

Statistical Theory on the Functional Form of Cloud Particle Size Distributions

WEI WU^a

*Department of Atmospheric Sciences, University of Illinois at Urbana–Champaign, Urbana, Illinois,
and National Center for Atmospheric Research, Boulder, Colorado*

GREG M. MCFARQUHAR

*Cooperative Institute for Mesoscale Meteorological Studies, and School of Meteorology,
University of Oklahoma, Norman, Oklahoma*

(Manuscript received 30 May 2017, in final form 21 May 2018)

ABSTRACT

Several functional forms of cloud particle size distributions (PSDs) have been used in numerical modeling and remote sensing retrieval studies of clouds and precipitation, including exponential, gamma, lognormal, and Weibull distributions. However, there is no satisfying theoretical explanation as to why certain distribution forms preferentially occur instead of others. Intuitively, the analytical form of a PSD can be derived by directly solving the general dynamic equation, but no analytical solutions have been found yet. Instead of a process-level approach, the use of the principle of maximum entropy (MaxEnt) for determining the theoretical form of PSDs from the perspective of system is examined here. MaxEnt theory states that the probability density function with the largest information entropy among a group satisfying the given properties of the variable should be chosen. Here, the issue of variability under coordinate transformations that arises using the Gibbs–Shannon definition of entropy is identified, and the use of the concept of relative entropy to avoid these problems is discussed. Focusing on cloud physics, the four-parameter generalized gamma distribution is proposed as the analytical form of a PSD using the principle of maximum (relative) entropy with assumptions on power-law relations among state variables, scale invariance, and a further constraint on the expectation of one state variable (e.g., bulk water mass). The four-parameter generalized gamma distribution is very flexible to accommodate various type of constraints that could be assumed for cloud PSDs.

1. Introduction

Various functional forms of cloud particle size distributions (PSDs), such as exponential (Marshall and Palmer 1948), gamma (e.g., Borovikov et al. 1963; Ulbrich 1983), lognormal (e.g., Feingold and Levin 1986; Tian et al. 2010), and Weibull distributions (e.g., Zhang and Zheng 1994; Liu et al. 1995), have been used in numerical models and remote sensing retrieval algorithms. These functional forms of the cloud PSDs and the choice of free parameters characterizing the distribution have been typically determined based on what provides the best match to in situ observations. The scaling technique, as an alternative approach to describe cloud PSDs based on observational data, has been used

recently to derive parameters characterizing a PSD by assuming a limited number of degrees of freedom and a “universal distribution” without stating its exact functional form (e.g., Testud et al. 2001; Lee et al. 2004). Without considering the number of degrees of freedom needed to characterize a PSD, determining the functional form of the universal distribution used in the scaling approach is a challenging question. Although many different functional forms of cloud PSDs have been proposed, no study has yet provided an adequate theoretical explanation as to why a certain functional form is preferred over another. Therefore, the choice of functional form varies from study to study, complicating the comparison of PSD parameters derived from different field campaigns and from model parameterization schemes. It is not known if the choice of functional form should vary with environmental conditions.

A theoretical way to find an analytical form of a PSD is to solve the general dynamic equation describing the particle system, given by

^a Current affiliation: Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma.

Corresponding author: Wei Wu, weiwu@ou.edu

DOI: 10.1175/JAS-D-17-0164.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) (www.ametsoc.org/PUBSReuseLicenses).

$$\begin{aligned} \frac{\partial n(v,t)}{\partial t} = & -n(v,t) \int_0^{+\infty} K(v,u)n(u,t) du + \frac{1}{2} \int_0^v K(u,v-u)n(u,t)n(v-u,t) du + \int_0^{+\infty} L(v,u)n(u,t) du \\ & - \frac{n(v,t)}{v} \int_0^v uL(v,u) du + SC(v,t) - SK(v,t), \end{aligned} \quad (1)$$

where $n(v, t)$ is the number distribution function for particles with volume v at time t ; $K(v, u)$ and $L(u, v)$ are the collection kernel and breakup kernel for particles with volumes v and u ; $SC(v, t)$ is the source term; and $SK(v, t)$ is the sink term. All variables used in this paper are also defined in the [appendix](#). This form of the equation can be used for several different types of particles in a mixed particle system, such as ice particles with varying shapes and liquid particles. For the particle system of a single species (e.g., purely liquid clouds), one equation is sufficient. Unfortunately, Eq. (1) can only be solved analytically for constant, additive, or multiplicative kernels. Therefore, even for the simplest case of liquid clouds without nucleation, sedimentation, and breakup, no analytic form for a cloud PSD has been found when a geometric collection kernel is used ([Drake 1972](#)). When more complex processes acting in ice or mixed-phase clouds are considered (e.g., sublimation, aggregation, melting, riming, and deposition), the equation is even more difficult to solve, and an analytic solution cannot be contemplated at this time. Because analytic solutions have not been possible, numerical methods have been used to determine PSDs in bin-resolved models. However, the derived PSDs are very sensitive to even the representation of processes in liquid-phase clouds, such as the choice of raindrop breakup kernel ([Srivastava 1971, 1982](#); [List and McFarquhar 1990](#); [Hu and Srivastava 1995](#); [McFarquhar 2004](#)), with the collision-induced breakup parameterization determining the shape of modeled PSD. There are sensitivities to the representation of even more processes for ice or mixed-phase clouds.

A statistical theory is another viable way to determine the form of PSDs. Here, the mass or size of every particle is considered as a random variable acting under the influence of stochastic processes from a statistical perspective, even though each individual particle follows physical laws. An example is the use of statistical mechanics in the field of the thermodynamics where every molecule follows physical laws, but the collections of molecules are described by statistical properties (e.g., the temperature represents the average kinetic energy of molecules). One promising statistical theory for determining cloud PSDs is the principle of maximum entropy (MaxEnt; [Jaynes 1957a,b](#)). MaxEnt theory states that for a group of probability density functions (PDFs)

that satisfy given properties of the variable, the PDF with largest information entropy for this variable should be chosen. Thus, a uniform distribution function (most uncertain) is selected if no other properties are specified. But if the mean of the distribution is prescribed, the exponential distribution is the most probable distribution, following the same logic as used to derive the Maxwell–Boltzmann distribution in statistical mechanics. If both the mean and variance are prescribed, the most probable distribution is the normal distribution. The concept of MaxEnt has been used widely in physics (e.g., [Rose et al. 1990](#); [Antoniazzi et al. 2007](#)), mechanical engineering (e.g., [Sellens and Brzustowski 1985](#); [Li et al. 1991](#); [Berger et al. 1996](#)), image processing (e.g., [Wernecke and D’Addario 1977](#); [Skilling and Bryan 1984](#)), machine learning (e.g., [Rosenfeld 1996](#); [Berger et al. 1996](#)), ecology (e.g., [Phillips et al. 2004, 2006](#); [Banavar et al. 2010](#)), economics (e.g., [Cozzolino and Zahner 1973](#); [Buchen and Kelly 1996](#)), and even in atmospheric sciences for representing cloud microphysics (e.g., [Zhang and Zheng 1994](#); [Liu et al. 1995](#); [Yano et al. 2016](#)) and turbulent flows (e.g., [Majda and Wang 2006](#); [Craig and Cohen 2006](#); [Verkley and Lynch 2009](#); [Verkley 2011](#); [Verkley et al. 2016](#)). Its use in the study of spray PSDs in mechanical and material engineering is closely related to its use in the study of cloud PSDs. [Li and Tankin \(1987\)](#), [Dumouchel \(2006\)](#), and [Lecompte and Dumouchel \(2008\)](#) employed MaxEnt to derive analytical forms of spray PSDs, and [Déchelette et al. \(2011\)](#) have a comprehensive review on the application of MaxEnt to spray PSDs. Some applications of statistical mechanics may not state the principle of MaxEnt explicitly, but similar methods have been employed by [Griffith \(1943\)](#) to explain the particle size distribution in a comminuted system and by [Lienhard \(1964\)](#) to explain the unit hydrograph in hydrology. Thus, they are considered the same approach.

The problem of determining PSDs in cloud physics is very similar to the problems in these other fields. For numerical models simulating clouds with bulk microphysics schemes, only a number of moments of the PSD are predicted. For example, many schemes prognose the mass and number concentration. Other moments of a PSD are then calculated using the assumed form of the PSD and assumptions about various constants describing these distribution forms ([Thompson et al. 2004](#); [Morrison et al. 2005](#); [Seifert and Beheng 2006](#);

Morrison and Milbrandt 2015). These other moments include radar reflectivity and extinction. Thus, for developing parameterizations of cloud microphysics, there are some constraints on the properties of PSDs, exactly the type of scenario where MaxEnt can be used. Using MaxEnt, Zhang and Zheng (1994) and Liu et al. (1995) introduced the Weibull distribution as the analytical form of PSDs assuming constraints on the surface area and mass, respectively. Their derived PSD forms differ on the parameters characterizing the Weibull distribution due to their different assumptions. Yano et al. (2016) extended the assumptions on the PSDs to include constraints on the mean maximum dimension and sedimentation flux of droplets, and they examined the impact of these assumptions using idealized simulations, laboratory, and observational datasets. All prior studies applying MaxEnt to cloud PSDs used the Gibbs–Shannon form of entropy. However, the Gibbs–Shannon entropy is not invariant under coordinate transformation, and therefore different PSD forms can be derived using the same assumptions, as discussed in detail in section 3. To solve these problems, a different formalism of entropy is needed, as Jaynes (1963, 1968) noted.

This paper applies the form of (relative) entropy proposed by Jaynes (1963, 1968) to cloud PSDs. The problem of Gibbs–Shannon entropy is discussed in section 3 after a brief review of MaxEnt in section 2. Based on the form of (relative) entropy and several plausible assumptions about the cloud system, the four-parameter generalized gamma distribution is proposed as the most reasonable analytical form of cloud PSDs in section 4. The properties of the generalized gamma distribution are summarized in section 5. The applications of the four-parameter generalized gamma distribution to in situ observed liquid and ice cloud PSDs are investigated in section 6. The principle findings of the study and directions for future work are summarized in section 7.

2. MaxEnt and its rationale for cloud physics

MaxEnt theory was first proposed by Jaynes (1957a,b) to explain the classical Maxwell–Boltzmann distribution. The same principle has also been applied to Fermi–Dirac statistics and Bose–Einstein statistics and nonequilibrium statistical mechanics (Jaynes 1963, 1968; Dougherty 1994; Banavar et al. 2010). In statistical mechanics, it is assumed that if there are N_i particles in the i th energy state E_i , the total energy of the system E is given by

$$E = \sum_{i=1}^n N_i E_i, \quad (2)$$

where there are n total energy states with the total number of particles in the ensemble N , given by the

summation of all particles in each energy state expressed by

$$N = \sum_{i=1}^n N_i. \quad (3)$$

The number of microscopic configurations in which the N particles can be distributed over the n different energy states W is given by

$$W = \frac{N!}{N_1! N_2! \cdots N_n!}. \quad (4)$$

Boltzmann defined the entropy as $S_B = k_B \ln(W)$, where k_B is the Boltzmann constant (Pathria and Beale 2011). Greater W means a larger number of microscopic configurations of the N particles distributed over the n different energy states. S_B monotonically increases with W and is a measure of disorder: the greater the number of microscopic configurations in the system, the more uncertain the system can be. Using Sterling's formula

$$\ln(n!) = n \ln(n) - n + O[\ln(n)], \quad (5)$$

Boltzmann's entropy becomes

$$\begin{aligned} S_B &= k_B \ln(W) = k_B \left[\ln(N!) - \sum_{i=1}^n \ln(N_i!) \right] \\ &\approx -k_B N \sum_{i=1}^n \frac{N_i}{N} \ln\left(\frac{N_i}{N}\right) = -k_B N \sum_{i=1}^n p_i \ln(p_i) = NS, \end{aligned} \quad (6)$$

where $p_i = N_i/N$ is the probability of particles in the i th energy state, and $S = -k_B \sum_{i=1}^n p_i \ln(p_i)$ is Gibbs' form of entropy, which is the same form as Shannon's information entropy, except for the inclusion of the Boltzmann constant (Shannon 1948). Assuming that there is a solution, denoted by \bar{N}_i (or \bar{p}_i), that maximizes W and therefore S , it can be proven using Eq. (6) and the definition of Boltzmann entropy that

$$\frac{W_{\max}}{W} = e^{(N/k_B)(S_{\max} - S)}. \quad (7)$$

Since N is a very large number, and k_B is a very small number in the context of statistical mechanics, W_{\max} will be much larger than any other W achieved with other N_i , indicating any other p_i that deviates from \bar{p}_i has significantly fewer microscopic configurations. For example, for a mole of gas, there are N_A (Avogadro constant, $6.02 \times 10^{23} \text{ mol}^{-1}$) particles, and the large ratio of W_{\max} to other W rules out the possibility of other distributions. This is an important property of entropy.

The derivation of \bar{N}_i or \bar{p}_i is an optimization problem, $\arg \max_{N_i} \ln(W)$ subject to Eqs. (2) and (3), that can be derived using the method of Lagrange multipliers, where

$$d \ln(W) - \lambda_0 \left(\sum_{i=1}^n N_i - N \right) - \lambda_1 \left(\sum_{i=1}^n N_i E_i - E \right) = 0, \quad (8)$$

where λ_0 and λ_1 are the Lagrange multipliers. By using Stirling's approximation, Eq. (8) becomes

$$\begin{aligned} \sum_{i=1}^n -\ln(N_i) dN_i - \lambda_0 \left(\sum_{i=1}^n dN_i \right) - \lambda_1 \left(\sum_{i=1}^n dN_i E_i \right) \\ = \sum_{i=1}^n [-\ln(N_i) - \lambda_0 - \lambda_1 E_i] dN_i = 0, \end{aligned} \quad (9)$$

so that

$$\bar{p}_i = \frac{\bar{N}_i}{N} = C e^{-\lambda_1 E_i}, \quad \text{where } C = C_0 e^{-\lambda_0}, \quad (10)$$

which is the Maxwell–Boltzmann distribution (Pathria and Beale 2011). The Lagrange multipliers λ_0 and λ_1 can be obtained by substituting N_i in Eqs. (2) and (3).

Based on the above arguments, Jaynes (1957a,b) argued that for a group of PDFs that satisfy the given properties of a variable x , the PDF with largest information entropy (Shannon 1948) should characterize the variable, with statistical mechanics being just one example of this principle applied to an ideal gas. The methodology can be generalized using a continuous distribution to characterize the variable

$$\begin{aligned} \arg \max_{P(x)} - \int_0^{\infty} P(x) \ln P(x) dx \\ \text{subject to } (nc + 1) \text{ constraints: } \int_0^{\infty} f_k(x) P(x) dx = F_k, \end{aligned} \quad (11)$$

where $P(x)$ is the probability that state variable x will occur, and the $nc + 1$ constraints are expressed in the form of fixed expectation of $f_k(x)$ with $k = 0, 1, 2, \dots, nc$. For $k = 0$, $f_0(x) = 1$ and $F_0 = 1$ are chosen as the normalization conditions for the PDF. Since the zeroth constraint is valid for every PDF, only nc other constraints need to be given explicitly. Therefore, the number of given constraints is denoted as nc . The value of nc is determined by the knowledge of the system that is being considered and can vary according to the behavior of the particular system that is being modeled. To get the maximum of $S = -\int_0^{\infty} P(x) \ln P(x) dx$ with these constraints, the method of Lagrange multipliers can be

applied as before with the discrete sums so that the Lagrange function $L(x, \lambda_1, \lambda_2, \dots, \lambda_k)$ is expressed by

$$\begin{aligned} L(x, \lambda_1, \lambda_2, \dots, \lambda_k) \equiv - \int_0^{\infty} P(x) \ln P(x) \\ - \sum_{k=0}^{nc} \lambda_k \left[\int_0^{\infty} f_k(x) P(x) dx - F_k \right], \end{aligned} \quad (12)$$

where $k = 0, 1, 2, \dots, nc$. Then, the general result can be solved so that

$$P(x) = \frac{1}{Z(\lambda_1, \lambda_2, \dots, \lambda_{nc})} \exp \left[- \sum_{k=0}^{nc} \lambda_k f_k(x) \right], \quad (13)$$

where the partition function $Z(\lambda_1, \lambda_2, \dots, \lambda_{nc}) = \int \exp \left[- \sum_{k=0}^{nc} \lambda_k f_k(x) \right] dx \lambda_k f_k(x)$.

Following this technique, the exponential distribution can be derived as the maximum entropy distribution if only the mean of the variable is known. The Weibull distribution can be further derived as the maximum entropy distribution if the mean of the power function of the variable is known. If both the mean and variance of a variable are known, the normal distribution will be the maximum entropy distribution. Similarly, the lognormal distribution will be derived if the mean and variance of the logarithm of the variable are known. Kapur (1989) describes commonly used PDFs and their constraints.

Statistical mechanics can be used to define the properties of a cloud just as it is used to define the properties of an ideal gas. Just as in thermodynamics, where there are variables describing the microscopic and macroscopic states of the ideal gas, there are variables describing the microscopic and macroscopic properties of clouds as discussed previously, and the use of random variables is convenient so that statistical mechanics can be applied. The macroscopic states are mainly defined by the total number of cloud particles, the cumulative extinction (projected area) of all cloud particles, the bulk liquid or ice water content of all particles in a distribution, and other bulk microphysical properties. The microscopic states are described by the size, area, and mass of the individual hydrometeors. This is analogous to the case of an ideal gas, where even though there are numerous realizations of velocity for each individual gas molecule, the Maxwell–Boltzmann distribution, which has the largest entropy, has the largest number of microscopic configurations of molecule speeds and hence characterizes the distribution. Similarly, for clouds, the PDF with the maximum entropy also has the largest number of microscopic configurations of cloud particles distributed over different sizes. A key question in the application of statistical mechanics to distributions of

cloud particles is how many particles are needed to make the method robust because there are inevitably fewer cloud particles than gas molecules. If it is assumed that the total number concentration is N_t , then the total number of cloud particles in a sample volume V is $N = N_t V$. The volume should be sufficiently large to make N large, but at the same time not so large to exceed the typical volume of a cloud or a scale where there is a lot of horizontal or vertical inhomogeneity. Here, a unit cloud volume (V) of $100\text{ m} \times 100\text{ m} \times 10\text{ m} = 10^5\text{ m}^3$ is proposed as large enough. Assuming a concentration of $N_t \approx 100\text{ cm}^{-3}$, then $N = N_t V = 10^{13}$ should be big enough to make the derivation solid. This volume is also small enough, compared to typical model grid volume or radar sample volumes.

In cloud physics, the number distribution function is expressed as $N(D)$, which can be normalized by $N_t = \int_0^\infty N(D) dD$ to define the number distribution probability density function expressed by

$$P(D) = \frac{N(D)}{N_t}. \tag{14}$$

Thus, the MaxEnt approach can be applied in the study of cloud PSDs, and its use in cloud physics has been discussed by Zhang and Zheng (1994), Liu et al. (1995), and Yano et al. (2016). However, there are problems directly applying MaxEnt to cloud PSDs, as discussed in section 3. Previous studies chose the particle maximum dimension (D) or particle mass (m) as the state variable x , with all assuming two constraints: 1) the constraint of total particle number concentration and 2) a constraint of mean maximum dimension (Yano et al. 2016), total surface area (Zhang and Zheng 1994), total bulk water content (Liu et al. 1995), or mass flux (Yano et al. 2016). The constraints apply to bulk properties, which are integrations of particle properties over size. The derived PSD forms maximizing the entropy are then special cases of Eq. (13), with $nc = 1$, expressed by

$$P(x) = \frac{1}{Z(\lambda)} \exp[-\lambda_1 f_1(x)], \tag{15}$$

where x could be D or m , and $f_1(x)$ could be D , A , m , or $m\nu$ (ν is the fall speed of a particle). The $f_1(x)$ is typically written as a power-law function of x . An example for $f_1(x)$ is αD^β , where x is D . Note that the PDF as a function of D and the PDF as a function of any other variable x can be converted, so that the number distribution function can be expressed:

$$N(D) = N_t P(D) = N_t P(x) \frac{dx}{dD} = \frac{N_t}{Z(\lambda)} \exp[-\lambda_1 f_1(x)] \frac{dx}{dD}. \tag{16}$$

Usually, any other state variable x and the particle maximum dimension D are assumed to be related through a power law (e.g., $x = aD^b$), so that Eq. (16) can be rewritten as

$$N(D) = \frac{N_t a b}{Z(\lambda)} D^{b-1} \exp[-\lambda_1 f_1(aD^b)]. \tag{17}$$

3. Problems using Gibbs–Shannon entropy and the concept of relative entropy

Equation (17) is a general solution for the functional form of cloud PSDs maximizing the entropy content as long as one constraint is given explicitly. However, it can be proven that a different PSD form can be derived using the same constraint when using the Gibbs–Shannon entropy. For example, below, it is shown that the same constraints used in Liu et al. (1995) can be employed to derive a different PSD than the one derived in Liu et al. (1995). It should be noted that all the forms derived in Zhang and Zheng (1994) and Yano et al. (2016) suffer the same problem. Assuming that the total bulk number concentration N_t and total bulk water mass content (TWC) are constraints, and using mass m as the variable characterizing particles, Liu et al. (1995) showed that the MaxEnt distribution was given by

$$N(m) = C_1 \exp(-\lambda_1 m), \tag{18}$$

where $C_1 = N_t^2/\text{TWC}$ and $\lambda_1 = N_t/\text{TWC}$ are the distribution parameters. This distribution can be rewritten in terms of D using an assumed mass–dimensional relation $m = \alpha D^\beta$ as

$$N(D) = \bar{C}_1 D^{\beta-1} \exp(-\bar{\lambda}_1 D^\beta), \tag{19}$$

where $\bar{C}_1 = \alpha \beta N_t^2/\text{TWC}$ and $\bar{\lambda}_1 = \alpha N_t/\text{TWC}$ are the distribution parameters.

However, if the PDF is characterized in terms of D instead, and the same two constraints are applied, as expressed by

$$\int_0^\infty N(D) dD = \int_0^\infty N_t P(D) dD = N_t, \tag{20}$$

$$\int_0^\infty \alpha D^\beta N(D) dD = \int_0^\infty \alpha D^\beta N_t P(D) dD = \text{TWC}, \tag{21}$$

the MaxEnt distribution becomes

$$N(D) = \tilde{C}_1 \exp(-\tilde{\lambda}_1 D^\beta), \tag{22}$$

where $\tilde{C}_1 = [N_t \beta (\alpha N_t / \beta \text{TWC})^{1/\beta}] \Gamma(1/\beta)$ and $\tilde{\lambda}_1 = \alpha N_t / \beta \text{TWC}$ are the distribution parameters. By comparing Eqs. (19)

and (22), it is found that two different analytical forms of PSDs are derived using the same assumption. In fact, a different analytical form of the PSD can be derived whenever the state variable x characterizing the cloud particles changes. This is because the Gibbs–Shannon entropy is not invariant under the transformation of variables (Jaynes 1963, 1968). Thus, Jaynes (1963, 1968) proposed another definition of entropy, typically called relative entropy that makes entropy invariant under variable transformations. The relative entropy is mathematically sound and physically meaningful, as discussed below.

The definition of relative entropy proposed by Jaynes (1963, 1968) S_r is expressed by

$$S_r = - \int_0^{\infty} P(x) \ln \frac{P(x)}{I(x)} dx, \quad (23)$$

where $I(x)$ is called the invariant measure, or a prior distribution that represents an initial guess of what the distribution should be. This relative entropy S_r has also been called Kullback–Leibler divergence, which is a measure of how a PDF diverges from a prior distribution. When $I(x)$ is a uniform distribution, the definition of relative entropy is identical to the Gibbs–Shannon entropy minus a constant. For systems where coordinate transforms are important, the uniform distribution is not a good prior distribution. Therefore, a form that is invariant under coordinate transforms needs to be used for the generalized development. It can be shown that S_r is invariant under coordinate transformation [$x \rightarrow y$, where $y = g(x)$] because

$$S_r = - \int_0^{\infty} P'(y) \ln \frac{P'(y)}{I'(y)} dy = - \int_0^{\infty} P(x) \ln \frac{P(x)}{I(x)} dx, \quad (24)$$

with $P'(y) = P(x) dx/dy$ and $I'(y) = I(x) dx/dy$.

To maximize S_r with given constraints, the method of Lagrange multipliers is again used so that

$$L \equiv - \int_0^{\infty} P(x) \ln \frac{P(x)}{I(x)} - \sum_{k=1}^{\text{nc}} \lambda_k \left[\int_0^{\infty} f_k(x) P(x) dx - F_k \right], \quad (25)$$

where $k = 0, 1, 2, \dots, \text{nc}$, and the maximum (relative) entropy distribution is solved in the form

$$P(x) = \frac{1}{Z(\lambda_1, \lambda_2, \dots, \lambda_{\text{nc}})} I(x) \exp \left[- \sum_{k=0}^{\text{nc}} \lambda_k f_k(x) \right], \quad (26)$$

where the new partition function is

$$Z(\lambda_1, \lambda_2, \dots, \lambda_{\text{nc}}) = \int_0^{\infty} I(x) \exp \left[- \sum_{k=0}^{\text{nc}} \lambda_k f_k(x) \right] dx.$$

4. Application to cloud PSDs

The relative entropy is invariant under coordinate transformations, and the distribution derived maximizing this definition of entropy is consistent with the same constraint regardless of the variable used to characterize the PDF. However, before the theory can be applied to any system, the appropriate constraints and invariant measure $I(x)$ are needed. These can only be obtained from an understanding of the system studied. To apply the theory to cloud physics, the first step is to determine the constraints for a cloud. Yano et al. (2016) used observed and simulated datasets to evaluate constraints of mean maximum dimension, bulk extinction, bulk water content, and bulk mass flux. This paper does not examine the use of different constraints, as did Yano et al. (2016), but instead focuses on the general application of the new definition of entropy. Unlike the Gibbs–Shannon entropy used in previous studies, the choice of state variable x is not important for S_r , as the invariant measure $I(x)$ will adjust accordingly. In this study, the particle maximum dimension (D) is chosen as the state variable x of the cloud. The number distribution function $N(D)$, following Eq. (26), can thus be expressed by

$$N(D) = \frac{N_t}{Z(\lambda_1, \lambda_2, \dots, \lambda_{\text{nc}})} I(D) \exp \left[- \sum_{k=0}^{\text{nc}} \lambda_k f_k(D) \right], \quad (27)$$

where the number of constraints is usually equal or larger than 1. Equation (27) is the general form of $N(D)$, and the number of constraints, as well as the corresponding constraint function $f_k(D)$, need to be assumed to derive a specific form for $N(D)$. For this study, one constraint is assumed. In future studies, more than one constraint can be used if the solution using one constraint is not well validated against observations. If it is assumed that there is just one constraint and this constraint function can be represented as a power law with particle maximum dimension [$f_1(D) = aD^b$] following Zhang and Zheng (1994), Liu et al. (1995), and Yano et al. (2016), Eq. (27) then becomes

$$N(D) = \frac{N_t}{Z(\lambda_1)} I(D) e^{-\lambda_1 a D^b} = \frac{N_t}{Z(\lambda_1)} I(D) e^{-\lambda D^b}, \quad (28)$$

where $\lambda = \lambda_1 a$.

The next step applying MaxEnt theory is to determine the invariant measure $I(D)$, which must be provided by a knowledge of the underlying physics. Jaynes (1968) provided guidelines to choose the invariant measure based on the transformation group, and Jaynes (1973) showed an example using the transformation group. The basic idea is that the shape of the invariant measure should be invariant in two different systems. In particular

for this case, the shape of the invariant measure should not change with the volume of cloud studied. Assume two volumes of the same cloud: cloud A with total volume V_A and cloud B , a subset of the cloud A , with total volume $V_B = \kappa_1 V_A$, $\kappa_0 < \kappa_1 \leq 1$. Hereafter, the properties of volume A and volume B are denoted with the subscripts A and B , respectively. Here, κ_1 cannot be too small since a large number of particles is needed for the application of statistical mechanics; hence, κ_0 is used for a lower bound instead of 0. Here, $\kappa_0 = 0.1$ should be large enough, which will give the number of cloud particles in volume B around 10^{12} if volume A is as discussed in section 2.

For volume V_A , the total mass is $\text{TWC} \times V_A$. Therefore, no particles larger than $D_{\max A}$ are possible, where $\alpha D_{\max A}^\beta = \text{TWC} \times V_A$ with α and β the m - D relation parameters with $m = \alpha D^\beta$, the mass of an individual cloud particle. Thus, $I_A(D) = 0$ for $D > D_{\max A}$. For cloud A , the prior probability is $I_A(D)$ with $\int_0^{D_{\max A}} I_A(D) dD = 1$. For cloud B , the volume will be $V_B = \kappa_1 V_A$, the maximum particle size will be $D_{\max B} = \kappa D_{\max A}$ ($\kappa_1 = \kappa_1^{1/\beta}$), and the prior probability $I_B(D)$ satisfies $\int_0^{D_{\max B}} I_B(D) dD = 1$. A new-scaled dimensionless variable $x = D/D_{\max A}$ is defined to scale $I_A(D)$ into the range $[0, 1]$ so that

$$\begin{aligned} \int_0^{D_{\max A}} I_A(D) dD &= \int_0^1 I_A(xD_{\max A}) d(xD_{\max A}) \\ &= \int_0^1 D_{\max A} I_A(xD_{\max A}) dx \\ &= \int_0^1 f_A(x) dx = 1, \end{aligned} \tag{29}$$

where $f_A(x) = D_{\max A} I_A(xD_{\max A})$ and, similarly, $f_B(y) = D_{\max B} I_B(yD_{\max B})$, where $y = D/D_{\max B}$. Because of scale invariance, the scaled PDFs $f_A(x)$ and $f_B(y)$ over the same range of $[0, 1]$ should be the same, meaning that

$$\begin{aligned} f_A(x) &= f_B(x) \\ \text{or } D_{\max A} I_A(xD_{\max A}) &= D_{\max B} I_B(xD_{\max B}) \\ \text{or } I_A(D) &= \kappa I_B(\kappa D). \end{aligned} \tag{30}$$

This is the scale invariance that the cloud system must satisfy in order for two different volumes to have the same shape of invariant measure. Following the formula for conditional probability, for any D that is within the range of $[0, \kappa D_{\max A}]$, it can be shown that

$$I_A(D) = I_B(D) \int_0^{\kappa D_{\max A}} I_A(u) du. \tag{31}$$

Equation (31) is the standard conditional probability formula and will hold whether or not any transformation invariance is assumed.

Combining the invariance requirement Eq. (30) and the conditional probability relation Eq. (31), it is determined that

$$\kappa I_A(\kappa D) = I_A(D) \int_0^{\kappa D_{\max A}} I_A(u) du. \tag{32}$$

Differentiating Eq. (32) with respect to κ gives another form of Eq. (32):

$$I_A(\kappa D) + \kappa \frac{\partial I_A(\kappa D)}{\partial D} D = I_A(D) I_A(\kappa D_{\max A}) D_{\max A}. \tag{33}$$

This equation is still hard to solve since it involves both cloud A and cloud B and, therefore, one parameter κ . Remember that cloud B is a subset of cloud A . By setting $\kappa = 1$ to make cloud A and cloud B the same, the equation for one cloud (cloud A is chosen here, but it is the same to choose cloud B to solve first) yields

$$\begin{aligned} I_A(D) + \frac{\partial I_A(D)}{\partial D} D &= I_A(D) I_A(D_{\max A}) D_{\max A} \\ \rightarrow \frac{\partial I_A(D)}{\partial D} D &= [I_A(D_{\max A}) D_{\max A} - 1] I_A(D). \end{aligned} \tag{34}$$

Solving the differential equation Eq. (34), it can be shown that the most general solution is

$$I_A(D) = \frac{\mu + 1}{D_{\max A}^{\mu+1}} D^\mu, \tag{35}$$

where μ , defined as $I_A(D_{\max A}) D_{\max A} - 1$, is a constant in the range of $-1 < \mu < \infty$. The constant μ cannot be further determined by scale invariance. Using Eq. (30), the invariant measure of cloud B is

$$I_B(D) = \frac{\mu + 1}{(\kappa D_{\max A})^{\mu+1}} D^\mu = \frac{\mu + 1}{D_{\max B}^{\mu+1}} D^\mu. \tag{36}$$

The form of invariant measure provided by Eq. (35) [or Eq. (36)] satisfies translational, rotational, and scale transformations, typical transformations between coordinate systems as suggested by Jaynes (1968). In this case, PSDs are described with a single dimension in a space related to particle maximum dimension, so no rotational transformation exists. Since no spatial variables are involved, the PDF does not change when the coordinate system is translated. Scale transformation is the last transformation to be satisfied. For two coordinate systems R and S , the length relates by $\bar{\kappa}$ with $D_R = \bar{\kappa} D_S$, the invariant measure for R is $I_R(D_R)$ with $\int_0^{D_{\max R}} I_R(D_R) dD_R = 1$, and the invariant measure for S is $I_S(D_S)$ with $\int_0^{D_{\max S}} I_S(D_S) dD_S = 1$. Since it is the same cloud observed, the relation $I_R(D_R) dD_R = I_S(D_S) dD_S$

holds, which means $\bar{\kappa}I_R(\bar{\kappa}D_S) = I_S(D_S)$. Equation (35) clearly satisfies this relation. So far, the invariant measure provided by Eq. (35) satisfies all the Abelian group transformations proposed by Jaynes (1968).

If Eq. (35) is assumed to represent the invariant measure, combined with Eq. (28), the final $N(D)$ is the four-parameter generalized (or modified) gamma distribution, given by

$$N(D) = N_0 D^\mu e^{-\lambda D^b}, \quad (37)$$

where $N_0 = N_r C / Z(\lambda_1)$.

It should be noted that the derived PSD forms from Zhang and Zheng (1994), Liu et al. (1995), and Yano et al. (2016) are all special cases of Eq. (37), so that this study is consistent with, but more general than, previous studies. It is also clear now why the two approaches to derive the PSD form in section 3 generate different results. Equations (19) and (22) differ by D^μ , which is the invariant measure. The first approach assumed a uniform-invariant measure over particle mass, and the second assumed a uniform-invariant measure over particle size, and $dm/dD = \alpha \beta D^{\beta-1}$ is the difference.

5. Properties of generalized gamma distribution

The properties of the generalized gamma distribution are summarized in this section. The generalized (or modified) gamma distribution is a general form of a PSD, which can be simplified to an exponential, gamma, or Weibull distribution in special cases. To the authors' knowledge, the generalized gamma distribution was first proposed by Amoroso (1925) to study income distribution and later independently proposed by Nukiyama and Tanasawa (1939) for fitting the size distribution of sprays particles in mechanical and material engineering. Stacy (1962) studied the mathematical properties of the generalized gamma distribution, and the properties relevant to cloud PSDs are summarized here.

The cumulative distribution function for the generalized (modified) gamma distribution in the form of Eq. (37) is

$$F(D; N_0, \mu, \lambda, b) = \frac{N_0}{b\lambda \Gamma\left(\frac{\mu+1}{b}\right)} \gamma\left(\frac{\mu+1}{b}, \lambda D^b\right), \quad (38)$$

where $\gamma(s, x) = \int_0^x t^{s-1} e^{-t} dt$ is the lower incomplete gamma function. The n th moment can be calculated as

$$M_n = E(D^n) = \frac{N_0}{b\lambda^{(\mu+1+n)/b+1}} \Gamma\left(\frac{\mu+1+n}{b}\right). \quad (39)$$

For any variable x that is related to D through a power law (e.g., $x = cD^d$), it can also be represented by a generalized gamma distribution with the form

$$N(x) = \frac{N_0}{c^{(\mu+1)/d}} x^{(\mu+1)/d-1} e^{-(\lambda/c^{b/d})(x^{b/d})}. \quad (40)$$

One main benefit of the four-parameter generalized gamma distribution is that it is invariant under coordinate transformations when characterizing a PSD. The same form applies to all power-law variables, such as particle maximum dimension, area, and mass. The lognormal distribution also has this property, and this is one of the reasons Feingold and Levin (1986) recommended the lognormal distribution for PSDs. This property is not shared by the exponential distribution, gamma distribution, or Weibull distribution. For example, Seifert and Beheng (2006) assumed the commonly used three-parameter gamma distribution over mass, which will turn into a four-parameter generalized gamma distribution over particle size. A second benefit is that the generalized gamma distribution can also simplify to a gamma distribution, Weibull distribution, or even exponential distribution under certain circumstances. Third, the physical meaning of distribution parameters is clearer than parameters used in some empirical distribution functions used in previous studies. Because of the properties mentioned above, Maur (2001) and Petty and Huang (2011) also proposed the use of generalized (or modified) gamma distribution without stating the underlying physical basis.

6. Testing with in situ observed liquid and ice PSDs

In this section, in situ observed PSDs are fit to different analytical forms, including the gamma, Weibull, lognormal, and generalized gamma distributions. The fitting in this section is used to test the application of four-parameter generalized gamma distribution in real clouds.

An in situ dataset collected by a two-dimensional cloud probe (2DC) and high sample volume spectrometer (HVPS) during the Midlatitude Continental Convective Clouds Experiment (MC3E; Jensen et al. 2016) is used for the fitting. Wu and McFarquhar (2016) describe how the data were collected and how the binary data were processed to generate cloud PSDs. Two different distributions were used in the analysis: a 1-min time period in liquid clouds and another 1-min period in ice clouds. The particle images were all manually checked to make sure no mixed-phase time periods existed in these two time periods. Liquid PSDs measured between 1320:00 and 1320:59 UTC at a temperature of around 4°C are averaged, and the best fits to the different analytical functions listed in the legend of Fig. 1 were performed. Following McFarquhar et al. (2015), the fitting technique

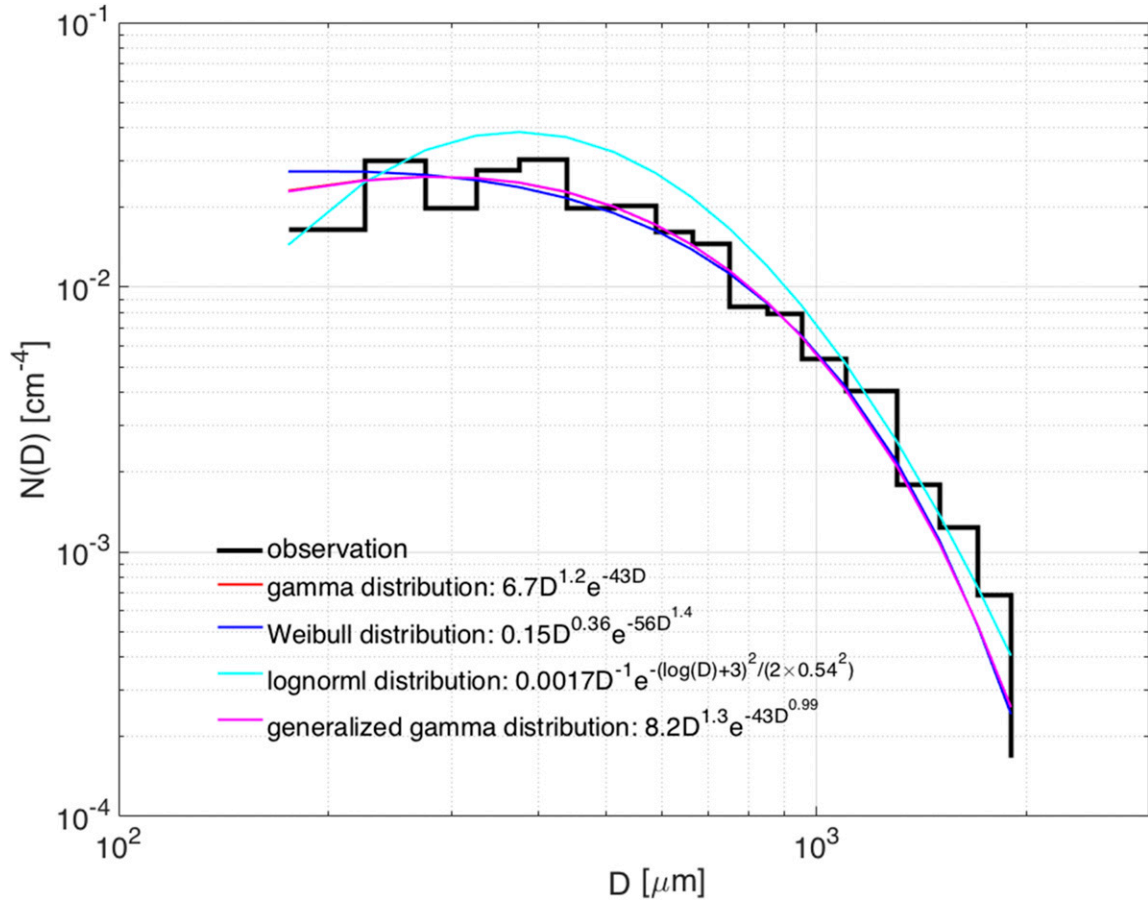


FIG. 1. Sample in situ liquid PSD $N(D)$ as function of D (black) and fitted for gamma distribution (red), Weibull distribution (blue), lognormal distribution (cyan), and generalized gamma distribution (purple). The red curve is right under the purple curve. The fitted parameters are listed in the legend.

minimized the χ^2 difference between the fit and observed moments of $N(D)$, defined by

$$\chi^2 = \sum_{i=1}^{nm} \left(\frac{M_{\text{fit},i} - M_{\text{obs},i}}{\sqrt{M_{\text{fit},i} M_{\text{obs},i}}} \right)^2, \quad (41)$$

where $M_{\text{obs},i}$ is the i th moment of the observed PSD, and $M_{\text{fit},i}$ is the i th moment of the fit PSD calculated using the assumed PSD form. Here the zeroth, third, and sixth moments corresponding to total number concentration, bulk liquid water content, and radar reflectivity were used in the fitting procedure to determine the parameters describing the gamma, Weibull, and lognormal distributions. To determine the parameters of the generalized gamma distribution, the first moment, representing the mean particle size, was also used because four moments are required to describe the four parameters of the generalized gamma distribution. All fit functions had $\chi^2 \leq 0.001$ in Eq. (41), showing all fits

provide good agreement between fit and measured moments. Further, the fit gamma, Weibull, and generalized gamma distributions all appear similar to the observed PSD, while the lognormal fit seems to deviate further from the observed PSDs. The fit generalized gamma distribution has a b parameter very close to 1 (0.99), so the fit curve is very close to the gamma distribution. This implies that the mean maximum dimension is the constraint for the liquid clouds in this time period.

Fits to the PSD measured in ice clouds from 1555:00 to 1555:59 UTC at a temperature of around -10°C from the same flight were also conducted with the zeroth, second, and fourth moments used to determine the fit parameters. For ice clouds, these approximately correspond to the total number concentration, bulk ice water content, and radar reflectivity, respectively. Similarly, an additional moment, the first moment, was used to find the generalized gamma distribution fit parameters. Figure 2 shows the results of the fits that

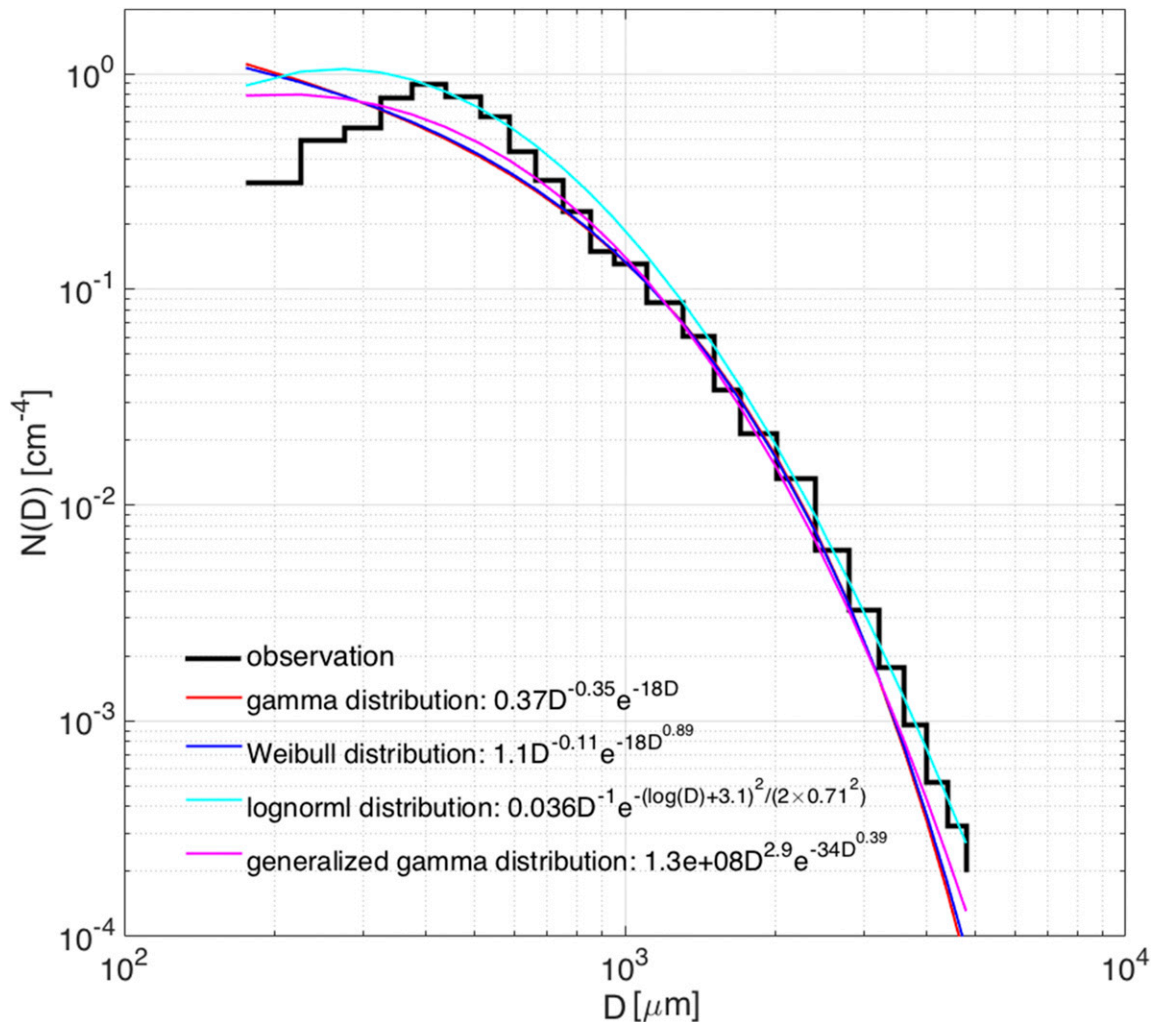


FIG. 2. As in Fig. 1, but for ice PSDs.

were performed. The b parameter in the generalized gamma distribution is 0.39, and the fit curve is closer to the observed PSDs, compared to the gamma distribution and Weibull distribution.

7. Conclusions and discussion

Several analytical forms of cloud PSDs, such as exponential and gamma distribution functions, have been assumed in numerical models and remote sensing retrievals in past studies. However, no satisfying physical basis has yet been provided for why any of these functions characterize PSDs. The use of the principle of maximum entropy (MaxEnt) to find analytical forms of PSDs was examined here, building upon its use in prior studies (Zhang and Zheng 1994; Liu et al. 1995; Yano et al. 2016). The main findings of this study are summarized as follows:

- 1) The definition of relative entropy $S_r = -\int_0^\infty P(x) \ln[P(x)/I(x)] dx$, which is invariant under coordinate transformations, was used to resolve an inconsistency in previous studies. The previous use of Gibbs–Shannon entropy allowed different PSDs to be derived using the same constraint by simply using a different state variable x .
- 2) The definition of relative entropy used in this study to determine a physical basis for a cloud PSD requires an assumption about an invariant measure $I(D)$, which is obtained from a physical understanding of the system studied. Here, it was shown that $I(D)$ can be obtained if invariance regarding group transformation is assumed.
- 3) Assuming that the microscopic state variables that characterize the properties of cloud particles (e.g., particle maximum dimension, area, mass, and fall speed) are related to each other through power laws,

it was shown that if one constraint related to any state variable was assumed, a four-parameter generalized gamma distribution can be derived. The state variable that needs to be used as a constraint is not yet well determined.

- 4) It was shown that if one state variable follows the generalized gamma distribution, all state variables having power-law relations with the state variable must also follow the generalized gamma distribution.
- 5) Directly fitting in situ observed PSDs using data obtained from optical array probes (OAPs) generates reasonable fits to the observed PSDs for all the analytical forms of PSD, even though the fit of generalized gamma distribution is slightly better. Because of the discrete nature of observed PSDs and large uncertainties for OAPs, parameters derived by directly fitting have large uncertainties.

Although the MaxEnt approach provides a physical basis for the form of the generalized four-parameter gamma distribution, it does not determine the values of parameters (N_0 , μ , λ , and b). These can only be determined using observational datasets. Among the four parameters, b is particularly interesting, since it implicitly implies what the constraint for the system is. [Yano et al. \(2016\)](#) provide a good approach to examine the assumptions of constraint (and therefore the value of b) using observational data.

It should be noted that the generalized gamma distribution is derived when only one constraint of the power function of particle dimension is used. It is possible that more than one constraint exists or that the constraint functions $f_k(D)$ cannot be represented as power laws. Either way, the more general form of cloud PSD [Eq. (26)] can be used in such circumstances. The full potential of the MaxEnt for cloud physics applications will be realized after more understanding of the physical systems is gained. The development of idealized models to simulate the evolution of cloud particles can also provide another perspective, from which the application of MaxEnt may provide more theoretical basis on the appropriate constraint for the system that should be used.

Acknowledgments. The authors are supported by the office of Biological and Environmental Research (BER) of the U.S. Department of Energy Atmospheric Systems Research Program through Grants DE-SC0014065 and DE-SC0016476 (through UCAR subcontract SUBAWD000397) and by the National Science Foundation (NSF) under Grants AGS-1213311 and AGS-1762096. The discussions with Hugh Morrison and Lulin Xue and the comments of three anonymous reviewers improved the quality of this paper considerably.

APPENDIX

List of Variables and Their Definitions

The variables used in this study are defined and summarized below.

α	Prefactor of m - D relations
β	Power factor of m - D relations
$\gamma(s, x)$	Lower incomplete gamma function
$\Gamma(x)$	Gamma function
κ	Scale factor between two lengths in two clouds
κ_0	Lower limit of κ
κ_1	Scale factor between two volumes in two clouds
$\bar{\kappa}$	Scale factor between two coordinate systems
λ	The slope parameter in generalized gamma distribution in Eq. (37)
λ_1	The Lagrange multiplier for the first constraint
$\bar{\lambda}_1$	The Lagrange multiplier relating to λ_1 by $\bar{\lambda}_1 = \alpha\lambda_1$
$\tilde{\lambda}_1$	The Lagrange multiplier relating to λ_1 by $\tilde{\lambda}_1 = (\alpha/\beta)\lambda_1$
λ_k	The Lagrange multiplier for the k th constraint
μ	The shape parameter in generalized gamma distribution in Eq. (37)
ρ	Particle density
χ^2	The measure of goodness for a fit in chi-square statistic
a	Prefactor of a general power-law relations
A	The projected area of a cloud particle
b	Power factor of a general power-law relation in generalized gamma distribution in Eq. (37)
C	The constant that relates to λ_1 through $C = C_0 \exp(-\lambda_0)$ in Eq. (10)
C_0	The constant in Eq. (10)
C_1	Constant in Eq. (18)
\bar{C}_1	Constant in Eq. (19)
\tilde{C}_1	Constant in Eq. (22)
D	The maximum dimension of a cloud particle
e	Euler's number, which equals approximately 2.71 828
E_i	The i th kinetic energy state
E	Total kinetic energy of the particle system
$f_A(x)$	The scaled invariant measure for $I_A(x)$

$f_B(x)$	The scaled invariant measure for $I_B(x)$
$f_k(x)$	The k th constraint as a function of x
F_k	The expected value of $f_k(x)$
IWC	Ice water content
$I(x)$	The invariant measure
$I_A(x)$	The invariant measure for cloud A
$I_B(x)$	The invariant measure for cloud B
$I_R(x)$	The invariant measure for coordinate system R
$I_S(x)$	The invariant measure for coordinate system S
k	Constraint number
$L(x, \lambda_1, \lambda_2, \dots, \lambda_n)$	Lagrangian function
LWC	Liquid water content
m	The mass of a cloud particle
n	Number of energy state in the ideal gas system
nc	The number of constraints
nm	The number of moments used for fitting
$M_{\text{obs},i}$	The i th moment of the observed PSD
$M_{\text{fit},i}$	The i th moment of the fit PSD
N	Total number of ideal gas molecules
N_0	Generalized gamma distribution parameter in Eq. (37)
$N(D)$	Number distribution function over size
$N(m)$	Number distribution function over mass
N_i	Total number of ideal gas molecules in energy state E_i
N_t	Total number concentration
P_i	Probability of ideal gas particles in energy state E_i
$P(x)$	Probability of x state
S	Gibbs–Shannon entropy
S_B	Boltzmann entropy
S_r	Relative entropy
TWC	Total water content
ν	Cloud particle fall speed
W	The multiplicity representing the number of microscopic configurations
x	A random state variable that describes the cloud particle
$Z(\lambda_1, \lambda_2, \dots, \lambda_n)$	Partition function

REFERENCES

- Amoroso, L., 1925: Ricerche intorno alla curva dei redditi. *Ann. Mat. Pura Appl.*, **2**, 123–159, <https://doi.org/10.1007/BF02409935>.
- Antoniazzi, A., D. Fanelli, J. Barré, P.-H. Chavanis, T. Dauxois, and S. Ruffo, 2007: A maximum entropy principle explains quasi-stationary states in systems with long-range interactions: The example of the Hamiltonian Mean Field model. *Phys. Rev. E*, **75**, 011112, <https://doi.org/10.1103/PhysRevE.75.011112>.
- Banavar, J. R., A. Maritan, and I. Volkov, 2010: Applications of the principle of maximum entropy: From physics to ecology. *J. Phys. Condens. Matter*, **22**, 063101, <https://doi.org/10.1088/0953-8984/22/6/063101>.
- Berger, A. L., V. J. D. Pietra, and S. A. D. Pietra, 1996: A maximum entropy approach to natural language processing. *Comput. Lang.*, **22** (1), 39–71.
- Borovikov, A. M., I. I. Gaivoronskii, E. G. Zak, V. V. Kostarev, I. P. Mazin, V. E. Minervin, A. K. Khrgian, and S. M. Shmeter, 1963: *Fizika Oblakov (Cloud Physics)*. Israel Program for Scientific Translations, 392 pp.
- Buchen, P. W., and M. Kelly, 1996: The maximum entropy distribution of an asset inferred from option prices. *J. Financ. Quant. Anal.*, **31**, 143–159, <https://doi.org/10.2307/2331391>.
- Cozzolino, J. M., and M. J. Zahner, 1973: The maximum-entropy distribution of the future market price of a stock. *Oper. Res.*, **21**, 1200–1211, <https://doi.org/10.1287/opre.21.6.1200>.
- Craig, G. C., and B. G. Cohen, 2006: Fluctuations in an equilibrium convective ensemble. Part I: Theoretical formulation. *J. Atmos. Sci.*, **63**, 1996–2004, <https://doi.org/10.1175/JAS3709.1>.
- Déchelette, A., E. Babinsky, and P. Sojka, 2011: Drop size distributions. *Handbook of Atomization and Sprays*, N. Ashgriz, Ed., Springer, 479–495, https://doi.org/10.1007/978-1-4419-7264-4_23.
- Dougherty, J. P., 1994: Foundations of non-equilibrium statistical mechanics. *Phil. Trans. Roy. Soc.*, **346A**, 259–305, <https://doi.org/10.1098/rsta.1994.0022>.
- Drake, R., 1972: A general mathematics survey of the coagulation equation. *Topics in Current Aerosol Research*, G. M. Hidy and J. R. Brock, Eds., Pergamon Press, 201–376.
- Dumouchel, C., 2006: A new formulation of the maximum entropy formalism to model liquid spray drop-size distribution. *Part. Part. Syst. Charact.*, **23**, 468–479, <https://doi.org/10.1002/ppsc.200500989>.
- Feingold, G., and Z. Levin, 1986: The lognormal fit to raindrop spectra from frontal convective clouds in Israel. *J. Climate Appl. Meteor.*, **25**, 1346–1363, [https://doi.org/10.1175/1520-0450\(1986\)025<1346:TLFTRS>2.0.CO;2](https://doi.org/10.1175/1520-0450(1986)025<1346:TLFTRS>2.0.CO;2).
- Griffith, L., 1943: A theory of the size distribution of particles in a comminuted system. *Can. J. Res.*, **21**, 57–64, <https://doi.org/10.1139/cjr43a-005>.
- Hu, Z., and R. Srivastava, 1995: Evolution of raindrop size distribution by coalescence, breakup, and evaporation: Theory and observations. *J. Atmos. Sci.*, **52**, 1761–1783, [https://doi.org/10.1175/1520-0469\(1995\)052<1761:EORSDB>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<1761:EORSDB>2.0.CO;2).
- Jaynes, E. T., 1957a: Information theory and statistical mechanics. *Phys. Rev.*, **106**, 620, <https://doi.org/10.1103/PhysRev.106.620>.
- , 1957b: Information theory and statistical mechanics. II. *Phys. Rev.*, **108**, 171, <https://doi.org/10.1103/PhysRev.108.171>.
- , 1963: Information theory and statistical mechanics (notes by the lecturer). *Statistical Physics*, K. W. Ford, Ed., Brandeis University Summer Institute in Theoretical Physics, Vol. 3, W. A. Benjamin, 181–218.
- , 1968: Prior probabilities. *IEEE Trans. Syst. Sci. Cyb.*, **4**, 227–241, <https://doi.org/10.1109/TSSC.1968.300117>.
- , 1973: The well-posed problem. *Found. Phys.*, **3**, 477–492, <https://doi.org/10.1007/BF00709116>.
- Jensen, M. P., and Coauthors, 2016: The Midlatitude Continental Convective Clouds Experiment (MC3E). *Bull. Amer. Meteor. Soc.*, **97**, 1667–1686, <https://doi.org/10.1175/BAMS-D-14-00228.1>.

- Kapur, J. N., 1989: *Maximum-Entropy Models in Science and Engineering*. John Wiley & Sons, 635 pp.
- Lecompte, M., and C. Dumouchel, 2008: On the capability of the generalized gamma function to represent spray drop-size distribution. *Part. Part. Syst. Charact.*, **25**, 154–167, <https://doi.org/10.1002/ppsc.200701098>.
- Lee, G. W., I. Zawadzki, W. Szyrmer, D. Sempere-Torres, and R. Uijlenhoet, 2004: A general approach to double-moment normalization of drop size distributions. *J. Appl. Meteor.*, **43**, 264–281, [https://doi.org/10.1175/1520-0450\(2004\)043<0264:AGATDN>2.0.CO;2](https://doi.org/10.1175/1520-0450(2004)043<0264:AGATDN>2.0.CO;2).
- Li, X., and R. S. Tankin, 1987: Droplet size distribution: A derivation of a Nukiyama-Tanasawa type distribution function. *Combust. Sci. Technol.*, **56**, 65–76, <https://doi.org/10.1080/00102208708947081>.
- , L. Chin, R. S. Tankin, T. Jackson, J. Stutrud, and G. Switzer, 1991: Comparison between experiments and predictions based on maximum entropy for sprays from a pressure atomizer. *Combust. Flame*, **86**, 73–89, [https://doi.org/10.1016/0010-2180\(91\)90057-I](https://doi.org/10.1016/0010-2180(91)90057-I).
- Lienhard, J. H., 1964: A statistical mechanical prediction of the dimensionless unit hydrograph. *J. Geophys. Res.*, **69**, 5231–5238, <https://doi.org/10.1029/JZ069i024p05231>.
- List, R., and G. M. McFarquhar, 1990: The role of breakup and coalescence in the three-peak equilibrium distribution of raindrops. *J. Atmos. Sci.*, **47**, 2274–2292, [https://doi.org/10.1175/1520-0469\(1990\)047<2274:TROBAC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1990)047<2274:TROBAC>2.0.CO;2).
- Liu, Y., Y. Laiguang, Y. Weinong, and L. Feng, 1995: On the size distribution of cloud droplets. *Atmos. Res.*, **35**, 201–216, [https://doi.org/10.1016/0169-8095\(94\)00019-A](https://doi.org/10.1016/0169-8095(94)00019-A).
- Majda, A., and X. Wang, 2006: *Non-Linear Dynamics and Statistical Theories for Basic Geophysical Flows*. Cambridge University Press, 551 pp.
- Marshall, J. S., and W. M. K. Palmer, 1948: The distribution of raindrops with size. *J. Meteor.*, **5**, 165–166, [https://doi.org/10.1175/1520-0469\(1948\)005<0165:TDORWS>2.0.CO;2](https://doi.org/10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2).
- Maur, A. A., 2001: Statistical tools for drop size distributions: Moments and generalized gamma. *J. Atmos. Sci.*, **58**, 407–418, [https://doi.org/10.1175/1520-0469\(2001\)058<0407:STFDSD>2.0.CO;2](https://doi.org/10.1175/1520-0469(2001)058<0407:STFDSD>2.0.CO;2).
- McFarquhar, G. M., 2004: A new representation of collision-induced breakup of raindrops and its implications for the shapes of raindrop size distributions. *J. Atmos. Sci.*, **61**, 777–794, [https://doi.org/10.1175/1520-0469\(2004\)061<0777:ANROCB>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<0777:ANROCB>2.0.CO;2).
- , T.-L. Hsieh, M. Freer, J. Mascio, and B. F. Jewett, 2015: The characterization of ice hydrometeor gamma size distributions as volumes in N_0 - λ - μ phase space: Implications for microphysical process modeling. *J. Atmos. Sci.*, **72**, 892–909, <https://doi.org/10.1175/JAS-D-14-0011.1>.
- Morrison, H., and J. A. Milbrandt, 2015: Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. Part I: Scheme description and idealized tests. *J. Atmos. Sci.*, **72**, 287–311, <https://doi.org/10.1175/JAS-D-14-0065.1>.
- , J. Curry, and V. Khvorostyanov, 2005: A new double-moment microphysics parameterization for application in cloud and climate models. Part I: Description. *J. Atmos. Sci.*, **62**, 1665–1677, <https://doi.org/10.1175/JAS3446.1>.
- Nukiyama, S., and Y. Tanasawa, 1939: An experiment on the atomization of liquid: 3rd report, on the distribution of the size of drops. *Trans. Japan Soc. Mech. Eng.*, **5**, 131–135, <https://doi.org/10.1299/kikai1938.5.131>.
- Pathria, R., and P. Beale, 2011: *Statistical Mechanics*. 3rd ed. Academic Press, 744 pp.
- Petty, G. W., and W. Huang, 2011: The modified gamma size distribution applied to inhomogeneous and nonspherical particles: Key relationships and conversions. *J. Atmos. Sci.*, **68**, 1460–1473, <https://doi.org/10.1175/2011JAS3645.1>.
- Phillips, S. J., M. Dudík, and R. E. Schapire, 2004: A maximum entropy approach to species distribution modeling. *Proc. 21st Int. Conf. on Machine Learning*, Banff, Alberta, Canada, ACM, 655–662.
- , R. P. Anderson, and R. E. Schapire, 2006: Maximum entropy modeling of species geographic distributions. *Ecol. Modell.*, **190**, 231–259, <https://doi.org/10.1016/j.ecolmodel.2005.03.026>.
- Rose, K., E. Gurewitz, and G. C. Fox, 1990: Statistical mechanics and phase transitions in clustering. *Phys. Rev. Lett.*, **65**, 945, <https://doi.org/10.1103/PhysRevLett.65.945>.
- Rosenfeld, R., 1996: A maximum entropy approach to adaptive statistical language modelling. *Comput. Speech Lang.*, **10**, 187–228, <https://doi.org/10.1006/csla.1996.0011>.
- Seifert, A., and K. Beheng, 2006: A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description. *Meteor. Atmos. Phys.*, **92**, 45–66, <https://doi.org/10.1007/s00703-005-0112-4>.
- Sellens, R., and T. Brzustowski, 1985: A prediction of the drop size distribution in a spray from first principles. *Atomisation Spray Technol.*, **1**, 89–102.
- Shannon, C. E., 1948: A mathematical theory of communication. *Bell Syst. Tech. J.*, **27**, 379–423, 623–656, <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
- Skilling, J., and R. Bryan, 1984: Maximum entropy image reconstruction: General algorithm. *Mon. Not. Roy. Astron. Soc.*, **211**, 111–124, <https://doi.org/10.1093/mnras/211.1.111>.
- Srivastava, R., 1971: Size distribution of raindrops generated by their breakup and coalescence. *J. Atmos. Sci.*, **28**, 410–415, [https://doi.org/10.1175/1520-0469\(1971\)028<0410:SDORGB>2.0.CO;2](https://doi.org/10.1175/1520-0469(1971)028<0410:SDORGB>2.0.CO;2).
- , 1982: A simple model of particle coalescence and breakup. *J. Atmos. Sci.*, **39**, 1317–1322, [https://doi.org/10.1175/1520-0469\(1982\)039<1317:ASMOPC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1982)039<1317:ASMOPC>2.0.CO;2).
- Stacy, E. W., 1962: A generalization of the gamma distribution. *Ann. Math. Stat.*, **33**, 1187–1192, <https://doi.org/10.1214/aoms/1177704481>.
- Testud, J., S. Oury, R. A. Black, P. Amayenc, and X. Dou, 2001: The concept of “normalized” distribution to describe raindrop spectra: A tool for cloud physics and cloud remote sensing. *J. Appl. Meteor.*, **40**, 1118–1140, [https://doi.org/10.1175/1520-0450\(2001\)040<1118:TCOND>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<1118:TCOND>2.0.CO;2).
- Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.*, **132**, 519–542, [https://doi.org/10.1175/1520-0493\(2004\)132<0519:EFOWPU>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2).
- Tian, L., G. M. Heymsfield, L. Li, A. J. Heymsfield, A. Bansemmer, C. H. Twohy, and R. C. Srivastava, 2010: A study of cirrus ice particle size distribution using TC4 observations. *J. Atmos. Sci.*, **67**, 195–216, <https://doi.org/10.1175/2009JAS3114.1>.
- Ulbrich, C. W., 1983: Natural variations in the analytical form of the raindrop size distribution. *J. Climate Appl. Meteor.*, **22**, 1764–1775, [https://doi.org/10.1175/1520-0450\(1983\)022<1764:NVITAF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1983)022<1764:NVITAF>2.0.CO;2).
- Verkley, W., 2011: A maximum entropy approach to the problem of parametrization. *Quart. J. Roy. Meteor. Soc.*, **137**, 1872–1886, <https://doi.org/10.1002/qj.860>.
- , and P. Lynch, 2009: Energy and enstrophy spectra of geostrophic turbulent flows derived from a maximum entropy principle. *J. Atmos. Sci.*, **66**, 2216–2236, <https://doi.org/10.1175/2009JAS2889.1>.

- , P. Kalverla, and C. Severijns, 2016: A maximum entropy approach to the parametrization of subgrid processes in two-dimensional flow. *Quart. J. Roy. Meteor. Soc.*, **142**, 2273–2283, <https://doi.org/10.1002/qj.2817>.
- Wernecke, S. J., and L. R. D’Addario, 1977: Maximum entropy image reconstruction. *IEEE Trans. Comput.*, **26**, 351–364, <https://doi.org/10.1109/TC.1977.1674845>.
- Wu, W., and G. M. McFarquhar, 2016: On the impacts of different definitions of maximum dimension for nonspherical particles recorded by 2D imaging probes. *J. Atmos. Oceanic Technol.*, **33**, 1057–1072, <https://doi.org/10.1175/JTECH-D-15-0177.1>.
- Yano, J.-I., A. J. Heymsfield, and V. T. Phillips, 2016: Size distributions of hydrometeors: Analysis with the maximum entropy principle. *J. Atmos. Sci.*, **73**, 95–108, <https://doi.org/10.1175/JAS-D-15-0097.1>.
- Zhang, X., and G. Zheng, 1994: A simple droplet spectrum derived from entropy theory. *Atmos. Res.*, **32**, 189–193, [https://doi.org/10.1016/0169-8095\(94\)90059-0](https://doi.org/10.1016/0169-8095(94)90059-0).