because small uncharged aerosols may pass through the classifier at very low rotational speeds when the centrifugal force is low (*e.g.,* when the CPMA operates at high mass set points and low resolutions). Consequently, an AAC (Cambustion Ltd., Cambridge, UK) was used upstream of the CPMAs to remove small particles that were outside the CPMA transfer function which could have passed through the CPMA if they were uncharged. The AAC was set to operate at low resolutions (*e.g.,* R_s ranging between 1.5 and 4), so that a wide range of particle sizes could pass through it and enter the first CPMA. Furthermore, the AAC aerodynamic diameter set points were selected so that the mode of AAC-classified distributions would always be slightly smaller than the mass equivalent diameters of the first CPMA mass set points. This AAC set point selection was undertaken to maximize the number of aerosols entering the TCPMA system while minimizing the portion of multiply charged particles smaller than that of singly charged ones. The size distributions of AAC-classified aerosols were then measured by scanning mobility particle sizers (SMPS, TSI Inc.) to be used in the data convolution explained in Section 4.2.

In the measurements taken at flow rates of 0.3 or 1.5 LPM, the TCPMA aerosol flow rate was controlled by an internal flow controller within the condensation particle counter (CPC; TSI, Model 3775). For the measurements with higher flow rates (*i.e.,* 4 and 8 LPM), the CPC was operated at a flow rate of 1.5 LPM, and a vacuum pump and a mass flow controller (MFC; Alicat Scientific, MCV–Series: Vacuum Gas Mass Flow Controllers, serial number 175509) were employed together to regulate makeup flows of 2.5 and 6.5 LPM, resulting in the net flow of 4 and 8 LPM through the TCPMA system, respectively. Additionally, because the AAC cannot operate

at flow rates higher than 3 LPM, the aerosol flow from the AAC was supplemented with filtered compressed air controlled with a mass flow controller (MFC; Alicat Scientific, mc-series, serial number 270591) to maintain the desired flows. A static mixer (Ko-flo Corporation, 1-800 Static, model number 3/8-40-3-12-2) was then used to generate a well-mixed aerosol flow prior to measurement by the TCPMA system.

Prior to TCPMA measurement, the aerodynamically classified aerosols were electrically charged by passing them through a radioactive Kr-85 bipolar diffusion charger (TSI Inc., Model 3077A). The first CPMA (Cambustion Ltd., serial number C220) was operated at a fixed mass set point of m^*_1 = 0.05, 0.1, 1, 10, or 100 fg, while the downstream CPMA (Cambustion Ltd., serial number C313) was stepped (m^*_2) through the range of classified particles, and the number concentration of the twice-classified particles (N_2) was measured with the CPC. At the beginning and end of each TCPMA measurement, the second CPMA was bypassed to measure the steady-state number concentration of aerosols passing through the upstream CPMA (N_1) during the measurement.

The CPMAs were examined at five mass set points between 0.05 – 100 fg, four flow rates ranging from 0.3 – 8 LPM, and five resolutions (R_m) of 2, 3, 6, 10 or 15 if the operating conditions fell within the CPMA's operating window (*i.e.,* it was not possible to test certain conditions because the voltage or rotational speed of the CPMA would have exceeded the maximum or minimum). There were some measurements that were within the operating envelope of the CPMAs; however, the experiments could not be completed because high motor temperature errors that occurred. Each TCPMA configuration was repeated three times in order to assess the repeatability of the measurement See Table A1 for a complete list of all data

collected in the study.

4.3.2 Stationary loss measurements

In addition to the CPMA losses characterized as height factor from the TCPMA measurements, the losses were measured as the transmission efficiency of the CPMA (N_2/N_1) when the CPMA was not operating (*i.e.,* no rotational speed and no voltage). Figure 4.3*b* shows a schematic of the stationary loss measurement setup. In this measurement, the SMPS was used to measure the number concentration of particles upstream (N_1) and downstream (N_2) of the CPMA when it was not operating. These measurements approximate the TCPMA case of $R_m = 0$, which were made at flow rates of 0.3, 1.5, and 4 LPM.

4.4 Results and Discussion

Figure 4.4 depicts the concentration ratios (N_2/N_1) derived from the TCPMA measurements (*i.e.,* convolution of the two CPMA transfer functions) conducted at a flow rate of 0.3 LPM for five CPMA mass set points ranging from 0.05 to 100 fg, with a resolution of R_m= 6, and compares them with their theoretical counterparts, which are the convolution of two theoretical transfer functions: (*i*) triangular model, (*ii*) trapezoidal, and (*iii*) diffusion model, shown as dashed green, solid red, and blue lines, respectively. Compared to their ideal counterparts, the observed N_2/N_1 ratios generally had lower amplitudes and a slight offset, and they were either slightly wider for low mass set points, *e.g.,* $m^* = 0.05$ and 0.1 fg, or narrower for high mass set points. It can be seen that the convolution of the theoretical models could not accurately represent the measured N_2/N_1. In contrast, the convolution of two non-

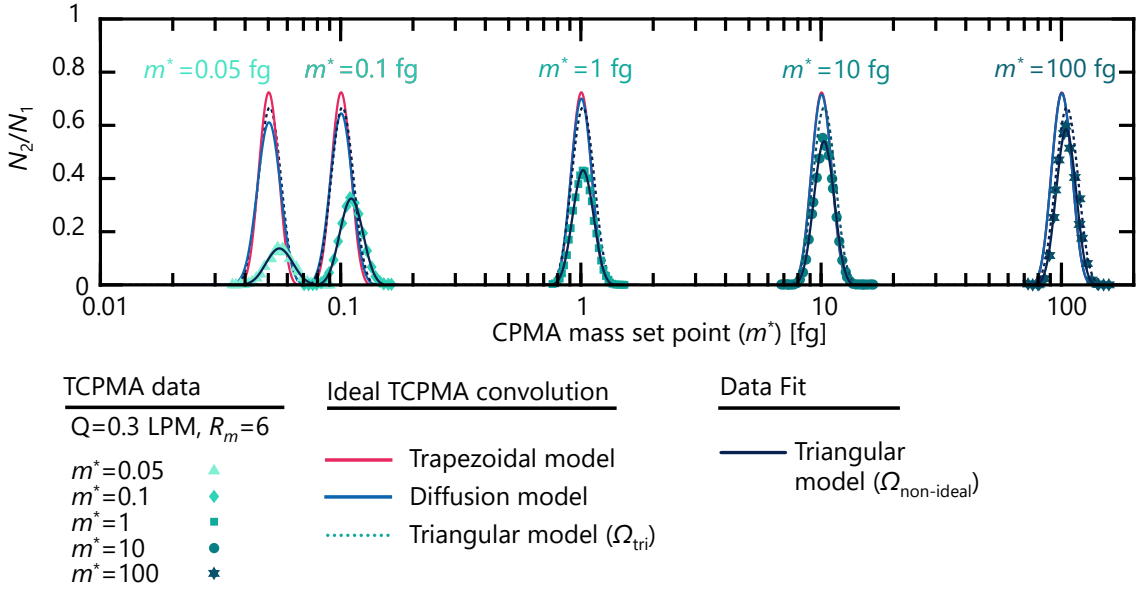

Figure 4.4: TCPMA measurements conducted at a flow rate of 0.3 LPM for five CPMA mass set points (0.05 – 100 fg), with a resolution of R_m= 6.

idealized CPMA triangular functions fits the data well (solid black lines), indicating that a reasonable approximation of the actual CPMA transfer function had been obtained by using the triangular transfer function with three adjustable factors (λ, μ and m_{12}^*). The following sections describe the results of the three factors for the rest of the data set.

4.4.1 Set point agreement (m_{12}^*)

Figure 4.5*a* illustrates the mass set point agreement between the two CPMAs. The values in this figure represent the relative offset between the mass set points of the two CPMAs, such that a value of 1 indicates that the mass set points of the two CPMA are equivalent. The offset between the actual and ideal N_2/N_1 ratios results from any discrepancies in the mass set points of the two CPMAs that might have

been caused by differences in voltage and rotational speed calibrations or differences in classifier radii. It is observed that the offset has a small dependence on mass set

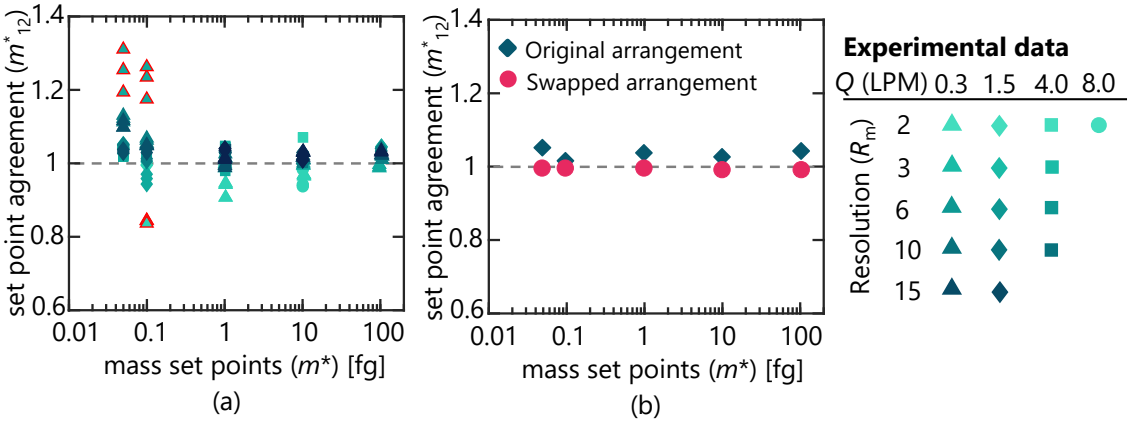

Figure 4.5: (a) The the mass set point agreement derived from TCPMA measurements for different flow rates ranging from 0.3 to 8 LPM, at five CPMA mass set points between 0.05 and 100 fg, with five different resolutions ranging in R_m from 2 to 15. (b) Comparison between the mass set point agreement derived from the original arrangement of TCPMA system and the swapped arrangement at a particular measurement configuration of $Q = 1.5$ LPM and $R_m = 6$.

point and resolution, as reflected in color in Figure 4.5*a*, such that higher CPMA resolutions tend to result in a smaller offset between the CPMA's mass set points at larger mass set points. In the current study, the measurements with mass set point agreements greater than 1.15 or lower than 0.85, have been regarded as outliers as they do not conform to the main assumption underlying TCPMA measurements (running two *identical* CPMAs in series). The outliers are identified by red lines encircling the data points in Figure 4.5*a*, which are associated with low mass set point measurements conducted at a low flow rate of $Q = 0.3$ LPM and resolutions of R_m =2 and 3. Neglecting the outliers, the average mass set point agreement between the CPMAs is $m^*_{12} = 1.02 \pm 0.03$, indicating sound reproducibility among the CPMAs.

This was tested by swapping the positions of the CPMAs as shown in Figure 4.5*b*, which shows the mass set point agreements of the original CPMA arrangement and when the CPMA position is swapped for mass set points ranging between 0.05 fg to 100 fg, with $Q = 1.5$ LPM and $R_m = 6$. The nature of the experiment calls for symmetry around 1 (here with an average error of ∼1.5 %), such that when CPMA positions are swapped, the magnitude of the offset is expected to remain the same while shifting in opposite direction (inversed). This implies that swapping the CPMA positions does not have a major impact on the results.

4.4.2 Width factor (μ)

Figure 4.6*a*–*c* illustrate the width factors measured by TCPMA method as a function of the CPMA mass set point for the flow rates of $Q = 0.3$, 1.5, 4 and 8 LPM, respectively. Figure 4.6*a*–*c* shows the CPMA resolution in gradations of color to highlight its effect on the width factor variation. It is observed that the dependence of width factors on particle mass appears to be notable only at a flow rate of 0.3 LPM and low mass set points (see Figure 4.6*a*). On the other hand, the width factor has little dependence on the particle mass at higher flow rates. These observations are expected since particle transport is influenced by diffusion when the flow rate and the particle mass are sufficiently low. This finding was verified by three-way ANOVA analysis, which was used to quantitatively evaluate the significance of the CPMA settings (m^*, R_m, Q) on the width factor. The results are listed in Table 4.1, sorted by p-value for three groups of data: (*i*) all measurements, (*ii*) low flow measurements ($Q = 0.3$ LPM), and (*iii*) high flow measurements ($Q > 0.3$ LPM). The effects of any parameter on an adjustable factor are more significant when the p-value is smaller.

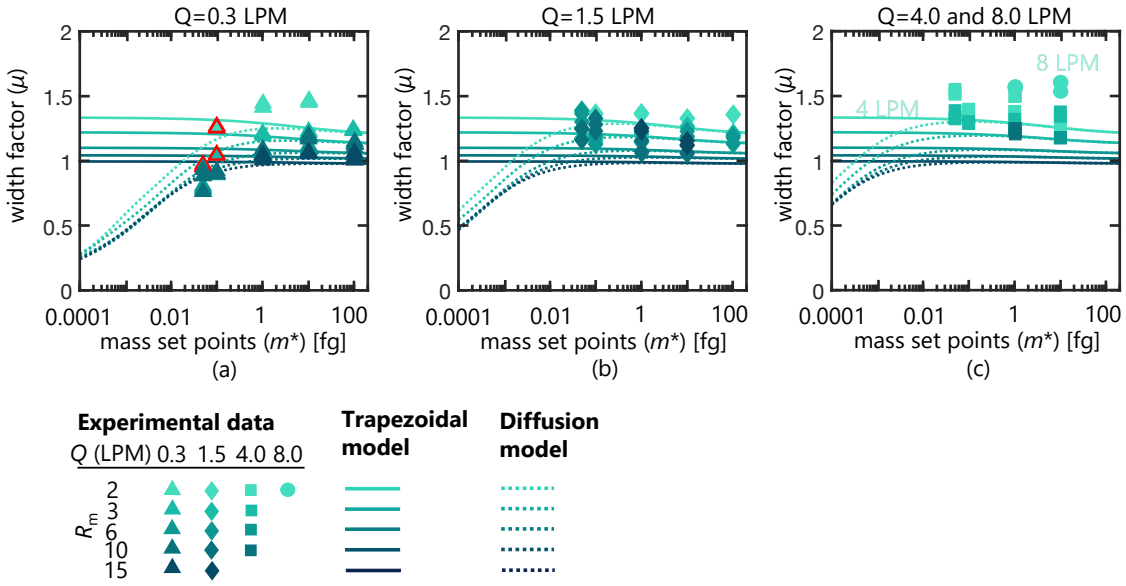

Figure 4.6: The width factor derived from TCPMA measurements and synthetic data generated from the convolution of trapezoidal and diffusion models. The measured and the corresponding theoretical width factor derived for different flow rates ranging from 0.3 to 8 LPM, at five CPMA mass set points between 0.05 and 100 fg, with five different resolutions ranging in R_m from 2 to 15.

The ANOVA analysis performed on data points at all flow rates (first column of Table 4.1) indicates that Q was the most effective input parameter to explain the variations in width factor (p-value $\sim 10^{-16}$). Comparing the ANOVA analysis using only the low-flow data (second column in the Table) and measurements at flow rates above 0.3 LPM (third column in the Table), reveals that the mass set point (m^*) only became significant at the low flow rate setting. It is seen in Figure 4.6a that the width factor dropped to less than one for low mass set points ($m^* < 1$ fg) when the CPMA was run at low flow rates ($Q = 0.3$ LPM), which means that the CPMA transfer function was broader than the ideal transfer function. This is expected since the particles will have a longer residence time in the classifier when the sample flow

Table 4.1: Three-way ANOVA table indicating the significance of the CPMA sampling configurations (*i.e.,* m^*, R_m, Q) to the width and height factors. The analysis was conducted using the logarithm of mass set points, and all parameters were treated as continuous.

ANOVA Analysis												
Width factor									Height factor			
All flow rates (Q= 0.3 — 4 LPM)			Low flow rate (Q= 0.3 LPM)			High flow rates (Q= 1.5 , 4 LPM)			All TCPMA data			
	F	p-value		F	p-value		F	p-value		F	p-value	
Q	85.0	1.49×10^{-16}	m^*	45.78	6.23×10^{-9}	R_m	30.7	2.73×10^{-7}	m^*	255	3.63×10^{-35}	
R_m	44.7	3.44×10^{-10}	R_m	41.9	1.95×10^{-8}	Q	9.89	2.23×10^{-3}	R_m	82.7	3.19×10^{-16}	
m^*	7.65	6.32×10^{-3}	Q	—	—	m^*	8.14	5.33×10^{-3}	Q	4.21	4.18×10^{-2}	

rate is low allowing small particles to diffuse more efficiently. As shown in Figure 4.6*a–c*, the width factor is generally observed to be higher than one and increases with flow rate. This means that the actual CPMA transfer function was narrower than the idealized triangular transfer function. Referring back to Figure 4.2*a*, the actual transfer function, which is likely better represented with the trapezoidal or diffusion model, is expected to be narrower that the triangular transfer function (except for a flow rate of Q =0.3 LPM cases where diffusion is significant, resulting in a broader transfer function).

These observations can be supported by theoretical models. Figure 4.6*a–c* also illustrate the width factor derived from fitting synthetic data fabricated by the convolution of the theoretical trapezoidal and diffusion transfer functions as solid and dotted lines, respectively, in color gradations that correspond to the variation in resolution at a given CPMA flow rate. In agreement with the experimental observations, Figure 4.6*a–c* demonstrates that the theoretical width factor (diffusion model)

decreases at low masses only at low flow rates for the range of data measured, with some functional dependence on the CPMA resolution, and is constant with particle mass at high flows. Additionally, the ANOVA analysis using only data at flow rates above 0.3 LPM (third column of the Table) demonstrates that the CPMA resolution to a great extent explain a functional variation in the width factor (Q, and m^* had negligible impacts) at high-flow measurements. According to theory, the convolution of two trapezoidal transfer functions results in width factors greater than one and is a function of resolution, such that it decreases with resolution and gets close to one at higher resolutions, which is in close agreement with the measurements shown in panels of Figure 4.6*a*–*c*. However, it is evident that the values of theoretical width factors do not fully correspond to experimental observations. This suggests that there are some other mechanisms inside and outside the classification region that may be important but are not considered in the transfer function models. For instance, the theoretical models neglect entrance effects and assume the flow is fully developed and laminar at the entrance of the classification region when in reality, there is an entry length for fully developed flow. Moreover, the models assume an even distribution of particles across the gap, but flow effects may result in non-uniform concentration at the entrance. Therefore, since current theoretical models of the transfer function cannot accurately model the resolution of the CPMA, thus an empirical model is needed.

An empirical model was found using multivariate non-linear curve-fitting of the width factor data. An iteratively reweighted least squares (Dumouchel, O'brien, et al., 1989; Holland and Welsch, 1977) routine was implemented in Matlab® with the `nlinfit` function using the Welsch weight function. From this fitting procedure, the

following equation was derived

$$\begin{aligned}
&\mu\left(m^*, R_m, Q\right) = \frac{a_l}{1+\exp\left[-b_l \cdot \left(\log_{10}\left(m^*\right)+\mathrm{c}_l\right)\right]} \pm 0.105; \\
&a_l = \left(0.0258Q + 1.31R_m^{-2.57} + 1.12\right), \\
&b_l = \left(4.07R_m + 0.967\right), \\
&c_l = 4.63R_m + 0.428.
\end{aligned} \tag{4.13}$$

where Q, R_m, and m^* are the volumetric flow rate in LPM, CPMA resolution, and the CPMA mass set point in fg; respectively. Figure 4.7*a–c* show the measured width factor and the results of the fitted model shown as solid lines at the four different flow rates. The length of the solid lines represent the operating region of the CPMA such that the CPMA cannot be operated above or below the mass set point at the flow and resolution indicated. It is worth mentioning that although it is theoretically possible to run the CPMA for particle masses smaller than 0.05 fg at a flow of 0.3 LPM, in reality, the particle losses are so high that reliable data could not be collected [2]. Figure 4.7*d* shows the distribution of measured versus predicted values of width factor along with a 95% prediction interval, which is defined as an interval which contains 95% of the data (±0.105). The parity plot shows that the majority of points lie near the diagonal, which is a positive indication that the fit model was acceptable. The determined prediction interval was added to Equation 4.4.2 as an upper and lower bound to illustrate the accuracy of the fit. For reference, the root-mean-square error for the fit is 0.0651.

[2] As shown later in Figure 4.8*a*, the stationary transmission efficiency for particles below 0.05 fg in a single CPMA at Q =0.3 LPM is $\sim$ 0.2, which would be nearly two times higher for two CPMAs in the tandem arrangement.

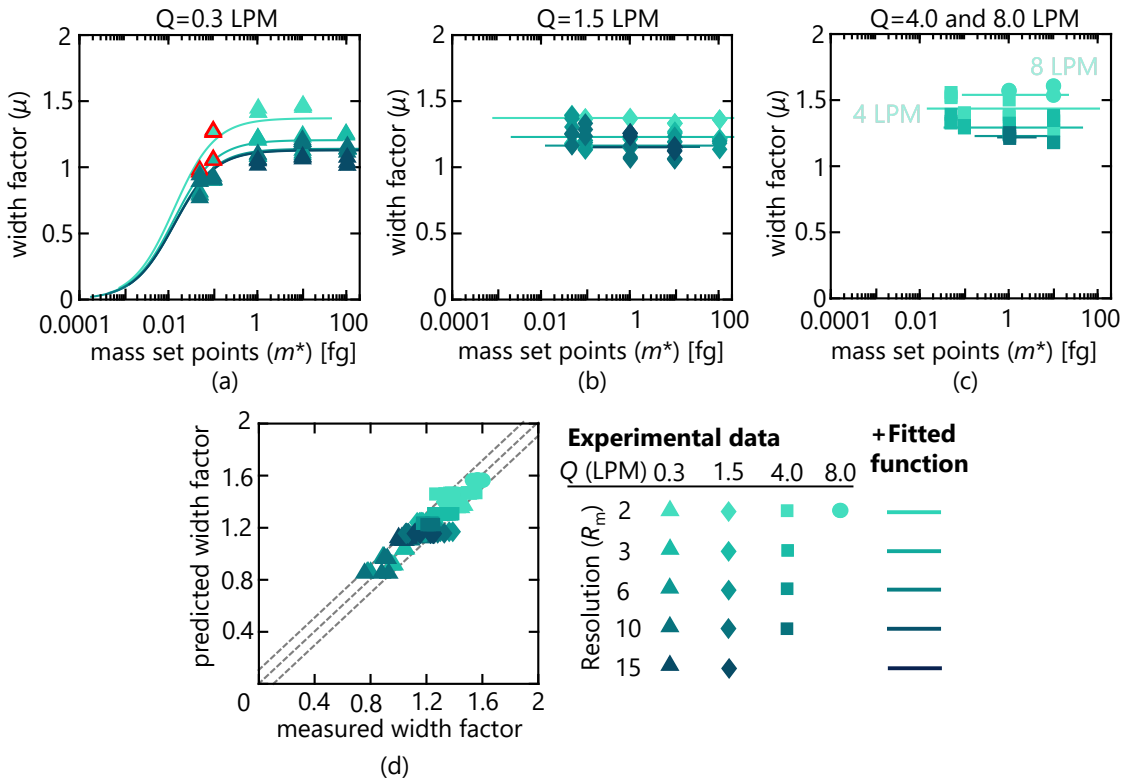

Figure 4.7: The width factor derived from TCPMA measurements for (a) $Q = 0.3$ LPM, (b) $Q = 1.5$ LPM, and (c) $Q = 4.0$ and $Q = 8.0$ LPM, at five CPMA mass set points between 0.05 and 100 fg, with five different resolutions ranging in R_m from 2 to 15. The corresponding results of fitted model (Equation 4.4.2) to distribution of measured width factor are illustrated as solid lines at a corresponding flow rate and CPMA resolution. (d) Parity plot of the width factor correlation (Equation 4.4.2) based on the TCPMA measurements.

4.4.3 Height factor (λ)

Figure 4.8*a–c* illustrate the height factor derived from TCPMA measurements as a function of mass set point for the flow rates of $Q = 0.3$, 1.5, 4 and 8 LPM, respectively. The height factor is a measure of the fractional loss of particles within the CPMA with the ideal transfer function having a value of one and a lower value

indicating lower transmission efficiency. In addition to the height factors derived from the TCPMA measurements, the losses were measured when the CPMA was not operating (no rotational speed and no voltage) using SMPS measurements upstream and downstream of the CPMA. This is an approximation of the case of $R_m = 0$, and the measurements were made with flow rates of 0.3, 1.5, and 4 LPM. This piece of information provides information at small particle masses where TCPMA measurements were not possible. The CPMA resolution is shown in gradations of color in Figure 4.8*a–c* to illustrate its effects on height factor. A distinct slope in Figure 4.8*a–c* indicates that height factors are highly dependent on particle mass, such that the height factor decreased with the CPMA mass set point. This trend is supported by the ANOVA findings shown in Table 4.1 (fourth column), which shows that a variation of the height factor was predominantly determined by the mass set point of the CPMA. Also, ANOVA suggests that the CPMA resolution also contributed to the variation of the height factor to a considerable extent. This functional dependence on CPMA resolution can be seen as a vertical layer of color in Figure 4.8*a–c*, such that higher resolutions (moving from light to dark color) result in a lower height factor. Additionally, as can be seen from the ANOVA results in Table 4.1, the p-value for flow rate is close to than 0.05, implying that the loss factor did not functionally vary with CPMA flow rate for the range of set points tested however this data set did not include the stationary measurements at mass set points lower than 0.05 fg, which clearly show a flow rate dependence.

Part of the losses in the CPMA can be attributed to inertial impaction, which is caused by the centrifugal forces acting on the particles as they pass through the CPMA. At higher CPMA resolution, the rotational speed of the cylinders for the

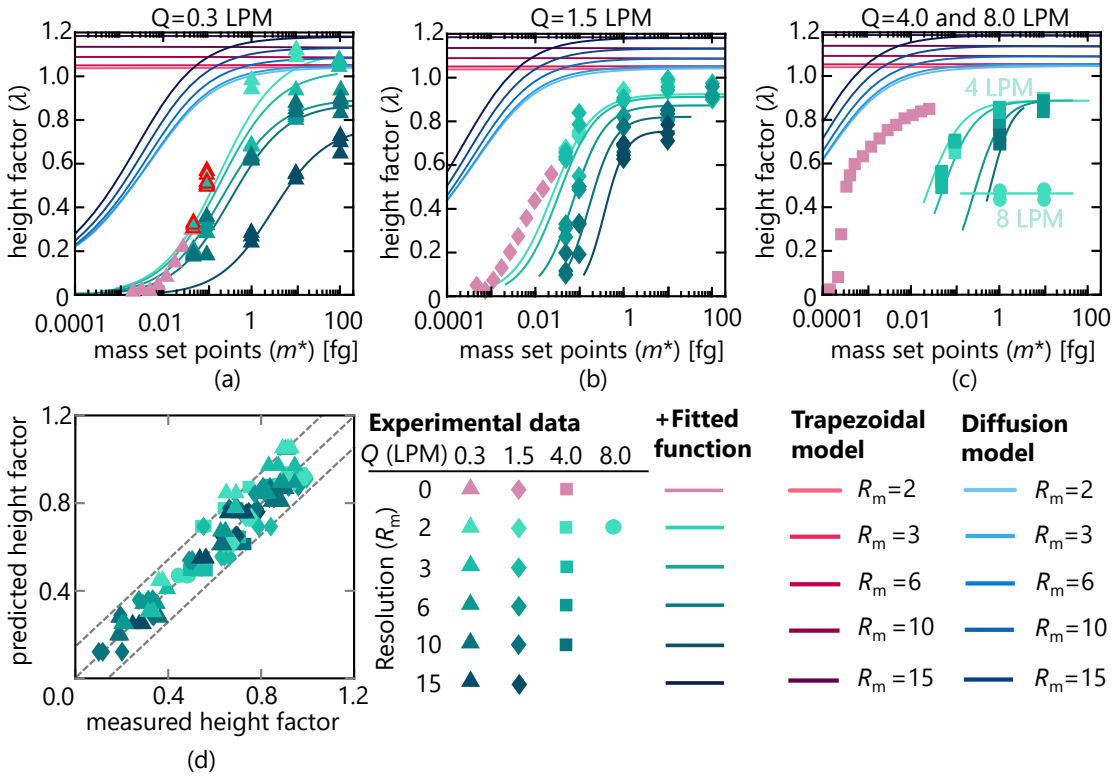

Figure 4.8: The height factor derived from the stationary loss measurements, TCPMA measurements, and corresponding synthetic data as a function of mass set point for flow rates of (a) Q=0.3 LPM, (b) Q=1.5 LPM, and (c) Q=4.0 and 8.0 LPM. The results of fitted model (Equation 4.4.3) to distribution of measured height factor are illustrated as solid lines at a corresponding flow rate and CPMA resolution. (d) Parity plot of the height factor correlation (Equation 4.4.3) based on the TCPMA measurements.

same mass set point is higher, so higher impaction loss are expected for particles moving through the turns before and after the classification region. Impaction, however, cannot account for the lower height factor of lower mass set points at a given resolution. At smaller particle masses losses are caused by diffusion which maybe be enhanced by turbulent flow outside of the classification region.

Figure 4.8*a*–c also depict the height factor derived from the synthetic data fabricated by convolution of theoretical trapezoidal and diffusion transfer functions for flow rates of Q=0.3, 1.5, 4 and 8 LPM, at CPMA mass set points ranging from 0.0001 to 100 fg, with five different resolutions ranging in R_m from 2 to 15. The theoretical height factors are greater than one (except where particle diffusion dominates, *e.g.,* at low flow rate and low mass set points), which is caused by the larger area of the theoretical models compared to the idealized triangular transfer function. It is evident that theoretical models do not correspond to experimental observations. The substantial discrepancy between the theoretical and measured height factors implies that there might be some effects, both inside and outside of the classification region, which have been neglected while developing the models. In addition to the assumptions mentioned in the previous section, the theoretical models assume no forces act on particles outside of the classification regions. In contrast, particles are subjected to both centrifugal and electrostatic forces while moving through the turns before and after the classification region. Therefore, the current theoretical models of the transfer function cannot accurately describe the loss of the CPMA, and an empirical model is necessary.

Accordingly, as with the width factor, data on height factors derived from TCPMA measurements for each flow rate were fitted to a single form of multivariate nonlinear model using multivariate nonlinear regression. The stationary data were incorporated into the fit to constrain the fit at small particle masses where TCPMA measurements

were impossible. The fitted relation found was,

$$\begin{aligned}
&\lambda\left(m^{*}, R_{m}\right)|_{Q=\text{const}} = \frac{a_{l}}{1+\exp\left[-b_{l}\cdot\left(\log_{10}\left(m^{*}\right)+\mathrm{c}_{l}\right)\right]} \pm 0.143; \\
&a_{l} = \left(a_{1}R_{m}^{3} + a_{2}R_{m}^{2} + a_{3}R_{m} + a_{4}\right), \\
&b_{l} = \left(b_{1}R_{m} + b_{2}\right), \\
&c_{l} = c_{1}\exp\left(c_{2}R_{m}\right) + c_{3}.
\end{aligned} \tag{4.14}$$

where R_m is the CPMA resolution, and m^* is the CPMA mass set point in fg; respectively. Table 4.2 shows the required coefficients of Equation 4.4.3 for each flow rate. Figure 4.8*a–c* show the results of fitted model to distribution of measured height factor as solid lines at three different flow rates of 0.3 to 8 LPM. The length of solid lines represents the theoretical operating region of the CPMA. As can be seen in Figure 4.8, there is a very narrow operational range in which the CPMA can operate at $Q = 8$ LPM, and the height factor remains constant throughout that range. Therefore, the regression coefficients in Equation 4.4.3 were set in such a way that the fit model gives a constant result. It is significant to note that this fit is only valid over the range of data collected, and the height factor for flow rates other than those measured in this study can be derived by interpolation.

Figure 4.8*d* shows the distribution of measured versus predicted values of height factor based on the TCPMA measurements for all three flow rates along with a 95% prediction interval (±0.143). The parity plot indicates that most points are near a diagonal, which suggests the model was fitted correctly. In the case of the fit, the root-mean-square error was found to be 0.0876.

Table 4.2: The regression coefficients of the fitted expression for height factor at three flow rates of Q =0.3, 1.5, and 8 LPM.

	Regression Coefficients								
	a_1	a_2	a_3	a_4	b_1	b_2	c_1	c_2	c_3
$Q_{0.3}$	-0.000605	0.0170	-0.159	1.373	0.00970	1.49	-0.0125	0.305	0.730
$Q_{1.5}$	0.000	0.000	-0.0130	0.950	0.151	1.89	2.37	-0.0636	-0.482
$Q_{4.0}$	0.000	0.000	0.000970	0.882	0.190	1.99	3.28	-0.300	0.000
$Q_{8.0}$	0.000	0.000	0.000	0.938	0.000	0.000	0.000	0.000	0.000

4.5 Conclusions

This study examined the transfer function of the CPMA by using a tandem CPMA configuration. Following the methodology described, a triangular transfer function was used to approximate the CPMA transfer function, and deviations from the idealized triangular transfer function were determined by three factors: (*i*) the mass set point agreement, (*ii*) the width factor, and (*iii*) the losses factor. The measurements with mass set point agreements greater than 1.15 or lower than 0.85, have been considered as outliers in this study as they do not conform to the main assumption underlying TCPMA measurements (*i.e.,* running two *identical* CPMAs in series). Excluding the outliers, the average mass set point agreement between the CPMAs is $m_{12}^{*} = 1.02 \pm 0.03$, indicating sound reproducibility among the CPMAs. In accordance with the TCPMA measurements, the width factor was typically found to be marginally higher than one ($\mu > 1$), indicating that the CPMA transfer function was narrower than the idealized triangular transfer function. Nevertheless, a broader

CPMA transfer function ($\mu < 1$) was observed when the CPMA was run at low mass set points (*i.e.,* $m^* < 1$ fg) and low flow rates ($Q = 0.3$ LPM). Actual CPMA transfer functions differed most significantly from idealized triangular transfer functions in terms of the height factor, which was influenced by the mass set point, resolution, and flow rate. The height factor decreased with a decrease in the CPMA mass set point and flow rate and increased with a decrease in the resolution. Multivariate non-linear fitting models were fitted to measured data in order to capture functional variations in the CPMA width and height factors.

Chapter 5

Comparing single-particle soot photometer mixing state measurement methods

5.1 Introduction

As described in Section 1.2.1 BC mixing states can be derived from SP2 measurements using different methods suggested in the literature, including (*i*) leading-edge only (LEO), (*ii*) incandescence lag-time, and (*iii*) tandem centrifugal particle mass analyzer (CPMA)-SP2 measurement. This study builds upon Chapter 2 work to provide the first comprehensive comparison of SP2-based methods to determine the mixing state of BC. Specifically, we compare the leading-edge-only (LEO), incandescence lag-time, and CPMA-SP2 methods in terms of detection size range, temporal resolution, counting statistics, and associated uncertainties. The normalized-derivative method is not explicitly considered, since it is similar to the leading-edge-only method. We evaluate these methods using atmospheric measurements taken in Kamloops, British Columbia, Canada during episodes dominated by predominately

urban and highway soot particles (low coating content) or moderate to heavy wildfire smoke particulate (moderate to high coating content).

5.2 Experimental methods

5.2.1 Measurement location

The measurements were carried out in a urban setting with the city of Kamloops, British Columbia, Canada (50°39′58.4″ N, 120°21′45.5″ W). The site was located 2.2 km from an air quality station operated by the British Columbia Ministry of Environment and Climate Change and $\sim$ 0.7 km from the Trans-Canada highway, which is a major corridor for heavy-duty diesel vehicles carrying freight. Experiments took place on July 21 and July 22, 2021; however, three periods of time are used in this work as examples to compare the measurement methods. The examples include periods where ambient particulate levels were at: *i*) Case I; low concentrations and mostly thinly coated rBC particles, presumably composed of mostly urban and highway emissions and may also contain some wildfire smokes (July 21, 11:54 am to 1:44; PM2.5 concentrations of 1.7 to 4 μg/m^3 as measured at the nearby air quality station), *ii*) Case II; moderate concentrations and mixture of thinly and thickly coated rBC particles due to urban and highway emissions and smoke from nearby forest fires having bimodal particle size/mass distribution (July 22, 11:31 am to 12:41 pm; PM2.5 concentrations of $\sim$ 104 to 81 μg/m^3), and iii) Case III; high concentrations and mostly thickly coated rBC due to wildfire smoke (July 22, 10:23 to 11:29 am; PM2.5 concentrations of 122 to 104 μg/m^3).

5.2.2 Measurement system

Figure 5.1 shows the method used to measure ambient particles. Ambient particles are drawn through $\sim$ 2.5 m of silicone conductive tubing and an X-ray aerosol charger (Model 3088, TSI Inc) where they received a bipolar equilibrium charge distribution. The particles were then directly sampled by the SP2 or classified by mass-to-charge ratio by the CPMA (Cambustion Ltd.) before being measured by SP2. The SP2 flow rate was maintained at 0.12 L/min during the experiments. However, to lessen particle diffusional losses (in the sampling line and the CPMA), a additional pump and critical orifice was used to maintain a flow of 1.5 L/min from the ambient inlet to the SP2 inlet.

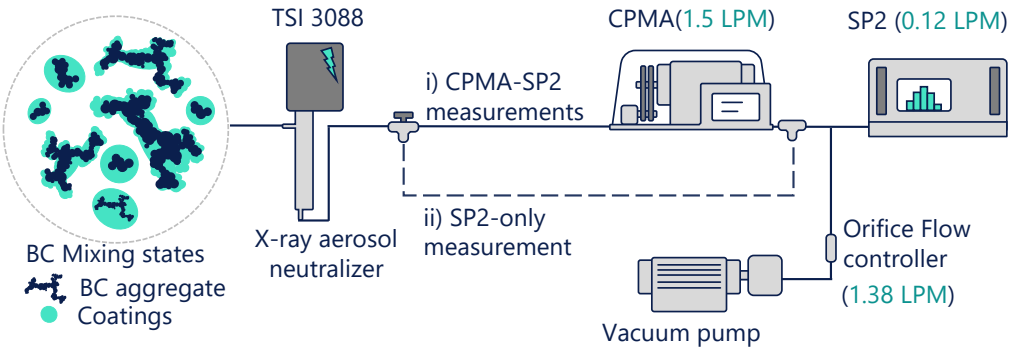

Figure 5.1: A schematic of the measurement system.

For CPMA-SP2 measurements, the CPMA was stepped through mass set points over the range of m^* =0.2–100 fg, with 5 CPMA set points per decade, and a CPMA resolution of 9. In general, 3 to 5 minutes were spent at each CPMA set point so that the total measurement time was limited to about an hour. This procedure led to about 1000 or more BC-containing particles being counted by the SP2 at each CPMA set point (the exception being CPMA set points near the limits of

the distribution where particle counts were very low). These experimental settings (resolution, number of CPMA set points, SP2 counts) were chosen based on the optimized settings suggested by Naseri et al. (2021b) for the particle concentrations and measurement time frames in this work. The SP2-only measurements were taken at the beginning, end, and occasionally between CPMA-SP2 scans.

5.2.3 SP2 calibration

The SP2 incandescence signal was calibrated using soot from a miniature inverted soot generator (Argonaut Scientific Corp.; Kazemimanesh et al., 2019; Moallemi et al., 2019) denuded with a catalytic stripper (Model CS08, Catalytic Instruments; Swanson and Kittelson, 2010). The denuded soot was classified with a CPMA over a wide range of set points and measured by the SP2, similar to the procedure used by Irwin et al., 2013. Since volatile material is removed from the rBC particles, the calibration particles have no coating by definition and the CPMA mass is equal to the rBC mass ($m_{\mathrm{p}} = m_{\mathrm{rBC}}$).[1] The scattering signal was calibrated using polystyrene latex (PSL) sphere standards (Thermo Scientific 3000 Series) with a diameter of 300 nm. The PSI SP2 toolkit, version 4.114, was used to obtain the calibration curves for each signal (Moteki and Kondo, 2007).

[1] In this context, the masses measured for incandescing particles using this calibration data are to be interpreted as rBC equivalent mass.

5.3 Data analysis methods

5.3.1 SP2-only measurements

5.3.1.1 Leading-edge only

Section 1.2.1 provides an explanation of the assumptions that underlie the LEO approach, as well as how it measures coating thickness (t_{coating}). To measure coating thickness/mass of rBC-containing particles using the Equation 1.1, the following values, which are often used in the literature, are used in this work as inputs to LEO approach unless otherwise stated: ρ_{rBC}= 1800 kg/m^3 (Bond and Bergstrom, 2006; Corbin et al., 2018; Liu et al., 2020; Moteki, Kondo, and Nakamura, 2010; Ouf et al., 2019; Park et al., 2004), ρ_{coating} = 1000 kg/m^3 (Ditas et al., 2018; Liu et al., 2019), $n_{\text{rBC}} = 2.26 + i\ 1.26$ (Dahlkötter et al., 2014; Laborde et al., 2013; Moteki, Kondo, and Nakamura, 2010; Taylor et al., 2015; Zanatta et al., 2018), and $n_{\text{coating}} = 1.5 + i\ 0$ (Laborde et al., 2013; Liu et al., 2015; Nakayama et al., 2010; Yuan et al., 2021).

5.3.1.2 Lag-time analysis

Section 1.2.1.3 provides an explanation of how Lag-time analysis categorizes BC-containing particles and provides a general picture of the mixing state of BC. In the present study, we used a time-lag threshold of $\Delta\tau$ = 2 μs based on the observed bimodal distribution of lag times in our data.

5.3.2 CPMA-SP2

Chapter 2 describes in detail how the CPMA-SP2 system measures BC mixing states. In this study, the model of Sipkens, Olfert, and Rogak (2020c) (Case 1C) was used to calculate the CPMA transfer function and the model of Wiedensohler (1988) was

used for the charge fraction. The inversion used a solution resolution of 64 bins per decade, a Bayesian model for regularization parameter selection, and the exponential distance method for imposing distribution smoothness on the solution.

5.3.3 Comparison of CPMA-SP2 and SP2-only data

We analyzed the CPMA-SP2 and SP2-only results according to their traditional methods of presentation below. We also converted between the two methods, to allow a direct comparison of differences in the two sets of results. To this end, a conversion from either rBC mass (m_{rBC}) and coating mass (m_{coating}) to rBC mass-equivalent diameter (d_{rBC}) and coating thickness (t_{coating}) or vice versa is required for comparison. The core-shell type morphology with core and shell material densities assumptions are required to convert CPMA-SP2 measurements (*i.e.,* m_{p} and m_{rBC}) to coating thickness (t_{coating}) and rBC mass equivalent diameter (d_{rBC}). The d_{rBC} is calculated by Equation 1.2, and t_{coating} can be derived from

$$t_{\mathrm{coating}} = \frac{1}{2}\left(\left(\frac{6\left(m_{\mathrm{p}} - m_{\mathrm{rBC}}\right)}{\pi\rho_{\mathrm{coating}}}\right)^{\frac{1}{3}} - d_{\mathrm{rBC}}\right), \tag{5.1}$$

where ρ_{coating} is coating density and is assumed to be 1000 kg/m^3, to match the assumptions made in the LEO analysis (*c.f.,* Section 5.3.1.1).This assumption is evaluated later in Section 5.4.3. While CPMA-SP2 needs no optical model assumptions to derive coating thickness, the LEO model, in addition to the core-shell Mie model and the material densities, requires assumptions for rBC and coating refractive indices.

5.4 Results and Discussion

5.4.1 LEO detection limits

Figure 5.2 shows two-dimensional distributions of Case II measured by the CPMA-SP2 (green images; panels a and b) and LEO (blue images; panels c and d) methods. The distributions are presented both in terms of coating thickness versus rBC mass-equivalent diameter (Figure 5.2*a* and *c*) and in rBC mass versus total particle mass (Figure 5.2*b* and *d*). The rBC mass and total mass are both measured directly by the CPMA-SP2 method, while the coating thickness and mass equivalent diameter are typically shown in the LEO literature. As such, we have included both types of plots for comparison.

Figure 5.2*a* and *c* demonstrate the relationships between rBC diameter and coating thickness derived from the CPMA-SP2 and the LEO analysis, respectively. The detection limits of the LEO method are indicated by red lines in Figure 5.2. The light scattering detection (LSD) and broadband incandescence detection (BID) limits are defined as the lowest and highest amount of light that the optics and detectors can collect and will vary between SP2 units and models. For the SP2 used here the limits are equivalent to about 150 nm — 430 nm in optical diameter, and 74 nm — 254 nm in rBC mass-equivalent diameter (0.38 and 15.36 fg), respectively. Lines *iii* and *iv* in Figure 5.2 are simply the low and high BID limits of the SP2. Line *ii* is the upper coating detection limit of the LEO method and it is limited by the saturation of the scattering detector. Since LEO is performed at a small fraction of the total scattering signal for a given particle, it increases the detection limit (*ii*) for sizing of non-absorbing particles. The upper coating thickness limit of LEO (line *ii*), was

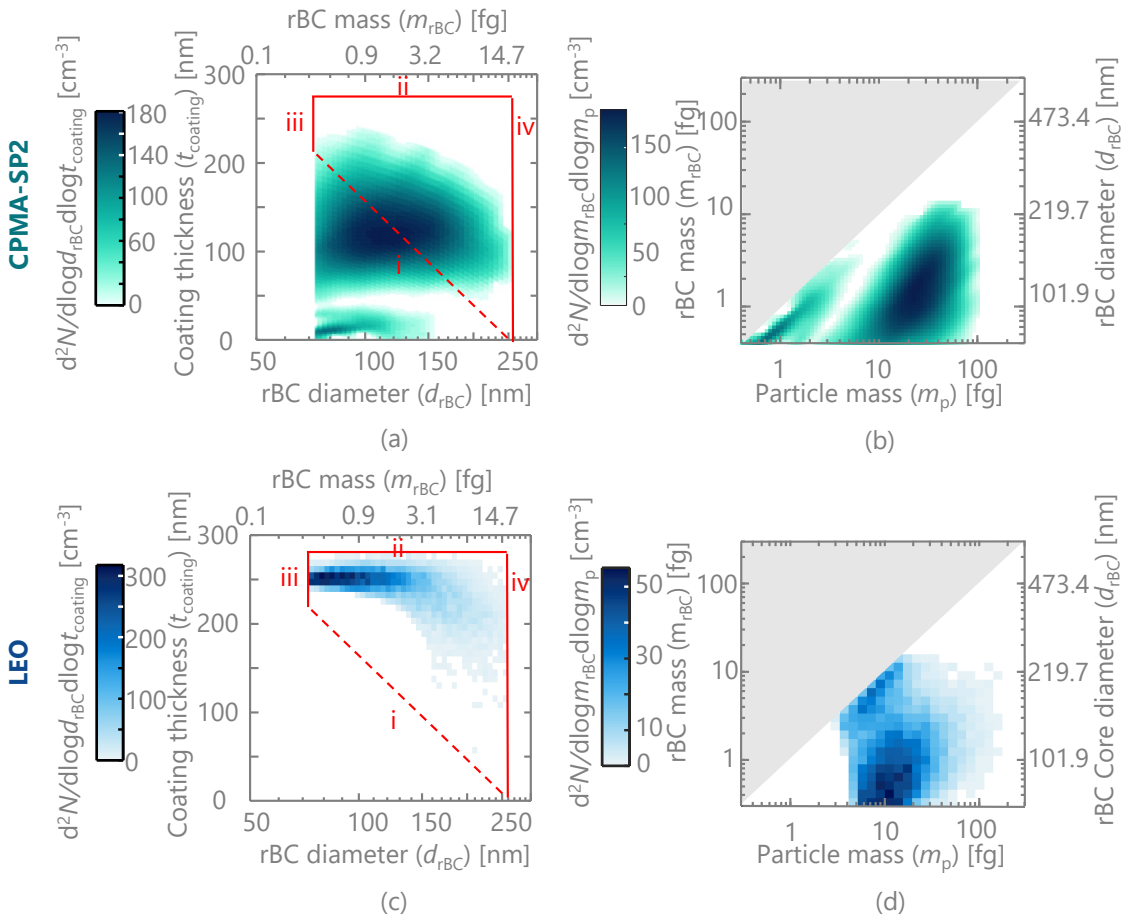

Figure 5.2: Distributions of rBC diameter and coating thickness (a, c), and total particle and rBC mass (b, d) of the same wildfire plume. Variants of the CPMA-SP2 measurements in panels (a) and (b) are shown in green. The variants of LEO analysis in panels (c) and (d) are shown in blue. Red lines in panels a and c indicate the detection limits, with *iii* and *iv* showing the low and high broadband incandescence detection limits, respectively, for both LEO and CPMA-SP2 methods, *i* and *ii* showing the LEO lower and higher coating detection limits.

$\sim$ 285 nm due to saturation of the scattering detector. The low coating thickness limit of LEO (Line *i*) is due to a combination of both lower LSD and BID limits. Because coating thickness is defined from the sum of rBC diameter and optical di-

ameter (Equation 1.1), the limit i means that the lowest coating thickness detectable increases with decreasing rBC diameter. This increase occurs because the additional light scattered by the rBC core raises the scattering signal above its detection limit. The exact boundary of the lower coating thickness limit of LEO is difficult to determine, especially because not only the LSD but also the split detector must be above the LOD. However, the endpoints of the limit are simply the lower coating thickness limit of LEO at the BID limits (Line i) (*i.e.,* lowest detectable coating thickness at low BID limit, and zero coating thickness at high BID limit) and is shown as a linear dashed line in the figure.

To make a general comparison between the LEO results and those derived directly by the CPMA-SP2 measurements without making any morphological and density assumptions (Figure 5.2b), the distribution of Figure 5.2c was mapped onto m_{p}—m_{rBC} mass space (Figure 5.2d) by rearranging Equation 5.1 to find m_{p}. It is not physically possible to have a particle in which the rBC mass exceeds the total mass of the particle; thus the reconstruction elements for which m_{p} is greater than m_{rBC} were prohibited during the formal inversion and are greyed out in the figure. A comparison of the particle mass range (m_{p}) of distributions represented in Figure 5.2b and d measured by the CPMA-SP2 and the LEO measurements, respectively, demonstrates that the LEO reconstruction range is far more limited than that of the CPMA-SP2 method. This deficiency is rooted in LEO's structural dependence on the scattering signal, which results in a high LOD for m_{p}, which can make using the LEO model inefficient when dealing with particles with low rBC mass and thin- to moderate-coatings. This point can be made clearer by a closer look at the mixing states of uncoated to heavily coated rBC particles from LEO and CPMA-SP2 measurements. Figure 5.3

shows three examples of m_{p}— m_{rBC} distributions, representing rBC-containing particles with mostly thin to no coatings (Case I, Figure 5.3*a* and *d*), as compared to a mixture of rBC particles with no coating to moderate and heavy coatings (Case II, Figure 5.3*b* and *e*), and rBC-containing particles mostly with heavy coatings (Case III, Figure 5.3*c* and *f*). In these plots, the main diagonal corresponds to the mass

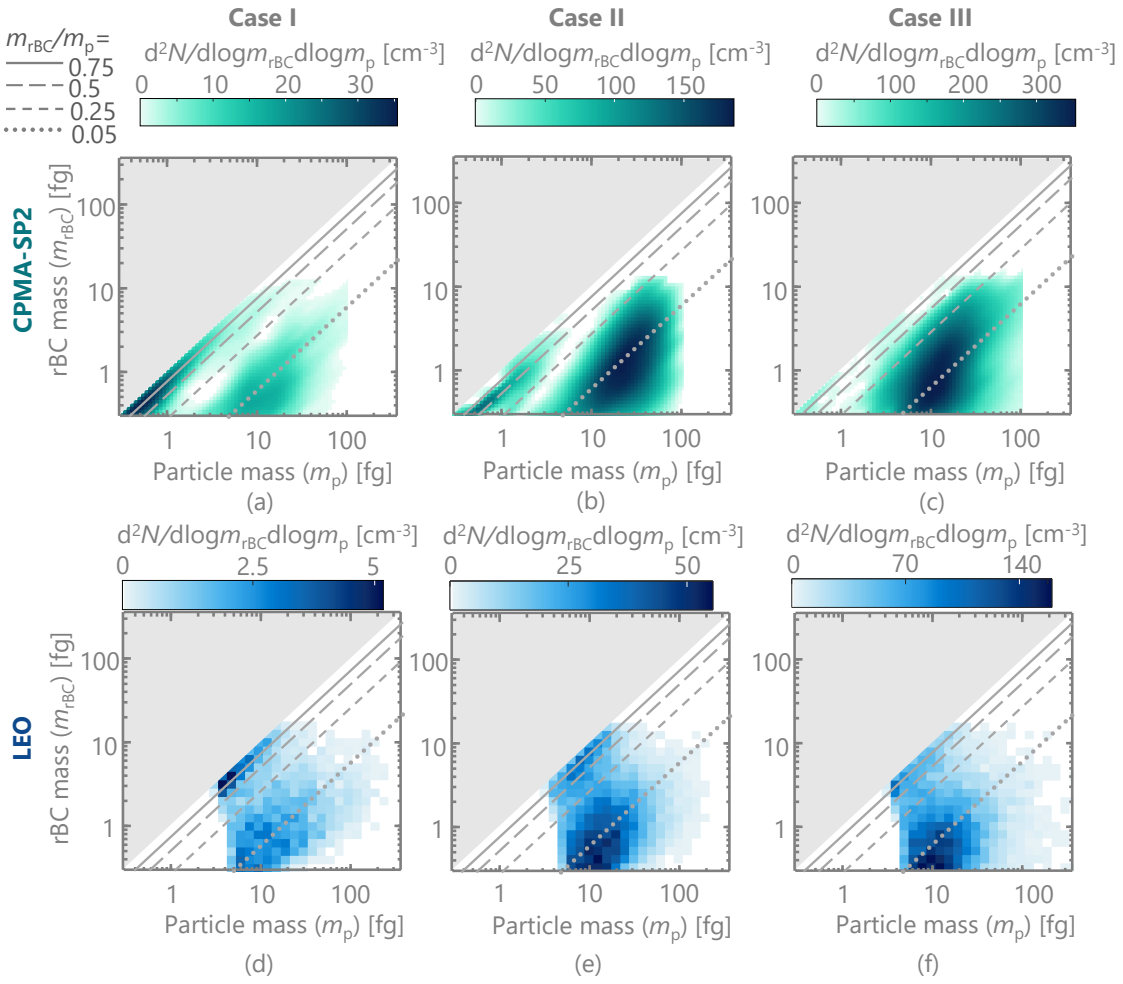

Figure 5.3: The total particle-rBC mass distributions (m_{p}—m_{rBC}) of cases I, II and III measured by (a-c) the CPMA-SP2 measurements and (d-f) the LEO analysis.

fraction of rBC of one, *i.e.,* $m_{\mathrm{rBC}}/m_{\mathrm{p}} = 1$, with any line parallel to it corresponding to lines of constant rBC mass fraction of less than one, representing coated particles (lines of rBC mass fractions of $m_{\mathrm{rBC}}/m_{\mathrm{p}}$ =0.75, 0.5, 0.25 and 0.05 are shown in Figure 5.3). Consequently, a part of distributions that are clustered along the diagonal line represents rBC-containing particles with thinly to no coatings. Overall, Figure 5.3*a*-*c* shows bimodal-bivariate distributions in which the distribution with a smaller particle mass mode corresponds to rBC particles with thin coatings (small rBC to coating mass ratio; $m_{\mathrm{coating}}/m_{\mathrm{rBC}} \ll 1$), and the second mode represents rBC particles with moderate to heavy coatings ($m_{\mathrm{coating}}/m_{\mathrm{rBC}} \gg 1$). As the relative concentration of rBC particles with thin and moderate to heavy coatings varies mostly due to changes in wildfire conditions, moving from the left distributions to the right ones in Figure 5.3, the second mode becomes wider and dominates the first mode.

It can be seen from Figure 5.3 that considerably more information on the mixing states of rBC-containing particles can be measured from CPMA-SP2 measurements than from the LEO analyses, such that LEO is biased towards the rBC with heavy coatings. The reason for this disparity is that rBC particles with thin-to-moderate may not be measured by the LEO method because of the lower LEO detection limit (line *i*). Thus, LEO analysis may only describe a small fraction of rBC-containing particle population.

Apart from a more limited detection range of LEO, a comparison of distributions represented in Figure 5.2*a* and *c*, show that LEO data is concentrated near the higher LSD limits (*ii*) in Figure 5.2*c*; however, the CPMA-SP2 results suggest there are no particles there. This indicates that the LEO method over-predicts the coating thickness of rBC-containing particles, which stems from the assumptions made in

LEO analysis, *e.g.,* the core-shell morphology. Besides, there are also some rBC particles with a total particle mass of $\sim$ 4–20 fg that is clustered near the 1:1 line in each case, which is an artifact discussed in detail in Section 5.4.3. Additionally, a closer look at Figures 5.2 and 5.3 shows that the number concentration of the LEO analysis is much lower that the CPMA-SP2. These two issues are discussed in the following sections.

5.4.2 LEO counting statistics limitation

Figure 5.4 shows the number concentration distributions of particle mass ($\mathrm{d}N/\mathrm{dlog}m_{\mathrm{p}}$) and rBC mass ($\mathrm{d}N/\mathrm{dlog}m_{\mathrm{rBC}}$), respectively, for the LEO and CPMA-SP2 methods which are found by integrating the two-dimensional distributions of Figure 5.2*b* and *d* over m_{p},

$$\frac{\mathrm{d}N}{\mathrm{dlog}m_{\mathrm{rBC}}} = \int_0^\infty \frac{\partial^2 N}{\partial \log m_{\mathrm{p}} \partial \log m_{\mathrm{rBC}}} \mathrm{dlog}m_{\mathrm{p}}, \tag{5.2}$$

and m_{rBC},

$$\frac{\mathrm{d}N}{\mathrm{dlog}m_{\mathrm{p}}} = \int_0^\infty \frac{\partial^2 N}{\partial \log m_{\mathrm{p}} \partial \log m_{\mathrm{rBC}}} \mathrm{dlog}m_{\mathrm{rBC}}. \tag{5.3}$$

According to Figure 5.4, the maximum number concentration of rBC-containing particles detected by LEO analysis is only $\sim$ 50% of that directly measured from the incandescence signal of the SP2, because of the higher LOD of the scattering signal.

There are three common types of rBC-containing particles with reliable incandescence signals whose scattering signals escape detection by LEO analysis. The first is rBC-containing particles with negligible scattering signals, which result in noisy fits and unreliable sizing (below line *i* in Figure 5.2). The second is rBC-containing particles with very large scattering signals, in which the LSD is saturated from the

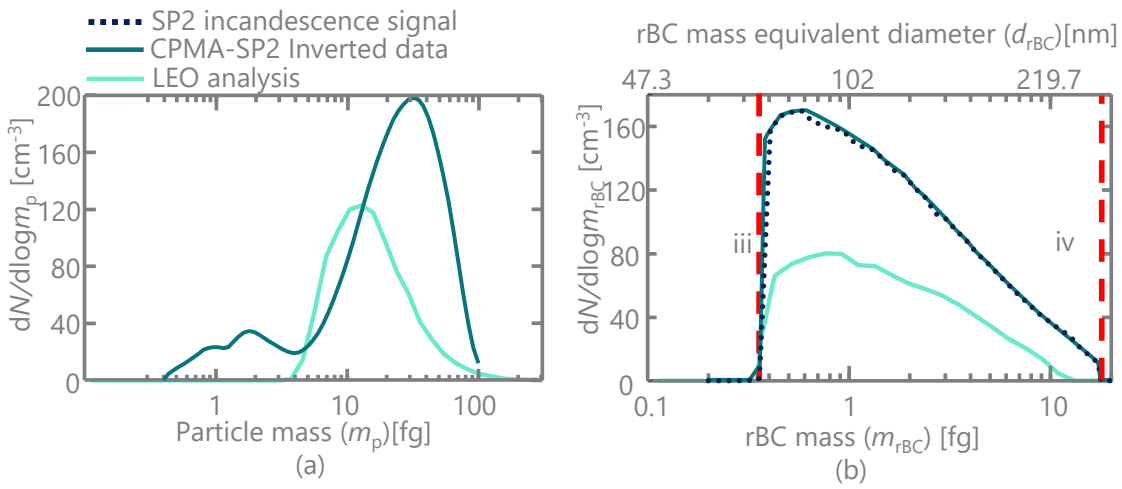

Figure 5.4: Comparison of the number concentration of rBC-containing particles of Case II as a function of total particle mass (a) and (b) as a function of rBC particle mass measured by the LEO analysis with the one directly measured from the incandescence signal of the SP2. The boundaries of low and high BID limits are indicated by red dashed lines of *iii* and *iv*, respectively.

leading edge, making the fits unreliable (above line *ii* in Figure 5.2). The exact threshold for this effect may vary between instruments, and these particles can be excluded from the LEO analysis (Gao et al., 2007). Finally, the third type is those rBC particles that evaporate before their position in the laser beam (*i.e.,* notch position) can be detected appropriately by the avalanche photodiode (*split detector*). The split detector is used in the SP2 measurements to constrain the Gaussian fits used in the LEO analysis in such a way that when rBC-containing particles cross through the laser beam, the light scattered by them shifts from one element of the detector to another while it is inverted on one side. As particles move through the center of the detector, this inversion in scattering signal creates a clear notch that is used to infer the position of rBC-containing particles inside the laser beam, as the

distance between the notch and the peak intensity of the laser changes only when the instrument optics are realigned during servicing (*i.e.,* it remains constant during the measurements). The effect of excluding these rBC-containing particles from the LEO analysis can be seen in Figure 5.4, where the concentration of rBC-containing particles having valid scattering data and are measurable by the LEO is compared with the one directly measured by the incandescence signal of the SP2. Overall, the concentration of LEO rBC-containing particles is considerably lower than the one measured by the CPMA-SP2 method, as discussed above. The second reason is that the CPMA-SP2 measures only incandescence signals and the rBC-containing particles excluded in the LEO analysis remain in the CPMA-SP2 analysis. Moreover, due to constraints placed on the inversion of the CPMA-SP2 measurements, the $\mathrm{d}N/\mathrm{d}\log m_{\mathrm{rBC}}$ distributions measured by the SP2-only and the CPMA-SP2 are almost the same. Overall, we conclude that LEO method only provides mixing state information on a small and biased subset of the entire rBC population.

5.4.3 Evaluation of the accuracy of LEO coating thickness measurement

A comparison of the particle mass (m_{p}) distributions illustrated in Figure 5.2 to 5.4 shows that the LEO particle mass modes are noticeably larger than the corresponding CPMA-SP2 ones. The accuracy of LEO coating thickness measurements, and the assumptions used to calculate it, can be tested by measuring coated rBC particles of known mass. This was done using the CPMA to select particles (known particle mass) and conducting LEO analysis on the CPMA-classified particles. In such an experiment, the particle mass determined by LEO analysis should fall within the

mass range given by the CPMA set-point (m^*) (*i.e.,* $m_\mathrm{p} \in qm^* \times (1 \pm 1/R_\mathrm{m})$; where R_m is the CPMA resolution). LEO analysis was performed on CPMA-classified particles at a mass set point of $m^* = 14.38$ fg with R_m=9 (a set point where a large number of coated particles are present) for the Case I and Case III. Figure 5.5*a* and *b* shows the normalized m_p— m_rBC distributions of CPMA-classified particles, along with their corresponding marginal distributions (*i.e.,* $\mathrm{d}N/\mathrm{dlog}m_\mathrm{p}$, Equation 5.3), that were determined by LEO for rBC particles mostly with thin or no coating (Case I) and moderate-to-heavy coatings (Case III) using a coating refractive index of $n_\mathrm{coating} = 1.5$ and ρ_coating=1000 kg/m^3. The distribution of CPMA-classified particles should exhibit two modes if the rBC-containing particles were homogeneous in chemical composition (*e.g.,* of the same type), and LEO analysis, and its corresponding assumptions, were perfectly correct: (*i*) a mode with a higher concentration which would equal the CPMA mass set point (shown as dashed line in Figure 5.5) and (*ii*) a lower-concentration mode at twice the CPMA mass set point representing doubly-charged particles ($q = 2$). However, it is observed that for both cases, not only is there another mode at smaller particle masses (Mode A), but also the expected modes (Mode B and C) occur at particle masses which are higher than they should be.

Figure 5.5a and 5.5b show that Mode A is associated with particles that have a large rBC mass and a small amount of coatings. The particle mass of these rBC particles is underestimated by LEO since there are some particles with a total mass below 12.8 fg which can not physically pass through the CPMA at the given operating condition (the transfer function width is 12.8 to 16.0 fg, shown as a red shaded region in Figure 5.5). Additionally, it is seen that the peak of Mode B for Case I

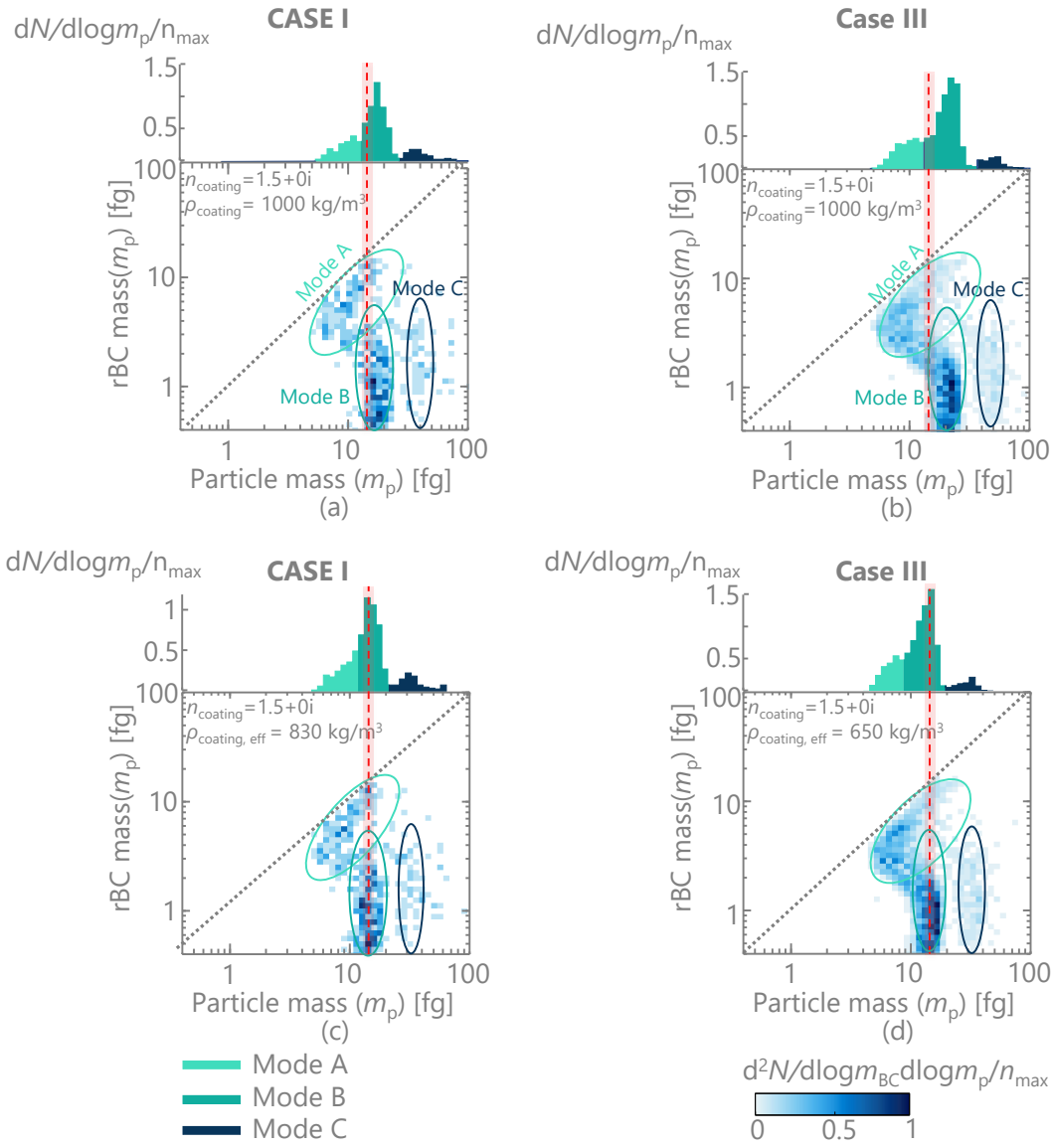

Figure 5.5: The normalized m_{p}— m_{rBC} distributions of mass-classified particles at m^{*} = 14.38 fg that were analyzed by LEO for rBC-containing particles with (a,c) thin or no coatings (Case I) and (b,d) moderate-to-heavy coatings (Case III) using the coating refractive index of n_{coating} = 1.5 and (a,b) ρ_{coating}=1000 kg/m^3 or the coating effective densities of $\rho_{\mathrm{eff,c}}$=830 and 650 kg/m^3 derived for (c) Case I and (d) Case III at the given coating refractive index, respectively.

is slightly larger than the CPMA set point ($\sim$ 6%) while the peak of Mode B for Case III is $\sim$ 53% larger than the CPMA set point and lies outside the physically-possible mass classification range of the CPMA transfer function at the given set point. Consequently, the LEO method overestimates the particle mass (*i.e,* coating mass) of particles that are highly coated (Mode B of Case III), while it underestimates the particle mass of thinly coated rBC (Mode A). This suggests that the error in the LEO method might have been due to the physical properties of the coatings (*e.g.,* coating density and refractive index) used in Mie theory or due to the shell-core assumption.

To explore the role of LEO's property assumptions, Equation 5.1 was solved for the coating density such that Mode B, determined by the LEO analysis at a given coating refractive index, matched the CPMA mass set point. Figure 5.5*b* and *d* show the adjusted m_{p}— m_{rBC} LEO distributions of CPMA-classified particles using effective coating densities at a coating refractive index of $n_{\mathrm{coating}} = 1.5$. As shown in Figure 5.5*c* and *d*, an effective coating density of 830 kg/m^3 was needed to adjust the peak of Mode B to the correct particle mass for Case I, while an effective coating density of 650 kg/m^3 was needed for Case III. However, using this effective coating density overcompensates Mode A and shifts it to even lower masses. Thus, the LEO results could only be partially corrected by this approach; Mode B and Mode C are forced to be in the expected places while Mode A still appears at physically-impossible particle masses. This suggests that particles in Mode A are physically different than Mode B, meaning they could have different coating or core physical properties (refractive index and density). This might be expected as some rBC particles could be mature soot generated by combustion engines or high-temperature

biomass combustion while the other could be immature soot from wildfire biomass pyrolysis.

This analysis method was extended to include changing both the effective coating density and the coating refractive index to align the LEO-analyzed particle mass from Mode B with the actual particle mass as shown in Figure 5.6 for the three cases. Case I yields, on average, a $\sim 30\%$ higher effective coating density than Cases II and III. The reason for this considerable difference could be because there were

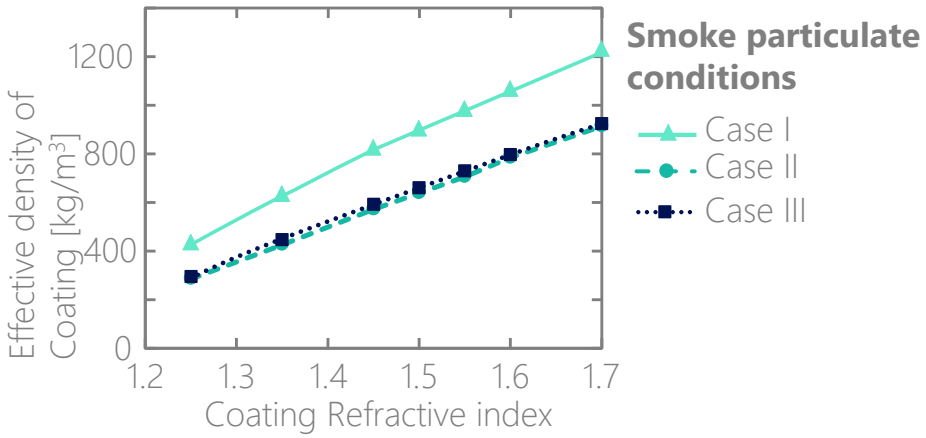

Figure 5.6: (a) The coating effect of the refractive index of coating materials in Mie calculations on the effective coating densities that match the particle mass modes of various rBC-containing particles retrieved by the LEO method at a CPMA mass set point of 14.38 fg.

different types of rBC-containing particles in each case. Many factors, like the type of biomass and its moisture contents, and any influence on the pre-ignition pyrolysis phase would change the physical and chemical composition, and consequently, the material density of organics emitted from the combustion (Hu et al., 2021a). Also, the particle concentrations in Case I are substantially lower that the other cases, which

likely results in a higher fraction of rBC particles originating from Diesel engines rather than wildfire smoke, which likely have different physical and optical properties. It is also seen that for Case II and III, which are primarily composed of thickly coated rBC particles presumably caused by wildfire smoke, the coating densities determined at a given range of coating refractive index are quite low, ranging from 350 to 900 kg/m^3 over a range of coating refractive index of 1.25 to 1.7. However, studies have shown that the material density of organics in wildfire smoke ranges from 800 to 1600 kg/m^3 (Hu et al., 2021a; Li et al., 2016; Nakao et al., 2013; Turpin and Lim, 2001). These densities are as low as the effective densities measured for soot agglomerates with highly open structures, which is extremely unlikely to be the case for soot coatings. This implies that the range of effective densities shown here are not the material density of the coating, but rather an effective density that reconciles the errors in the LEO analysis. Alternatively, these effective densities may imply tarball agglomerates were present in the sample (Girotto et al., 2018; Yuan et al., 2020).

This study found that the LEO method was not successful in measuring the particle mass (or coating mass) of rBC-containing particles. The error in LEO measurements becomes even larger when the sample populations are composed of a variety of rBC-containing particles with their own optical and morphological characteristics. While the LEO analysis is considerably sensitive to such assumptions and inputs, the CPMA-SP2 measurement, on the other hand, is independent of the optical and morphological properties of the particles.

5.4.4 Lag-time limitations in coating estimation

Lag-time analysis of SP2 data is used qualitatively to classify rBC particles as thickly- or thinly-coated. The range and accuracy of the lag-time classification can be assessed by comparing it to CPMA-SP2 data. Figure 5.7*a-c* shows the normalized number concentration distributions of lag-times as a function of rBC mass for the three sample cases. Colors represent the number fraction of particles within a finite rBC mass range and lag-time ($\mathrm{dlog}m_{\mathrm{rBC}}$ and $\mathrm{d}\Delta\tau$). As shown in Figure 5.7*a-c* the measured particles exhibit a distribution of lag-times, spanning negative and positive values, and with multiple modes (most clearly seen in Case I). A positive lag-time occurs when the scattering signal reaches a maxima before the incandescence signal maxima. This is interpreted as a considerable loss of coating material due to the heat absorbed by the rBC particles (*i.e.,* the particles were thickly-coated). This type of particle appeared as a distinct mode in Figure 5.7*a-c* centered around $\Delta\tau \sim 6\mu s$ and covering almost the entire rBC mass range. Lag-times near zero are attributed to rBC-containing particles with thin or no coatings. As expected, this type of particle is more evident in Case I, which consists mostly of thinly coated rBC particles, and exhibits a distinct mode in Figure 5.7*a*. This mode spans only over the range of rBC mass higher than 1 fg as rBC particles with thin or no coating do not scatter enough light to be detected. As a result, the lag-time method, similar to the LEO method, has a limited detection range. There is also another scenario where the incandescence signal peaks before that of the scattering signal, resulting in negative lag-times. Negative lag-times can be caused by the fracturing of particles which have a surface-attached rBC component (*e.g.,* not a core-shell morphology) (Dahlkötter

et al., 2014; Moteki and Kondo, 2007; Moteki, Kondo, and Adachi, 2014; Sedlacek III et al., 2012, 2015). In this case, the non-BC components pass through the laser beam and evaporate only partially or not at all, resulting in light scattering after the rBC has incandesced. Consequently, these particles are erroneously categorized as thinly coated.

The time-lag threshold of $\Delta\tau = 2$ μs, which is shown as gray dashed lines in Figure 5.7*a-c*, is used to categorize the rBC and the results are shown in Figure 5.7*d-f*. Figure 5.7*d-f* illustrates the number fractions of particles with *thick* coatings (lag-time >2 μs) versus *thin-to-moderate* coatings (lag-time <2 μs) for an rBC mass range $m_{\mathrm{rBC}} \sim$ 1 fg —15.4 fg (rBC mass-equivalent diameter d_{rBC} ~102 nm—254 nm). Additionally, there were some particles that could not be assigned to either category. The number fraction of these particles is attributed to particles with reliable incandescence signals whose scattering signals cannot be reliably detected or saturated by the detector. The indicated number fractions of particles with thick or thin-to-moderate coating are calculated with respect to either *i*) the total number of rBC particles detected over the course of measurement time (as it is done in this paper), or *ii*) the total number of rBC particles whose scattering signals are not less than limit of quantification as done in the PSI SP2 toolkit lag-time analysis. Therefore, the latter approach excludes a portion of rBC-containing particles from the lag-time analysis and estimates the number fractions of rBC with thick and thin-to-moderate coating at a given rBC mass based on the remainder. For meaningful binning, it is imperative that the number fraction of rBC particles with unreliable scattering signals be small in the lag-time analysis otherwise the lag-time-related uncertainties become so large because a significant number of particles are not identified with either category. According to

Figure 5.7*d-f*, the fraction of particles that could not be assigned to either category by the PSI SP2 toolkit remains below 40 % over the entire rBC mass range. It is evident, however, that when all rBC particles are taken into consideration, the proportion of uncategorized rBC particles rises to a significant level, such that over ∼60% of particles in Cases II and III were unclassified due to detector saturation. The unclassified particles could be of either type, thus the uncertainty in the number fraction of each type will be determined by assigning all the unclassified particles to one bin or the other. Figure 5.7*g-i* show the number fraction of particle types where the limits of the shaded region represent the maximum and minimum number fraction possible. The figure shows there is a substantial uncertainty above the values indicated by the lag-time analysis, such that the only time in which the uncertainties of particles do not overlap, and thus one can distinguish between particles with thick and thin-to-moderate coating, is when the number of particles with invalid scattering signals is below $\sim 10\%$, which occurred only in Case I when the rBC mass range was between 7 fg and 15 fg.

Figure 5.7*j-i* shows the normalized number distributions of coating mass fraction ($f_c = m_{coating}/m_p$) as a function of rBC mass for all three cases measured by the CPMA-SP2 system. These distributions are derived by transforming the m_p—m_{rBC} distributions of Figure 5.3 into coating mass fraction — rBC mass space. In contrast to lagtime analysis: i) the amount of coating is quantified, ii) the range is not limited by the scattering detector, iii) and particles are not misclassified if the rBC is attached to the side of the coating fraction.

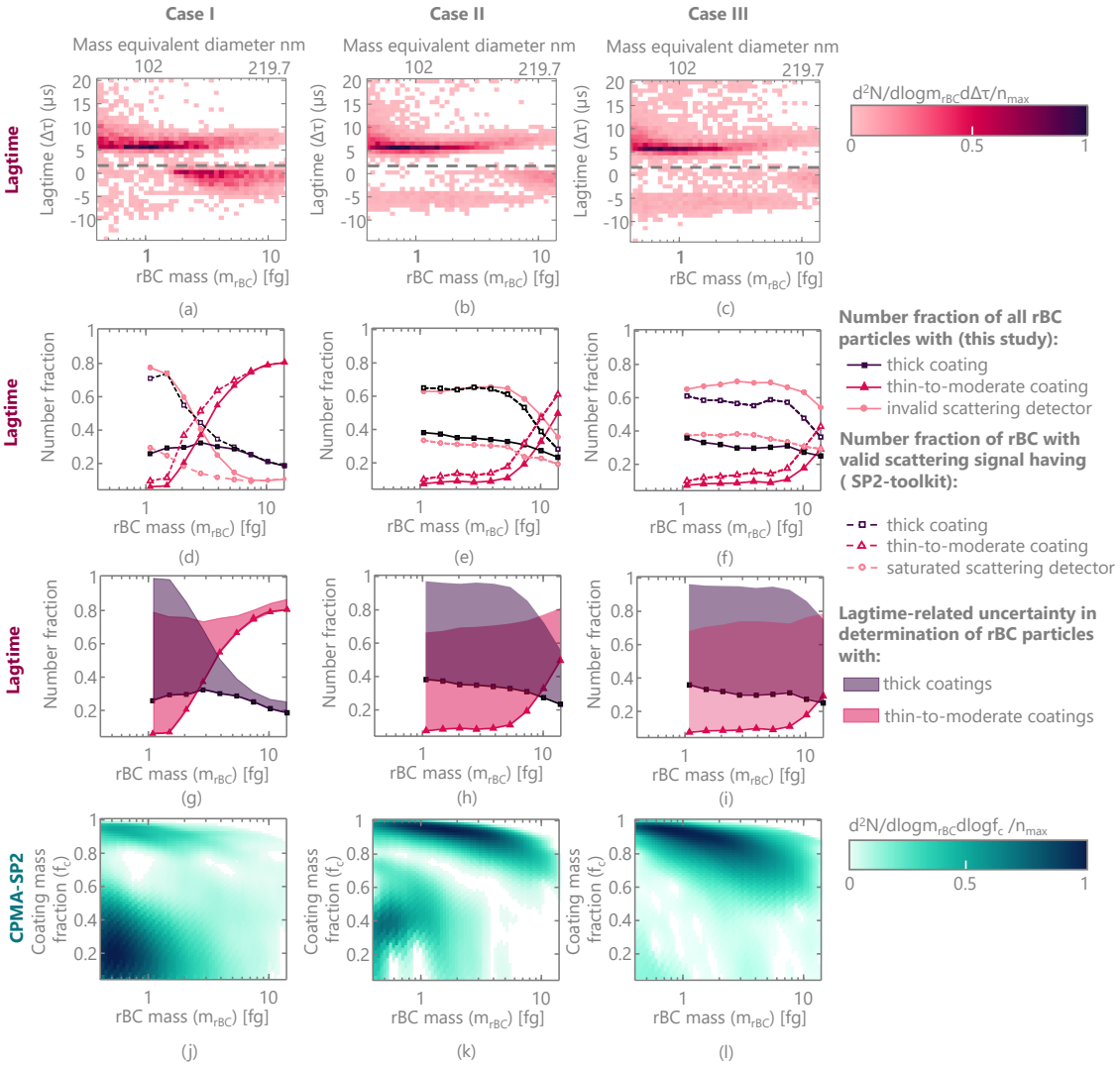

Figure 5.7: Normalized number concentration distributions of lag-times as a function of rBC mass for (a) Case I, (b) Case II and (c) Case III. The time-lag threshold of $\Delta\tau = 2\ \mu s$ is shown as a dashed line in panels (a) to (c). (d-f) The comparison of number fractions of rBC particles with thick coatings and thin-to-moderate coatings, as determined by SP2 lag-time analysis, for the measured rBC particle populations. (g-i) Lagtime-related uncertainty in determination of rBC particles with thick and thin-to-moderate coatings. (j-l) The distribution of coating mass fraction ($f_c = m_{\text{coating}}/m_{\text{p}}$) as a function of rBC mass of particle populations measured by the CPMA-SP2 measurement system.

5.4.5 Temporal resolution of SP2 measurement methods

Figure 5.8 shows the differences between the m_{p}—m_{rBC} distribution of smoke particulates of Case III derived from a ∼80-minute CPMA-SP2 measurement (Figure 5.8a) and that obtained by ∼ 1-minute SP2 LEO measurements performed (Figure 5.8b) before (t= 0 min), (Figure 5.8c) during (t= 30 min; between the CPMA set points of 5.2 fg and 8.6 fg), and (Figure 5.8d) after the CPMA-SP2 scan (t= 80 min).

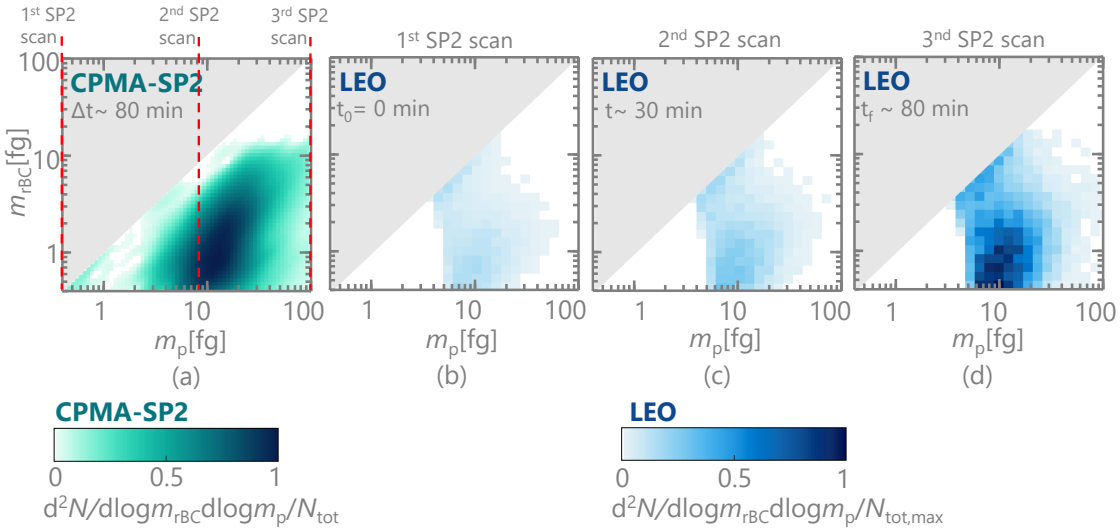

Figure 5.8: Comparison of the average m_{p}—m_{rBC} distribution of smoke particulates of Case III derived from (a) CPMA-SP2 measurement and the average distribution obtained by instantaneous LEO measurements performed (b) before, (c) during, and (d) after the CPMA-SP2 scan.

The CPMA-SP2 measurement was performed in such a way that the CPMA mass set point was scanned in an ascending order (*i.e.,* CPMA was stepped from the low CPMA mass set point of 0.4 fg to the high mass set point of 100 fg over ∼80 minutes). The m_{p}—m_{rBC} distribution in Figure 5.8a is normalized by the total number concentration of rBC-containing particles derived from the CPMA-SP2 measure-

ment, whereas the m_{p}—m_{rBC} distributions in Figure 5.8b-d are normalized by the maximum total number of particles found in the LEO measurements. The number concentration of rBC-containing particles increased by ~ 3.7 times over the first 30 min and $\sim$5.6 times over the entire time of the CPMA-SP2 scan. The CPMA-SP2 inversion assumes that the number concentration of particles is constant with respect to time over the course of the CPMA-SP2 scan. As is the case here, this assumption is not always true. Since the CPMA set points were stepped upward in mass, the number concentration of low-mass particles is underestimated, while the number concentration of high-mass particles is overestimated relative to time-averaged LEO measurements.

The compromise of the CPMA-SP2 approach is that it is slower, typically requiring over 30 minutes per scan, compared to a few minutes for SP2-only measurements (both durations depend on the number concentration of particles). Therefore, for measurements of samples that are highly transient, as in the case of high-altitude research aircraft the CPMA-SP2 method would require a sample reservoir.

5.5 Conclusions

The present study examined the performance of different SP2 measurement methods in measuring rBC mixing states using example data with a wide range of mixing states. The rBC mixing states were characterized with a particular focus on the tandem CPMA-SP2 measurements and SP2-only measurements using LEO analysis that map out the rBC-containing particle characteristics in two-dimensional mass space. Comparison of the results of the LEO analysis and the CPMA-SP2 indicates that the LEO reconstruction range is very limited, because it relies on the

less-sensitive light-scattering signal as well as the position-sensitive light-scattering detector. These limitations, along with the assumptions on which the LEO is based (*e.g.,* core-shell morphology and assumed density for coating materials), bring about biased results and make the LEO method a qualitative measure, for a narrow range of the population. The results of this study indicate that the CPMA-SP2 measurement outperforms SP2-only measurements in terms of accuracy. The compromise of the CPMA-SP2 method is that it is not a real-time measurement, requiring about 30 minutes per scan. Rapidly changing aerosol samples should therefore be captured in a reservoir prior to CPMA-SP2 measurement, to ensure valid results. In addition, as a result of the high level of uncertainty in classifying rBC particles with invalid scattering signals, the lag-time analysis could not distinguish between the relative number fraction of rBC particles with thick and thin-to-moderate coating in the example data presented.

Chapter 6

Conclusion

6.1 Summary

In previous chapters, an improved morphology-independent measurement was explained for determining the coating mass of atmospheric BC particles. This method employs CPMA and SP2 in tandem, coupled with novel inversion schemes, to measure the two-variable distribution of total particle mass and rBC mass, $\frac{\partial^2 N}{\partial \log m_{\mathrm{p}} \partial \log m_{\mathrm{rBC}}}$, from which the distributions of coating mass or coating mass ratio on each rBC particle can also be calculated ($\frac{\partial^2 N}{\partial \log m_{\mathrm{coating}} \partial \log m_{\mathrm{rBC}}}$). Furthermore, by performing iterative integration of the two-dimensional distributions, a single variable of mass and number concentrations, as well as the total mass and number concentration of rBC-containing particles, can be obtained. Since the inversion of CPMA-SP2 data is required to account for the CPMA-SP2 transfer function and multiply charged particles, as well as the physical constraints associated with this measurement system, extensive effort has been dedicated to developing a novel inversion scheme.

The novel inversion scheme introduced the deconvolution scheme appropriately

accommodating the physical constraint that limits rBC mass not to exceed the total particle mass. A regularization method must consider this constraint during the inversion to account for the sharp edge that occurs at m_{rBC} equals m_{p}, which results in established regularization methods underperforming and possibly smoothing the distribution over this border. Furthermore, the proposed method addresses the problem that the mobility of black carbon particles is unknown yet essential for kernel construction by using an innovative approach to constrain the inversion in such a way that the marginal distribution resulting from the CPMA-SP2 inversion corresponds to the distribution of the black carbon particles measured directly by SP2 when CPMA is bypassed (*i.e.,* $\mathrm{d}N/\mathrm{dlog}m_{\mathrm{rBC}}$). The introduced inversion scheme is examined by fabricating synthetic rBC mass — particle mass distributions (*i.e.,* phantoms) and discuss regularization approaches of the following categories: (*i*) unregularized least-squares analysis, (*ii*) Twomey-type approaches, (*iii*) maximum entropy regularization, (*iv*) Tikhonov regularization, and (*v*) exponential distance regularization. The results reveal that the exponential distance method outperforms the other regularization methods tested.

Determining the optimal sampling settings in a collection of the CPMA SP2 system response, which is described in Chapter 3, was the next step to improve the accuracy of reconstruction of *i.e.,* m_{p}— m_{BC} distribution. Monte Carlo sampling was used to assess the sensitivity of reconstruction error to experimental inputs across a range of distribution widths, and recommendations were made. Recommendations vary based on the concentration of the aerosol, the time available for the measurement, and the width of the distribution, yet general recommendations for optimum CPMA-SP2 sampling configuration to conduct the measurement on typical aged at-

mospheric aerosols include: (*i*) large numbers of SP2 bins per decade ($n_s^{'} = 64$), (*ii*) a preference for higher SP2 sampling counts per CPMA setpoint ($N_s > 10^4$), (*iii*) a moderate number of CPMA setpoints per decade ($n_c^{'} = 3$ to 8), (*iv*) CPMA resolution around 1.0, and (*v*) a high CPMA flow rate (1.5 L/min). These settings generally result in reconstruction accuracies below 3%.

Since the experimentally evaluated CPMA transfer function can significantly enhance CPMA data analysis routines, especially in the case of CPMA-SP2 data inversion, the tandem CPMA methodology, described in Chapter 3, was utilized to characterize the CPMA transfer function. The CPMA transfer function is approximated with a triangular transfer function and experimentally evaluated by the tandem CPMA measurement technique. Deviation of the idealized transfer function from the actual case is measured by two adjustable factors; (*i*) the width factor (μ), and (*ii*) the transmission efficiency (*i.e.*, the loss factor, λ). These factors are quantified in relation to CPMA configurations (*i.e.* CPMA mass setpoints, resolutions, and operating flow rates). Finally, multivariate non-linear fitting models were fitted to measured data to capture functional variations in the CPMA width and loss factors. The resulting correlations were presented accordingly after passing the goodness of fit test.

In Chapter 5, the performance of the proposed inversion scheme was examined on real-world mixing states of BC aerosols caused by seasonal wildfires at a regional site (Kamloops, about 350 km northeast of Vancouver, British Columbia, Canada) and compared with widely-accepted measurement methods in measuring rBC mixing states. The rBC mixing states were characterized with a particular focus on the tandem CPMA-SP2 and SP2-only measurements, including LEO and lagtime analysis.

The measurement methods were compared in terms of: detection size range, temporal resolution, counting statistics, and errors associated with the measurements. The CPMA-SP2 measurement is less uncertain than the other two methods in the overall comparison. Moreover, because the CPMA-SP2 measurement system can measure coatings on much smaller particles, it has a much larger detection range. However, CPMA-SP2 measurements have a poorer temporal resolution than SP2-only measurements, which can provide real-time, but qualitative measurements.

6.2 Future Work

Based on this work, three future research projects could be developed. According to the comparison between LEO and CPMA-SP2 approaches using wildfire measurements, there may not be consensus on whether the significant lack of accuracy in LEO results is a structural issue that cannot be corrected, even if all inputs are assigned accurately. Comparing these two methods using measurements made in a controlled laboratory environment can provide a suitable platform for an accurate and thorough assessment. Accordingly, smog chamber experiments can be designed to simulate atmospheric processes of forming secondary organic aerosols (SOAs) through photochemical reactions in a controlled laboratory environment and yet still provide flexibility in varying input parameters of the study, while the physical characteristics (*i.e.,* density and refractive index) of rBC and coating components can be measured throughout the measurements. Second, the same procedure followed for CPMA-SP2 measurements could be used for other tandem instrument arrangements to derive optimal sampling configurations as a function of aerosol concentration and desired measurement time. For instance, the optimum sampling configuration can be derived

for the CPMA-DMA system, which is used to derive a two-dimensional distribution of particle mass and mobility diameter ($m\mathrm{p} - d_{\mathrm{m}}$) distributions. Lastly, the CPMA transfer function is a function of the mobility diameter, and varying effective particle density (often resulting from varying particle morphology) will affect the particle's equivalent mobility diameter. For BC aggregates, more specifically, accurate mobility is unknown, and it is common to use an empirical mass-mobility equation to convert the particle mass to a mobility diameter whenever the CPMA transfer function is evaluated. This assumption will impact the width and, in the case of diffusion, the amplitude of the CPMA transfer. Consequently, for non-spherical particles like BC aggregates, employing tandem CPMA measurements for characterizing the CPMA transfer function would be challenging. Instead, one might take advantage of SP2 in tandem with CPMA and use thermally denuded BC particles to characterize the CPMA transfer function for BC aggregates.

Bibliography

Abbatt, JPD, AKY Lee, and JA Thornton (2012). "Quantifying trace gas uptake to tropospheric aerosol: recent advances and remaining challenges". In: *Chemical Society Reviews* 41.19, pp. 6555–6581.

Afroughi, Mohammad Javad et al. (2019). "Properties of carbon black produced by the thermal decomposition of methane in the products of premixed flames". In: *Journal of Aerosol Science* 131, pp. 13–27.

Agarwal, Jugal K and Gilmore J Sem (1980). "Continuous flow, single-particle-counting condensation nucleus counter". In: *Journal of Aerosol Science* 11.4, pp. 343–357.

Amato, Umberto et al. (1995). "Objective algorithms for the aerosol problem". In: *Applied optics* 34.24, pp. 5442–5452.

Andreae, Meinrat O and A Gelencsér (2006). "Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols". In: *Atmospheric Chemistry and Physics* 6.10, pp. 3131–3148.

Andreae, MO and DJESR Rosenfeld (2008). "Aerosol–cloud–precipitation interactions. Part 1. The nature and sources of cloud-active aerosols". In: *Earth-Science Reviews* 89.1-2, pp. 13–41.

Bashurova, VS et al. (1991). "Determination of atmospheric aerosol size distribution functions from screen diffusion battery data: mathematical aspects". In: *Journal of aerosol science* 22.3, pp. 373–388.

Belge, Murat, Misha E Kilmer, and Eric L Miller (1998). "Simultaneous multiple regularization parameter selection by means of the L-hypersurface with applications to linear inverse problems posed in the wavelet transform domain". In: vol. 3459. SPIE, pp. 328–336.

Belge, Murat, Misha E Kilmer, and Eric L Miller (2002). "Efficient determination of multiple regularization parameters in a generalized L-curve framework". In: *Inverse problems* 18.4, p. 1161.

Bell, David M et al. (2017). "The properties and behavior of α-pinene secondary organic aerosol particles exposed to ammonia under dry conditions". In: *Physical Chemistry Chemical Physics* 19.9, pp. 6497–6507.

Bellouin, Nicolas et al. (2005). "Global estimate of aerosol direct radiative forcing from satellite measurements". In: *Nature* 438.7071, pp. 1138–1141.

Beranek, Josef, Dan Imre, and Alla Zelenyuk (2012). "Real-time shape-based particle separation and detailed in situ particle shape characterization". In: *Analytical chemistry* 84.3, pp. 1459–1465.

Bertsekas, DP and JN Tsitsiklis (2002). "The bivariate normal distribution". In: *Introduction to Probability*, pp. 247–253.

Birmili, WFADSLMO et al. (1997). "Determination of differential mobility analyzer transfer functions using identical instruments in series". In: *Aerosol Science and Technology* 27.2, pp. 215–223.

Birmili, Wolfram et al. (2007). "Particle penetration through a 300 m inlet pipe for sampling atmospheric aerosols from a tall meteorological tower". In: *Aerosol science and technology* 41.9, pp. 811–817.

Bond, Tami C and Robert W Bergstrom (2006). "Light absorption by carbonaceous particles: An investigative review". In: *Aerosol Science and Technology* 40.1, pp. 27–67.

Bond, Tami C et al. (2013). "Bounding the role of black carbon in the climate system: A scientific assessment". In: *Journal of geophysical research: Atmospheres* 118.11, pp. 5380–5552.

Bondy, Amy L et al. (2018). "The diverse chemical mixing state of aerosol particles in the southeastern United States". In: *Atmospheric chemistry and physics* 18.16, pp. 12595–12612.

Brauer, Michael et al. (2016). "Ambient air pollution exposure estimation for the global burden of disease 2013". In: *Environmental science & technology* 50.1, pp. 79–88.

Broda, KN et al. (2018). "A novel inversion method to determine the mass distribution of non-refractory coatings on refractory black carbon using a centrifugal particle mass analyzer and single particle soot photometer". In: *Aerosol Science and Technology* 52.5, pp. 567–578.

Buckley, David T et al. (2017). "A corrected two dimensional data inversion routine for tandem mobility-mass measurements". In: *Journal of Aerosol Science* 114, pp. 157–168.

Burnett, Richard T et al. (2014). "An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure". In: *Environmental health perspectives* 122.4, pp. 397–403.

Calvetti, Daniela, Gene Howard Golub, and Lothar Reichel (1999). "Estimation of the L-curve via Lanczos bidiagonalization". In: *BIT Numerical Mathematics* 39.4, pp. 603–619.

Cappa, Christopher D et al. (2012). "Radiative absorption enhancements due to the mixing state of atmospheric black carbon". In: *Science* 337.6098, pp. 1078–1081.

Carfora, MF, F Esposito, and C Serio (1998). "Numerical methods for retrieving aerosol size distributions from optical measurements of solar radiation". In: *Journal of aerosol science* 29.10, pp. 1225–1236.

Cheng, YF et al. (2012). "Size-resolved measurement of the mixing state of soot in the megacity Beijing, China: diurnal cycle, aging and parameterization". In: *Atmospheric Chemistry and Physics* 12.10, pp. 4477–4491.

Collins, Don R et al. (2004). "The scanning DMA transfer function". In: *Aerosol Science and Technology* 38.8, pp. 833–850.

Corbin, Joel C et al. (2018). "Brown and black carbon emitted by a marine engine operated on heavy fuel oil and distillate fuels: optical properties, size distributions, and emission factors". In: *Journal of Geophysical Research: Atmospheres* 123.11, pp. 6175–6195.

Crump, James G and John H Seinfeld (1982). "A new algorithm for inversion of aerosol size distribution data". In: *Aerosol Science and Technology* 1.1, pp. 15–34.

Cultrera, Alessandro and Luca Callegaro (2016). "A simple algorithm to find the L-curve corner in the regularization of inverse problems". In: *CoRR*.

Dahlkötter, Florian et al. (2014). "The Pagami Creek smoke plume after long-range transport to the upper troposphere over Europe–aerosol properties and black carbon mixing state". In: *Atmospheric Chemistry and Physics* 14.12, pp. 6111–6137.

Dastanpour, Ramin et al. (2017). "Variation of the optical properties of soot as a function of particle mass". In: *Carbon* 124, pp. 201–211.

DeCarlo, Peter F et al. (2004). "Particle morphology and density characterization by combined mobility and aerodynamic diameter measurements. Part 1: Theory". In: *Aerosol Science and Technology* 38.12, pp. 1185–1205.

DeMott, Paul J et al. (2010). "Predicting global atmospheric ice nuclei distributions and their impacts on climate". In: *Proceedings of the National Academy of Sciences* 107.25, pp. 11217–11222.

DeMott, Paul J et al. (2011). "Resurgence in ice nuclei measurement research". In: *Bulletin of the American Meteorological Society* 92.12, pp. 1623–1635.

Ditas, Jeannine et al. (2018). "Strong impact of wildfires on the abundance and aging of black carbon in the lowermost stratosphere". In: *Proceedings of the National Academy of Sciences* 115.50, E11595–E11603.

Drakaki, Eleni, Clio Dessinioti, and Christina V Antoniou (2014). "Air pollution and the skin". In: *Frontiers in Environmental Science* 2, p. 11.

Dumouchel, William, Fanny O'brien, et al. (1989). "Integrating a robust option into a multiple regression computing environment". In: American Statistical Association Alexandria, pp. 297–302.

Elfving, Tommy (1980). "On some methods for entropy maximization and matrix scaling". In: *Linear algebra and its applications* 34, pp. 321–339.

Farmer, Delphine K, Christopher D Cappa, and Sonia M Kreidenweis (2015). "Atmospheric processes and their controlling influence on cloud condensation nuclei activity". In: *Chemical Reviews* 115.10, pp. 4199–4217.

Fissan, HDFPSDD et al. (1996). "Experimental comparison of four differential mobility analyzers for nanometer aerosol measurements". In: *Aerosol Science and Technology* 24.1, pp. 1–13.

Foi, Alessandro et al. (2008). "Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data". In: *IEEE Transactions on Image Processing* 17.10, pp. 1737–1754.

Gao, RS et al. (2007). "A novel method for estimating light-scattering properties of soot aerosols using a modified single-particle soot photometer". In: *Aerosol Science and Technology* 41.2, pp. 125–135.

Ghazi, Rouzbeh et al. (2013). "Mass, mobility, volatility, and morphology of soot particles generated by a McKenna and inverted burner". In: *Aerosol Science and Technology* 47.4, pp. 395–405.

Girotto, Giulia et al. (2018). "Fractal-like tar ball aggregates from wildfire smoke". In: *Environmental Science & Technology Letters* 5.6, pp. 360–365.

Gong, Xianda et al. (2016). "Size distribution and mixing state of black carbon particles during a heavy air pollution episode in Shanghai". In: *Atmospheric Chemistry and Physics* 16.8, pp. 5399–5411.

Gordon, Richard, Robert Bender, and Gabor T Herman (1970). "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography". In: *Journal of theoretical Biology* 29.3, pp. 471–481.

Graves, Brian et al. (2015). "Characterization of particulate matter morphology and volatility from a compression-ignition natural-gas direct-injection engine". In: *Aerosol Science and Technology* 49.8, pp. 589–598.

Graves, Brian et al. (2020). "Plasma production of nanomaterials for energy storage: continuous gas-phase synthesis of metal oxide CNT materials via a microwave plasma". In: *Nanoscale* 12.8, pp. 5196–5208.

Graves, Brian Mackenzie, Charles Robert Koch, and Jason Scott Olfert (2017). "Morphology and volatility of particulate matter emitted from a gasoline direct injection engine fuelled on gasoline and ethanol blends". In: *Journal of Aerosol Science* 105, pp. 166–178.

Hallquist, Mattias et al. (2009). "The formation, properties and impact of secondary organic aerosol: current and emerging issues". In: *Atmospheric chemistry and physics* 9.14, pp. 5155–5236.

He, Cenlin et al. (2014). "Black carbon radiative forcing over the Tibetan Plateau". In: *Geophysical Research Letters* 41.22, pp. 7806–7813.

He, Cenlin et al. (2016). "Microphysics-based black carbon aging in a global CTM: constraints from HIPPO observations and implications for global black carbon budget". In: *Atmospheric Chemistry and Physics* 16.5, pp. 3077–3098.

Healy, Robert M et al. (2014). "Single particle diversity and mixing state measurements". In: *Atmospheric Chemistry and Physics* 14.12, pp. 6289–6299.

Holland, Paul W and Roy E Welsch (1977). "Robust regression using iteratively reweighted least-squares". In: *Communications in Statistics-theory and Methods* 6.9, pp. 813–827.

Houghton, John T et al. (1995). *Climate change 1994: radiative forcing of climate change and an evaluation of the IPCC 1992 IS92 emission scenarios*. Cambridge University Press.

Houghton, John Theodore (1992). *Climate change 1992*.

Houghton, John Theodore, Geoffrey J Jenkins, and Jim J Ephraums (1990). "Climate change: the IPCC scientific assessment". In: *American Scientist;(United States)* 80.6.

Hu, Dawei et al. (2021a). "Physical and chemical properties of black carbon and organic matter from different sources using aerodynamic aerosol classification". In: *Atmos. Chem. Phys* 21.21, pp. 1–50.

Hu, Dawei et al. (2021b). "Vertical profile of particle hygroscopicity and CCN effectiveness during winter in Beijing: insight into the hygroscopicity transition threshold of black carbon". In: *Faraday Discussions* 226, pp. 239–254.

Hummes, Detlef et al. (1996). "Experimental determination of the transfer function of a differential mobility analyzer (DMA) in the nanometer size range". In: *Particle & particle systems characterization* 13.5, pp. 327–332.

Irwin, M et al. (2013). "Evaluation of a Heated-Inlet for Calibration of the SP2". In: *Aerosol Science and Technology* 47.8, pp. 895–905.

Jacobson, Mark Z (2001). "Strong radiative heating due to the mixing state of black carbon in atmospheric aerosols". In: *Nature* 409.6821, pp. 695–697.

Jacobson, Mark Z (2006). "Effects of externally-through-internally-mixed soot inclusions within clouds and precipitation on global climate". In: *The Journal of Physical Chemistry A* 110.21, pp. 6860–6873.

Jayne, John T et al. (2000). "Development of an aerosol mass spectrometer for size and composition analysis of submicron particles". In: *Aerosol Science & Technology* 33.1-2, pp. 49–70.

Johnson, Tyler J et al. (2014). "Steady-state measurement of the effective particle density of cigarette smoke". In: *Journal of Aerosol Science* 75, pp. 9–16.

Johnson, Tyler J et al. (2015a). "Hygroscopic effects on the mobility and mass of cigarette smoke particles". In: *Journal of Aerosol Science* 86, pp. 69–78.

Johnson, Tyler J et al. (2015b). "Transient measurement of the effective particle density of cigarette smoke". In: *Journal of Aerosol Science* 87, pp. 63–74.

Johnson, Tyler J et al. (2018). "Measuring aerosol size distributions with the aerodynamic aerosol classifier". In: *Aerosol Science and Technology* 52.6, pp. 655–665.

Junge, Christian (1952). *Die konstitution des atmosphärischen aerosols*. Selbstverl. d. Meteorolog. Amtes f. Nordwestdtschl.

Kandlikar, Milind and Gurumurthy Ramachandran (1999). "Inverse methods for analysing aerosol spectrometer measurements: a critical review". In: *Journal of Aerosol Science* 30.4, pp. 413–437.

Karlsson, Martin NA and Bengt G Martinsson (2003). "Methods to measure and predict the transfer function size dependence of individual DMAs". In: *Journal of Aerosol Science* 34.5, pp. 603–625.

Kazemimanesh, Mohsen et al. (2019). "A novel miniature inverted-flame burner for the generation of soot nanoparticles". In: *Aerosol Science and Technology* 53.2, pp. 184–195.

Kazemimanesh, Mohsen et al. (2021). "Particulate emissions from turbulent diffusion flames with entrained droplets: A laboratory simulation of gas flaring emissions". In: *Journal of Aerosol Science* 157, p. 105807.

Kelly, WP and Peter H McMurry (1992). "Measurement of particle density by inertial classification of differential mobility analyzer–generated monodisperse aerosols". In: *Aerosol science and technology* 17.3, pp. 199–212.

Khalizov, Alexei F, Miguel Cruz-Quinones, and Renyi Zhang (2010). "Heterogeneous reaction of NO2 on fresh and coated soot surfaces". In: *The Journal of Physical Chemistry A* 114.28, pp. 7516–7524.

Khalizov, Alexei F et al. (2013). "Role of OH-initiated oxidation of isoprene in aging of combustion soot". In: *Environmental science & technology* 47.5, pp. 2254–2263.

Knutson, EO and KT Whitby (1975). "Aerosol classification by electric mobility: apparatus, theory, and applications". In: *Journal of Aerosol Science* 6.6, pp. 443–451.

Kokhanovsky, Alexander (2019). *Springer Series in Light Scattering: Volume 4: Light Scattering and Radiative Transfer*. Springer.

Kulmala, Markku et al. (2012). "Measurement of the nucleation of atmospheric aerosol particles". In: *Nature protocols* 7.9, pp. 1651–1667.

Kuwata, M and Y Kondo (2009). "Measurements of particle masses of inorganic salt particles for calibration of cloud condensation nuclei counters". In: *Atmospheric Chemistry and Physics* 9.16, pp. 5921–5932.

Kuwata, Mikinori (2015). "Particle classification by the tandem differential mobility analyzer–particle mass analyzer system". In: *Aerosol Science and Technology* 49.7, pp. 508–520.

Laborde, M et al. (2012). "Sensitivity of the Single Particle Soot Photometer to different black carbon types". In: *Atmospheric Measurement Techniques* 5.5, pp. 1031–1043.

Laborde, M et al. (2013). "Black carbon physical properties and mixing state in the European megacity Paris". In: *Atmospheric chemistry and physics* 13.11, pp. 5831–5856.

Lamb, Kara D (2019). "Classification of iron oxide aerosols by a single particle soot photometer using supervised machine learning". In: *Atmospheric Measurement Techniques* 12.7, pp. 3885–3906.

Lawson, Charles L and Richard J Hanson (1995). *Solving least squares problems*. SIAM.

Lee, AKY et al. (2015). "Mixing state of carbonaceous aerosol in an urban environment: single particle characterization using the soot particle aerosol mass spectrometer (SP-AMS)". In: *Atmospheric Chemistry and Physics* 15.4, pp. 1823–1841.

Lee, Wei-Liang et al. (2017). "Impact of absorbing aerosol deposition on snow albedo reduction over the southern Tibetan plateau based on satellite observations". In: *Theoretical and Applied Climatology* 129.3, pp. 1373–1382.

Leighton, TG (1994). "The Acoustic Bubble.— Academic". In: *Press, London*, pp. 234–243.

Lent, A (1976). "Maximum entropy and multiplicative ART". In.

Lesnic, D, L Elliott, and DB Ingham (1995). "The inverse determination of the particle size distribution from diffusion battery data". In: *Journal of Aerosol Science* 8.26, p. 1319.

Li, Chunlin et al. (2016). "Physiochemical properties of carbonaceous aerosol from agricultural residue burning: Density, volatility, and hygroscopicity". In: *Atmospheric Environment* 140, pp. 94–105.

Li, Weiling, Lin Li, and Da-Ren Chen (2006). "A new deconvolution scheme for the retrieval of true DMA transfer function from tandem DMA data". In: *Aerosol science and technology* 40.12, pp. 1052–1057.

Liou, KN et al. (2014). "Stochastic parameterization for light absorption by internally mixed BC/dust in snow grains for application to climate models". In: *Journal of Geophysical Research: Atmospheres* 119.12, pp. 7616–7632.

Liu, Benjamin YH and David YH Pui (1974). "A submicron aerosol standard and the primary, absolute calibration of the condensation nuclei counter". In: *Journal of Colloid and Interface Science* 47.1, pp. 155–171.

Liu, Dantong et al. (2015). "The effect of complex black carbon microphysics on the determination of the optical properties of brown carbon". In: *Geophysical Research Letters* 42.2, pp. 613–619.

Liu, Dantong et al. (2017). "Black-carbon absorption enhancement in the atmosphere determined by particle mixing state". In: *Nature Geoscience* 10.3, pp. 184–188.

Liu, Hang et al. (2019). "Effective densities of soot particles and their relationships with the mixing state at an urban site in the Beijing megacity in the winter of 2018". In: *Atmospheric Chemistry and Physics* 19.23, pp. 14791–14804.

Liu, Hang et al. (2020). "Mixing characteristics of refractory black carbon aerosols at an urban site in Beijing". In: *Atmospheric Chemistry and Physics* 20.9, pp. 5771–5785.

Liu, Hang et al. (2022). "Mixing characteristics of black carbon aerosols in a coastal city using the CPMA-SP2 system". In: *Atmospheric Research* 265, p. 105867.

Markowski, Gregory R (1987). "Improving Twomey's algorithm for inversion of aerosol measurement data". In: *Aerosol science and technology* 7.2, pp. 127–141.

Marsden, Nicholas A et al. (2018). "Online differentiation of mineral phase in aerosol particles by ion formation mechanism using a LAAP-TOF single-particle mass spectrometer". In: *Atmospheric Measurement Techniques* 11.1, pp. 195–213.

Martinsson, Bengt G, Martin NA Karlsson, and Göran Frank (2001). "Methodology to estimate the transfer function of individual differential mobility analyzers". In: *Aerosol Science & Technology* 35.4, pp. 815–823.

McMeeking, GR et al. (2010). "Black carbon measurements in the boundary layer over western and northern Europe". In: *Atmospheric Chemistry and Physics* 10.19, pp. 9393–9414.

Miljevic, Branka et al. (2012). "Restructuring of carbonaceous particles upon exposure to organic and water vapours". In: *Journal of Aerosol Science* 47, pp. 48–57.

Moallemi, Alireza et al. (2019). "Characterization of black carbon particles generated by a propane-fueled miniature inverted soot generator". In: *Journal of Aerosol Science* 135, pp. 46–57.

Momenimovahed, A and JS Olfert (2015). "Effective density and volatility of particles emitted from gasoline direct injection vehicles and implications for particle mass measurement". In: *Aerosol Science and Technology* 49.11, pp. 1051–1062.

Moteki, Nobuhiro and Yutaka Kondo (2007). "Effects of mixing state on black carbon measurements by laser-induced incandescence". In: *Aerosol Science and Technology* 41.4, pp. 398–417.

Moteki, Nobuhiro and Yutaka Kondo (2008). "Method to measure time-dependent scattering cross sections of particles evaporating in a laser beam". In: *Journal of aerosol science* 39.4, pp. 348–364.

Moteki, Nobuhiro, Yutaka Kondo, and Kouji Adachi (2014). "Identification by single-particle soot photometer of black carbon particles attached to other particles: Experimental basis and ambient measurements". In: p. 2529.

Moteki, Nobuhiro, Yutaka Kondo, and Shin-ichi Nakamura (2010). "Method to measure refractive indices of small nonspherical particles: Application to black carbon particles". In: *Journal of Aerosol Science* 41.5, pp. 513–521.

Murphy, DM and DS Thomson (1997). "Chemical composition of single aerosol particles at Idaho Hill: Negative ion measurements". In: *Journal of Geophysical Research: Atmospheres* 102.D5, pp. 6353–6368.

Nakao, Shunsuke et al. (2013). "Density and elemental ratios of secondary organic aerosol: Application of a density prediction method". In: *Atmospheric Environment* 68, pp. 273–277.

Nakayama, Tomoki et al. (2010). "Laboratory studies on optical properties of secondary organic aerosols generated during the photooxidation of toluene and the ozonolysis of α-pinene". In: *Journal of Geophysical Research: Atmospheres* 115.D24.

Naseri, A et al. (2021a). "An improved inversion method for determining two-dimensional mass distributions of non-refractory materials on refractory black carbon". In: *Aerosol Science and Technology* 55.1, pp. 104–118.

Naseri, A et al. (2021b). "Optimized instrument configurations for tandem particle mass analyzer and single particle-soot photometer experiments". In: *Journal of Aerosol Science*, p. 105897.

Nilsson, Patrik T et al. (2015). "In-situ characterization of metal nanoparticles and their organic coatings using laser-vaporization aerosol mass spectrometry". In: *Nano Research* 8.12, pp. 3780–3795.

O'Brien, Rachel E et al. (2015). "Chemical imaging of ambient aerosol particles: Observational constraints on mixing state parameterization". In: *Journal of Geophysical Research: Atmospheres* 120.18, pp. 9591–9605.

Olfert, Jason and Steven Rogak (2019). "Universal relations between soot effective density and primary particle size for common combustion sources". In: *Aerosol Science and Technology* 53.5, pp. 485–492.

Olfert, Jason S et al. (2017). "Effective density and volatility of particles sampled from a helicopter gas turbine engine". In: *Aerosol Science and Technology* 51.6, pp. 704–714.

Olfert, JS (2005). "A numerical calculation of the transfer function of the fluted centrifugal particle mass analyzer". In: *Aerosol science and technology* 39.10, pp. 1002–1009.

Olfert, JS and Nick Collings (2005). "New method for particle mass classification—the Couette centrifugal particle mass analyzer". In: *Journal of Aerosol Science* 36.11, pp. 1338–1352.

Olfert, JS, JPR Symonds, and N Collings (2007). "The effective density and fractal dimension of particles emitted from a light-duty diesel vehicle with a diesel oxidation catalyst". In: *Journal of Aerosol Science* 38.1, pp. 69–82.

Olfert, JS et al. (2006). "The experimental transfer function of the Couette centrifugal particle mass analyzer". In: *Journal of Aerosol Science* 37.12, pp. 1840–1852.

Onasch, TB et al. (2012). "Soot particle aerosol mass spectrometer: development, validation, and initial application". In: *Aerosol Science and Technology* 46.7, pp. 804–817.

Oudin, Anna et al. (2016). "Association between neighbourhood air pollution concentrations and dispensed medication for psychiatric disorders in a large longitudinal cohort of Swedish children and adolescents". In: *BMJ open* 6.6, e010004.

Ouf, F-X et al. (2019). "True density of combustion emitted particles: A comparison of results highlighting the influence of the organic contents". In: *Journal of Aerosol Science* 134, pp. 1–13.

Pachauri, Rajendra K et al. (2014). *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Ipcc.

Park, Kihong et al. (2004). "Measurement of inherent material density of nanoparticle agglomerates". In: *Journal of Nanoparticle Research* 6.2, pp. 267–272.

Penner, Joyce E, Xiquan Dong, and Yang Chen (2004). "Observational evidence of a change in radiative forcing due to the indirect aerosol effect". In: *Nature* 427.6971, pp. 231–234.

Petzold, Andreas et al. (2013). "Recommendations for reporting" black carbon" measurements". In: *Atmospheric Chemistry and Physics* 13.16, pp. 8365–8379.

Prather, Kimberly A, Courtney D Hatch, and Vicki H Grassian (2008). "Analysis of atmospheric aerosols". In: *Annual Review of Analytical Chemistry* 1.1, p. 485.

Qi, Ling et al. (2017). "Effects of the Wegener–Bergeron–Findeisen process on global black carbon distribution". In: *Atmospheric Chemistry and Physics* 17.12, pp. 7459–7479.

Quant, FR et al. (1992). "Performance of condensation particle counters with three continuous-flow designs". In: *Journal of Aerosol Science* 23, pp. 405–408.

Quiros, David C et al. (2015). "Particle effective density and mass during steady-state operation of GDI, PFI, and diesel passenger cars". In: *Journal of Aerosol Science* 83, pp. 39–54.

Raatikainen, Tomi et al. (2017). "Size-selected black carbon mass distributions and mixing state in polluted and clean environments of northern India". In: *Atmospheric Chemistry and Physics* 17.1, pp. 371–383.

Ramanathan, Veerabhadran and Gregory Carmichael (2008). "Global and regional climate changes due to black carbon". In: *Nature geoscience* 1.4, pp. 221–227.

Ravishankara, AR, Yinon Rudich, and DJ Wuebbles (2015). "Physical chemistry of climate metrics". In: *Chemical reviews* 115.10, pp. 3682–3703.

Rawat, Vivek K et al. (2016). "Two dimensional size–mass distribution function inversion from differential mobility analyzer–aerosol particle mass analyzer (DMA–APM) measurements". In: *Journal of Aerosol Science* 92, pp. 70–82.

Reddington, Carly L et al. (2013). "The mass and number size distributions of black carbon aerosol over Europe". In: *Atmospheric Chemistry and Physics* 13.9, pp. 4917–4939.

Riccobono, Francesco et al. (2014). "Oxidation products of biogenic emissions contribute to nucleation of atmospheric particles". In: *Science* 344.6185, pp. 717–721.

Robinson, Dorothy L (2020). "Accurate, low cost PM2. 5 measurements demonstrate the large spatial variation in wood smoke pollution in regional Australia and improve modeling and estimates of health costs". In: *Atmosphere* 11.8, p. 856.

Sakai, Stephen and David Rothamer (2017). "Effect of ethanol blending on particulate formation from premixed combustion in spark-ignition engines". In: *Fuel* 196, pp. 154–168.

Schill, Gregory P et al. (2018). "Use of the single particle soot photometer (SP2) as a pre-filter for ice nucleation measurements: Effect of particle mixing state and determination of SP2 conditions to fully vaporize refractory black carbon". In: *Atmospheric Measurement Techniques* 11.5, pp. 3007–3020.

Schwarz, Joshua P et al. (2006). "Single-particle measurements of midlatitude black carbon and light-scattering aerosols from the boundary layer to the lower stratosphere". In: *Journal of Geophysical Research: Atmospheres* 111.D16.

Schwarz, JP et al. (2010). "The detection efficiency of the single particle soot photometer". In: *Aerosol science and technology* 44.8, pp. 612–628.

Sedlacek III, Arthur J et al. (2012). "Determination of and evidence for non-core-shell structure of particles containing black carbon using the Single-Particle Soot Photometer (SP2)". In: *Geophysical Research Letters* 39.6.

Sedlacek III, Arthur J et al. (2015). "Investigation of refractory black carbon-containing particle morphologies using the single-particle soot photometer (SP2)". In: *Aerosol Science and Technology* 49.10, pp. 872–885.

Sedlacek III, Arthur J et al. (2018). "Formation of refractory black carbon by SP2-induced charring of organic aerosol". In: *Aerosol Science and Technology* 52.12, pp. 1345–1350.

Seinfeld, John H and Spyros N Pandis (1998). "From air pollution to climate change". In: *Atmospheric chemistry and physics* 1326.

Shah, Sandip D and David R Cocker (2005). "A fast scanning mobility particle spectrometer for monitoring transient particle size distributions". In: *Aerosol Science and Technology* 39.6, pp. 519–526.

Shapiro, M et al. (2012). "Characterization of agglomerates by simultaneous measurement of mobility, vacuum aerodynamic diameter and mass". In: *Journal of Aerosol Science* 44, pp. 24–45.

Sharma, Sangeeta et al. (2017). "An evaluation of three methods for measuring black carbon in Alert, Canada". In: *Atmospheric Chemistry and Physics* 17.24, pp. 15225–15243.

Sipkens, TA, JS Olfert, and SN Rogak (2020a). "Inversion methods to determine two-dimensional aerosol mass-mobility distributions: A critical comparison of established methods". In: *Journal of Aerosol Science* 140, p. 105484.

Sipkens, TA, JS Olfert, and SN Rogak (2020b). "Inversion methods to determine two-dimensional aerosol mass-mobility distributions II: Existing and novel Bayesian methods". In: *Journal of Aerosol Science* 146, p. 105565.

Sipkens, Timothy A, Jason S Olfert, and Steven N Rogak (2020c). "New approaches to calculate the transfer function of particle mass analyzers". In: *Aerosol Science and Technology* 54.1, pp. 111–127.

Sipkens, Timothy A et al. (2017). "General error model for analysis of laser-induced incandescence signals". In: *Applied optics* 56.30, pp. 8436–8445.

Sipkens, Timothy A et al. (2021). "Using two-dimensional distributions to inform the mixing state of soot and salt particles produced in gas flares". In: *Journal of Aerosol Science*, p. 105826.

Sorensen, Christopher M et al. (2011). "Optical measurement techniques: fundamentals and applications". In: *Aerosol measurement: Principles, techniques, and applications*, pp. 269–312.

Sram, R (1999). *Impact of air pollution on reproductive health.*

Stephens, Michelle, Nelson Turner, and Jon Sandberg (2003). "Particle identification by laser-induced incandescence in a solid-state laser cavity". In: *Applied Optics* 42.19, pp. 3726–3736.

Stratmann, FTDH et al. (1997). "Differential electrical mobility analysis: A theoretical study". In: *Aerosol Science and Technology* 26.4, pp. 368–383.

Subramanian, R et al. (2010). "Black carbon over Mexico: the effect of atmospheric transport on mixing state, mass absorption cross-section, and BC/CO ratios". In: *Atmospheric Chemistry and Physics* 10.1, pp. 219–237.

Swanson, Jacob and David Kittelson (2010). "Evaluation of thermal denuder and catalytic stripper methods for solid particle measurements". In: *Journal of Aerosol Science* 41.12, pp. 1113–1122.

Talukdar, Suddha S and Mark T Swihart (2003). "An improved data inversion program for obtaining aerosol size distributions from scanning differential mobility analyzer data". In: *Aerosol Science and Technology* 37.2, pp. 145–161.

Tavakoli, F and JS Olfert (2013). "An instrument for the classification of aerosols by particle relaxation time: theoretical models of the aerodynamic aerosol classifier". In: *Aerosol Science and Technology* 47.8, pp. 916–926.

Taylor, JW et al. (2015). "Assessment of the sensitivity of core/shell parameters derived using the single-particle soot photometer to density and refractive index". In: *Atmospheric Measurement Techniques* 8.4, pp. 1701–1718.

Thompson, Alan M and Jim Kay (1993). "On some Bayesian choices of regularization parameter in image restoration". In: *Inverse problems* 9.6, p. 749.

Toner, Stephen M et al. (2008). "Using mass spectral source signatures to apportion exhaust particles from gasoline and diesel powered vehicles in a freeway study using UF-ATOFMS". In: *Atmospheric Environment* 42.3, pp. 568–581.

Turpin, BJ and HJ Lim (2001). "Contributions to PM2. 5 mass concentrations: Revisiting common assumptions for estimating organic mass". In: *Aerosol Sci. Technol* 35, pp. 602–610.

Twomey, S (1975). "Comparison of constrained linear inversion and an iterative nonlinear algorithm applied to the indirect estimation of particle size distributions". In: *Journal of Computational Physics* 18.2, pp. 188–200.

Ubogu, Emamode A et al. (2018). "Determination of the effective density and fractal dimension of PM emissions from an aircraft auxiliary power unit". In: *Journal of Environmental Sciences* 74, pp. 11–18.

Voutilainen, A, V Kolehmainen, and JP Kaipio (2001). "Statistical inversion of aerosol size measurement data". In: *Inverse Problems in Engineering* 9.1, pp. 67–94.

Wallace, Lance (2006). "Indoor sources of ultrafine and accumulation mode particles: size distributions, size-resolved concentrations, and source strengths". In: *Aerosol science and technology* 40.5, pp. 348–360.

Wang, Qiyuan et al. (2014a). "Black carbon aerosol characterization in a remote area of Qinghai–Tibetan Plateau, western China". In: *Science of the Total Environment* 479, pp. 151–158.

Wang, Qiyuan et al. (2014b). "Mixing state of black carbon aerosol in a heavily polluted urban area of China: implications for light absorption enhancement". In: *Aerosol Science and Technology* 48.7, pp. 689–697.

Wang, Shih Chen and Richard C Flagan (1990). "Scanning electrical mobility spectrometer". In: *Aerosol Science and Technology* 13.2, pp. 230–240.

Wang, Yanfei (2008). "An efficient gradient method for maximum entropy regularizing retrieval of atmospheric aerosol particle size distribution function". In: *Journal of Aerosol Science* 39.4, pp. 305–322.

Wang, Yanfei et al. (2006). "Regularized inversion method for retrieval of aerosol particle size distribution function in W 1, 2 space". In: *Applied optics* 45.28, pp. 7456–7467.

Weber, RJ et al. (1996). "Measured atmospheric new particle formation rates: Implications for nucleation mechanisms". In: *Chemical Engineering Communications* 151.1, pp. 53–64.

Weingartner, E, H Burtscher, and U Baltensperger (1997). "Hygroscopic properties of carbon and diesel soot particles". In: *Atmospheric Environment* 31.15, pp. 2311–2327.

Wiedensohler, A (1988). "An approximation of the bipolar charge distribution for particles in the submicron size range". In: *Journal of Aerosol Science* 19.3, pp. 387–389.

Wiedensohler, AEMC et al. (1986). "Investigation of the bipolar charge distribution at various gas conditions". In: *Journal of aerosol science* 17.3, pp. 413–416.

Willis, Matthew (2000). "Algebraic reconstruction algorithms for remote sensing image enhancement". PhD thesis. Brigham Young University. Department of Electrical and Computer Engineering.

Winkler, Peter (1973). "The growth of atmospheric aerosol particles as a function of the relative humidity—II. An improved concept of mixed nuclei". In: *Journal of Aerosol Science* 4.5, pp. 373–387.

Wirgin, Armand (2004). "The inverse crime". In: *arXiv preprint math-ph/0401050*.

Wolfenbarger, J Kenneth and John H Seinfeld (1990). "Inversion of aerosol size distribution data". In: *Journal of Aerosol Science* 21.2, pp. 227–247.

Wu, Yunfei et al. (2016). "Effect of ambient humidity on the light absorption amplification of black carbon in Beijing during January 2013". In: *Atmospheric Environment* 124, pp. 217–223.

Wu, Yunfei et al. (2017). "Size distribution and source of black carbon aerosol in urban Beijing during winter haze episodes". In: *Atmospheric Chemistry and Physics* 17.12, pp. 7965–7975.

Wu, Yunfei et al. (2021). "Size-resolved refractive index of scattering aerosols in urban Beijing: A seasonal comparison". In: *Aerosol Science and Technology* 55.9, pp. 1070–1083.

Yee, E (1989). "On the interpretation of diffusion battery data". In: *Journal of aerosol science* 20.7, pp. 797–811.

Yu, Chenjie et al. (2020). "Characterising mass-resolved mixing state of black carbon in Beijing using a morphology-independent measurement method". In: *Atmospheric Chemistry and Physics* 20.6, pp. 3645–3661.

Yuan, Jinfeng et al. (2021). "Variability in the mass absorption cross section of black carbon (BC) aerosols is driven by BC internal mixing state at a central European background site (Melpitz, Germany) in winter". In: *Atmospheric Chemistry and Physics* 21.2, pp. 635–655.

Yuan, Qi et al. (2020). "Evidence for large amounts of brown carbonaceous tarballs in the Himalayan atmosphere". In: *Environmental Science & Technology Letters* 8.1, pp. 16–23.

Zanatta, Marco et al. (2018). "Effects of mixing state on optical and radiative properties of black carbon in the European Arctic". In: *Atmospheric Chemistry and Physics* 18.19, pp. 14037–14057.

Zhang, Yuxuan et al. (2016). "Measuring the morphology and density of internally mixed black carbon with SP2 and VTDMA: new insight into the absorption

enhancement of black carbon in the atmosphere". In: *Atmospheric Measurement Techniques* 9.4, pp. 1833–1843.

Zhang, Yuxuan et al. (2018a). "Amplification of light absorption of black carbon associated with air pollution". In: *Atmospheric Chemistry and Physics* 18.13, pp. 9879–9896.

Zhang, Yuxuan et al. (2018b). "Reduction in black carbon light absorption due to multi-pollutant emission control during APEC China 2014". In: *Atmospheric Chemistry and Physics* 18.14, pp. 10275–10287.

Zhang, Yuxuan et al. (2018c). "Sizing of Ambient Particles From a Single-Particle Soot Photometer Measurement to Retrieve Mixing State of Black Carbon at a Regional Site of the North China Plain". In: *Journal of Geophysical Research: Atmospheres* 123.22, pp. 12–778.

Zhu, Yifang et al. (2002). "Study of ultrafine particles near a major highway with heavy-duty diesel traffic". In: *Atmospheric environment* 36.27, pp. 4323–4335.

Zikova, N and V Zdimal (2012). "Relative humidity, precipitation and fog influence on SMPS spectra". In: p. 10616.

Appendix A: Inversion Schemes

A.1 Iterative regularization methods

Iterative regularization methods have been broadly employed in the aerosol literature for a one dimensional aerosol inversion problem, *,e.g.,* Twomey, 1975. Rawat et al. (2016) developed and employed the improved version of the Twomey method to obtain two-dimensional mass-mobility distributions, and Broda et al. (2018) applied this method to derive the two-dimensional mass distribution of non-refractory coatings on refractory black carbon.

An initial guess is required for the initiation of iterative regularization methods like the Twomey-type algorithms or MART regularization methods. Here in this paper, the methodology introduced by Broda et al. (2018) is modified and employed to obtain an adequate initial guess. The initial guess is acquired by dividing the CPMA-SP2 data by the analytical solution to the integral of a triangular CPMA transfer function ($\tilde{\beta}$) and by the charge fraction ($\bar{f}$) of singly charged particles ($\Phi = 1$);

$$\mathbf{x}^0 = \frac{\mathbf{b}}{\bar{f}\left(m_\mathrm{p}, \Phi = 1\right)\tilde{\beta}\left(m_\mathrm{p}, \Phi = 1\right)} \tag{A.1}$$

The fabricated data $\mathbf{b}$ derived from Phantom 1 to 5 are depicted in Figure A.1.

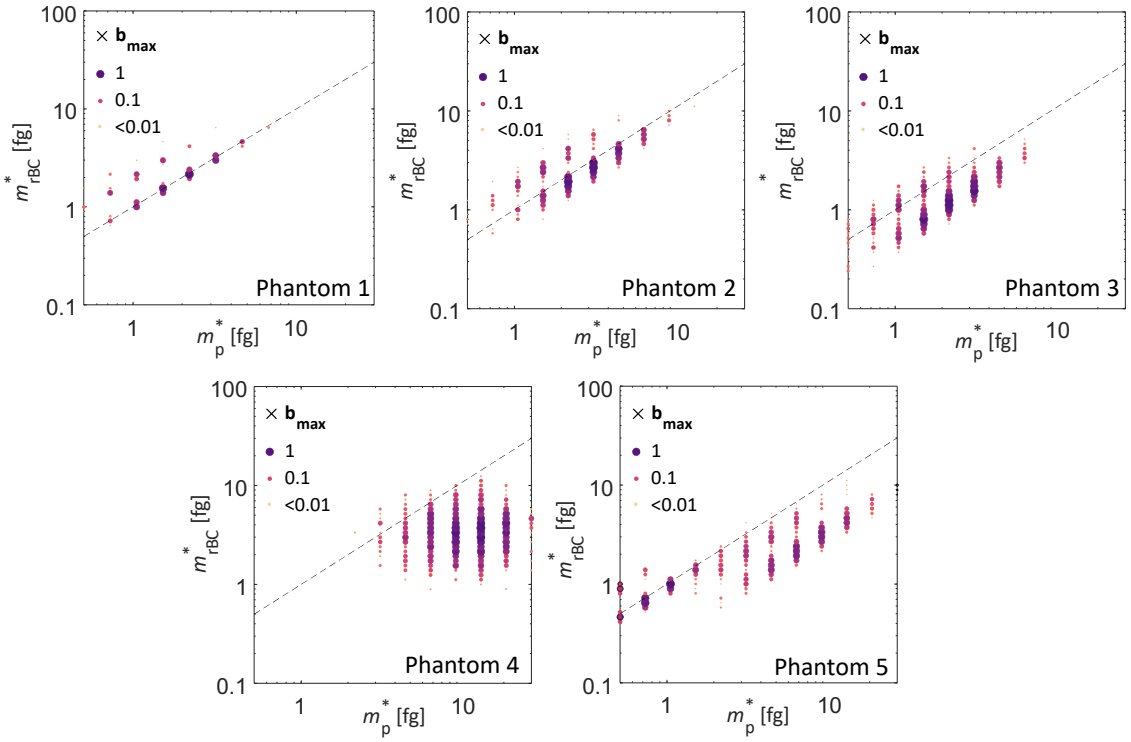

Figure A.1: Normalized synthetic data, generated by implementing the forward approach on each of the phantoms and corrupting with Poisson-Gaussian noise. Dots are located at the intersection of the CPMA and SP2 setpoints, while the size and color indicate the relative number of particles measured at that location. Multiple modes correspond to the multiple integer charge states ($\Phi = 1$, 2, or 3) considered in the present paper. The dotted-dash lines represent the location where m_{rBC} equals m_{p}.

Additionally, to have the initial guess consistent with the dimension $\mathbf{x}$, linear interpolation from the quantity calculated by the Equation (A.1) is required (Buckley et al., 2017). Also, as mentioned in Section 2.3, it is easy to use the SP2 to obtain the total particle number concentration and a correction factor (F) could be defined as, $F = N_{\mathrm{SP2}}/N$, to make the initial guess more accurate. Figure A.2 shows the initial guess for all of the phantoms considered in this study. Results of Figure 2.4 quantitatively indicates the accuracy of the initial guesses with relatively no computational time.

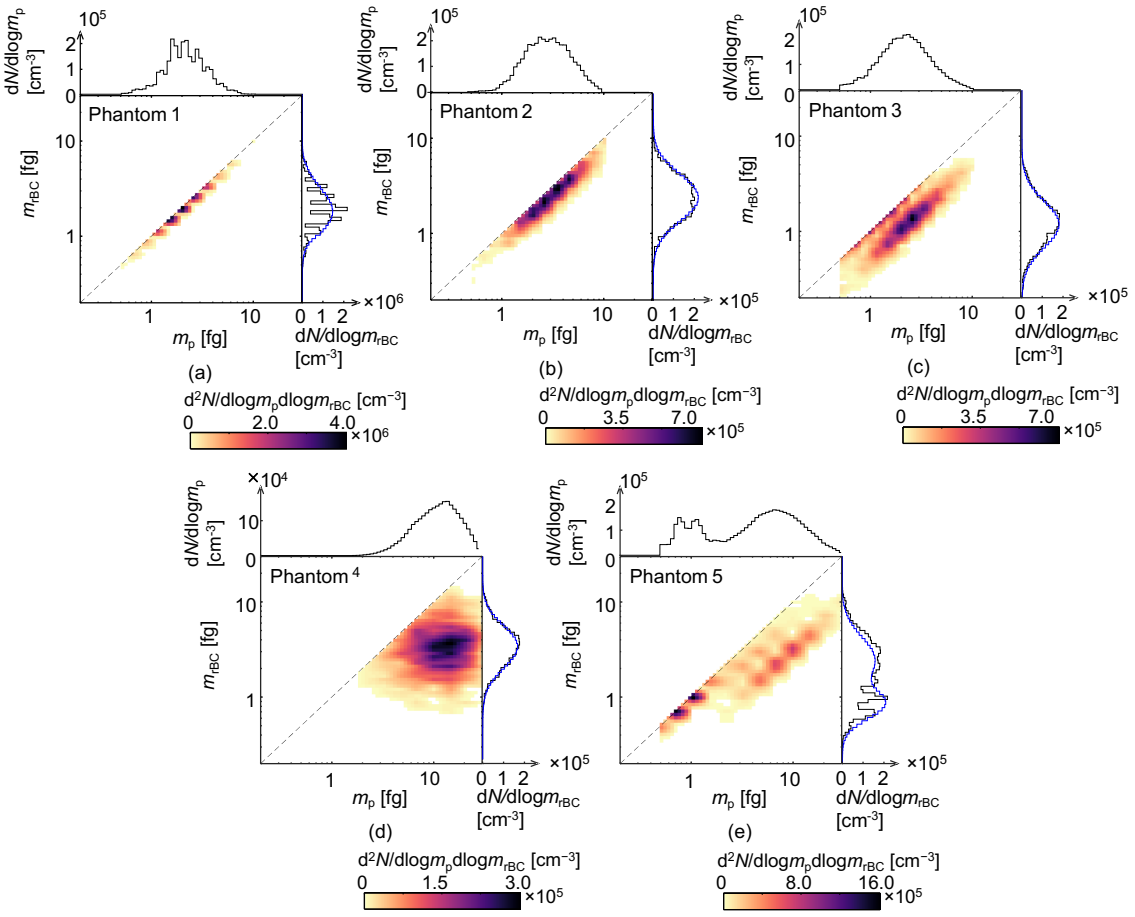

Figure A.2: Initial guesses obtained from synthetic data. The blue lines are the actual marginal number concentration distribution of refractory black carbon ($\mathrm{d}N/\mathrm{dlog}m_{\mathrm{rBC}}$) of the original phantom.

The general shapes of the initial guesses, especially for broad distributions, are relatively similar to what the Twomey algorithm, described later in this section, can retrieve at the end of consecutive iterations. Also, it can be seen that there are clouds of data near the main distributions which are related to the multiply-charged particles classified by the CPMA. This deficiency is expected because the initial guess is evaluated for only singly charged particles ($\Phi = 1$). Besides, it is noted that as distributions get broader, the accuracy of initial guess improves.

A.1.1 Twomey's original method

Twomey, 1975 original algorithm uses an updating scheme of

$$x_j^{k+1} = \left[1 + \left(\frac{b_{\mathrm{i}}}{\mathbf{a}_i \mathbf{x}^k} - 1\right) A_{ij}\right] x_j^k, \tag{A.2}$$

where $x_j^{k+1} \in \mathbb{R}^{N_j^2}$ represent a set of updated values obtained from multiplying the two-dimensional distribution (the initial guess or the one calculated in the previous iteration) to the kernel of the system. Here, $\boldsymbol{a}_i$ is the i^{th} row of kernel matrix, $\mathbf{A}$, scaled such that its element have a maximum value of unity; $|A_{ij}| \leq 1$. The kernel matrix should be normalized by its maximum value to avoid returning negative values (Twomey, 1975). The updating rule of Equation (A.2) should be repeated $n_b = n_i n_j$ times to update the entire $\mathbf{x}$ for one iteration. Also, note that the accuracy of the Twomey algorithm depends strongly on the initial guess. Figure A.3 shows the distribution retrieved by the Twomey method for phantoms 1 to 5 .

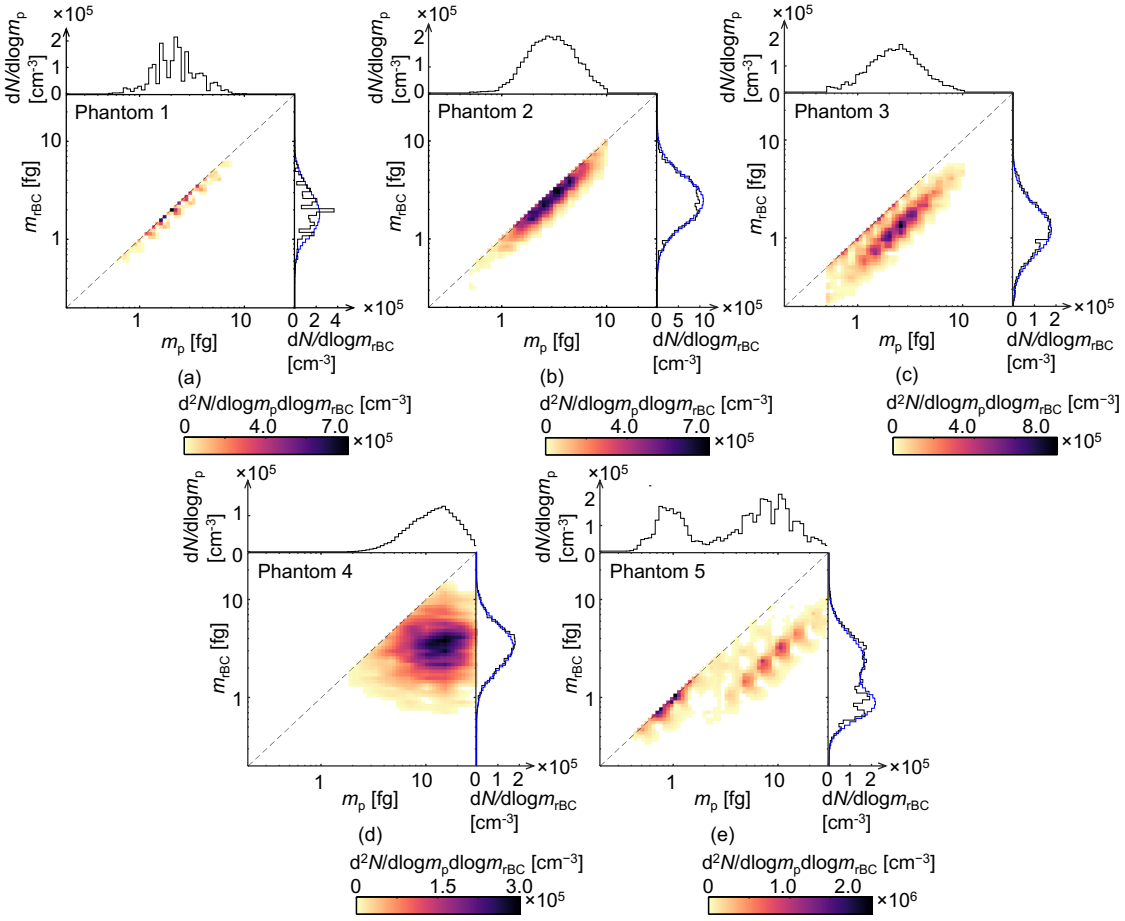

Figure A.3: Distributions retrieved by the Twomey method for synthetic data fabricated from phantoms 1 to 5 . The blue lines are the actual marginal number concentration distribution of refractory black carbon ($dN/d\log m_{\mathrm{rBC}}$) of the original phantom.

The results show relatively noisy solutions. The noisy solutions are an outcome of the rank-deficient problem as the collected CPMA-SP2 data is sparse in comparison with the reconstruction grid, and consequently, it results in poor reconstructions away from the sampled data points. The solutions get worst if the number of CPMA setpoints decreases relative to the reconstruction resolution.

A.1.2 Twomey-Markowski regularization method

Some modifications of the Twomey method have been proposed to deal with this problem. Markowski (1987) proposed the most widespread variant of Twomey's, here referred to as the Twomey-Markowski regularization method following Sipkens, Olfert, and Rogak (2020a). The Twomey- Markowski regularization method solves the *speckle* problem caused by the original Twomey regularization approach by adding a smoothing step before and after each Twomey updating step. Markowski proposed and applied the updating step for one-dimensional size distributions by using a locally-weighted average smoothing practice:

$$x_j^{k+1} = \begin{cases} 0.75x_j^k + 0.25x_{j+1}^k & j = 1 \\ 0.5x_j^k + 0.25x_{j-1}^k + 0.25x_{j+1}^k & 1 < j < N_j. \\ 0.75x_j^k + 0.25x_{j-1}^k & j = N_j \end{cases} \tag{A.3}$$

The Markowski smoothing method can be extended to the two-dimensional distribution by including all surrounding cells and reweight the adjacent cells in keeping with the number of nodes. Nonetheless, the reweight of peripheral cells in a two-dimensional case needs careful attention, and simple reweights, *i.e.,* allocating half of the value of the central cell to itself and assigning the rest to the adjacent cells, precisely like what Markowski did for one-dimensional case, leads into exceedingly

smooth results. Rawat et al. (2016) modified the original Markowski smoothing function to make it applicable for a two-variable distribution. Broda et al. (2018) adjusted Rawat et al.'s (2016) smoothing function by assigning a weight function according to a Gaussian distribution for smoothing function to account for the distance between bins of data. Also, they modified the smoothing function to make it agree with the physics of the no-negative coating. Buckley et al. (2017) modified Rawat et al. (2016) smoothing function by introducing the smoothing factor, S_{f}, to regulate the weight assigned to the peripheral cells, directly regulating the degree to which smoothing is applied. Buckley et al. (2017) proposed the smoothing function for the central cells ($2 < i < N_i - 1; 2 < j < N_j - 1$) as

$$x_{i,j}^{k+1} = \frac{0.5x_{i,j}^{k+1} + S_{\mathrm{f}}\left(x_{i+1,j}^{k} + x_{i,j+1}^{k} + x_{i,j+1}^{k} + x_{i+1,j+1}^{k} + x_{i-1,j}^{k} + x_{i,j-1}^{k} + x_{i+1,j-1}^{k} + x_{i-1,j+1}^{k}\right)}{0.5 + 8S_{\mathrm{f}}} \tag{A.4}$$

Besides, the smoothing function for the cells located at the edge of two-dimensional distributions was proposed to be:

$$\begin{aligned} x_{1,j}^{k+1} &= 0.75x_{1,j}^{k} + 0.25x_{2,j}^{k} \\ x_{i,1}^{k+1} &= 0.75x_{i,1}^{k} + 0.25x_{i,2}^{k} \\ x_{N_{i,j}}^{k+1} &= 0.75x_{N_{i,j}}^{k} + 0.25x_{N_{i,j},j}^{k}, \quad \text{and} \\ x_{i,N_j}^{k+1} &= 0.75x_{i,N_j}^{k} + 0.25x_{2,N_j}^{k}. \end{aligned} \tag{A.5}$$

Nonetheless, for the CPMA-SP2 application, the smoothing function proposed by Buckley et al. (2017) needs to be modified to be inline with the physics of no-negative coating. This modification, or in other words, imposing constraint, is only considered for those cells on the diagonal of two-dimensional distribution: For internal cells on

the diagonal,

$$x_{i,j}^{k+1} = \frac{0.5x_{i,j}^{k+1} + S_{\mathrm{f}}\left(x_{i+1,j}^{k} + x_{i,j+1}^{k} + x_{i-1,j}^{k} + x_{i,j-1}^{k} + x_{i-1,j+1}^{k}\right)}{0.5 + 5S_{\mathrm{f}}} \tag{A.6}$$

and for the first and the last cell on the diagonal,

$$\begin{aligned} x_{1,1}^{k+1} &= x_{1,1}^{k} \\ x_{N_i,N_j}^{k+1} &= 0.75x_{N_i,N_j}^{k} + 0.25x_{N_{i-1},N_j}^{k}. \end{aligned} \tag{A.7}$$

Figure A.4 depicts the distribution retrieved by the Twomey-Markowski regularization scheme for the phantom 1 to 5. It is noted that the Twomey-Markowski regularization scheme, to some extent, enhanced the quality of deconvolution of the Twomey scheme by decreasing the oscillations between nearby cells. Besides, it is noted that as distributions get broader, the performance of Twomey-Markowski enhances, *e.g.,* the error or reconstruction of Phantom 1, 3 and 4 by the Twomey-Markowski decreases respectively, from 67 to 30 and 17%. Despite this improvement and avoiding producing overly smooth results by using the optimal value of S_{f} in the present study, specifying S_{f} is so tricky for experimental cases and makes the use of the Twomey-Markowski method a bit challenging.

A.1.3 Maximum entropy regularization

A few studies in the aerosol literature have considered Maximum-entropy-type methods to solve ill-posed problems (*e.g.,* Amato et al., 1995; Sipkens, Olfert, and Rogak, 2020a; Wang, 2008; Yee, 1989). This method regularizes the problem by simultaneously minimizing the information entropy:

$$H = -\int_0^\infty \int_0^\infty N\left(m_{\mathrm{p}}, m_{\mathrm{rBC}}\right) \ln N\left(m_{\mathrm{p}}, m_{\mathrm{rBC}}\right) \mathrm{dln}m_{\mathrm{p}} \mathrm{dln}m_{\mathrm{rBC}} \tag{A.8}$$

or in the discrete form,

$$H(x) = -\sum_{j=1}^{N_x} x_j \ln x_j \cdot (\Delta \ln m_{\mathrm{p}})_j \cdot (\Delta \ln m_{\mathrm{rBC}})_j \,. \tag{A.9}$$

The regularization of ill-posed problems can be done by linking it to the entropy information of the solution. One way of incorporating entropy information into the deconvolution is to use the entropy information as penalty term, $\Phi(\mathbf{x}) = -\mathrm{H}(\mathbf{x})$ in Tikhonov methods (discussed in the next section):

$$\begin{aligned} \mathbf{x}^{\mathrm{ME}} &= \arg\min_x \left\{ \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \lambda \Phi(\mathbf{x}) \right\} \\ &= \arg\min_x \left\{ \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \lambda \sum_{j=1}^{N_x} x_j \ln x_j \cdot (\Delta \ln m_{\mathrm{p}})_j \cdot (\Delta \ln m_{\mathrm{rBC}})_j \right\} \end{aligned} \tag{A.10}$$

In the above equation, λ is a regularization parameter. A nonlinear solver, along with two types of constraints, should be considered to minimize the above expression. The convergence rate of the above approach, due to non-linearity of iterative solvers combined with the constrained solution, is extremely low. Therefore, some studies (Amato et al., 1995; Carfora, Esposito, and Serio, 1998), suggest not to implement this technique to regularize any ill-posed problem. The convergence of this approach gets worse for two-dimensional cases, like those in the present study. Hence, the above-explained approach has not been considered in the present paper. Nevertheless, another type of maximum entropy regularization method, known as the multiplicative algebraic reconstruction technique (MART) Elfving, 1980; Lent, 1976, in which the regularized solution is obtained by minimizing the argument of

$H(\mathbf{x})$ instead of minimizing the argument mentioned in Equation (A.10) is:

$$\mathbf{x}^{\mathrm{ME}} = \arg\min_x\{H(\mathbf{x})\} \quad \text{s.t.} \; = \begin{cases} \mathbf{Ax} = \mathbf{b} \\ \text{Constraints} \equiv \begin{cases} \text{non-negative particle concentration} \\ \text{no negative coating} \end{cases} \end{cases} \tag{A.11}$$

Introduced by Gordon, Bender, and Herman (1970), the MART method requires regularization parameters which are challenging to determine. Also, it is a multiplicative method, like the Twomey method, which keeps updating an initial guess. The updating rule of MART method is:

$$x_j^{k+1} = \left(e^{\sum_i f_{\mathrm{r}} A_{ij} d_i^k}\right) x_j^s = \left(\frac{b_i}{\boldsymbol{a}_i \mathbf{x}^k}\right)^{f_{\mathrm{r}} A_{ij}} x_j^k, \tag{A.12}$$

where $\mathbf{a}_j^{\mathrm{T}}$ the j^{th} row of $\mathbf{A}$ and f_{r} is a relaxation factor, which should be determined to maximize the entropy information. According to the study of Lent (1976), f_{r} equals to unity leads to the convergence of the maximum entropy information. Also, d_i^k is the fidelity term, here is taken as $\ln\left(b_i/\boldsymbol{a}_i \mathbf{x}^k\right)$. The iterative procedure of the MART algorithm is quite similar to the original Twomey's method. Similar to Twomey-type methods, not only does it need the initial guess linear interpolated from the quantity calculated by Equation (A.1), but also it starts iterating over every element in $\mathbf{b}$. The speed of the MART algorithm can be improved considerably by the *block* implementation of the MART algorithm (Willis, 2000), though both executions lead to the same result:

$$\mathbf{x}^{\mathrm{k+1}} = \left(e^{f_{\mathrm{r}} A^T d^k}\right) \odot \mathbf{x}^{\mathrm{k}} \tag{A.13}$$

where $\odot$ is the elementwise product. As the results of MART and Twomey look quite similar, the graphical results of the MART method are not presented.

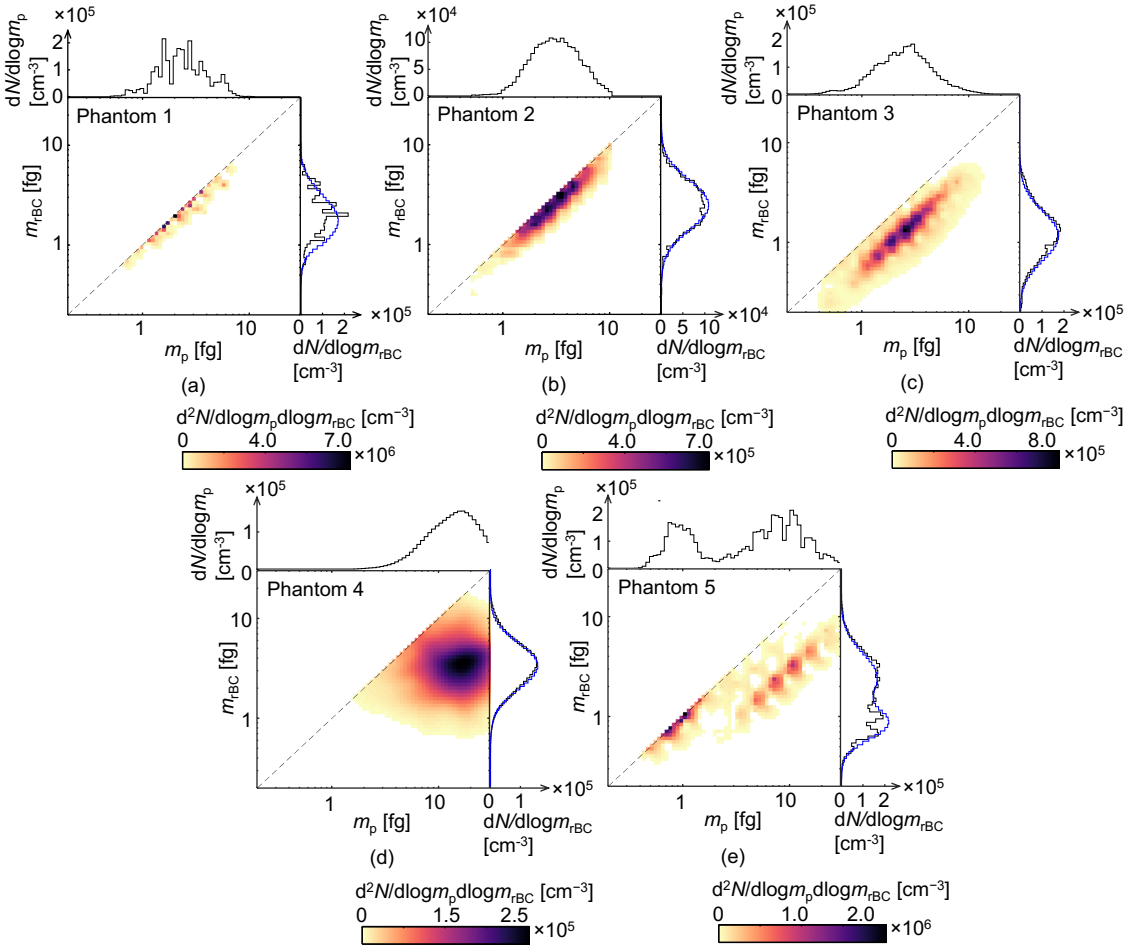

Figure A.4: Distributions retrieved by the Twomey-Markowski methods for synthetic data fabricated from phantoms 1 to 5 . Blue lines are the actual marginal number concentration distribution of refractory black carbon ($dN/dlogm_{rBC}$) of the ground truths.

A.1.4 MART with the smoothing function

Similar to the Twomey-Markowski algorithm, the MART solution can be improved by implementing a smoothing step. Consequently, the smoothing function already presented for the Twomey-Markowski algorithm was also employed for the MART algorithm to examine the potential advantage of adding smoothing steps to the MART algorithm over the TwomeyMarkowski algorithm. Figure A.5 shows the results of MART with Markowski smoothing. As shown in Figure 2.4, MART-Markowski (*i.e.,* MART with Markowski smoothing) and Twomey-Markowski have the same accuracy; however, the convergence rate is almost eight times faster with MART-Markowski.

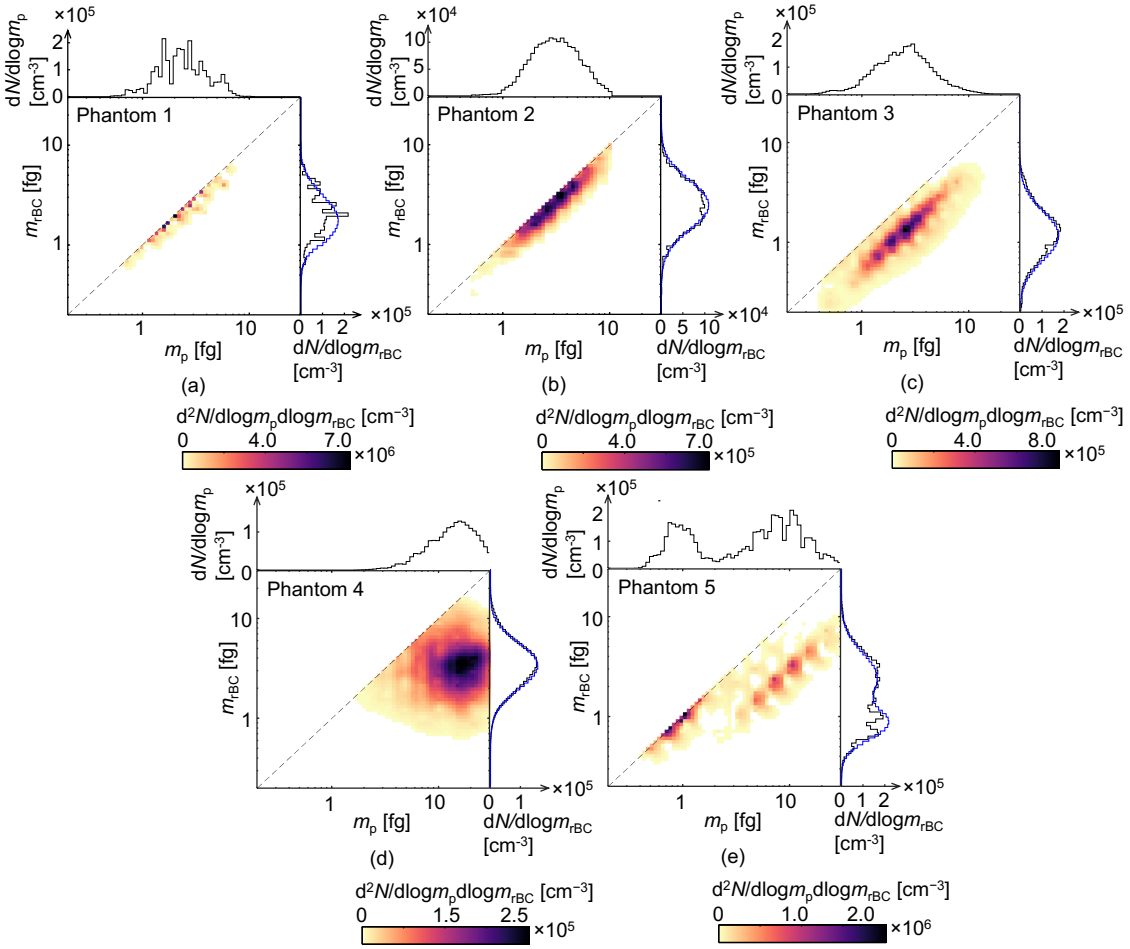

Figure A.5: Distributions retrieved by the MART with the smoothing approach for synthetic data fabricated from phantoms 1 to 5 . Blue lines are the actual marginal number concentration distribution of refractory black carbon ($dN/d\log m_{\mathrm{rBC}}$) of the ground truths.

A.2 Tikhonov regularization methods

The least-squares problem expressed in Equation 2.31 is presented in the stacked form of:

$$\begin{bmatrix} \sqrt{\lambda}\mathbf{A}^1 \\ \mathbf{C} \\ \sqrt{\alpha}\mathbf{L}^{\mathrm{TK}} \end{bmatrix} \mathbf{x}^{\mathrm{TK}} = \begin{bmatrix} \sqrt{\mathbf{b}} \\ \mathbf{d} \\ \mathbf{0} \end{bmatrix} \tag{A.14}$$

This new form of a least-squares problem shows two sets of equations are added: $n_{j,0}$ equations of $\boldsymbol{C}\mathbf{x}^{\mathrm{TK}} = \boldsymbol{d}$, for considering the marginal distribution prior, and $\left(n_{j,0}^2 + n_{j,0}\right)/2$ equations of $\sqrt{\alpha}\mathbf{L}^{\mathrm{TK}}\mathbf{x}^{\mathrm{TK}} = \mathbf{0}$, for considering smoothing prior information. Additionally, Equation (A.14) can also be rearranged to be solved algebraically for any Tikhonov scheme:

$$\begin{aligned} {\mathbf{x}_{\alpha,\lambda}}^{\mathrm{TK}} &= \left(\begin{bmatrix} \sqrt{\lambda}\mathrm{A} \\ \boldsymbol{C} \\ \sqrt{\alpha}\mathbf{L}^{\mathrm{TK}} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \sqrt{\lambda}\mathrm{A} \\ \boldsymbol{C} \\ \sqrt{\alpha}\mathbf{L}^{\mathrm{TK}} \end{bmatrix} \right)^{-1} \begin{bmatrix} \sqrt{\lambda}\mathrm{A} \\ \boldsymbol{C} \\ \sqrt{\alpha}\mathbf{L}^{\mathrm{TK}} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \sqrt{\lambda}\mathbf{b} \\ \boldsymbol{d} \\ \mathbf{0} \end{bmatrix} \\ &= \left[\lambda\mathbf{A}^{\mathrm{T}}\mathbf{A} + \boldsymbol{C}^{\mathrm{T}}\mathbf{A} + \alpha\mathbf{L}^{\mathrm{TK}}\mathbf{L}^{\mathrm{TK}}\right]^{-1} \left[\lambda\mathbf{A}^{\mathrm{T}}\mathbf{b} + \boldsymbol{C}^{\mathrm{T}}\boldsymbol{d}\right]. \end{aligned} \tag{A.15}$$

However, since the non-negative particle concentration constraint should be considered in retrieving the distributions, iterative solvers, *e.g.,* Lagrange's method, along with regularization searching schemes, , here the L-curve and Bayesian approaches, should be used in practice to derive the optimum reconstruction. General information regarding L-curve and Bayesian methods is also provided in section A.2.2.

A.2.1 Tikhonov matrix

The form and design of the discretized differential operator, $\mathbf{L}^{\mathrm{TK}}$, depends on the order of Tikhonov applied for regularizing the problem. In the literature (Bashurova et al., 1991; Crump and Seinfeld, 1982; Lesnic, Elliott, and Ingham, 1995; Talukdar and Swihart, 2003; Wang et al., 2006; Wolfenbarger and Seinfeld, 1990), three rules (*i.e.,* 0th-, 1st-, and 2nd-order) of the Tikhonov method are typically used where each of them has a damping effect on the experimental data noise or, in other words, the level of smoothness (*e.g.,* the first-order Tikhonov (FOT) scheme tries to minimize the differences between the nearby cells, the second-order Tikhonov (SOT) scheme seeks to lessen second-order differences between neighboring cells).

The basics of the Tikhonov matrix can be understood through finite-difference in the two dimensions, though to help understanding how the $\mathbf{L}^{\mathrm{TK}}$ would be defined, the $\mathbf{x}^{\mathrm{TK}}$ with the dimensions of 4×4 corresponding to the upper triangle elements of 10×1, (see Figure A.6a) is considered and the $\mathbf{L}^{\mathrm{TK}}$ is calculated based on the first order forward difference. The colored boxes and their corresponding dotted-dash lines in Figure A.6a are presented to schematically show which elements are used in implementing the two-dimensional form of the first-order forward difference for some sample element of $\mathbf{x}^{\mathrm{TK}}$ to derive the $\mathbf{L}^{\mathrm{TK}}$ matrix (see Figure A.6b).

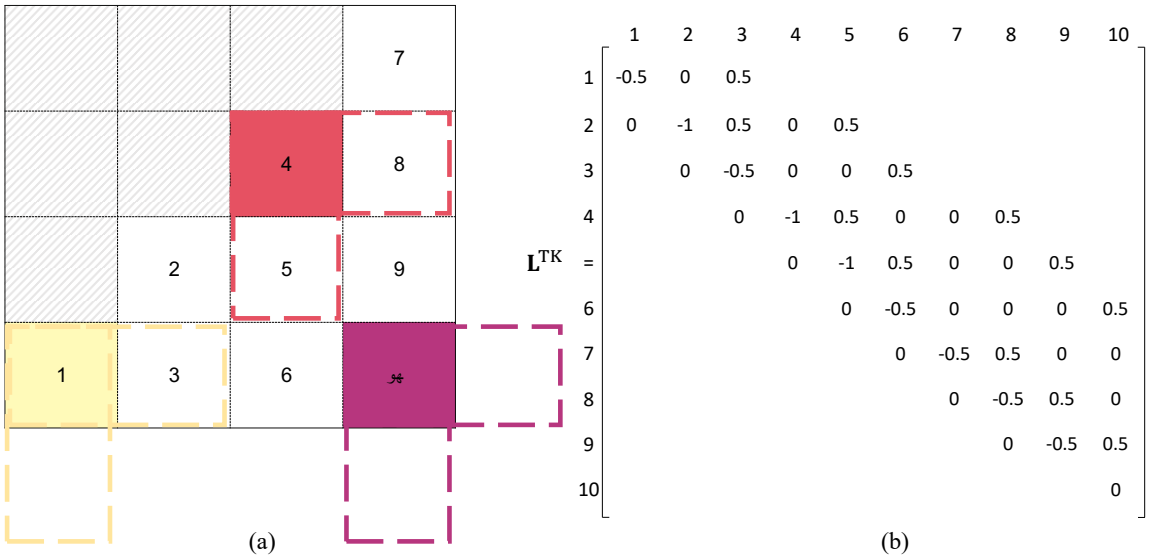

Figure A.6: (a) Vector representation of $\mathbf{x}^{TK}$ with colored boxes to schematically show which elements are used in implementing the two-dimensional form of forward-difference for some sample elements, element number 1,2 , and 10 , of $\mathbf{x}^{TK}$ to derive the Tikhonov matrix of the first-order forward difference. The colored dash lines represent the row and column directions of the twodimensional forward difference of each sample cell of the $\mathbf{x}^{TK}$, highlighted with the same solid colors.

In the two-dimensional form, one cell for each direction is needed to implement the first-order forward difference for each cell of the $\mathbf{x}^{TK}$, *e.g.,* cell number 3 in column direction and cell number 5 in the row direction are required to implement the first-order finite-difference for cell number 2. However, boundary cells may have one (*e.g.,* cell number 1) or two missing cells (*e.g.,* cell number 10) in the forward directions. The $\mathbf{L}^{TK}$ (Figure A.6b) is generated in a way that every row should sum to zero, and in the case that one of adjacent element in the forward direction is missing (on either side), this means that either the diagonal must be decreased or the off-diagonal increased. There would be a slight difference depending on which one of these two schemes is chosen, and in the present paper, the former scheme is selected

to minimize the effect of the slope at the boundaries, such that the absolute value of diagonals, the targeted elements, decreases from 1 to 0.5 for boundary elements (Figure A.6).

A.2.2 L-curve and Bayesian methods to inversion and regularization parameter selection

In all Tikhonov regularization approaches considered in the present study, an iterative nonnegative least-squares approach was employed to solve the system of linear equations introduced by Equation (2.27) and the L-curve and Bayesian approaches were considered to derive an optimal regularization parameter.

The concept of the L-curve method for obtaining the optimal regularization parameter relies on the fact that the Tikhonov regularized solution would be achieved by making a balance between having a small residual $\left(\mathbf{A}\mathbf{x}_{\alpha}^{\mathrm{TK}} - \mathbf{b}\right)$ and having the L^2-norm of the penalty term, *i.e.,* $\alpha \left\|\mathbf{L}^{\mathrm{Tk}}\mathbf{x}\right\|_2^2$, as small as possible (see Equation (2.27). Thereby, the regularization parameter would act as a tuning parameter and relates the residual and L^2-norm of the penalty term by an L-shaped curve. The optimal regularization parameter would exist at the corner of the L-shaped curve since both the residual and the L^2-norm of the penalty terms are simultaneously at the closest location to zero. Although, the visual inspection of the L-shaped curve is rather straightforward, implementing automated methods (*e.g.,* Calvetti, Golub, and Reichel, 1999 and Cultrera and Callegaro, 2016) to accurately attain the corner in the presence of noise can be challenging. Finding an optimal regularization parameter would become worse when several types of constraints should also be considered in solving a system of linear equations, which would lead to multiple irregularities in

the L-shaped curve. The details of the automated version of the L-curved method (Cultrera and Callegaro, 2016) developed and used in the present study are not described here due to the sake of conciseness, and only a summary of the Bayesian approach is explained in the following.

The Bayesian scheme, like all statistical models, is based on some random quantities following some form of statistical distributions. It treats the data and quantities of desired the two-dimensional distributions, as random variables. Additionally, it seeks to relate these to random variables by means of likelihood, *e.g.,* the probabilities describing the errors existing between the model and the data, $\boldsymbol{\varepsilon} = \mathbf{Ax} - \mathbf{b}$. In the Bayesian scheme, Bayes' equation, is used to derive the prior information,

$$p_{\mathrm{po}}(\boldsymbol{x} \mid \boldsymbol{b}) = \frac{p_{\mathrm{li}}(\boldsymbol{b} \mid \boldsymbol{x}) p_{\mathrm{pr}}(\boldsymbol{x})}{p(\boldsymbol{b})} \propto p_{\mathrm{li}}(\boldsymbol{b} \mid \boldsymbol{x}) p_{\mathrm{pr}}(\boldsymbol{x}) \tag{A.16}$$

Here, $p_{\mathrm{po}}(\mathbf{x} \mid \mathbf{b})$ is the posterior distribution or, in other words, the inclusive solution of the problem; $p_{pr}(\mathbf{x})$ is the prior, and $p(\mathbf{b})$ is the evidence. Also, $p_{\mathrm{li}}(\mathbf{bx})$ is the likelihood defined based on the understanding of the error distribution between the model and the data, ε. In the case of the CPMA-SP2, in which single soot particles are measured by the SP2, the error distribution (the likelihood) is expected to follow the Poisson distribution (Voutilainen, Kolehmainen, and Kaipio, 2001). Priors are selected based on the applications, and in case of aerosol inversion, the Bayesian interpretation of Tikhonov regularization is often used, retrieving the same distributions as we already discussed. However, there are quite a few deficiencies for Tikhonov regularization priors, most remarkably that all points in the reconstruction area are equally weighted in regularization. Thereby, we use another type of Gaussian prior, which regularizes the solutions regionally, and it is known as *exponential*

distance prior. The general form of the Gaussian prior used along with the Bayesian method is

$$\begin{aligned} p_{\mathrm{pr}}(\boldsymbol{x}) &= \det\left(2\pi\boldsymbol{\Sigma}^{pr}\right)^{-1/2}\exp\left(-\frac{1}{2}\left(x-x_0\right)^{\mathrm{T}}\left(\boldsymbol{\Sigma}^{pr}\right)^{-1}\left(x-x_0\right)\right) \\ &= \det\left(2\pi\left(\mathbf{L}^{\mathrm{pr}}\right)^{\mathrm{T}}\mathbf{L}^{\mathrm{pr}}\right)\exp\left(-\frac{1}{2}\left\|\mathbf{L}\mathbf{p}^{\mathrm{pr}}\left(x-x_0\right)\right\|_2^2\right) \end{aligned} \tag{A.17}$$

where Σ^{pr} is the prior covariance and $\mathbf{L}^{\mathrm{pr}}$ is the Cholesky factorization of the inverse covariance matrix of prior $\left(\mathbf{L}^{\mathrm{pr}}=\mathrm{chol}\left[\left(\boldsymbol{\Sigma}^{\mathbf{pr}}\right)^{-1}\right]\right.$; and x_0 is a rough estimation of desired quantities (more discussion regarding the prior is provided later in this section). Considering the proper form of both likelihood and prior, the optimum solution, statistically regularized solution, would be the one that maximizes the posterior distribution, known as the maximum a posteriori estimate:

$$x^{\mathrm{MAP}}=\arg\max_{\boldsymbol{x}}\left\{p_{\mathrm{po}}(\boldsymbol{x}\mid\boldsymbol{b})\right\}=\arg\max_{\boldsymbol{x}}\left\{p_{\mathrm{1i}}(\boldsymbol{b}\mid\boldsymbol{x})p_{\mathrm{pr}}(\boldsymbol{x})\right\}. \tag{A.18}$$

which can also be represented to the CPMA-SP2 system of equations as

$$\begin{aligned} x^{\mathrm{MAP}} &= \arg\min_{\mathbf{x}}\left\{\lambda\left\|\mathbf{L}^b\left(\mathbf{A}\mathbf{x}^{\mathrm{TK}}-\mathbf{b}\right)\right\|_2^2+\left\|\mathbf{L}^d\left(\boldsymbol{C}\mathbf{x}^{\mathrm{TK}}-\mathbf{d}\right)\right\|_2^2+\alpha\left\|\mathbf{L}^{\mathrm{pr}}\left(\mathbf{x}-\mathbf{x}_0\right)\right\|_2^2\right\} \\ &= \arg\min_{x}\left\{\left\|\begin{array}{c}\mathbf{L}^b\left(\mathbf{A}\mathbf{x}^{\mathrm{TK}}-\mathbf{b}\right)\\ \mathbf{L}^d\left(\boldsymbol{C}\mathbf{x}^{\mathrm{TK}}-\mathbf{d}\right)\\ \alpha\mathbf{L}^{\mathrm{pr}}\left(\mathbf{x}-\mathbf{x}_0\right)\end{array}\right\|_2^2\right\} \end{aligned} \tag{A.19}$$

which is similar to what was discussed in the Tikhonov regularization methods. The most direct way of solving the above system of equations would be linear algebra. However, as the nonnegative particle concentration constraint should be considered, using iterative solvers is inevitable (*e.g.,* the interior-point optimization algorithm). In this study, using the `lsqnonneg` function from Matlab® (Lawson and Hanson,

1995) used as the iterative solver to achieve the solutions. More detail regarding the Bayesian approach could be found in the work of Sipkens, Olfert, and Rogak (2020a,b). Figure A.7 depicts the regularized distributions reconstructed by the Tikhonov zeroth order method, along with implementing the Bayesian method to find the optimum regularization parameter.

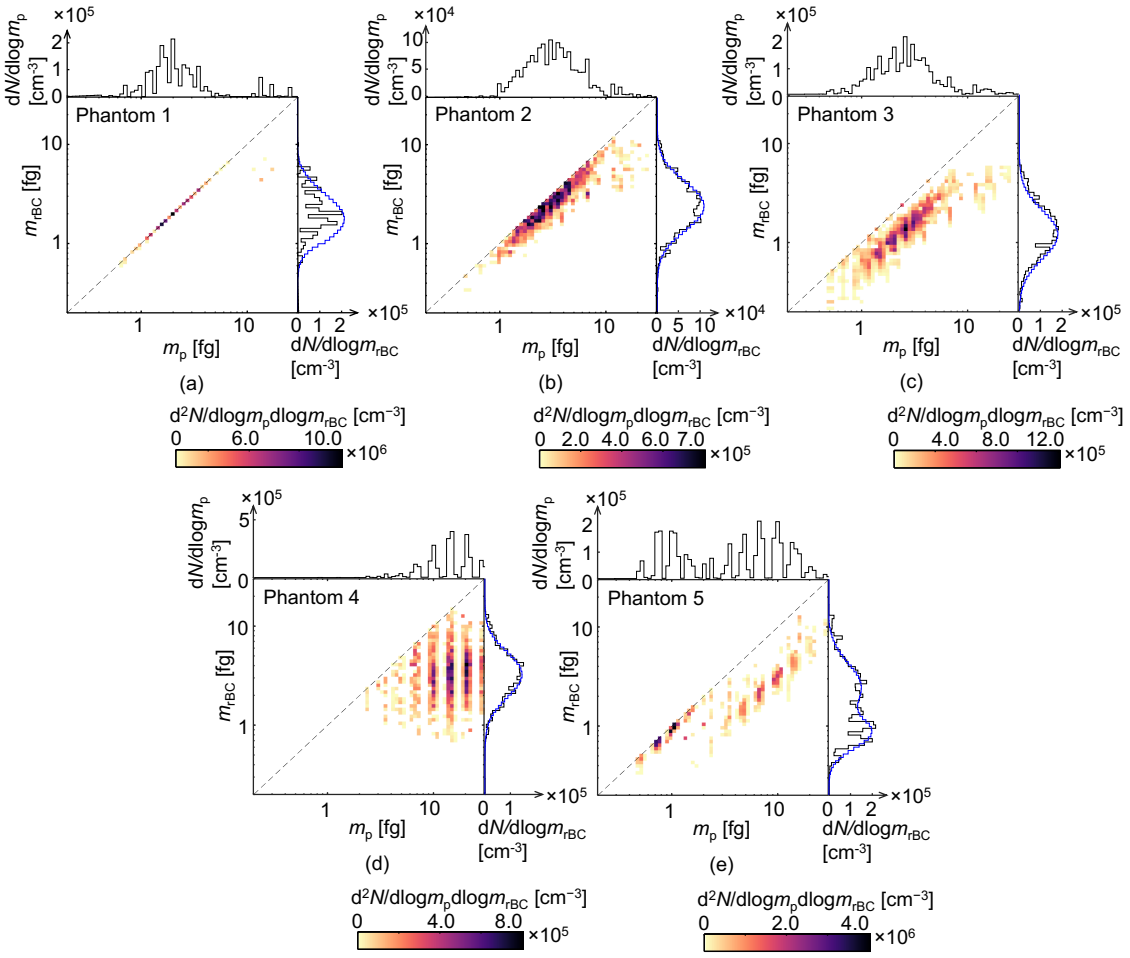

Figure A.7: Distributions retrieved by the Tikhonov 0th-order regularization method for synthetic data fabricated from phantoms 1 to 5. The blue line is the actual marginal number concentration distribution of refractory black carbon ($\mathrm{d}N/\mathrm{dlog}\,(m_{\mathrm{rBC}})$ of ground truths.

It can be seen that the Tikhonov zeroth-order scheme does not impose smoothness on the solution and it only tries to minimize the solution norm. The scheme was only able to control the large fluctuations in the unregularized least-squares solution, and the solutions retrieved by the Tikhonov zeroth-order scheme seem jagged rather than being smooth. In addition, a graphical comparison between Tikhonov zeroth-order scheme and Towomey and MART with smoothing function shows that the latter methods reconstructed two-dimensional distributions to a greater extent. Consequently, to enhance the regularization results of the Tikhonov method, some level of smoothness on the solutions is considered as prior information about x values. The degree to which each x value is smoothed depends on the order of the Tikhonov method. In the present study, the first and second-order Tikhonov schemes are considered as higher-order Tikhonov regularization methods to retrieve the solutions. Figure 2.4 indicates that although the convergence rate of the L-curve method is nearly two times faster than the Bayesian method, implementing the L-curve method for the first and second-order Tikhonov causes respectively 15% and $\sim$ 17% higher Euclidian error than using the Bayesian approach for finding the regularization parameter. This is because the L-curve method underestimates the value of the optimal regularization parameter at which the optimal reconstruction would be attained, leading to under-regularized solutions, which is shown in Figure A.8. Consequently, the L-curve method is not considered as the best searching scheme for the regularization parameter in comparison with the Bayesian method.

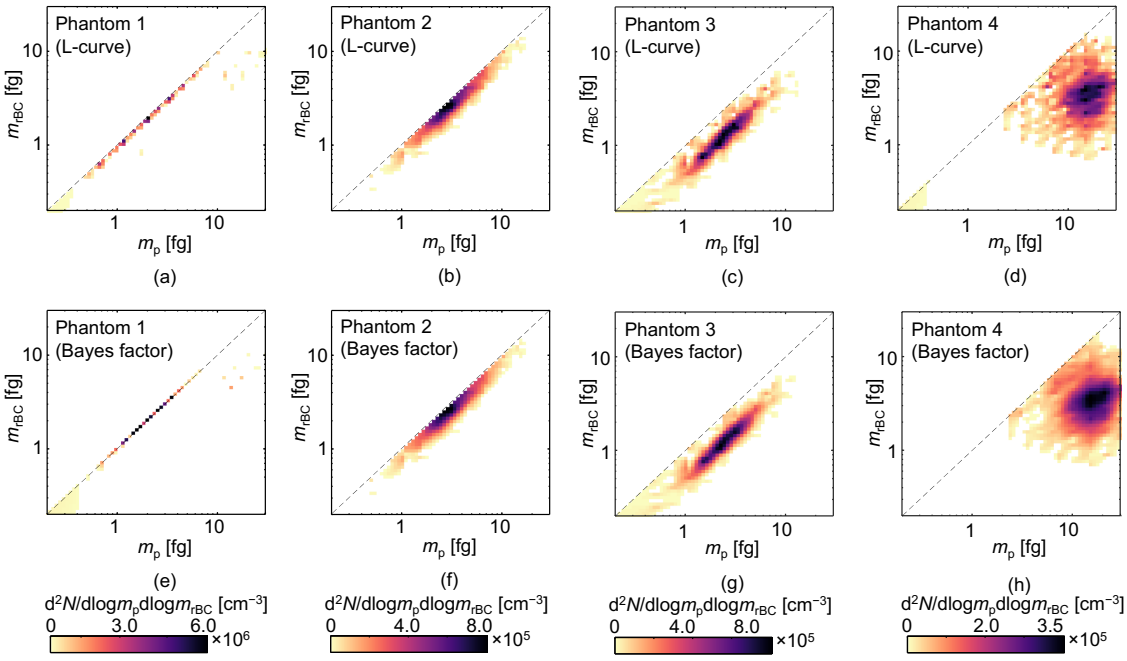

Figure A.8: Distributions retrieved by implementing the first-order Tikhonov regularization method along with (top) L-curve method and (bottom) Bayesian method for synthetic data fabricated from phantoms 1, 2, 3, and 4. The blue line is the actual marginal number concentration distribution of refractory black carbon (dN/dlogm_{rBC}) of ground truths.

The considerable reduction in the corresponding relative Euclidean error and oscillations in solutions of higher-order Tikhonov schemes in comparison with the other inversion methods discussed up to here is mostly because of imposing the smoothing prior to the structure of the solution making the higher-order Tikhonov schemes able to fill the cells containing no data by acquiring information from neighboring cells (especially for those states having bare soot, where m_{rBC} equals m_{P}). Furthermore, the novel approach used in the present paper to accommodate the sharp edge where m_{rBC} equals m_{p}, is able to retrieve the values on the sharp edge without performing smoothing, and it can also reconstruct the distribution away from the edge with

the proper level of smoothing. However, it is seen that black carbon (Phantom 1) satisfactorily (as also shown in Figure A.8).

A.3 Exponential Distance Prior

In comparison with the Tikhonov regularization prior, which considers the equal weight for all points in the reconstruction, the correlation matrix, $\mathbf{\Gamma}^{\mathrm{D}}$, only applies regularization locally. This is the case since, in the correlation matrix, a high degree of correlation is expected when the Euclidean distance between the elements is small, and the correlation between elements in the reconstruction would quickly decline as the distance grows. As mentioned in Section 2.4, when dealing with unimodal distributions, *i.e.,* phantoms 1 to 4, the width of the SP2-only measurement can be used to properly estimate the optimum correlation lengths and define the correlation matrices, albeit the true correlation matrices are used in the present study to derive the optimal distributions. The optimal solutions retrieved by the exponential distance method in reconstructing the unimodal distributions (*i.e.,* phantoms 1 to 4) are demonstrated in Fig. A.9. However, defining the correlation lengths and off-diagonals of $\mathbf{\Gamma}^{\mathrm{D}}$ for multimodal distributions, *i.e.,* Phantom 5, is more challenging. Using the width of one of the modes or simply averaging values over multiple modes may over-regularize narrow distributions, making them wider, and under-regularized broad distributions, resulting in excessive noise. For instance, in the case of the bimodal distribution of Phantom 5, $\mathbf{\Gamma}^{\mathrm{D}}$ can be approximated by different approaches listed in Table S1, and the optimal solution is obtained when the elements of $\mathbf{\Gamma}^{\mathrm{D}}$ are obtained by doubling the arithmetic average of the correlation matrices of modes 1 and 2 of Phantom 5.

Fig. A.10 depicts the distributions retrieved by the exponential distance method, while different approaches listed in Table S1 are employed to estimate $\mathbf{\Gamma}^{\mathrm{D}}$.

Table A.1: Relative error of reconstruction of Phantom 5, obtained by implementing the exponential distance method, where $\mathbf{\Gamma}^{\mathrm{D}}$ defined based on different approaches.

Phantom	$m_{rBC,g}$(fg)	$m_{p,g}$(fg)	σ_{rBC}	σ_{p}	R_{12}	m_{p}[fg] range	m_{rBC}[fg] range
1) Uncoated soot particles	2.0	2.0	0.204	0.204	1	0.1-20	0.1-20
2) Uncoated soot mixed with slightly coated soot particles	2.69	3.0	0.24	0.23	0.95	0.1-20	0.1-20
3) Moderately coated soot particles	2.30	1.22	0.24	0.23	0.95	0.1-20	0.1-20
4) Heavily coated soot particles	3.38	15	0.23	0.23	0.01	0.1-20	0.1-20
5) Bimodal uncoated and coated soot particles	0.87, 2.56	0.90,8.00	0.15,0.24	0.15,0.23	0.96,0.97	0.1-20	0.1-20

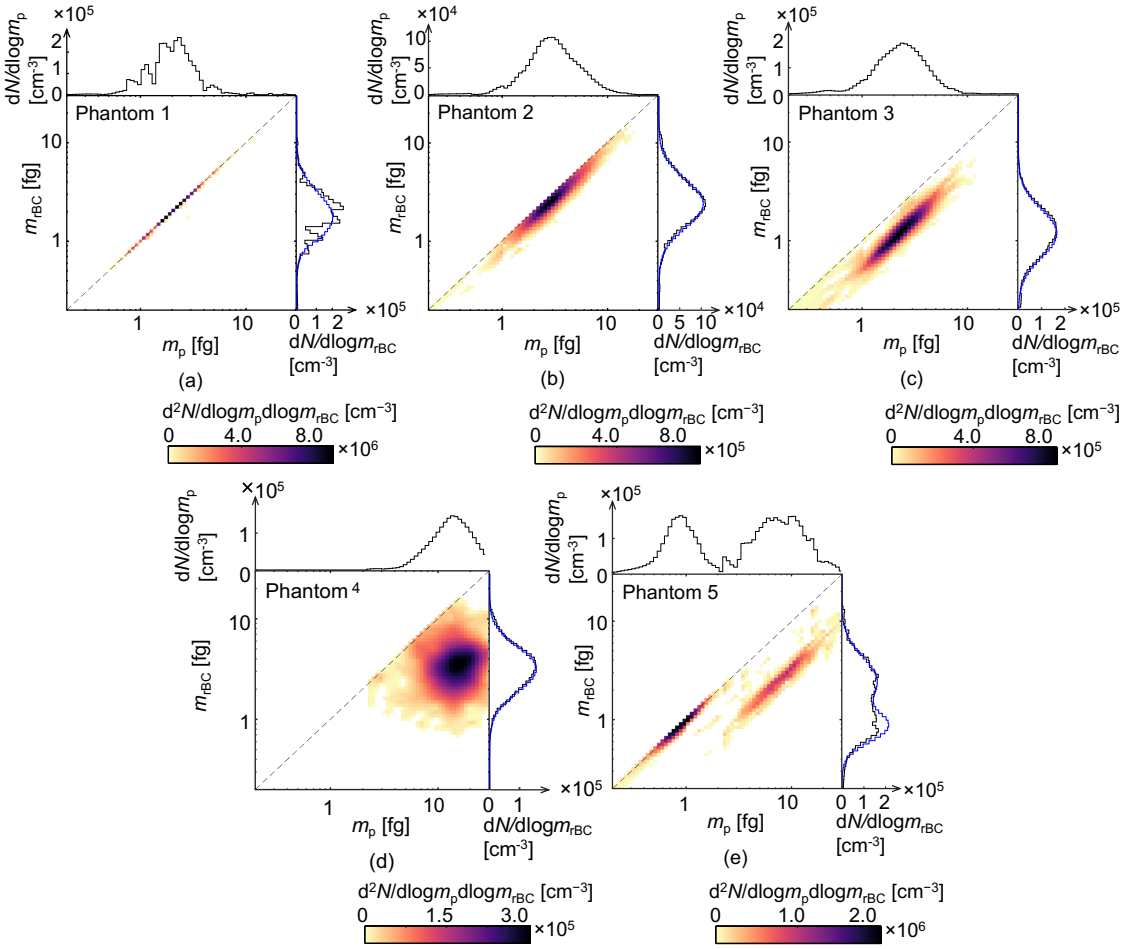

Figure A.9: Distributions retrieved by the exponential distance method for synthetic data fabricated from 1 to 5. Blue lines are the actual marginal refractory black carbon ($\mathrm{d}N/\mathrm{dlog}m_{\mathrm{rBC}}$) of the original phantom.

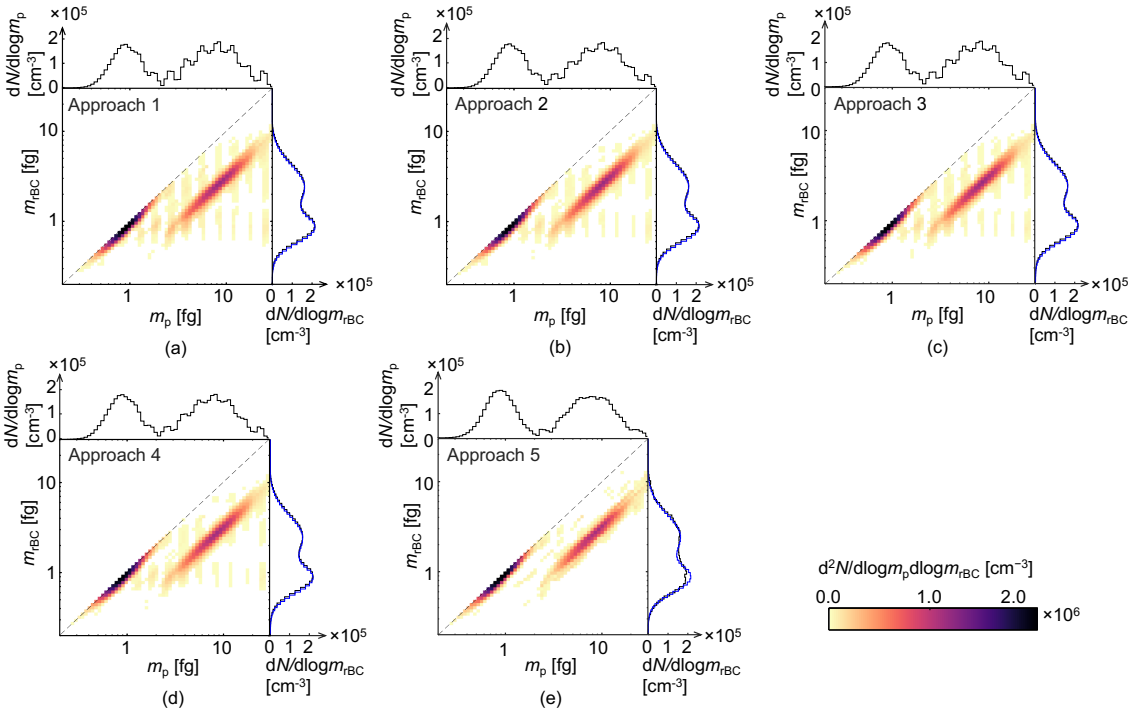

Figure A.10: Distributions retrieved by the exponential distance method, the exponential distance method, while different approaches listed in Table A.1 are employed to estimate $\mathbf{\Gamma}^{\mathrm{D}}$, for synthetic data fabricated from Phantom 5. Blue lines are the actual marginal refractory black carbon ($\mathrm{d}N/\mathrm{dlog}m_{\mathrm{rBC}}$) of the original phantom.

Although, a linear combination of correlation matrices of modes 1 and 2 of Phantom 5 results in the optimal solution, this approach cannot be generalized to other multimodal distributions. However, this finding does provide a basis for future studies involving more multimodal distributions, different in shape, location, arrangement concentration, etc. Also, the quantitative comparison of exponential distance and the Tikhonov methods in retrieving each $m_{\mathrm{p}} - m_{\mathrm{rBC}}$ distributions with considering both smoothing and marginal distribution priors are presented in Fig. A.11.

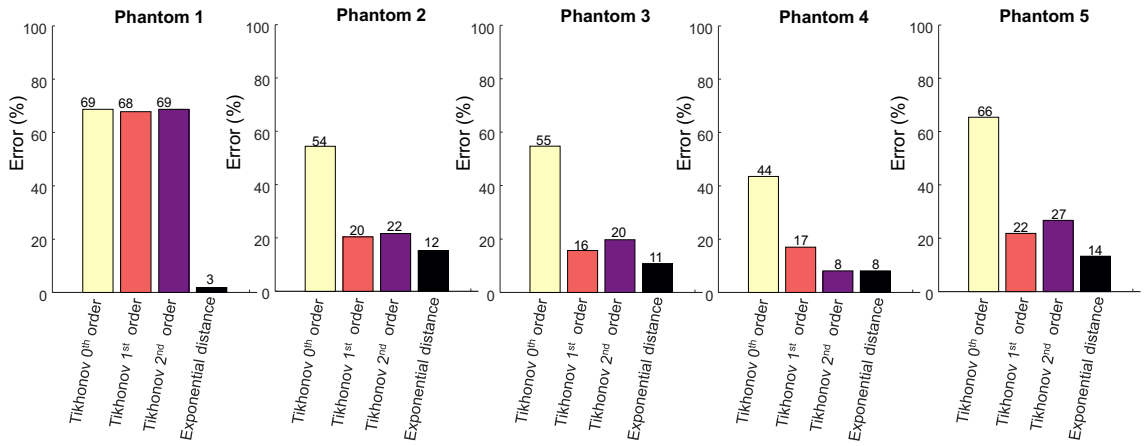

Figure A.11: The relative Euclidean error of reconstructions regularized by the exponential distance method and the Tikhonov methods with considering both smoothing and marginal distribution priors.

References

Bertsekas, DP and JN Tsitsiklis (2002). "The bivariate normal distribution". In: *Introduction to Probability*, pp. 247–253.

Sipkens, TA, JS Olfert, and SN Rogak (2020b). "Inversion methods to determine two-dimensional aerosol mass-mobility distributions II: Existing and novel Bayesian methods". In: *Journal of Aerosol Science* 146, p. 105565.

Appendix B: An alternative parameterization for bivariate lognormal distributions

B.1 An alternative parameterization for bivariate lognormal distributions

Following Sipkens, Olfert, and Rogak (2020b), the above representation can also be phrased in terms of a conditional geometric standard deviation, $\sigma_{\mathrm{rBC|p}}$. This quantity represents the width of vertical slices through the 2D distributions, as shown in Figure B.1.

For very narrow distributions, where $R_{12} = 1, \sigma_{\mathrm{rBC|p}} = 1$; for wide distributions, where $R_{12} = 0, \sigma_{\mathrm{rBC|p}} = \sigma_{\mathrm{rBC}}$. As such, this parameter can replace R_{12} in the definition of the bivariate lognormal distribution to give,

$$p\left(m_{\mathrm{p}}, m_{\mathrm{rBC}}\right) = \frac{N_{\mathrm{tot}}}{2\pi m_{\mathrm{p}} m_{\mathrm{rBC}} \ln\sigma_{\mathrm{p}} \ln\sigma_{\mathrm{rBC|p}}} \exp\left\{ -\frac{1}{2}\left(\frac{\ln m_{\mathrm{p}} - \ln m_{\mathrm{p,g}}}{\ln\sigma_{\mathrm{p}}}\right)^2 - \frac{1}{2}\left(\frac{\ln m_{\mathrm{rBC}} - \ln\left(m_{\mathrm{rBC,g}} \mid m_{\mathrm{p}}\right)}{\ln\sigma_{\mathrm{rBC|p}}}\right)^2 \right\},$$

where $m_{\mathrm{rBC,g}} \mid m_{\mathrm{p}}$ is the conditional geometric mean rBC mass, and $\sigma_{\mathrm{rBC|p}}$ is the conditional geometric standard deviation, which be derived from Σ in Equation (2.21)

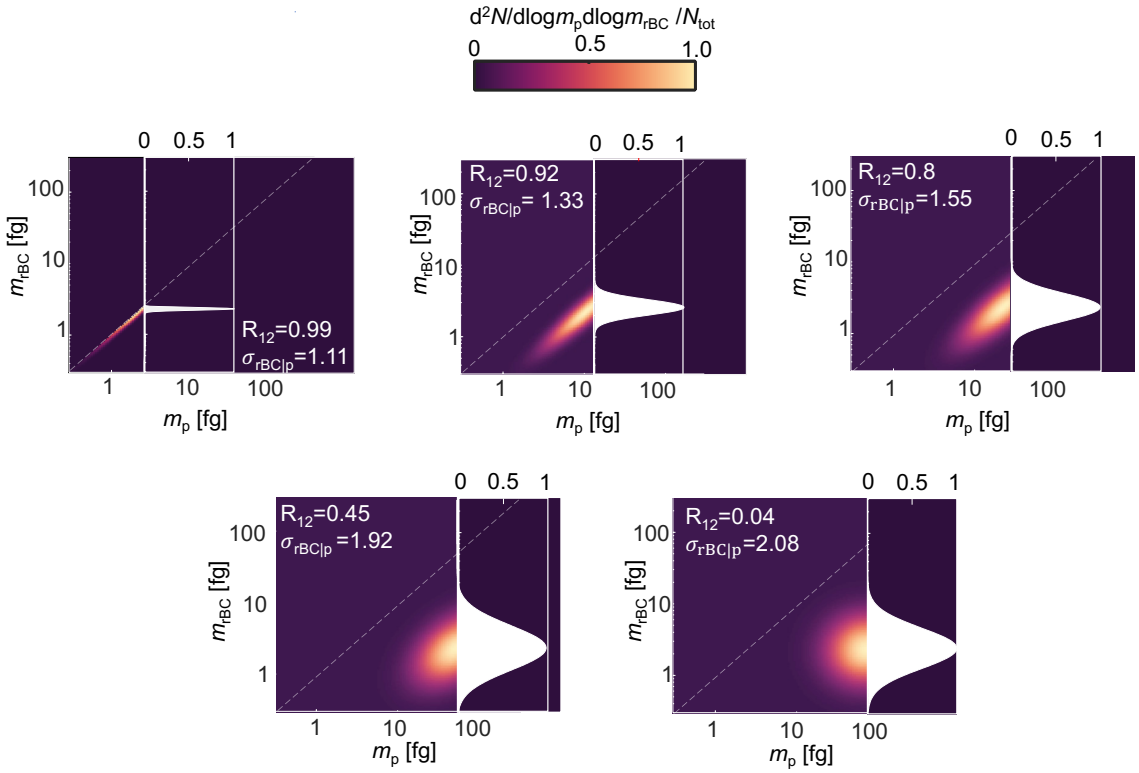

Figure B.1: Different non-dimensionalized phantoms with the conditional distributions.

using (Bertsekas and Tsitsiklis, 2002),

$$\log_{10} \sigma_{rBC|p} = \sqrt{\Sigma_{1,1} \left(1 - R_{1,2}^2\right)}.$$

Figure B.2 shows this relationship, where the condition geometric standard deviation changes from $\sigma_{\mathrm{rBC|p}} = 1.0$ when $R_{12} = 1$ to $\sigma_{\mathrm{rBC|p}} = \sigma_{\mathrm{rBC}}$ when $R_{12} = 0$.

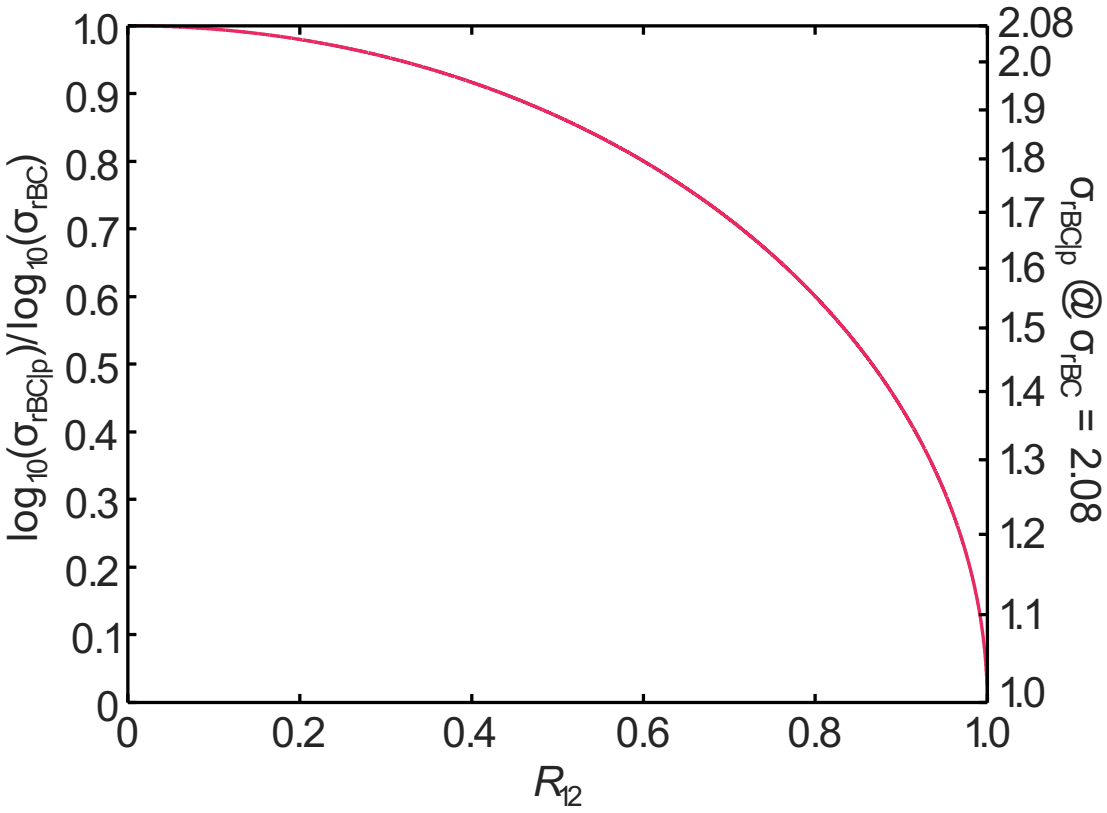

Figure B.2: Relationship between the correlation, R_{12}, and the conditional distribution geometric standard deviation, $\sigma_{\mathrm{rBC|p}}$

Appendix C: Trends in optimum Tikhonov regularization parameter with the distribution width

C.1 Trends in optimum Tikhonov regularization parameter with the distribution width

Figure C.1 indicates a positive correlation between the optimal Tikhonov regularization parameter and the width of the distribution. While larger regularization parameters are often associated with lower errors, these large regularization parameters are only reasonable for broad phantoms. Generally, broader distributions respond better to prior information (via the Tikhonov regularization scheme) that encodes smoothness. By contrast, very narrow phantoms will have their edges smoothed away by the prior, requiring lower regularization parameters to preserve sharpness. At the same time, the smaller regularization parameter will increase the amount of noise that propagates forward into the reconstructions, such that the minimum achievable error is larger. Some of this may be overcome using a different prior, *e.g.,* exponential distance prior (Chapter 2) or a total variation prior.

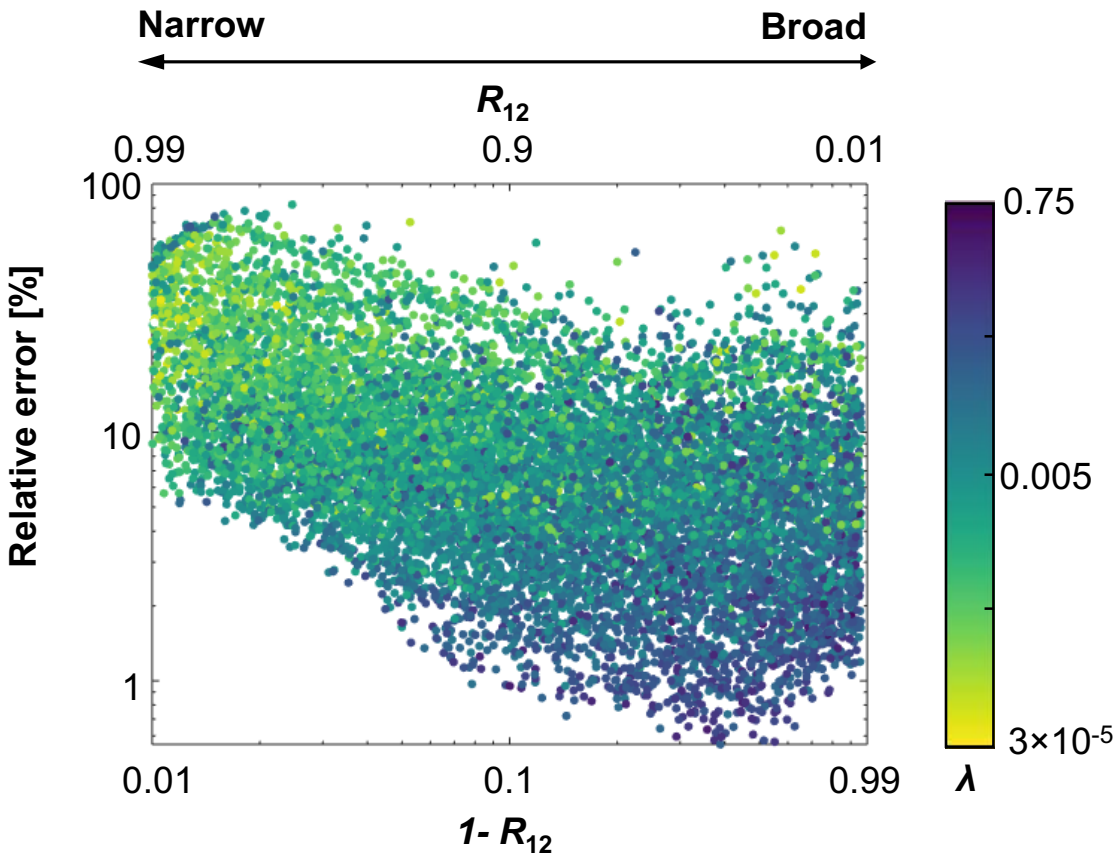

Figure C.1: The sensitivity of reconstruction accuracy to R_{12} indices (*i.e.*, distributions' broadness) and colored with the regularization parameter.

Appendix D: Effects of multiple charges and summary of data

D.1 Effects of multiple charges on TCPMA data

This study evaluates the CPMA transfer function on the assumption that the aerosol particles are only singly charged; however, the charge neutralizer employed is capable of imparting multiple charges to the particles. Assuming the number of charge state $\varphi = 1$ to 5 on the aerosols, a fraction of multiply charged aerosols entering the TCPMA measurement can be calculated as

$$F_{\varphi>+1} = 1 - F_{+1} \approx 1 - \frac{N_{m^*_{\varphi=1}} \times f\left(\varphi = 1, d_{\varphi=1}\right)}{\sum_{\varphi=2}^{5} N_{m^*_{\varphi}} \times f\left(\varphi, d_{\varphi}\right)}, \tag{D.1}$$

where F_{+1} is the number fraction of singly charged aerosols, and $N_{m^*_{\varphi}}$ is the number concentration of aerosols with a mass of $m^*_{\varphi} = \varphi \times m^*_1$ leaving the AAC. Here, d_{φ} is the equivalent mobility diameter of particles having the mass of m^*_{φ}, and $f\left(z_{\varphi}, d_{\varphi}\right)$ represents the charge fraction of aerosols of a certain size (d_{φ}) and charge state (φ). Traditionally, the charging function is calculated using Wiedensohler, 1988 as

$$f\left(z_k, d\right) = \frac{1}{\left[2\pi\sigma_{\mathrm{z}}(d)\right]^{1/2}} \exp\left\{-\frac{1}{2}\frac{\left[z_k - \sigma_{\mathrm{z}}(d)\ln\left(Z_{\mathrm{Z}}\right)\right]^2}{\sigma_{\mathrm{z}}(d)}\right\} \tag{D.2}$$

where Z_z is the ion mobility ratio which is approximated by ~ 0.875 (Wiedensohler et al., 1986). Also,

$$\sigma_Z(d) = \frac{2\pi\varepsilon_0 d k_B T}{e^2}, \tag{D.3}$$

where e is the elementary electron charge ($e = 1.6 \times 10^{-19} c$), ε_0 is the dielectric constant, k_B is the Boltzmann constant, and T is the aerosol temperature.

D.2 Summary of TCPMA data

Table A1 illustrates a summary of the results measured from the TCPMA experiments, along with the fraction of multiply charged aerosols entering the measurement system.

The calculated likelihood of multiply-charged aerosols at the entrance of the first CPMA is about 2% on average, which is negligible.

Table A1: An overview of all tandem CPMA measurements as well as the fraction of multiply-charged aerosols entering the measurement system. The measurements marked by † are regarded as outliers and not included in the fits.

Outline of TCPMA data

	Q	R_m	m^*	μ	λ	m^*_{12}	$F_{\Phi>1}$
†	0.3	2	0.1	1.26 ± 0.01	0.54 ± 0.03	0.84 ± 0.00	0.00 ± 0.00
	0.3	2	1	1.42 ± 0.02	0.96 ± 0.02	0.93 ± 0.02	0.02 ± 0.00
	0.3	2	10	1.46 ± 0.01	1.12 ± 0.02	0.97 ± 0.00	0.06 ± 0.00
†	0.3	3	0.05	0.96 ± 0.01	0.32 ± 0.01	1.25 ± 0.06	0.00 ± 0.00
†	0.3	3	0.1	1.05 ± 0.01	0.51 ± 0.01	1.22 ± 0.04	0.01 ± 0.00
	0.3	3	1	1.20 ± 0.01	0.84 ± 0.05	1.00 ± 0.01	0.02 ± 0.00
	0.3	3	10	1.21 ± 0.02	0.88 ± 0.06	1.00 ± 0.00	0.05 ± 0.00
	0.3	3	100	1.24 ± 0.01	1.07 ± 0.02	0.99 ± 0.00	0.13 ± 0.00
	0.3	6	0.05	0.79 ± 0.01	0.20 ± 0.01	1.12 ± 0.01	0.01 ± 0.00
	0.3	6	0.1	0.90 ± 0.00	0.32 ± 0.03	1.06 ± 0.01	0.01 ± 0.00
	0.3	6	1	1.08 ± 0.01	0.64 ± 0.01	1.02 ± 0.01	0.03 ± 0.00
	0.3	6	10	1.10 ± 0.02	0.82 ± 0.01	1.01 ± 0.00	0.03 ± 0.00
	0.3	6	100	1.13 ± 0.00	0.90 ± 0.03	1.02 ± 0.01	0.00 ± 0.00
	0.3	10	0.05	0.86 ± 0.09	0.18 ± 0.01	1.10 ± 0.01	0.01 ± 0.00
	0.3	10	0.1	0.91 ± 0.01	0.24 ± 0.10	1.05 ± 0.00	0.01 ± 0.00
	0.3	10	1	1.06 ± 0.02	0.63 ± 0.01	0.99 ± 0.00	0.03 ± 0.00
	0.3	10	10	1.18 ± 0.02	0.86 ± 0.02	1.02 ± 0.00	0.17 ± 0.00
	0.3	10	100	1.13 ± 0.02	0.87 ± 0.03	1.02 ± 0.00	0.00 ± 0.00
	0.3	15	1	1.03 ± 0.02	0.27 ± 0.02	1.01 ± 0.00	0.01 ± 0.00
	0.3	15	10	1.06 ± 0.01	0.55 ± 0.02	1.03 ± 0.00	0.02 ± 0.00
	0.3	15	100	1.03 ± 0.03	0.70 ± 0.04	1.03 ± 0.00	0.00 ± 0.00
	1.5	2	0.05	1.38 ± 0.01	0.65 ± 0.02	1.04 ± 0.00	0.00 ± 0.00
	1.5	2	0.1	1.36 ± 0.01	0.74 ± 0.01	1.00 ± 0.00	0.00 ± 0.00
	1.5	2	1	1.36 ± 0.01	0.94 ± 0.00	1.00 ± 0.00	0.01 ± 0.00
	1.5	2	10	1.33 ± 0.00	0.98 ± 0.01	1.02 ± 0.00	0.03 ± 0.00
	1.5	2	100	1.36 ± 0.00	1.02 ± 0.03	1.01 ± 0.00	0.00 ± 0.00
	1.5	3	0.05	1.17 ± 0.01	0.64 ± 0.02	1.04 ± 0.01	0.01 ± 0.00
	1.5	3	0.1	1.15 ± 0.01	0.72 ± 0.15	0.96 ± 0.01	0.02 ± 0.00
	1.5	3	1	1.22 ± 0.01	0.90 ± 0.01	0.99 ± 0.00	0.02 ± 0.00
	1.5	3	10	1.26 ± 0.01	0.96 ± 0.03	1.03 ± 0.00	0.02 ± 0.00
	1.5	3	100	1.21 ± 0.01	0.93 ± 0.04	1.04 ± 0.00	0.00 ± 0.00
	1.5	6	0.05	1.35 ± 0.05	0.30 ± 0.03	1.05 ± 0.01	0.01 ± 0.00
	1.5	6	0.1	1.17 ± 0.04	0.50 ± 0.01	1.02 ± 0.03	0.00 ± 0.00
	1.5	6	1	1.15 ± 0.01	0.80 ± 0.03	1.03 ± 0.00	0.01 ± 0.00
	1.5	6	10	1.20 ± 0.02	0.81 ± 0.02	1.02 ± 0.00	0.06 ± 0.00
	1.5	6	100	1.17 ± 0.03	0.93 ± 0.02	1.04 ± 0.00	0.00 ± 0.00
	1.5	10	0.05	1.22 ± 0.05	0.14 ± 0.05	1.03 ± 0.00	0.01 ± 0.00
	1.5	10	0.1	1.28 ± 0.05	0.28 ± 0.08	1.04 ± 0.01	0.01 ± 0.00
	1.5	10	1	1.10 ± 0.05	0.71 ± 0.06	1.03 ± 0.00	0.01 ± 0.00
	1.5	10	10	1.06 ± 0.00	0.84 ± 0.01	1.03 ± 0.00	0.04 ± 0.00
	1.5	15	1	1.25 ± 0.00	0.67 ± 0.04	1.04 ± 0.00	0.01 ± 0.00
	1.5	15	10	1.17 ± 0.06	0.72 ± 0.07	1.02 ± 0.02	0.04 ± 0.03
	4	2	0.05	1.53 ± 0.02	0.55 ± 0.01	1.03 ± 0.00	0.01 ± 0.00
	4	2	0.1	1.39 ± 0.01	0.67 ± 0.03	1.02 ± 0.00	0.01 ± 0.00
	4	2	1	1.41 ± 0.07	0.81 ± 0.06	1.02 ± 0.00	0.01 ± 0.00
	4	2	10	1.31 ± 0.04	0.89 ± 0.01	1.02 ± 0.00	0.03 ± 0.00
	4	3	0.05	1.35 ± 0.04	0.52 ± 0.04	1.02 ± 0.00	0.03 ± 0.00
	4	3	0.1	1.29 ± 0.01	0.70 ± 0.01	1.02 ± 0.00	0.04 ± 0.00
	4	3	1	1.26 ± 0.01	0.84 ± 0.02	1.05 ± 0.00	0.02 ± 0.00
	4	3	10	1.36 ± 0.02	0.87 ± 0.01	1.02 ± 0.00	0.02 ± 0.00
	4	6	1	1.24 ± 0.01	0.76 ± 0.01	1.00 ± 0.00	0.03 ± 0.00
	4	6	10	1.18 ± 0.00	0.86 ± 0.03	1.00 ± 0.00	0.05 ± 0.00
	4	10	1	1.23 ± 0.02	0.70 ± 0.02	1.01 ± 0.00	0.02 ± 0.00
	8	2	1	1.56 ± 0.01	0.46 ± 0.02	0.99 ± 0.00	0.01 ± 0.00
	8	2	10	1.58 ± 0.04	0.47 ± 0.02	0.95 ± 0.00	0.02 ± 0.00