Maximum-Entropy Method Applied to Micro- and Nanolasers

Dissertation
zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

genehmigt durch die Fakultät für Naturwissenschaften der
Otto-von-Guericke-Universität Magdeburg

von M. Sc. Boris Melcher
geb. am 04. August 1990 in Kara-Balta

Gutachter: Prof. Dr. Jan Wiersig
Jun.-Prof. Dr. Doris Reiter

eingereicht am: 15. Dezember 2020
verteidigt am: 28. April 2021

Abstract

During the last decades, the principle of maximum entropy established itself as a fundamental principle of consistent reasoning and hence has been successfully applied in many fields inside and outside of physics. Whenever only partial information is available it is an ideal tool to derive full statistical distributions or, in the case of quantum-mechanical descriptions, the full density matrix in a least-biased manner by exactly reproducing the given a-priori information while being maximally non-committal otherwise.

In this thesis the maximum-entropy method is applied to the context of micro- and nanolasers. While tremendous improvements concerning the miniaturization and efficiency of these lasers have been achieved in the recent years, their light is still most often characterized in terms of the first statistical moments of the photon-number distribution, e.g., the light intensity and auto-correlation function. Lately, however, direct measurement of the full photon distribution became available experimentally, and theoretically its availability is highly desirable since it contains the full information about the system and therefore enables a deeper insight into the underlying physical processes.

The application of the maximum-entropy method in this work is threefold. Firstly, a birth-death model that describes the interaction of a cavity light-field with a quantum-dot system is considered as a benchmark. The full photon-number distribution which is available via conventional methods here is compared to the maximum-entropy distributions that are obtained when known photon moment values are used as input information. Secondly, a combination of the maximum-entropy method and equation-of-motion approaches is proposed where output quantities of the latter are used as input for the former. Apart from the access to the full photon-number distribution, the method also provides the possibility to define an unambiguous threshold even for cases where usual definitions fail.

Lastly, the maximum-entropy method is used as a stand-alone approach without the need for moment values from external sources. Instead, knowledge of the stationarity of several distinct observables is used as an input to derive the least biased steady state. By doing so, the many-particle hierarchy problem that arises in conventional equation-of-motion techniques is circumvented. Moreover, the numerical determination of the full density matrix is enabled and access to all relevant expectation values and the full statistics is obtained. The results are compared to the numerically exact solution of the von Neumann-Lindblad equation for a four-level system resonantly coupled to a single cavity mode. Moreover, usage of the maximum-entropy method as a trial-and-error approach to identify the relevant processes of quantum systems is exemplified.

Contents

1 Introduction

In the middle of the 19th century, a time when heat machines were already widely used, 28-year-old Sadi Carnot not only laid the foundation for the physical sub-field of thermodynamics but also was the first one to encounter the concept of entropy (Carnot, 1824; Volkenstein, 2009; Carnot, 2012). Afterwards, the attempt to describe the dynamics of heat phenomena quickly expanded to also include entropy's statistical nature as well as its role as an uncertainty measure. Many well-known scientists such as Clausius, Boltzmann, Maxwell and Gibbs in the earlier years and Shannon, von Neumann and Jaynes in the later years were involved in this process that finally also lead to the principle of maximum entropy (Clausius, 1850a; Clausius, 1850b; Clausius, 1865; Gibbs, 1902; Shannon, 1948; Brillouin, 1951; Jaynes, 1957a; Jaynes, 1957b; Jaynes, 1965; Ehrenberg, 1967; Garber, 1970; Tribus & McIrvine, 1971; Von Neumann, 1996; Lieb & Yngvason, 1999; Jaynes, 2003; Downarowicz, 2007; Griffin, 2008; Volkenstein, 2009; Pressé *et al.,* 2013; Uffink, 2017; Chehade & Vershynina, 2019). With the more recent discoveries, the procedure of entropy maximization is no longer merely seen as maximizing uncertainty or minimizing bias but rather as a way to draw consistent inferences from partial data (Griffin, 2008; Volkenstein, 2009; Pressé *et al.,* 2013; Uffink, 2017; Jizba & Korbel, 2019; Jizba & Korbel, 2020). With that, the maximum-entropy principle exceeded the realm of physics and established itself as a form of "extended logic"[1]. Because the maximum-entropy method is an ideal tool to handle incomplete information and deduce self-consistent conclusions without artificially including any bias, it has since then been used in various fields inside and outside of physics, such as astronomy, biology, economics, finance, and quantum chemistry among others (Pressé *et al.,* 2013; Zhou *et al.,* 2013; He & Kolovos, 2018; Jizba & Korbel, 2019; Tsallis, 2019; Chanda *et al.,* 2020).

In the same time that was required to arrive at today's notion of the principle of maximum entropy, an enormous information technological transformation took place. While in 1824, the year of Carnot's publication, great effort was put into the development of the first telegraph (Meadow, 2002), we are now living in the digital Zettabyte Era where the annual IP traffic exceeds at least one Zettabyte ($1\,\mathrm{ZB} = 10^9\,\mathrm{TB} = 10^{21}\,\mathrm{B}$) (Cisco, 2019; Ma & Oulton, 2019). Handling this huge amount of digital information accordingly consumes a substantial percentage of the worldwide electricity (Malmodin & Lundén, 2018; Ma & Oulton, 2019) and has a large impact on the environment with an estimated $2.3\,\%$ of

[1] Jaynes, 2003, p. xxii.

the global green house gas emission stemming from the information and communication technologies sector in 2020 (Global e-Sustainability Initiative, 2012). Fortunately, the ongoing effort to decrease the emission already resulted in a more optimistic trend with an estimation of a percentage of roughly $2.0\,\%$ for 2030 (Global e-Sustainability Initiative, 2015). Nevertheless, the search for energy-efficient technologies is still of major interest both ecologically and economically.

One of its many aspects is the ongoing miniaturization of lasing devices that are, for instance, used for data transmission between individual chips or different sections within chips (Ma & Oulton, 2019). Concerning laser systems a plethora of different aspects, e.g., the light-matter coupling strength (Vahala, 2003; He *et al.,* 2013; Dovzhenko *et al.,* 2018; Frisk Kockum *et al.,* 2019), the quality factor of the cavity mode (Vahala, 2003; Reitzenstein & Forchel, 2010; He *et al.,* 2013), and the power efficiency (Ma & Oulton, 2019) play a role and are currently exploited to improve the desired properties of those systems. As gain medium quantum-dots are among the structures considered most often (Noda, 2006; Norman *et al.,* 2019). In those nanometer-sized objects the charge carriers are locally confined in all three spatial dimensions which can lead to strong light-matter interaction and high gain (Noda, 2006; Reitzenstein, 2012; Norman *et al.,* 2019). Apart from that, the confining potential and the discrete energy levels of these artificial atoms can be tailored by, e.g., different choices of material, the quantum dot size, shape or by application of strain (Reitzenstein, 2012; Huo *et al.,* 2013; Huo *et al.,* 2014a; Huo *et al.,* 2014b; Goldmann, 2014; Schumann, 2016; Melcher, 2016; Norman *et al.,* 2019). The advances in the area of laser cavities, on the other hand, where improved manufacturing techniques allowed for increasingly higher quality factors and smaller mode volumes, further facilitated the possibility to build lasers where the better part of the spontaneous emission is directed into the laser mode (Vahala, 2003; Noda, 2006; Strauf *et al.,* 2006; Thyrrestrup *et al.,* 2010; He *et al.,* 2013; Ma & Oulton, 2019). This finally lead to the rapid approach towards devices with ultra-low laser threshold and even thresholdless lasers, where the jump in the input-output curve completely vanishes (Strauf *et al.,* 2006; Khajavikhan *et al.,* 2012; Prieto *et al.,* 2015; Ota *et al.,* 2017; Jagsch *et al.,* 2018; Ma & Oulton, 2019). More recently, two-dimensional materials, especially the class of transition metal dichalcogenides, have emerged as proposed gain materials (Wang *et al.,* 2012; Xiao *et al.,* 2012; Jariwala *et al.,* 2014; Bhimanapati *et al.,* 2015; Schaibley *et al.,* 2016; Zheng *et al.,* 2018; Lohof *et al.,* 2019). There, strong Coulomb interaction between charge carriers and the weak dielectrical screening make the electronic and optical properties highly susceptible to their environment (Mak *et al.,* 2010; Butler *et al.,* 2013; Steinhoff *et al.,* 2018a; Steinhoff *et al.,* 2018b; Erben *et al.,* 2018; Lohof *et al.,* 2018; Lohof *et al.,* 2019; Steinhoff *et al.,* 2020). By combining and/or stacking of different two-dimensional materials, their properties can be strongly manipulated (Geim & Grigorieva, 2013; Xu *et al.,* 2014; Bhimanapati *et al.,* 2015; Chen *et al.,* 2019; Alexeev *et al.,*

2019; Tran *et al.,* 2019). In combination with high-quality cavities these new materials are a topic of current investigation (Jiang *et al.,* 2017; Lohof *et al.,* 2019). Also, since the carrier densities of extended gain media strongly exceed the carrier densities of quantum dots, new operational regimes are to be expected (Lohof *et al.,* 2019).

Regardless of the particular properties of the laser system, experimentally the emitted light is oftentimes characterized by measurements of the light intensity and the second-order (and sometimes higher-order) auto-correlation functions (Schneebeli *et al.,* 2008; Leymann *et al.,* 2013b; Kazimierczuk *et al.,* 2015; Jahnke *et al.,* 2016; Strauß *et al.,* 2016; Schlottmann *et al.,* 2018). These measurements give the first statistical moments of the photon probability distribution. The light intensity is proportional to the first photon moment, i.e., the mean photon number $\langle n \rangle$. The second-order auto-correlation function contains information about the second photon moment $\langle n^2 \rangle$, which is related to the variance of the distribution, and higher-order auto-correlation functions contain the respective higher order moments. Depending on the specific values of these quantities, it is possible to draw conclusions about the emission dynamics and light properties. Also, more recently, direct measurements of the full photon probability distribution became available via the use of a transition-edge sensor (Schlottmann *et al.,* 2018; Schmidt *et al.,* 2020).

When it comes to the theoretical description of those laser devices, one common approach is to derive equations of motion for the quantities mentioned above (Gies *et al.,* 2007; Florian *et al.,* 2012; Leymann *et al.,* 2014; Leymann, 2015; Fanaei *et al.,* 2016; Jahnke *et al.,* 2016; Kreinberg *et al.,* 2017; Foerster, 2017; Moody *et al.,* 2018; Fanaei, 2019). There, due to many-particle interactions, an infinite hierarchy of coupled differential equations arises. What is referred to as the *many-particle hierarchy problem* is usually solved by application of factorization schemes such as the cluster-expansion method, where higher order quantities are either neglected or factorized such that the set of differential equations can be closed (Fricke, 1996; Leymann *et al.,* 2014; Leymann, 2015; Foerster, 2017). Although this kind of approach has proven to be successful in a variety of problems (Kira *et al.,* 1998; Köhler & Burnett, 2002; Hoyer *et al.,* 2003; Baer *et al.,* 2006; Feldtmann *et al.,* 2006; Gies *et al.,* 2007; Kapetanakis & Perakis, 2008; Ritter *et al.,* 2010; Kabuss *et al.,* 2012; Florian *et al.,* 2013; Leymann *et al.,* 2013b; Khanbekyan *et al.,* 2015; Leymann *et al.,* 2015; Jahnke *et al.,* 2016; Fanaei *et al.,* 2016), it also comes with several downsides. Apart from the fact that factorization and truncation might induce unphysical effects and therefore should be done very carefully (Leymann *et al.,* 2014), the equation-of-motion plus cluster-expansion method approaches only deliver partial information about the system and never the full probability distributions. Availability of the full distribution is highly desirable though, since in certain cases only the full probability distribution allows for an accurate physical understanding (Leymann *et al.,* 2013b; Schlottmann *et al.,* 2018). Furthermore, since photon-number distributions are also available experimentally nowadays, a comparison with theoretical

calculations could prove to be beneficial. The central question is therefore how to deduce full probability distributions from partial knowledge. The maximum-entropy method is proposed as a suitable approach here.

The main goal of this thesis is a systematic introduction of the maximum-entropy method to the field of micro- and nanolasers. This is done in three parts that make up the three main chapters (Chaps. 3, 4, and 5) of the thesis. Firstly, in Chap. 3 the maximum-entropy-method approach is benchmarked by considering a birth-death model that describes the interaction of a cavity light-field with a quantum-dot or atomic system. There, the full photon probability distribution is available by conventional methods and serves as a comparison. The photon moments of the original photon distribution are used as an input for the maximum-entropy method and convergence of the maximum-entropy distribution towards the original distribution is monitored for inclusion of different orders of photon moments. In addition to that, a new laser threshold definition is derived that is applicable even in cases where usual definitions fail. Furthermore, the light is characterized in terms of entropy, and general existence criteria of maximum-entropy distributions are exploited to derive generic features that a photon-number probability distribution must have to exhibit superthermal photon bunching.

Secondly, in Chap. 4 a combination of the maximum-entropy method with equation-of-motion approaches is proposed. There, two different models are considered that both produce photon moments as output quantities. These outputs are used as an input for the maximum-entropy method to derive the full photon-number distribution. General existence criteria are investigated and laser threshold comparisons as well as characterizations of the light by entropy are given. Additionally, a proposition is made to clear up some of the confusions concerning the laser threshold in extended gain media.

Thirdly, in Chap. 5 a fully quantized single-quantum-dot microcavity laser system is regarded, where the maximum-entropy-method results are compared to the conventional von Neumann-Lindblad-equation approach. Here, the full density matrix of the system is derived that delivers all relevant expectation values and auto-correlation functions. Furthermore, application of the maximum-entropy method as a stand-alone approach, i.e., without the need for moment values from external sources is exemplified. Instead, knowledge of the stationarity of several distinct observables is utilized and the equations of motion themselves are used as an input. The need to close or truncate the many-particle hierarchy is thereby circumvented. Lastly, it is demonstrated how the maximum-entropy method can be used to identify the physically relevant processes within the investigated systems.

The theoretical foundations are set in the following Chapter 2, where, after a brief historical overview, the maximum-entropy method is derived for both quantized and non-quantized descriptions.

Maximum-Entropy Method

2

2.1 Historical Overview

Since the concept of entropy looks back on a history of almost 200 years, this section is devoted to a rough overview ranging from the early discovery of entropy in the middle of the 19th century to more recent work that extends the Shannonian entropy definition to a much wider class of entropies. This section includes the statistical and information theoretical perspectives on entropy as well as the notions of entropy maximization as method of statistical inference and procedure of consistent reasoning. Furthermore, the extension of the principle of maximum entropy to a whole class of maximum-entropy principles is introduced. Finally, very recent work is briefly summarized that connects the various lines of reasoning and different approaches from information theory and statistical inference with non-additive entropies.

The discovery of entropy In 1824, at the age of 28, Sadi Carnot laid groundwork for the second law of thermodynamics with his only publication *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance*[1] (Carnot, 1824). In the midst of the industrial revolution, when steam engines were already widely used, "the physics of heat phenomena [...] was still awaiting its Newton"[2]. At that time, it was controversial whether heat phenomena were the result of kinetic energy of atoms or a weightless, invisible substance called *caloricum* (caloric) that "was contained in a gas much like juice in an orange. Squeeze an orange and you get orange juice. Compress a gas and caloric oozes out, that is, it heats up"[3]. Although Carnot himself relied on caloric theory (Carnot, 1824; Volkenstein, 2009), he revealed that the motive power is actually due to a transport process. In his publication Carnot writes:

[1] "Reflections on the motive power of fire, and on machines fitted to develop that power" (Carnot, 2012)
[2] Volkenstein, 2009, p. 4.
[3] Volkenstein, 2009, p. 4.

> La production de la puissance motrice est donc due, dans les machines à vapeur, non à une consommation réelle du calorique, mais à son transport d'un corps chaud à un corps froid, c'est-à-dire à son rétablissement d'équilibre.[4]

More than two decades later, in an attempt to mechanistically understand "die bewegende Kraft der Wärme"[5], Rudolf Clausius noticed that, in addition to heat, another quantity must have been transferred (Clausius, 1850a; Clausius, 1850b). After quickly rejecting his first idea to name it "Verwandlungsinhalt"[6], he deliberately coined the term "Entropie"[7] (entropy) by replacing the root of the Greek term for energy with τροπή (transformation) because, on the one hand, he found entropy to be such an important concept that he wanted a term that was phonetically close to *energy* (Clausius, 1865) and, on the other hand, it was supposed to be applicable in all new languages[8]. This finally lead to one of the possible formulations of the second law of thermodynamics (Clausius, 1850a; Clausius, 1850b; Lieb & Yngvason, 1999), which later was described as "the greatest scientific achievement of the nineteenth century"[9]. Clausius himself already saw the principle of entropy maximization as a fundamental principle of nature ("Die Entropie der Welt strebt einem Maximum zu"[10]).

The statistical nature of entropy Another two decades later, in 1873, James Clerk Maxwell, who initially was working together with Clausius on the kinetic theory of gases but later changed to the field of electrodynamics, reacted to Clausius's approach to thermodynamics in purely mechanical terms with "amusement mixed with a little smugness"[11]. He "regarded the second law as statistical [and] used his demon to illustrate the statistical nature of the second law"[12] (Ehrenberg, 1967). "[W]hen Maxwell returned to gas theory it was Boltzmann rather than Clausius who was closest to him and who served as his main stimulus."[13] It also was Ludwig Eduard Boltzmann who recognized the connection between entropy and the multiplicity of gas particles occupying small volumetric cells in phase space (Pressé *et al.,* 2013). He asserted that the occupation probabilities of the equilibrium state are those that maximize the entropy while simultaneously satisfying the constraint of a constant particle number and average energy (Pressé *et al.,* 2013). Through a variational principle this finally lead to what is nowadays known as the Maxwell-Boltzmann distribution (Pressé *et al.,*

[4] Carnot, 1824, p. 10. "The production of motive power is then due in steam-engines not to an actual consumption of caloric, but to its transportation from a warm body to a cold body, that is, to its re-establishment of equilibrium" (Carnot, 2012, p. 7).
[5] Clausius, 1850a, p. 368. "the moving force of heat" (author's translation)
[6] Clausius, 1865, p. 390. "transformational content" (author's translation)
[7] Clausius, 1865, p. 390.
[8] "damit sie unverändert in allen neuen Sprachen angewandt werden können" (Clausius, 1865, p. 390)
[9] Lieb & Yngvason, 1999, p. 4.
[10] Clausius, 1867, p. 17. "The entropy of the world tends to a maximum" (author's translation)
[11] Garber, 1970, p. 317.
[12] Garber, 1970, p. 318.
[13] Garber, 1970, p. 319.

2013). Jaynes et al. later described Boltzmann's reasoning as a "'superefficient' way to capture the essential mathematical ingredients"[14]. Pointing out that Boltzmann's approaches were often rather messy (Griffin, 2008; Uffink, 2017), they further commented: "Whether by luck or inspiration, he [Boltzmann] put into his equations only the dynamical information that happened to be relevant to the questions he was asking."[15] After his death Boltzmann's famous formula was engraved in his tombstone (Jaynes, 2003). Josiah Willard Gibbs, who realized that Boltzmann's approach neglected interactions between particles (Griffin, 2008), applied the method of ensembles to equilibrium statistical mechanics and arrived at his own entropy definition (Gibbs, 1902; Pressé *et al.,* 2013). Much later, Jaynes connected Boltzmann's and Gibbs's definition to the phenomenological entropy of thermodynamics (Jaynes, 1965), indeed revealing that Gibbs's definition yields the correct entropy, while Boltzmann's "yields an 'entropy' that is in error by a nonnegligible [sic] amount whenever interparticle forces affect thermodynamic properties"[16]. Beyond his ensemble approach, Gibbs already outlined the bigger picture and commented that statistical mechanics "seems eminently worthy of an independent development [independent of thermodynamics, author's note], both on account of the elegance and simplicity of its principles, and because it yields new results and places old truths in a new light in departments quite outside of thermodynamics."[17]

Entropy as a measure of uncertainty In 1932, John von Neumann worked on a mathematically grounded book for quantum mechanics. There, he generalized the concept of entropy to quantum systems (Von Neumann, 1996; Downarowicz, 2007; Chehade & Vershynina, 2019). Interestingly, although nowadays the von-Neumann entropy is often seen as "a natural generalization of the classical Shannon entropy"[18], Claude Elwood Shannon worked on information theory much later. In 1948, he published his work on the most efficient way to encode a message and better designs of communication channels (Shannon, 1948; Tribus & McIrvine, 1971). There, he axiomatically derived an information measure "independent of the means used to generate the information"[19] that was "fundamental in information science, just as the Pythagorean theorem is fundamental in geometry"[20]. It is reported that Shannon, who was not directly interested in thermodynamics, commented on his findings with:

> My greatest concern was what to call it. I thought of calling it "information," but the word was overly used, so I decided to call it "uncertainty." When I

[14] Pressé *et al.,* 2013, p. 1117.
[15] Pressé *et al.,* 2013, p. 1117.
[16] Jaynes, 1965, p. 391.
[17] Gibbs, 1902; Pressé *et al.,* 2013, p. 12.
[18] Chehade & Vershynina, 2019.
[19] Tribus & McIrvine, 1971, p. 179.
[20] Tribus & McIrvine, 1971, p. 180.

> discussed it with John von Neumann, he had a better idea. Von Neumann told me, "You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage."[21]

Later, Jaynes commented on the choice of name that the confusion of information entropy and experimental entropy is a "major occupational disease"[22] and that "[t]hey should never have been called by the same name"[23]. Nevertheless, great effort was put into connecting Shannon's entropy with the thermodynamic entropy. Among many others, Léon Nicolas Brillouin established the connection by using Maxwell's famous demon (Brillouin, 1951; Volkenstein, 2009). Arguing that the demon would need to gain information about the particle velocities by at least exchanging one quantum of light and consequently by needing a battery-powered flashlight that itself would lead to an increase in temperature, "he concluded that one bit of information requires $k \ln 2$ thermal entropy units [with Boltzmann constant k, author's note]"[24]. Consequently, "there is no conflict between abstract Shannon information and thermodynamic information, as long as the questions we ask are physically real questions."[25]

Entropy maximization as statistical inference Edwin Thompson Jaynes, who is presumably associated with the maximum-entropy method the most often, was the first one to consider "statistical mechanics as a form of statistical inference rather than as a physical theory"[26]. As a foundational prerequisite he argued that it is essential to "accept the von-Neumann-Shannon expression for entropy, very literally, as a measure of the amount of uncertainty represented by a probability distribution"[27]. By maximizing that uncertainty measure, i.e. the entropy, with respect to the known constraints, one acquires the otherwise least informative or most honest probability distribution (Jaynes, 1957a; Jaynes, 1957b; Griffin, 2008; Pressé *et al.,* 2013). Then, "it is found that the usual computational rules, starting with the determination of the partition function, are an immediate consequence of the maximum-entropy principle"[28] and that "[i]n the resulting 'subjective statistical mechanics,' the usual rules are thus justified independent of any physical argument"[29]. Jaynes emphasized that "the maximization of entropy is not an application of a law of physics, but

[21] Tribus & McIrvine, 1971, p. 180.
[22] Jaynes, 2003, p. 351.
[23] Jaynes, 2003, p. 351.
[24] Tribus & McIrvine, 1971, p. 184.
[25] Tribus & McIrvine, 1971, p. 183.
[26] Jaynes, 1957a, p. 620.
[27] Jaynes, 1957a, p. 629.
[28] Jaynes, 1957a, p. 620.
[29] Jaynes, 1957a, p. 620.

merely a method of reasoning which ensures that no unconscious arbitrary assumptions have been introduced."[30]

Since at that time, probability theory was based on two very different schools of thought, first the objective school of thought that saw the probability of an event as an objective property of that event, and second the subjective school that regarded probabilities as expressions of human ignorance (Jaynes, 1957a; Pressé *et al.,* 2013), Jaynes's approach was controversial (Pressé *et al.,* 2013). "Why should someone's state of knowledge have any bearing on physics?"[31] was one of the many objections. In one of his later publication, where Jaynes connected the various definitions and entropy concepts (Jaynes, 1965), he described entropy as an "anthropomorphic concept"[32] and further stressed that "it is a property, not of the physical system, but of the particular experiments you or I choose to perform on it"[33]. With this kind of interpretation he added a new perspective on the second law and on thermodynamics as a whole. He wrote that, "[w]ith such an interpretation the expression 'irreversible process' represents a semantic confusion; it is not the physical process that is irreversible, but rather our ability to follow it."[34] The concept of entropy maximization to him was fundamentally different from usual approaches. While "[a] physical theory asks bluntly, 'How does the system behave?' and seeks to answer it by deductive reasoning from the known laws of physics [...] [p]redictive statistical mechanics, instead of seeking the unattainable, asks a more modest question: 'Given the partial information that we do, in fact, have, what are the best predictions we can make of observable phenomena?'"[35] Consequently, the application of predictive statistical mechanics is in strong contrast to physical theories and can rather be seen as "extended logic"[36]. If the procedure of entropy maximization gives a definite prediction, i.e., a sharp probability distribution, it can be compared to experiments. Then, "experimental proof that a definite prediction is incorrect gives evidence of the existence of new laws of physics."[37] On the other hand, new "information is redundant [...] if it is what the theory would have predicted from the old information"[38]. Hence from Jaynes's perspective, "the most difficult problem of all is to learn how to state clearly *what is the specific question we are trying to answer?*"[39] Undoubtedly, Jaynes did great work by unifying the different approaches to the concept of entropy (Griffin, 2008). Nevertheless, the main point of criticism remained unanswered. Although the Shannon "entropy is the only function

[30] Jaynes, 1957a, p. 630.
[31] Pressé *et al.,* 2013, p. 1120.
[32] Jaynes, 1965, p. 627.
[33] Jaynes, 1965, p. 398.
[34] Jaynes, 1957b, p. 171.
[35] Jaynes, 1980, p. 593.
[36] Jaynes, 2003, p. xxii.
[37] Jaynes, 1957a, p. 627.
[38] Jaynes, 1980, p. 599.
[39] Jaynes, 1965, p. 398.

satisfying axioms that are accepted as requirements for an uncertainty measure"[40], there were "many scientists who [were] reluctant to use the procedure because of its reliance on the so-called subjective notion of missing information"[41] (Pressé *et al.,* 2013). The question remained "why maximize entropy; why not some other function?"[42]

Maximum-entropy principle without subjective measure of uncertainty In 1980, John Shore and Rodney Johnson succeeded in overcoming the need for a subjective measure of uncertainty (Shore & Johnson, 1980; Shore & Johnson, 1981; Pressé *et al.,* 2013). Their arguing was based on the notion that "if a problem can be solved in more than one way, the results should be consistent."[43] Hence, "[s]ince the maximum entropy principle is asserted as a general method of inductive inference, it is reasonable to require that different ways of using it to take the same information into account should lead to consistent results."[44] Based on that, they formulated four basic consistency axioms,

> *Uniqueness:* The result should be unique.
> *Invariance:* The choice of coordinate system should not matter.
> *System Independence:* It should not matter whether one accounts for independent information about independent systems separately in terms of different densities or together in terms of a joint density.
> *Subset Independence:* It should not matter whether one treats an independent subset of system states in terms of a separate conditional density or in terms of the full system density,[45]

and demonstrated that "[i]n a brilliantly simple way, entropy emerges as an elementary consequence of consistent reasoning"[46] and hence "we must use maximum entropy — or lay ourselves open to the charge of inconsistency"[47]. Shore and Johnson thereby shifted Shannon's axiomatic derivation of the information measure to an axiomatic derivation of the procedure of statistical inference itself. In the following years many publications followed with similar lines of reasoning, for instance, by requiring physical experiments to be reproducible (Tikochinsky *et al.,* 1984a; Tikochinsky *et al.,* 1984b). Furthermore, the principle of entropy maximization was extended to continuous probability distributions by using the concept of relative entropy (Shore & Johnson, 1981; Uffink, 1995). Despite the appeal of the simple reasoning, there also were critical voices (Uffink, 1995; Uffink, 1996;

[40] Shore & Johnson, 1980, p. 27.
[41] Tikochinsky *et al.,* 1984b, p. 1357.
[42] Shore & Johnson, 1980, p. 27.
[43] Shore & Johnson, 1980, p. 27.
[44] Shore & Johnson, 1980, p. 27.
[45] Shore & Johnson, 1980, p. 27.
[46] Skilling, 1984, p. 748.
[47] Skilling, 1984, p. 749.

Csiszar, 1991; Pressé *et al.,* 2013). Alternative rules for inference (Van Fraassen, 1981; Van Fraassen *et al.,* 1986; Csiszar, 1991) and generalizations of the entropy (Tsallis, 1988) were proposed and a general discussion about conceptual inadequacies of the approach of Shore and Johnson emerged (Brukner & Zeilinger, 2001; Tsallis, 2015; Pressé *et al.,* 2015).

A class of maximum-entropy principles A collection of the flaws of the previous findings were summed up by Jos Uffink in his paper titled "Can the Maximum Entropy Principle be Explained as a Consistency Requirement?"[48] published in 1995. He writes that "Jaynes' approach puts a heavy weight on the assumptions of Shannon's uniqueness theorem, as if they were implied by the ideal of consistency itself"[49] while describing Shore and Johnson's work as "the most sophisticated"[50] of the so far available theories. However, Uffink further states that "in their analysis a hidden requirement is made, additional to their explicitly stated ones, which can be expressed as the demand that when it is not given whether systems (or experiments) are to be regarded as dependent we are justified in believing that they are independent."[51] By revealing Shore and Johnson's unreasonably strong assumptions and flawed proofs, he showed that "a slightly smaller set of reasonable requirements are fulfilled if and only if the rule [of inference, author's note] belongs to a class of which the maximum entropy principle, as well as the alternative rules of Van Fraassen et al. are members. This is the class of rules that maximize a generalized entropy expression containing a free continuous parameter (the so-called Rényi entropies)."[52] Hence, by omitting Shore and Johnson's strong assumptions a whole "class of 'maximum (relative Rényi) entropy principles'"[53] is derived that "can be seen as a new 'continuum of inductive methods' [...] that generalize the maximum entropy method."[54]

Non-additive entropies In the following years, evidence increased that "the Boltzmann-Gibbs-von Neumann-Shannon logarithmic entropic functional [...] is inadequate for wide classes of strongly correlated systems"[55]. In fact, a plethora of different more exotic entropies, e.g., the q-entropy, Kaniadakis entropy, Borges-Roditi entropy, Curado entropy, Hanel-Thurner entropy, Tsekouras-Tsallis entropy, δ-entropy, Burg entropy, Jizba-Arimitsu entropy were used especially in ultra-low and ultra-high energy physics (Tsallis, 2019; Jizba & Korbel, 2019; Lima & Deppman, 2020). The crucial difference to the class

[48] Uffink, 1995, p. 223.
[49] Uffink, 1995, p. 253.
[50] Uffink, 1995, p. 253.
[51] Uffink, 1995, p. 253.
[52] Uffink, 1995, p. 225.
[53] Uffink, 1995, p. 254.
[54] Uffink, 1995, p. 254.
[55] Tsallis, 2015, p. 2853.

of Rényi entropies is that the new entropies no longer satisfy entropic additivity, that is $S(A + B) = S(A) + S(B)$ where S is the entropy and A and B are probabilistically independent subsystems. It is argued by Constantino Tsallis, who is a strong advocate of non-additive entropies (Tsallis, 1988; Tsallis, 2015; Tsallis, 2019), that this development is calling for a "paradigm shift"[56] concerning the additivity of entropy. He further argues that, similar to quantum mechanics where "a coordinate and its conjugate momentum [are] correlated in a well known subtle manner"[57] characterized by the Planck constant h, or special relativity where the vacuum speed of light c characterizes the "nontrivial mutual influence"[58] of space and time, "nonadditive entropies emerge from strong correlations (among the random variables of the system) which are definitely out of the SJ [Shore and Johnson (and Uffink), author's note] hypothesis."[59] He states that "this lack of independence can be characterized by some parameter(s) [κ], which can always be defined in such a way that independence is recovered when κ vanishes."[60] Certainly, it would not be the first time that the principle of maximum entropy aids to discover new physical truths. Jaynes remarked earlier that the principle "is most useful to us in just those cases where it fails to predict the correct experimental facts."[61] He further pointed out that "the first clues indicating the need for the quantum theory [...] were uncovered by a seemingly 'unsuccessful' application of the principle of maximum entropy [and that] we may expect that such things will happen again in the future"[62]. Consequently, Tsallis's "paradigm shift"[63] might indeed happen in the years to come.

Connection of information theory, statistical inference and non-additive entropies
One step in that direction was very recently done by Petr Jizba and Jan Korbel (Jizba & Korbel, 2019; Jizba & Korbel, 2020) who first expanded the class of admissible entropies to non-additive entropy concepts. They showed that "the Shore-Johnson axioms for the maximum entropy principle in statistical estimation theory account for a considerably wider class of entropic functionals than previously thought"[64], retracing their findings to the same hidden requirement that was also found by Uffink earlier. Jizba and Korbel argue that "SE [Shannon entropy] is a unique candidate for MEP [maximum-entropy principle] only when an extra desideratum is added to SJ [Shore-Johnson] axioms, namely, strong system independence [...]: Whenever two subsystems of a system are disjoint, we can treat the

[56] Tsallis, 2019, p. 1.
[57] Tsallis, 2015, p. 2853.
[58] Tsallis, 2015, p. 2854.
[59] Tsallis, 2015, p. 2854.
[60] Tsallis, 2015, p. 2854.
[61] Jaynes, 2003, p. 371.
[62] Jaynes, 2003, p. 371.
[63] Tsallis, 2019, p. 1.
[64] Jizba & Korbel, 2019, p. 1.

subsystems in terms of independent distributions."[65] However, "having no information about interaction encoded in constraints (i.e., having independent constraints) is not the same as having no correlations among systems."[66] To exemplify their findings they investigate a system where "Shannonian MEP predicts entanglement even if [...] there is a separable state that is fully compatible with the constraining data."[67] Furthermore, Jizba and Korbel were able to show the "[e]quivalence of information theory and statistical inference axiomatics"[68] and thus connect the Shannon axiomatics of information theory with the mathematical framework for statistical inference of Shore and Johnson. By a modification of the need for additivity of the entropic functional in the Shannon axiomatic framework they arrive at a general class of entropic functionals that satisfy the information theoretical framework. This class is the same that Jizba and Korbel derived for the Shore-Johnson axiomatic system of statistical inference and therefore, what the authors call the "Uffink class of entropies" is a wide class of entropies that connects statistical inference and information theory even for non-additive entropic functionals (cf. Sec. 2.2.1).

Summary and fields of applications The concept of entropy was initially introduced in the description of heat phenomena in the 19th century. Later on, Boltzmann, Maxwell and Gibbs added their statistical viewpoints arriving at a statistically motivated definition of entropy. With Shannon's work on information theory in the 1940s, apart from the thermodynamic and statistical notions of entropy, a third one as measure of uncertainty became apparent. Based upon this, Jaynes developed the maximum-entropy method as a universal procedure of statistical inference. In this way, the concept left the realm of physics and established itself as a form of "extended logic"[69]. With the advancements of Shore and Johnson in the 1980s, the maximum-entropy principle became independent of a figurative interpretation of entropy. Instead of just maximizing uncertainty or minimizing bias, the procedure of entropy maximization rather ensures consistent conclusions. Nevertheless, this does not make the notion of entropy as a measure of uncertainty less true. With Uffink, Jizba and Korbel's more recent work, this understanding was further expanded to other (non-additive) entropy functionals that lie beyond the Shannonian entropy definition. Regardless of the many unanswered questions that still exist, due to its generality, the principle of entropy maximization has been successfully applied in various fields inside and outside of physics. Some but not all of them are quantum information, complex networks, astronomy, (non-equilibrium) statistical physics, geophysics, biology, geography, economics, medical diagnosis, image processing, quantum chemistry, environmental studies, epidemiology,

[65] Jizba & Korbel, 2019, p. 2.
[66] Jizba & Korbel, 2019, p. 3.
[67] Jizba & Korbel, 2019, p. 3.
[68] Jizba & Korbel, 2020, p. 1.
[69] Jaynes, 2003, p. xxii.

cognitive sciences, ecology, engineering, finance, public health, (Pressé *et al.,* 2013; Zhou *et al.,* 2013; He & Kolovos, 2018; Jizba & Korbel, 2019; Tsallis, 2019; Chanda *et al.,* 2020) and presumably many more to come in the future.

2.2 Maximum-Entropy Method with Moment Constraints

The principle of entropy maximization is a key principle in the theory of statistical inference as well as information theory. In the former, emphasis is placed on the resulting maximum-entropy distribution while the entropy function itself plays a secondary role. In information theory (and thermodynamics), on the other hand, entropy and its properties are the main quantities of investigation. As indicated in the historical overview, today, the connection between both approaches is well-established and both lead to the same class of entropies. The procedure of entropy maximization gives a possibility to self-consistently deduce a full probability distribution out of partial information. Most commonly, only statistical moments of the distribution, i.e., mean value, variance, skewness and so forth, are available. The goal of this section is to determine the general form of a maximum-entropy distribution that is constrained with given moment values. To do so, the main results of (Jizba & Korbel, 2019; Jizba & Korbel, 2020) are summarized to derive the Shannonian entropy definition that is used throughout this thesis (Sec. 2.2.1). Then, the general form of the maximum-entropy distribution is obtained by using a variational principle in Sec. 2.2.2. Finally, bounds on the admissible a priori moment values are summarized for the special case of photon probability distributions (Sec. 2.2.3) and general remarks on numerical implementation of the maximum-entropy method are given in Sec. 2.2.4.

Parts of Sec. 2.2.3 have been published in (Gulyak *et al.,* 2018). Results concerning the Stieltjes moment problem were mainly developed by B. Gulyak. The thesis's author checked them for plausibility. All contributing authors discussed the results. Parts of the second part of Sec. 2.2.3 have been published in (Lettau *et al.,* 2018). Results concerning the upper bound on moments were mainly developed by T. Lettau and the thesis's author. All contributing authors discussed the results. Parts of Sec. 2.2.4 have been published in (Gulyak *et al.,* 2018). Results concerning the numerical procedures were mainly developed by B. Gulyak and the thesis's author. All contributing authors discussed the results.

2.2.1 Shannon Entropy

Two approaches lead to the same class of admissible entropies for the procedure of entropy maximization (Jizba & Korbel, 2019; Jizba & Korbel, 2020). In the first one, an axiomatic system is imposed on the entropy function, while in the second approach from statistical inference theory axioms for the inference procedure itself are given. In this section, the axiomatic systems are briefly stated for both cases. Moreover, the special cases that lead to the well-known Shannon entropy are recapitulated, ending with a collection of its main features.

Let A and B be two random variables with discrete values $\{A_i\}_{i=1}^n$ and $\{B_i\}_{i=1}^m$. The ensuing probability distributions are then denoted with $p = \{p_i\}_{i=1}^n$ and $q = \{q_i\}_{i=1}^m$ and the elementary-event, joint and conditional probabilities are

$$\begin{aligned} P(A_i) = p_i, \quad P(B_j) = q_j, \\ P(A_i, B_j) = r_{ij} \\ P(A_i|B_j) = r_{i|j} = \frac{r_{ij}}{q_j}. \end{aligned} \tag{2.1}$$

The axiomatic system imposed on entropy S is directly cited from (Jizba & Korbel, 2020):

1. "*Continuity:* Entropy is a continuous function with respect to all its arguments"[70].
2. "*Maximality:* Entropy is maximal for uniform distribution"[71].
3. "*Expandability:* Adding an elementary event with probability zero does not change the entropy"[72].
4. "*Composability:* Joint entropy for a pair (A, B) of random variables can be expressed as $S(A, B) = S(A|B) \otimes_f S(B)$, where $S(A|B)$ is conditional entropy satisfying consistency requirements"[73] listed below.

Here, the conditional entropy is given by $S(A|B) = \sum_i q_i S(A|B_i)$ and $\otimes_f$ denotes multiplication in generalized Kolmogorov-Nagumo arithmetic (cf. Jizba & Korbel, 2020 for details). Additional requirements are

I) "For two independent random variables A and B the joint entropy $S(A, B)$ should be *composable* from entropies $S(A)$ and $S(B)$"[74].

[70] Jizba & Korbel, 2020, p. 3.
[71] Jizba & Korbel, 2020, p. 3.
[72] Jizba & Korbel, 2020, p. 3.
[73] Jizba & Korbel, 2020, p. 3.
[74] Jizba & Korbel, 2020, p. 3.

II) "Conditional entropy should be *decomposable* into entropies of conditional distributions"[75].

As demonstrated in (Jizba & Korbel, 2020), this axiomatic system allows for a wide class of admissible entropies

$$S_q^f = f\left\{\exp_q\left[\sum_i p_i \log_q\left(\frac{1}{p_i}\right)\right]\right\} \tag{2.2}$$

where $f(x)$ is a generic strictly increasing function defined on $x \in [0, \infty)$ and $\exp_q$ and $\log_q$ are the q-deformed exponential and q-deformed logarithmic functions

$$\begin{aligned} \exp_q(x) &= [1 + (1-q)x]^{1/(1-q)}, \quad q \neq 1 \\ \log_q(x) &= \frac{x^{1-q} - 1}{1-q}, \quad q \neq 1 \end{aligned} \tag{2.3}$$

that turn into the usual exponential and logarithmic functions $\exp_1(x) = \exp(x)$ and $\log(x)_1 = \log(x)$ for the case $q \to 1$. If the composability axiom is restricted to the requirement for Shannon additivity (cf. Jizba & Korbel, 2020, p. 2)

4.$^{\mathrm{S}}$ *Shannon additivity:* $S(A, B) = S(A|B) + S(B) = S(B|A) + S(A)$,

the unique admissible entropy is the well-known Shannon entropy

$$S(p) = -\sum_i p_i \ln p_i. \tag{2.4}$$

The same class of entropies (2.2) arises from consistency axioms for the statistical inference procedure itself. In (Jizba & Korbel, 2019; Jizba & Korbel, 2020) the authors build upon the former results of Shore, Johnson (Shore & Johnson, 1980; Shore & Johnson, 1981) and Uffink (Uffink, 1995), imposing the following axiomatic system onto the procedure of inference (cited from Jizba & Korbel, 2020):

1. "*Uniqueness:* the result should be unique."[76]
2. "*Permutation invariance:* the permutation of states should not matter."[77]
3. "*Subset independence:* It should not matter whether one treats disjoint subsets of system states in terms of separate conditional distributions or in terms of the full distribution."[78]

[75] Jizba & Korbel, 2020, p. 3.
[76] Jizba & Korbel, 2020, p. 5.
[77] Jizba & Korbel, 2020, p. 5.
[78] Jizba & Korbel, 2020, p. 5.

4. "*System independence:* It should not matter whether one accounts for independent constraints related to independent systems separately in terms of marginal distributions or in terms of full system."[79]

5. "*Maximality:* In absence of any prior information, the uniform distribution should be the solution."[80]

If one replaces axiom 4. with the need for strong system independence (like it was unconsciously done by Shore and Johnson in their original paper)

4.$^{\text{SSI}}$ "*Strong system independence:* Whenever two subsystems of a system are disjoint, we can treat the subsystems in terms of independent distributions"[81]

the requirements are uniquely fulfilled by the Shannon entropy. But for the full axioms 1.–5. the admissible class of entropies for the procedure of entropy maximization is again given by Eq. (2.2), which demonstrates the equivalency of the information theoretical approach and the approach from statistical inference theory.

The present work confines itself to the Shannon entropy, leaving the study of more exotic entropies to future investigations. It is easily obtained by setting $q = 1$ and $f(x) = \ln x$ in the most general expression (2.2). Throughout this thesis, the base of the logarithm will be e, hence the entropy is measured in nats, but other bases could be chosen as well. Further, the convention $0 \ln 0 = 0$ applies. The most important properties of the Shannon entropy are given as follows (Gzyl, 1995; Vedral, 2002; Jaynes, 2003; Cover & Thomas, 2006; Gulyak, 2016):

- *Positivity:* $S(p) \geq 0$ with equality if and only if $p_j = 1$ and $p_{i \neq j} = 0$.
- *Maximum value:* $S(p) \leq \ln n$. Equality if and only if the probability distribution is uniform.
- *Permutation symmetry:* $S(p)$ is invariant under permutation of the probabilities. Let $p = \{p_1, p_2, \dots\}$ and $q = \{p_{\sigma(1)}, p_{\sigma(2)}, \dots\}$ and $\sigma : \mathbb{N} \to \mathbb{N}$ be a permutation of the indices, then $S(p) = S(q)$.
- *Concavity:* $S\left(\sum_{i=1}^{k} \beta_i p_i\right) \geq \sum_{i=1}^{k} \beta_i S(p_i)$ with k probability distributions p and $\sum_{i=1}^{k} \beta_i = 1$. $S(p)$ is a concave function of p, therefore, a local maximum is the global maximum.
- *Triangle inequality:* $S(A_1, A_2, \dots, A_n) \leq \sum_{i=1}^{n} S(A_i)$ with equality if and only if all A_i are independent.

[79] Jizba & Korbel, 2020, p. 5.
[80] Jizba & Korbel, 2020, p. 5.
[81] Jizba & Korbel, 2020, p. 6.

- *Conditioning reduces entropy:* $S(A|B) \leq S(A)$ with equality if and only if A and B are independent.

The first two properties correspond to the notion of entropy as a measure of uncertainty or missing information, since it equals zero when no uncertainty is present, allows for positive values only, and is maximal when the probability distribution contains no bias towards any outcome. Next, the entropy value only depends on the values of the probability function and not on the specific form of the distribution (permutation symmetry). The concavity ensures that once a local maximum of the entropy functional is found, it is guaranteed that it is also the global maximum. The triangle inequality states that whenever marginal distributions are considered instead of the full distribution, one loses information about the system (hence the entropy increases). The last property corresponds to the intuitive notion that knowing the random variable B can only lead to a decrease of the entropy of A, since through knowledge of B the information about A can only increase — or stay the same, if B contains no information about A and hence both are independent. In either case, no information about A can be lost through knowledge of B.

2.2.2 The Maximum-Entropy Distribution

The probability distribution with maximum entropy and accordingly least uncertainty can be derived from the Shannon entropy via a variational principle (Jaynes, 2003; Cover & Thomas, 2006; Gulyak, 2016). Let X_i with $i = 1, \ldots, k$ be k random variables with discrete values $\{x_{j,i}\}$ and let

$$\langle x_i \rangle = \sum_j x_{j,i} p_j = \mu_i, \quad i = 1, \ldots, k \tag{2.5}$$

be the k corresponding expectation values that take the known values μ_i. The range of j is determined by the specific range of the random variables. Additionally, the normalization condition

$$\sum_j p_j = 1 \tag{2.6}$$

has to hold for the probability distribution p. To maximize the Shannon entropy (2.4) under the moment constraints (2.5) and the normalization condition (2.6) one considers the Lagrange function

$$L[p] = -\sum_j p_j \ln p_j - \sum_{i=0}^{k} \lambda_i \left(\sum_j x_{j,i} p_j - \mu_i \right), \tag{2.7}$$

where λ_i denotes the k Lagrange multipliers and the normalization condition is included for the summand $i = 0$ with $x_0 = 1$ and $\mu_0 = 1$. For p to be a maximum-entropy distribution, the first variation of the Lagrange function has to vanish, therefore

$$\delta L[\hat{p}] \stackrel{!}{=} 0, \tag{2.8}$$

where $\hat{p}$ denotes the maximum-entropy distribution compatible with the given constraints. From now on a hat will be used to indicate distributions of maximum entropy. Evaluating Eq. (2.8) leads to

$$\hat{p}_j = \exp\left(-1 - \sum_{i=0}^{k} \lambda_i x_{j,i}\right), \tag{2.9}$$

which can be rewritten by using (2.6), finally leading to a general expression for the maximum-entropy distribution

$$\begin{aligned} \hat{p}_j &= \frac{1}{Z(\lambda)} \exp\left(-\sum_{i=1}^{k} \lambda_i x_{j,i}\right), \\ Z(\lambda) &= \sum_j \exp\left(-\sum_{i=1}^{k} \lambda_i x_{j,i}\right), \end{aligned} \tag{2.10}$$

where $Z(\lambda)$ is referred to as the partition function. The main task is to find suitable Lagrange multipliers λ_i such that the expectation values calculated from the maximum-entropy distribution $\hat{p}$ coincide with the given a priori values μ_i, that is

$$\sum_j x_{j,i} \hat{p}_j - \mu_i = 0, \quad i = 1, \ldots, k. \tag{2.11}$$

For relatively small problems this may be done analytically using the well-know relation

$$\mu_i = \langle x_i \rangle = -\frac{\partial}{\partial \lambda_i} \ln Z(\lambda) \tag{2.12}$$

between the Lagrange multipliers and the given expectation values (Jaynes, 1957a; Jaynes, 1957b; Jaynes, 2003), but in general, numerical methods need to be applied to determine the Lagrange multipliers (cf. Sec. 2.2.4). For a known maximum-entropy distribution the entropy can be determined by evaluating (2.4) or directly from the Lagrange multipliers and the partition function using

$$S = \sum_{i=1}^{k} \lambda_i \mu_i + \ln Z(\lambda) \tag{2.13}$$

(Jaynes, 1957a; Jaynes, 1957b; Jaynes, 2003). Finally, it should be noted that the existence of a maximum-entropy distribution (2.10) is not guaranteed for all possible moment values μ_i, which will be discussed in the following section.

2.2.3 Bounds on the Maximum-Entropy Distribution

Since for this thesis the maximum-entropy method is mainly applied to infer a photon probability distribution from the knowledge of photon moments, the general form of the maximum-entropy distribution derived in the previous section can be concretized. The probability of finding n photons in a light field mode is encoded in the photon-number distribution $p = \{p_n\}_{n=0}^{\infty}$ (the terms photon-number distribution, photon distribution, photon probability distribution, and photon statistics will be used synonymously here). Due to the bosonic character of the photons, the photon number n is not bound to a finite value. It is assumed that the numerical values of the first k moments $\langle n \rangle = \mu_1, \langle n^2 \rangle = \mu_2, \ldots, \langle n^k \rangle = \mu_k$ are given. Eq. (2.5) then becomes

$$\left\langle n^i \right\rangle = \sum_{n=0}^{\infty} n^i p_n = \mu_i, \quad i = 1, \ldots, k \tag{2.14}$$

and the maximum-entropy photon distribution $\hat{p}$ and the partition function $Z(\lambda)$ are given by

$$\begin{aligned} \hat{p}_n &= \frac{1}{Z(\lambda)} \exp\left(- \sum_{i=1}^{k} \lambda_i n^i \right), \\ Z(\lambda) &= \sum_{n=0}^{\infty} \exp\left(- \sum_{i=1}^{k} \lambda_i n^i \right). \end{aligned} \tag{2.15}$$

The distribution $\hat{p}$ is referred to as the maximum-entropy distribution of order k. Since the photon number is not bound to a finite value, the maximum-entropy distribution must exist on $\mathbb{N}_0 = \{0, 1, \ldots, \infty\}$. Consequently, the first necessary existence condition is a positive last Lagrange multiplier

$$\lambda_k \geq 0, \tag{2.16}$$

because otherwise the maximum-entropy distribution (2.15) would grow to infinity for $n \to \infty$ and would not be normalizable.

Secondly, not all possible moment constraints μ_i are allowed, because existence of a corresponding probability distribution is only guaranteed in certain cases. It is pointed out that the following findings apply generally; the distribution does not need to a maximum-entropy distribution. Since the photon distribution is bound to $\mathbb{N}_0$, the Stieltjes moment problem applies here (cf. Shohat & Tamarkin, 1970; Frontini & Tagliani, 1994), where one asks the question of existence and uniqueness of an inversion of a mapping from a measure $\psi(x)$ to its moments

$$\mu_i = \int_0^{\infty} x^i \, \mathrm{d}\psi(x) \, . \tag{2.17}$$

In the case of photon-number distributions, Eq. (2.17) translates to (Tagliani, 2000)

$$\mu_i = \sum_{i=0}^{\infty} n^i p_i. \tag{2.18}$$

For the infinite moment problem (i.e. when moment values up to infinite order are known), a necessary existence condition is that for all $m = 0, 1, \ldots, \infty$ the determinants of the moment Hankel matrices

$$\Delta_m^{(0)} = \begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_m \\ \mu_1 & \mu_2 & \cdots & \mu_{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_m & \mu_{m+1} & \cdots & \mu_{2m} \end{pmatrix}, \quad \Delta_m^{(1)} = \begin{pmatrix} \mu_1 & \mu_2 & \cdots & \mu_{m+1} \\ \mu_2 & \mu_3 & \cdots & \mu_{m+2} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{m+1} & \mu_{m+2} & \cdots & \mu_{2m+1} \end{pmatrix} \tag{2.19}$$

must be strictly positive (Shohat & Tamarkin, 1970; Frontini & Tagliani, 1994):

$$\det\left(\Delta_m^{(0)}\right) > 0 \quad \text{and} \quad \det\left(\Delta_m^{(1)}\right) > 0. \tag{2.20}$$

In the case of photon probability distributions the normalization condition $\mu_0 = 1$ holds and $\mu_i = \langle n^i \rangle$ are the corresponding photon moments. Consequently, the first two existence conditions

$$\begin{aligned} \det\left(\Delta_0^{(0)}\right) &= \mu_0 = 1 > 0 \\ \det\left(\Delta_0^{(1)}\right) &= \mu_1 = \langle n \rangle > 0 \end{aligned} \tag{2.21}$$

are trivially satisfied except for the case of zero photons in the mode. The condition $\det\left(\Delta_1^{(0)}\right) > 0$ implies a lower bound for the second photon moment

$$\left\langle n^2 \right\rangle < \langle n \rangle^2 . \tag{2.22}$$

Oftentimes one is more interested in the auto-correlation function $g^{(2)} = (\langle n^2 \rangle - \langle n \rangle) / \langle n \rangle^2$ (cf. Sec. 3.1.4 for further details), where Eq. (2.22) can be rewritten to

$$\left(1 - g^{(2)}\right) \langle n \rangle < 1, \tag{2.23}$$

which is automatically satisfied for $g^{(2)} \geq 1$ and gives an upper bound for the mean photon number in the other case

$$\langle n \rangle < \frac{1}{1 - g^{(2)}}, \quad \text{if} \quad g^{(2)} < 1. \tag{2.24}$$

In the finite moment problem, where only a finite number k of moments is given (which is the usual case), the number of conditions (2.20) reduces accordingly. They then give a possibility

to (at least numerically) check for the existence of a photon probability distribution. In this finite moment case the conditions (2.20) simultaneously guarantee existence of a maximum-entropy distribution (Tagliani, 2000).

In comparison to the discrete photon probability distribution regarded here, continuous maximum-entropy distributions are a subject studied far better. There, it can be shown that a maximum-entropy distribution of arbitrarily high order exists when certain criteria are met (Dowson & Wragg, 1973; Einbu, 1977; Tagliani, 1990) and that the sequence of maximum-entropy distributions that corresponds to a given sequence of moments converges (in terms of entropy) towards the probability distribution of the Stieltjes moment problem (Frontini & Tagliani, 1997; Milev & Tagliani, 2017). Fortunately, some findings are easily translatable to the discrete case as well. The results from (Einbu, 1977) were translated to the discrete case in (Lettau *et al.,* 2018) and give upper bounds for the allowed moment values which can be applied to successively check the existence of a maximum-entropy distribution. If a given maximum-entropy distribution of order $k-1$ is associated with moment values $\mu_1, \mu_2, \ldots, \mu_{k-1}$ and Lagrange multipliers $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$, then existence of the maximum-entropy distribution of order k is guaranteed if the next moment value μ_k does not exceed the upper bound μ_k^*:

$$\mu_k \leq \mu_k^*. \tag{2.25}$$

The upper bound μ_k^* is given by the moment of order k predicted by the maximum-entropy distribution of order $k-1$, that is

$$\mu_k^* = \sum_{n=0}^{\infty} \frac{n^k}{Z(\lambda)} \exp\left(-\sum_{i=1}^{k-1} \lambda_i n^i\right). \tag{2.26}$$

Since the first order maximum-entropy distribution always exists (cf. App. A.1), Eq. (2.26) provides a possibility to successively check whether the next order maximum-entropy distribution exists for a given moment value. It is important to mention that Eq. (2.26) makes no prediction for the maximum-entropy distribution of order $k+1$, therefore even if the maximum-entropy distribution of order k does not exist, the maximum-entropy distribution of order $k+1$ might nevertheless be well defined (cf. Sec. 3.1.3). As a consequence of Eq. (2.26), an upper bound for the auto-correlation function of a second order maximum-entropy distribution can be derived analytically [cf. App. A.1, Eq. (A.12)]

$$\left\langle n^2 \right\rangle \leq 2\left\langle n \right\rangle^2 + \left\langle n \right\rangle \quad \Leftrightarrow \quad g^{(2)} \leq 2. \tag{2.27}$$

This will be discussed in more detail in Sec. 3.2.1.

Because for a numerical implementation the maximum photon number must necessarily be fixed to a finite maximum value $n_{\max}$, the maximum-entropy distribution is fixed to an approximation space $\{0, 1, \ldots, n_{\max}\}$ there. In that situation, the upper bound (2.26) does no longer apply. But fortunately, the necessary and sufficient conditions for the existence of a maximum-entropy distribution on $\{0, 1, \ldots, n_{\max}\}$ coincide with the general conditions (2.20) for the finite moment problem (Tagliani, 2000; Gulyak *et al.,* 2018). Thus, it is more likely that a maximum-entropy distribution exists on the approximation space than on $\mathbb{N}_0$. Consequently, it should be mentioned that even though a maximum-entropy distribution might be found numerically on $\{0, 1, \ldots, n_{\max}\}$, it does not guarantee its existence on the infinite range $\{0, 1, \ldots, \infty\}$.

2.2.4 Implementation and Numerical Remarks

The main task to determine the maximum-entropy distribution is finding the Lagrange multipliers in Eq. (2.10). To do so numerically, a lot of procedures have been proposed and analyzed in the past (Martínez, 2000; Bandyopadhyay *et al.,* 2001; Wu *et al.,* 2001; Bandyopadhyay *et al.,* 2005; Abramov, 2006; Abramov, 2007; Abramov, 2009; Abramov, 2010; Hao & Harlim, 2018). The approach applied in this thesis is the one proposed in (Batou & Soize, 2013). The basic idea is to transform the problem of constrained entropy maximization [with constraints (2.5) and (2.6)] to an unconstrained minimization problem where the well-known Newton algorithm can be utilized (Wu *et al.,* 2001; Abramov, 2010; Batou & Soize, 2013). To achive that, the duality principle from optimization (cf. Boyd & Vandenberghe, 2004) is applied to derive the concave objective function $f(\boldsymbol{\lambda})$ such that its minimization leads to the optimal Lagrange multipliers $\hat{\boldsymbol{\lambda}} = \arg\min f(\boldsymbol{\lambda})$ (Templeman & Xingsi, 1985; Wu *et al.,* 2001; Abramov, 2010; Batou & Soize, 2013). It can be shown that, due to the advantageous properties of the Shannon entropy (cf. Sec. 2.2.1), solving the dual minimization problem is equivalent to solving the primal entropy-maximization problem (Templeman & Xingsi, 1985; Wu *et al.,* 2001). A derivation of the objective dual function $f(\boldsymbol{\lambda})$ can be found in (Templeman & Xingsi, 1985; Wu *et al.,* 2001) and reads

$$f(\boldsymbol{\lambda}) = \langle \boldsymbol{\lambda}, \boldsymbol{\mu} \rangle + \ln Z(\boldsymbol{\lambda}) \,, \tag{2.28}$$

where $\langle \boldsymbol{\lambda}, \boldsymbol{\mu} \rangle$ denotes the conventional inner product of vectors $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ and $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_k)$. Since the Hessian matrix $\mathbf{H}_f$ (see below) is positive definite, the function f is strictly convex and takes its unique minimum such that $\boldsymbol{\nabla} \boldsymbol{f}(\hat{\boldsymbol{\lambda}}) = \mathbf{0}$ holds.

Starting with an initial value $\boldsymbol{\lambda}^{(0)}$ an iterative Newton method for optimization with the update rule

$$\boldsymbol{\lambda}^{(m+1)} = \boldsymbol{\lambda}^{(m)} - \alpha \left[\mathbf{H}_f(\boldsymbol{\lambda}^{(m)})\right]^{-1} \cdot \boldsymbol{\nabla} f(\boldsymbol{\lambda}^{(m)}) \tag{2.29}$$

is used until convergence of the unconstrained problem is reached, i.e., $\left\|\boldsymbol{\nabla} f(\boldsymbol{\lambda}^{(m)})\right\| \leq \varepsilon$ with a tolerance ε close to zero. The iteration step is denoted by (m) and the product $\boldsymbol{d} = \left[\mathbf{H}_f(\boldsymbol{\lambda}^{(m)})\right]^{-1} \cdot \boldsymbol{\nabla} f(\boldsymbol{\lambda}^{(m)})$ yields the step direction while $0 < \alpha \leq 1$ denotes the step size. To determine the step direction it is numerically more efficient to solve the system of linear equations $\mathbf{H}_f(\boldsymbol{\lambda}^{(m)}) \cdot \boldsymbol{d} = \boldsymbol{\nabla} f(\boldsymbol{\lambda}^{(m)})$ instead of calculating the inverse of the Hessian matrix $\left[\mathbf{H}_f(\boldsymbol{\lambda}^{(i)})\right]^{-1}$. The gradient $\boldsymbol{\nabla} f$ and Hessian matrix $\mathbf{H}_f$ are analytically given by

$$\begin{aligned} \boldsymbol{\nabla} f(\boldsymbol{\lambda}) &= \boldsymbol{\mu} - \langle \boldsymbol{x} \rangle \,, \\ \mathbf{H}_f(\boldsymbol{\lambda}) &= \langle \boldsymbol{x} \otimes \boldsymbol{x} \rangle - \langle \boldsymbol{x} \rangle \otimes \langle \boldsymbol{x} \rangle \,. \end{aligned} \tag{2.30}$$

Here, $\otimes$ denotes the outer product and $\boldsymbol{x}$ is the vector of moments calculated with the current Lagrange multipliers $\boldsymbol{\lambda}^{(m)}$ from iteration step m. Expectation values have to be evaluated component-wise, hence $\langle \boldsymbol{x} \rangle = \left(\langle n \rangle, \langle n^2 \rangle, \ldots, \left\langle n^k \right\rangle\right)$.

Choice of the start value For application of Eq. (2.29) a suitable set of start Lagrange multipliers $\boldsymbol{\lambda}^{(0)}$ needs to be chosen. For most general purposes $\boldsymbol{\lambda}^{(0)} = \mathbf{0}$ which leads to a uniform probability distribution, is a decent choice and quick convergence is usually achieved. Another possibility is to choose Lagrange multipliers that correspond to a Gaussian distribution with the given moment values μ_i (Abramov, 2010). Especially for higher order problems a third strategy can be applied to notably improve convergence. Oftentimes, one varies a specific parameter (e.g. the pump rate P, cf. Sec. 3.1) to derive a set of maximum-entropy distributions for different parameter regimes. In such cases the Lagrange multipliers are usually continuous functions of the specific parameter. Hence, as soon as convergence is achieved for one specific value of the parameter P and the corresponding Lagrange multipliers $\hat{\boldsymbol{\lambda}}_P$ are found, a good starting point for a neighboring value $P \pm \Delta P$ is given by the same Lagrange multipliers $\boldsymbol{\lambda}^{(0)}_{P \pm \Delta P} = \hat{\boldsymbol{\lambda}}_P$. This way, the maximum-entropy distributions for a whole set of parameters can be found with just a few iterations.

Choice of the step size For usual applications it is sufficient to choose a fixed step size α. Nevertheless, convergence of the Newton method might be considerably improved if α is dynamically chosen for each step. Once a descent direction is found, one tries to find a step size that, on the one hand, is not too large for the function to increase again but, on the other hand, not small enough to require too many iterations. For each iteration, this can

be achieved by starting with an initial step size α (e.g. $\alpha = 1$) and reducing it via inexact backtracking linesearch until the Armijo condition

$$f(\boldsymbol{\lambda} + \alpha \boldsymbol{d}) \leq f(\boldsymbol{\lambda}) + c\alpha \boldsymbol{d}^{\intercal} \cdot \boldsymbol{\nabla} \boldsymbol{f}(\boldsymbol{\lambda}) \tag{2.31}$$

is met (Nocedal & Wright, 2006b; Skajaa, 2010), where $c \in [0, 1]$ is usually taken quite small (e.g. $c = 10^{-4}$). Figuratively speaking, the new function value should lie beneath the decreasing tangent line [decreasing since $\boldsymbol{d}^{\intercal} \cdot \boldsymbol{\nabla} \boldsymbol{f}(\boldsymbol{\lambda}) < 0$] at the current point. Because the Armijo condition is necessary but not sufficient for convergence, it is often required additionally that the strong or weak Wolfe condition (Wolfe, 1969; Wolfe, 1971; Nocedal & Wright, 2006b; Skajaa, 2010) hold. For the systems investigated in this thesis, however, requiring the Armijo condition already leads to proper convergence.

Rescaling of the input information One difficulty of the numerical search for the optimal Lagrange multipliers is that in general their numerical values are of very different magnitude. For instance, for the calculations of Sec. 3.1 the magnitude of the Lagrange multiplier λ_i that corresponds to the photon moment $\langle n^i \rangle$ is roughly of order 10^{i-1}. For the numerical search, though, it would be desirable to have Lagrange multipliers that are of the same magnitude. Although more sophisticated procedures were proposed (e.g. Abramov, 2010), a simple yet efficient approach is to rescale the input information

$$\tilde{\lambda}_i = \lambda_i s_i, \quad \tilde{x}_{j,i} = \frac{x_{j,i}}{s_i}, \quad \tilde{\mu}_i = \frac{\mu_i}{s_i}, \quad i = 1, \ldots, k. \tag{2.32}$$

Obviously, the maximum-entropy distribution and the partition sum (2.10) remain unchanged, and the conditional equation (2.11) becomes

$$s_i \left(\sum_j \tilde{x}_{j,i} \hat{p}_j - \tilde{\mu}_i \right) = 0, \quad i = 1, \ldots, k, \tag{2.33}$$

which is fulfilled for arbitrary choices of the scaling factors s_i. A numerically satisfying choice is usually given by

$$s_i = \max_j |x_{j,i}|, \quad i = 1, \ldots, k, \tag{2.34}$$

such that all rescaled values $\tilde{x}_{j,i}$ lie in the range $[-1, 1]$. This results in Lagrange multipliers $\tilde{\lambda}_i$ that do not differ that much in magnitude and convergence is greatly improved.

2.3 Maximum-Entropy Method for Quantum Systems

In quantum mechanics the density matrix is of central importance to describe the statistical state of a system. Similar to the probability distribution of the previous sections, the density matrix contains all information about the system and makes the probability for any outcome available (Nolting, 2013). Fortunately, the maximum entropy principle of the previous sections can straightforwardly be extended to the quantum case (Fick & Sauermann, 1990; Jaynes, 1957a; Jaynes, 1957b; Gulyak, 2016). It answers the question: "having given k pieces of incomplete information about a quantum system what would be a reasonable guess for the density matrix ρ?" The maximum entropy principle suggests to choose the one that has maximum entropy, and thus maximum indeterminateness or rather least bias, yet still accounts for all given constraints. As pointed out in the previous sections, the procedure of entropy maximization can also be seen a as basic self-consistency requirement when drawing inferences from stochastic data, because the procedure of entropy maximization is the only one that ensures self-consistent conclusions without introducing any bias. Sec. 2.3.1 introduces the quantum mechanical extension of the Shannon entropy, that is the von Neumann entropy and its main properties. After that, in Sec. 2.3.2 the concept of observation levels is presented and the density matrix with maximum entropy is derived, followed by a brief outline of what is known as the many-particle hierarchy problem in Sec. 2.3.3. Furthermore, a novel approach to treat it via the maximum-entropy principle is presented. There, instead of deriving equations of motion for the quantities of interest, the maximum-entropy method is applied to answer the question: "what is the system's least biased steady state?" Finally, numerical remarks are given in Sec. 2.3.4.

Parts of this section have been published in (Melcher *et al.,* 2019). The general theory was developed in discussions by all contributing authors of the publication.

2.3.1 Von Neumann Entropy

For quantum systems represented by the density matrix ρ, the measure of indeterminateness is given by the von Neumann entropy (Fick & Sauermann, 1990; Vedral, 2002)

$$S(\rho) = -\operatorname{Tr}(\rho \ln \rho), \tag{2.35}$$

where Tr denotes the trace, i.e., the sum of the diagonal elements. If the density matrix is given in its eigenbasis with eigenvalues $\nu_j \geq 0$ and $\sum_j \nu_j = 1$, one recovers the classical Shannon entropy

$$S(\rho) = -\sum_j \nu_j \ln \nu_j. \tag{2.36}$$

Since the von Neumann entropy is represented by a trace, it is invariant with respect to unitary transformations. Not unlike the features of the Shannon entropy, important features of the von Neumann entropy are (Fick & Sauermann, 1990; Vedral, 2002; Gulyak, 2016):

- *Positivity:* $S(\rho) \geq 0$ with equality if and only if $\nu_j = 1$ and $\nu_{i \neq j} = 0$, so the system is in a pure state.
- *Maximum value:* $S(\rho) \leq \ln n$ where n is the dimensionality of the underlying (finite) Hilbert space. Equality if and only if the system is in a completely undetermined state, i.e., $\rho = I_n / \operatorname{Tr}(I_n)$ with identity matrix I_n.
- *Transformation invariance:* $S(\rho)$ is invariant under change of basis with a unitary transformation U, i.e., $S(\rho) = S(U \rho U^\dagger)$.
- *Concavity:* $S\left(\sum_{i=1}^k \beta_i \rho_i\right) \geq \sum_{i=1}^k \beta_i S(\rho_i)$ with k density operators ρ and $\sum_{i=1}^k \beta_i = 1$. $S(\rho)$ is a concave function of ρ, consequently, a local maximum is the global maximum.
- *Triangle inequality:* $|S(\rho_A) - S(\rho_B)| \leq S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$ where ρ_A and ρ_B are the reduced density matrices from ρ_{AB}.
- *Constancy of time:* $\frac{\mathrm{d}}{\mathrm{d}t} S(\rho) = 0$ during a dynamical process in an isolated system.

Similar to Sec. 2.2.1 the first two properties justify the von Neumann entropy as a measure of uncertainty with zero uncertainty for a pure state and maximum uncertainty when the quantum state is maximally undetermined. The transformation invariance is the quantum mechanical equivalent to the permutation invariance of the classical Shannon entropy and the concavity ensures that once a local maximum is found it is automatically the global maximum. The right hand inequality of the triangle inequality states that the entropy of a composite system can only decrease in comparison to the sum of the individual systems where equality holds if and only if A and B are independent systems. For the left hand inequality, if the subsystems have different entropies then the smaller entropy can only partially cancel out the entropy of the other system and some part of the entropy must be left over when both systems are considered as a composite system. And finally, the last property states that no entropy can be lost or gained during any dynamical process as long as the system is isolated. The indeterminateness of the system remains constant unless a measurement is performed and new information is gained about the system's properties or unless the isolation is lifted and dissipative processes take place.

2.3.2 The Maximum-Entropy Density Matrix

The derivation of the maximum-entropy density matrix is closely related to the approach in Sec. 2.2.2. Here, it is assumed that the a priori information is encoded in expectation values

$$\langle A_i \rangle = \mathrm{Tr}(\rho A_i), \quad i = 1, \ldots, k \tag{2.37}$$

for operators A_i. Usually the A_i should be chosen selfadjointly (cf. App. A.2) and thus correspond to quantum mechanical observables. The given set of k expectation values and corresponding quantum operators

$$\{\alpha\} = \{A_i : i = 1, \ldots, k\}, \quad \mu_i = \langle A_i \rangle \tag{2.38}$$

is then referred to as *observation level* $\{\alpha\}$ (Fick & Sauermann, 1990). It also includes all possible linear combinations of observables within the observation level. Therefore, the results of the maximum-entropy method do not depend on the explicit form in which the information is presented but only on the information itself (Fick & Sauermann, 1990; Gulyak, 2016). In general, the observation level only consists of a small part of all possible observables of the system and the numerical values μ_i might either stem from experimental measurements or calculations from another theory. But it is also possible to circumvent the need for input from external sources which will be the topic of Sec. 2.3.3. Requiring the maximum-entropy density matrix $\hat{\rho}$ (again a hat is used to label the density matrix with maximum entropy) to fulfill the conditions

$$\begin{aligned} S(\hat{\rho}) &= \max \\ \mathrm{Tr}\,(\hat{\rho}) &= 1 \\ \mathrm{Tr}\,(\hat{\rho} A_i) &= \mu_i, \quad i = 1, \ldots, k, \end{aligned} \tag{2.39}$$

leads to the Lagrange function

$$L[\rho] = -\,\mathrm{Tr}\,(\rho \ln \rho) - \sum_{i=0}^{k} \lambda_i (\langle A_i \rangle - \mu_i), \tag{2.40}$$

whose first variation has to vanish for the maximum-entropy density matrix. This ultimately leads to (Jaynes, 1957a; Jaynes, 1957b; Fick & Sauermann, 1990; Gulyak, 2016)

$$\begin{aligned} \hat{\rho}_{\{\alpha\}} &= \frac{1}{Z(\lambda)} \exp\left(-\sum_{i=1}^{k} \lambda_i A_i \right), \\ Z_{\{\alpha\}}(\lambda) &= \mathrm{Tr}\left[\exp\left(-\sum_{i=1}^{k} \lambda_i A_i \right) \right]. \end{aligned} \tag{2.41}$$

Here, $\hat{\rho}_{\{\alpha\}}$ is referred to as density matrix with respect to observation level $\{\alpha\}$, while $Z_{\{\alpha\}}(\lambda)$ is the generalized partition function that ensures that $\mathrm{Tr}\left(\hat{\rho}_{\{\alpha\}}\right) = 1$ is obeyed. Again, the input information enters the maximum-entropy density matrix via the Lagrange multipliers $\lambda_i = \lambda_i(\mu_1, \ldots, \mu_k)$ that are functions of the given constraints. The main task is to find suitable Lagrange multipliers such that

$$\mathrm{Tr}\left[\hat{\rho}(\lambda) A_i\right] - \mu_i = 0, \quad i = 1, \ldots, k, \tag{2.42}$$

is fulfilled. Similar to the non-quantum case, the analytical expression connecting the expectation values with the Lagrange multipliers and the partition function (Jaynes, 1957a; Jaynes, 1957b; Fick & Sauermann, 1990)

$$\mu_i = \langle A_i \rangle = -\frac{\partial}{\partial \lambda_i} \ln Z_{\{\alpha\}}(\lambda) \tag{2.43}$$

holds. Furthermore, the entropy can directly be related to the Lagrange multipliers and the partition function (Jaynes, 1957a; Jaynes, 1957b; Fick & Sauermann, 1990)

$$S_{\{\alpha\}} = \sum_{i=1}^{k} \lambda_i \mu_i + \ln Z_{\{\alpha\}}(\lambda). \tag{2.44}$$

The index $\{\alpha\}$ is used as a reminder that there is not one unique entropy of the system but that it rather depends on the chosen observation level. Moreover, if one chooses to extend an observation level by adding further observables $\{\beta\} = \{B_j : j = 1, \ldots, l\}$, the available information about the system can only increase (or stay the same) and consequently, the entropy can only decrease or remain unchanged (Fick & Sauermann, 1990; Gulyak, 2016)

$$S_{\{\alpha,\beta\}} \leq S_{\{\alpha\}}. \tag{2.45}$$

If a substantial decrease of entropy is achieved by an extension of the observation level, then the added set of observables carries physically relevant information. No change of entropy, on the other hand, indicates irrelevant information. Whether an observation level is sufficient depends on the ability to predict expectation values of other observables

$$\langle F \rangle = \mathrm{Tr}\left(\rho F\right) \approx \mathrm{Tr}\left(\hat{\rho}_{\{\alpha\}} F\right) \tag{2.46}$$

with desired accuracy. Then, the observation level and therefore the choice of input information is physically reasonable. If the current observation level is not sufficient, the maximum-entropy principle gives a strong indication that some physically essential information about the system is not yet captured within the choice of observation level (Jaynes, 1957a; Jaynes, 1957b; Jaynes, 2003). This justifies the usage of the maximum-entropy

method as trial-and-error approach to identify which observables contain physically relevant information.

2.3.3 Many-Particle Hierarchy Problem and Least Biased Steady State

A common theme in the description of many-particle systems is the arising many-particle hierarchy problem, which is briefly summarized at this point before a different handling of this problem is proposed by application of the maximum-entropy method. In the Schrödinger picture the time evolution of the density operator ρ of a system with Hamiltonian H is given by the von Neumann equation (Fick & Sauermann, 1990)

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho = -\frac{\mathrm{i}}{\hbar}[H,\rho]. \tag{2.47}$$

Among many others [see e.g. (de Vega & Alonso, 2017)], one possibility to treat driving and dissipative processes (e.g. pumping processes, radiative losses, cavity losses, etc.) is to divide the whole system into two parts, S and R, where the former is the open system of interest and the latter is interpreted as a reservoir (Carmichael, 1999; Leymann, 2015; Foerster, 2017). The Hamiltonian can then be divided into the respective parts plus a third contribution describing the interaction between system and reservoir

$$H = H_S + H_R + H_{SR}. \tag{2.48}$$

Then, since the exact processes of the reservoir are of minor interest, one takes the partial trace over the reservoir part. With what is known as the Born-Markov approximation, i.e., assuming that the relaxation time of the reservoir is much smaller than the relaxtion time of the system of interest and furthermore, that the system of interest does not affect the reservoir, one arrives at the most general form of a trace preserving master equation, the von Neumann-Lindblad equation (Carmichael, 1999; Leymann, 2015; Foerster, 2017)

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_S = -\frac{\mathrm{i}}{\hbar}[H_S,\rho_S] + \sum_{\nu,\nu'}\vartheta_{\nu,\nu'}\Big(2L_\nu\rho_S L_{\nu'}^\dagger - L_\nu^\dagger L_{\nu'}\rho_S - \rho_S L_\nu^\dagger L_{\nu'}\Big). \tag{2.49}$$

Now, only the density matrix ρ_S of the open system of interest is considered. The commutator describes the reversible part of the time dynamics as in Eq. (2.47) while the irreversible one is given by the Lindblad part that describes the influence of the reservoir on the system of interest. The L_ν are Lindblad operators that act in the system's Hilbert space and $\vartheta_{\nu,\nu'}$ are the corresponding rates that depend on the bath's parameters (Carmichael, 1999; Leymann,

2015; Foerster, 2017). Since the entropy is constant in an isolated system (cf. Sec. 2.3.1) all changes of entropy during the time evolution are caused by the Lindblad part.

One common approach to work with Eq. (2.49) is to derive equations of motion for the desired quantities of interest by using the well-known relation $\langle A \rangle = \mathrm{Tr}(\rho A)$. In doing so, a set of differential equations

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \langle A \rangle &= \frac{\mathrm{i}}{\hbar} \langle [H_S, A] \rangle + \sum_{\nu,\nu'} \vartheta_{\nu,\nu'} \left\langle 2L_\nu^\dagger A L_{\nu'} - L_\nu^\dagger L_{\nu'} A - A L_\nu^\dagger L_{\nu'} \right\rangle \\ &= \langle \mathcal{L}(A) \rangle \end{aligned} \tag{2.50}$$

unfolds that describes the time evolution of several quantum mechanical expectation values (Leymann, 2015; Foerster, 2017). Here,

$$\mathcal{L}(A) = \frac{\mathrm{i}}{\hbar}[H_S, A] + \sum_{\nu,\nu'} \vartheta_{\nu,\nu'} \left(2L_\nu^\dagger A L_{\nu'} - L_\nu^\dagger L_{\nu'} A - A L_\nu^\dagger L_{\nu'} \right) \tag{2.51}$$

is a superoperator that summarizes the effect of the commutator and the Lindblad terms on the operator A (Fick & Sauermann, 1990). Due to the many-particle contributions (e.g. Coulomb interation, electron-photon interaction) in the Hamiltonian and the Lindblad terms, an infinite hierarchy of differential equations

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \langle 1 \rangle &= \langle \mathcal{L}(1) \rangle = \langle 2 \rangle \\ \frac{\mathrm{d}}{\mathrm{d}t} \langle 2 \rangle &= \langle \mathcal{L}(2) \rangle = \langle 3 \rangle \\ &\vdots \end{aligned} \tag{2.52}$$

unfolds, where expectation values of one-particle quantities $\langle 1 \rangle$ couple to expectation values of two-particle quantities $\langle 2 \rangle$ which continues to infinite order unless the hierarchy is truncated at a certain point. Specific examples can be found in for instance (Gies *et al.,* 2007; Leymann, 2015; Foerster, 2017; Grothe, 2018). For a methodical truncation approximation schemes such as the cluster expansion method (Fricke, 1996; Leymann *et al.,* 2014; Leymann, 2015; Foerster, 2017) are widely used. There, depending on the particular problem, higher-order expectation values are neglected, separated into lower-order expectation values, or a reformulation in terms of correlation functions takes place where again, higher-order correlation functions are either neglected or separated into lower-order correlation functions (Leymann, 2015; Foerster, 2017). Although this approach has been successfully applied to a variety of problems, e.g., the microscopical description of exciton dynamics in quantum wells (Hoyer *et al.,* 2003), photoluminescence (Kira *et al.,* 1998; Baer *et al.,* 2006; Feldtmann *et al.,* 2006; Florian *et al.,* 2013), ultracold Bose gases (Köhler & Burnett, 2002), spin dynamics (Kapetanakis & Perakis, 2008), cavity phonons (Kabuss

et al., 2012), quantum-dot lasers (Gies *et al.,* 2007; Ritter *et al.,* 2010; Leymann *et al.,* 2013b; Khanbekyan *et al.,* 2015; Leymann *et al.,* 2015; Jahnke *et al.,* 2016; Fanaei *et al.,* 2016), there are several downsides to those techniques. If not done carefully, the truncation can lead to unphysical behaviour such as the occurrence of negative values for the photon autocorrelation function (Leymann *et al.,* 2014). Consequently, the choice of truncation usually strongly depends on the investigated system and has to be tested carefully for plausibility (Leymann *et al.,* 2014; Grothe, 2018). Moreover, equation-of-motion techniques only provide moments and correlation functions but never the full statistics of quantum mechanical observables, let alone the full density matrix. Furthermore, the same moments and correlation function values are compatible to fundamentally different statistics, so access to the full statistics is necessary to guarantee a clear physical understanding (Leymann *et al.,* 2013b; Schlottmann *et al.,* 2018). Although it is possible to construct the full statistics out of moments alone (cf. Chaps. 3 and 4) the full density matrix is still not at hand and the problem of truncating the infinite hierarchy, that may induce unreasonable effects, remains.

Another possibility to handle Eq. (2.50) and the many-particle hierarchy is by application of the maximum-entropy method (Sobczyk & Trębicki, 1990; Sobczyk & Trębicki, 1993; Trębicki & Sobczyk, 1996; Sobczyk & Trębicki, 1999; Gulyak, 2016). To do so, one regards the system's properties in a steady-state situation, i.e., when switching-on processes have relaxed and the system is in a dynamic equilibrium. Then, gain and loss processes are balanced out. In that case, any observable A_i is constant in time. Consequently,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle A_i\rangle = \left\langle \frac{\mathrm{d}A_i}{\mathrm{d}t} \right\rangle = \mu_i = 0, \quad i = 1, \ldots, k \tag{2.53}$$

can be used as constraint, where the time evolution follows directly from Eq. (2.50). This approach translates to the question: "which is the least biased steady state?" Or put differently: "if the observables A_i are expected to be stationary, what is a consistent guess for the system's state ρ?" The maximum-entropy density matrix (2.41) is then given by

$$\begin{aligned} \hat{\rho}_{\{\alpha\}} &= \frac{1}{Z(\lambda)} \exp\left(-\sum_{i=1}^{k} \lambda_i \mathcal{L}(A_i)\right), \\ Z_{\{\alpha\}}(\lambda) &= \mathrm{Tr}\left[\exp\left(-\sum_{i=1}^{k} \lambda_i \mathcal{L}(A_i)\right)\right], \end{aligned} \tag{2.54}$$

where $\mathcal{L}(A_i)$ is the right-hand side of Eq. (2.50) without the expectation value. The expressions (2.54) can surely be evaluated numerically (cf. Sec. 2.3.4) and since all $\mu_i = 0$, the entropy expression (2.44) reduces handily to

$$S_{\{\alpha\}} = \ln Z_{\{\alpha\}}(\lambda). \tag{2.55}$$

This approach is in strong difference to the inclusion of explicit numerical values for moments $\langle A_i \rangle = \mu_i$. With (2.53) the maximum-entropy-method approach does not depend on explicit numerical values for $\langle A_i \rangle$. These do not need to be obtained by experiment or other calculations and the maximum-entropy method can be applied as a stand-alone approach. Instead of the explicit inclusion of moment values, they rather enter the maximum-entropy method implicitly by being a possible solution to the given equations of motion (Sobczyk & Trębicki, 1999). Another viewpoint is that the algebraic information encoded in the structure of the equations of motion (2.50) themselves is included as constraint (Gulyak, 2016). Note that there is neither any need to solve the equations of motion nor to close the set of equations of motion by truncating the many-particle hierarchy. Consequently, the many-particle hierarchy problem is circumvented and the neglect or factorization of expectation values is no longer needed. Instead, the choice of a sufficient observation level becomes the crucial point. But also in the worst case of a poorly chosen observation level, no unphysical predictions are expected since the procedure of entropy maximization ensures consistent inferences and might at worst only lead to predictions that have higher uncertainty than desired. Quite the contrary, this is one of the strengths of the maximum-entropy method, because an insufficient observation level implies that some part of physically relevant information is not yet included. Via trial and error new information can be probed until the relevant one, i.e., information that leads to a significant decrease of entropy, is found. Jaynes already advocated this kind of approach (Jaynes, 1957a; Jaynes, 1957b; Jaynes, 2003) since it might lead to the investigation of unknown physical effects. One example is elaborately studied in Chap. 5. An apparent downside of the maximum-entropy method approach is that, in contrast to equation-of-motion techniques, only insight into the steady state is gained but no temporal information is available. But fortunately, the steady state is often the more interesting case. Nevertheless, also approaches with time-dependent Lagrange multipliers that would also capture the time-dynamics have been proposed in the past (Trębicki & Sobczyk, 1996; Sobczyk & Trębicki, 1999).

2.3.4 Implementation and Numerical Remarks

Since the last section might have left the impression that the steady-state density matrix is obtained out of thin air, here, the concrete numerical procedure together with several remarks on convergence is given. As mentioned in Sec. 2.3.2 the main numerical task is to find optimal Lagrange multipliers $\hat{\boldsymbol{\lambda}}$ such that the expectation values computed with the maximum-entropy density matrix $\mathrm{Tr}\left[\hat{\boldsymbol{\rho}}(\hat{\boldsymbol{\lambda}})\mathbf{A}_i\right]$ are equal to the input values μ_i. Consequently,

$$\mathrm{Tr}\left[\hat{\boldsymbol{\rho}}(\hat{\boldsymbol{\lambda}})\mathbf{A}_i\right] - \mu_i = 0, \quad i = 1, \ldots, k, \tag{2.56}$$

should hold for all k constraints. This problem can be solved numerically with any suitable non-linear least-squares solver, but here, the Levenberg-Marquardt algorithm (Nocedal & Wright, 2006a) (which is included in standard packages for e.g. *Matlab* or *Python*) is highly recommended, since it has, at least for the problems discussed in this thesis, the best convergence properties. As proposed in Sec. 2.3.3 now the equations of motion (2.50) themselves are used as input. Therefore, it is needed to numerically calculate the quantity $\mathcal{L}(A)$ without the expectation value (Eq. 2.51). To express the quantum mechanical operators as matrices one should choose a suitable basis, for instance the Fock basis for bosonic particles, and calculate the matix elements via the well-known relation $(\mathbf{A})_{rs} = \langle\psi_r|A|\psi_s\rangle$ where $|\psi_r\rangle$ and $|\psi_s\rangle$ denote the basis states and A is the quantum mechanical operator of interest. In matrix representation it is a straightforward task to calculate the right hand side of Eq. (2.51) for a set of operators A_i which results in the matrices $\mathcal{L}(A_i) \rightarrow \mathbf{L}_i$. With these $\mathbf{L}_i$ and a suitable choice of starting Lagrange multipliers $\boldsymbol{\lambda}^{(0)}$ (the same strategies as presented in Sec. 2.2.4 also apply here) the initial density matrix $\hat{\boldsymbol{\rho}}^{(0)}$ can by calculated by application of Eq. (2.54):

$$\begin{aligned} \hat{\boldsymbol{\rho}}^{(m)} &= \frac{1}{Z(\boldsymbol{\lambda}^{(m)})} \exp\left(-\sum_{i=1}^{k} \lambda_i^{(m)} \mathbf{L}_i\right), \\ Z(\boldsymbol{\lambda}^{(m)}) &= \mathrm{Tr}\left[\exp\left(-\sum_{i=1}^{k} \lambda_i^{(m)} \mathbf{L}_i\right)\right]. \end{aligned} \tag{2.57}$$

The iteration step is denoted by $m = 0$. Essentially, the final maximum-entropy density matrix is given by the matrix exponential of the sum of the weighted inputs $\mathbf{L}_i$, where the weights are given by the optimal Lagrange multipliers. As in the non-quantum case (Sec. 2.2.4), a choice of $\boldsymbol{\lambda}^{(0)} = \mathbf{0}$ as start value corresponds to a uniform distribution with diagonal density matrix, since $\exp(\mathbf{0}_n) = \mathbf{I}_n$ where $\mathbf{0}_n$ and $\mathbf{I}_n$ are the square zero and identity matrix of dimensionalty n. Since the relation (2.53) is used as input information, all μ_i are equal to zero and the objective function needed for the iteration step m of the Levenberg-Marquardt algorithm is given by

$$\mathrm{Tr}\left[\hat{\boldsymbol{\rho}}^{(m)}(\hat{\boldsymbol{\lambda}^{(m)}})\mathbf{L}_i\right] = 0, \quad i = 1, \ldots, k. \tag{2.58}$$

The algorithm then delivers the Lagrange multipliers $\boldsymbol{\lambda}^{(m+1)}$ for the next iteration until finally convergence is achieved according to the internal termination criteria of the algorithm.

Rescaling of the input information As in Sec. 2.2.4, a rescaling of the input information is also highly beneficial in the quantum case. Here, the rescaled input

$$\tilde{\lambda}_i = \lambda_i s_i, \quad \tilde{\mathbf{A}}_i = \frac{\mathbf{A}_i}{s_i}, \quad \tilde{\mu}_i = \frac{\mu_i}{s_i}, \quad i = 1, \ldots, k \tag{2.59}$$

is used where the scaling factors are determined by a suitable matrix norm

$$s_i = \|\mathbf{A}_i\|, \quad i = 1, \ldots, k. \tag{2.60}$$

For determination of the least biased steady-state, the sought matrices are of course given with $\mathbf{A}_i = \mathbf{L}_i$. The numerically most beneficial choice of a matrix norm depends on the specific scenario but the row-sum norm and column-sum norm

$$\|\mathbf{A}\|_\infty = \max_r \sum_s |(\mathbf{A})_{rs}|, \quad \|\mathbf{A}\|_1 = \max_s \sum_r |(\mathbf{A})_{rs}| \tag{2.61}$$

give good improvements for all general purposes. Figuratively speaking, the sum of the absolute values of all matrix elements in a row (column) is computed and the scaling factor is taken as the maximum value of all row-sums (column-sums). Consequently, the rescaled input matrices $\tilde{\mathbf{A}}_i$ have a maximum row (column) sum of one. This leads to Lagrange multipliers whose magnitudes are closer together than without rescaling and convergence is greatly facilitated.

Computationally efficient evaluation of the objective function Especially when the size of the matrices $\mathbf{L}_i$ is large, for instance, when large photon numbers or multiple photon modes are considered, numerical evaluation of Eq. (2.58) can get demanding. The most demanding and time-consuming task is the evaluation of the matrix exponential in Eq. (2.57). Therefore, unnecessary multiple evaluation has to be avoided. This can be done by calculating the intermediate result

$$\mathbf{Z}_1 = \exp\left(-\sum_{i=1}^{k} \lambda_i \mathbf{L}_i\right) \tag{2.62}$$

out of which follow the partition sum and a further intermediate result

$$Z(\boldsymbol{\lambda}) = \mathrm{Tr}(\mathbf{Z}_1), \quad \mathbf{Z}_2 = \frac{\mathbf{Z}_1}{Z(\boldsymbol{\lambda})}. \tag{2.63}$$

Eq. (2.58) is then given by

$$\mathrm{Tr}\left(\mathbf{Z}_2 \mathbf{L}_i\right) = 0, \quad i = 1, \ldots, k, \tag{2.64}$$

such that the matrix exponential has to be performed only once for each function evaluation within the Levenberg-Marquardt algorithm. Surely, an efficient algorithm for the evaluation of the matrix exponential should be used (for instance, the implemented matrix exponential functions in *Matlab* or *Python*). Furthermore, expression (2.64) includes k matrix multiplications that are the second most demanding numerical task. Fortunately, since both $\mathbf{Z}_2$ and all $\mathbf{L}_i$ are self-adjoint, the trace of the matrix product can be rewritten into the grandsum (sum of all matrix elements) of the Hadamard (elementwise) matrix product:

$$\mathrm{Tr}(\mathbf{AB}) = \sum_{r,s=1}^{n} (\mathbf{A})_{rs}(\mathbf{B})^*_{rs}. \tag{2.65}$$

Since the matrix multiplication has (in the worst case) a computational complexity of $\mathcal{O}(n^3)$ while the Hadamard matrix produkt has one of $\mathcal{O}(n^2)$, a further speed improvement can be achieved. This is especially beneficial for large matrices or when the number of included input information k is large. Apart from that, it should be mentioned that for all numerical procedures sparse matrix formats should be used (which are readily available in for instance *Matlab* or standard *Python* packages).

3 Maximum-Entropy Method with Moment Constraints

In this two-part chapter basic features of the maximum-entropy method are investigated when the probability distribution is constrained with moment values in the sense of Sec. 2.2. As a critical prerequisite, the maximum-entropy method should be able to reconstruct a given probability distribution out of its statistical moments. To evaluate this, a benchmark model is considered that delivers the full photon-number distribution p_n and consequently gives access to the moments $\left\langle n^k \right\rangle$. These are then used as input for the maximum-entropy method and a maximum-entropy distribution $\hat{p}_n$ is determined that exactly reproduces the moment values in the least biased way. Varying the amount of input information, i.e., the number of included moments, the quality of the maximum-entropy distribution compared to the original distribution is investigated in Sec. 3.1. Secondly, the general existence criteria are exploited to identify the least biased probability distribution that can create superthermal values of the auto-correlation function $g^{(2)} > 2$ in Sec. 3.2.

3.1 Benchmark Model

As model of choice, the birth-death approach by Rice and Carmichael (Rice & Carmichael, 1994) that describes the interaction of a cavity light-field with an atomic system is considered. Although more than two decades old, this model is still widely referred to when it comes to laser systems in general and especially concerning the question of thresholdless lasing and laser transition (Vahala, 2003; Ulrich *et al.,* 2007; Gies *et al.,* 2007; Wiersig *et al.,* 2009; Dubin *et al.,* 2010; Auffèves *et al.,* 2010; Gartner, 2019; Lettau & Leymann, 2019; Wang *et al.,* 2020). Moreover, such rate equation approaches are widely used to be fitted to experimental data (e.g. Khajavikhan *et al.,* 2012; Prieto *et al.,* 2015; Takemura *et al.,* 2019) and for instance extract the β-factor, i.e., the fraction of spontaneous emission directed into the laser mode compared to the overall spontaneous emission, as well as in extended theoretical approaches to laser systems (Leymann *et al.,* 2013b).

After a brief introduction to the model in Sec. 3.1.1, the convergence of the maximum-entropy method of various orders is studied (Secs. 3.1.2 and 3.1.3). Then, since the maximum-entropy method not only yields the probability distribution but also the entropy as

well as Lagrange multipliers, the transition from non-lasing to lasing is reviewed in terms of these quantities and a new threshold definition is proposed in Sec. 3.1.4. Finally, the photon distribution is closely investigated and its entropy is compared to characteristic entropy values of purely thermal, Poissonian and Gaussian distributions in Sec. 3.1.5.

Parts of this section have been published in (Gulyak *et al.,* 2018). The maximum-entropy method calculations were mainly done by B. Gulyak an the thesis's author. All contributing authors discussed the results.

3.1.1 Birth-Death Model

The birth-death model is a phenomenological approach to describe the interaction of a single-mode cavity light-field with atomic (or quantum dot) systems via single electron interaction. Originally, it was developed to "analyze the dependence of the lasing threshold on the fraction, β, of spontaneous emission directed into the laser mode."[1] In spite of its simplicity, it is able to provide all statistical information needed to benchmark the maximum-entropy method approach. Although the authors of the original paper saw their model simply as "a translation of Einstein rate equation theory into probabilistic language"[2], it is also possible to derive it by adiabatically eliminating the off-diagonal matrix elements of the density matrix in a von Neumann-Lindblad equation approach while neglecting superradiance, photon absorption, and self-quenching effects (cf. Gartner & Halati, 2016).

However, here, neighboring states (n, N) and $(n \pm 1, N \pm 1)$ are phenomenologically coupled considering the processes listed in Tab. 3.1. There, n and N denote the number of photons inside the cavity and the number of excited emitters, respectively. The probability

[1] Rice & Carmichael, 1994, p. 4318.
[2] Rice & Carmichael, 1994, p. 4320.

Tab. 3.1.: Physical processes accounted for in the birth-death model. The neighboring states are coupled with the according transition rates. The number of excited emitters is given by N, while n denotes the number of photons inside the cavity. All rates are given in units of the total spontaneous emission rate τ_{sp}.

Rate	States	Process
P	$(n, N) \to (n, N+1)$	Pump
β	$(n, N) \to (n+1, N-1)$	Spontaneous emission into lasing mode
β	$(n, N) \to (n+1, N-1)$	Stimulated emission into lasing mode
$1-\beta$	$(n, N) \to (n, N-1)$	Spontaneous emission into non-lasing modes
κ	$(n, N) \to (n-1, N)$	Cavity losses

of the system's state being (n, N) is then given by $p_{n,N}$. Weighting all incoming rates positively and outgoing rates negatively leads to the birth-death master equation (cf. Rice & Carmichael, 1994)

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}p_{n,N} =& -\kappa[np_{n,N} - (n+1)p_{n+1,N}] \\ &- \beta[nNp_{n,N} - (n-1)(N+1)p_{n-1,N+1}] \\ &- \beta[Np_{n,N} - (N+1)p_{n-1,N+1}] \\ &- (1-\beta)[Np_{n,N} - (N+1)p_{n,N+1}] \\ &+ P[p_{n,N-1} - p_{n,N}].\end{aligned} \tag{3.1}$$

The form of the master equation ensures that the probability distribution stays positive and normalized (cf. Gulyak, 2016).

One key feature of the birth-death approach is that the well-known laser rate-equations can be derived by calculating equations of motion for the mean photon number in the lasing mode $\langle n \rangle$ and the mean number of excited emitters $\langle N \rangle$. By using

$$\left\langle n^i N^k \right\rangle = \sum_{n,N=0}^{\infty} n^i N^k p_{n,N} \tag{3.2}$$

and further neglecting photon emitter correlations by setting $\langle nN \rangle = \langle n \rangle \langle N \rangle$ the laser rate-equations

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle n \rangle &= -\kappa \langle n \rangle + \beta \langle n \rangle \langle N \rangle + \beta \langle N \rangle \\ \frac{\mathrm{d}}{\mathrm{d}t}\langle N \rangle &= -\langle N \rangle + P - \beta \langle n \rangle \langle N \rangle\end{aligned} \tag{3.3}$$

are reproduced (Yokoyama & Brorson, 1989; Rice & Carmichael, 1994; Gies *et al.,* 2009). However, for the benchmark, not the rate equations (3.3) but rather the master equation (3.1) is directly used to determine the steady state

$$\frac{\mathrm{d}}{\mathrm{d}t}p_{n,N} = 0 \tag{3.4}$$

of the system. To achieve that numerically, three different approaches are pointed out that, in principle, lead to the same result but may differ dramatically in terms of robustness and efficiency.

(i) Starting with a system in the ground state, that is $p_{0,0} = 1$, any suitable numerical integration routine, for instance, *ode45* in *matlab*, can be used to integrate until the system's state changes less than a given tolerance.

(ii) Because the master equation (3.1) is linear in $p_{n,N}$ it can conveniently be written as $\mathrm{d}p_{n,N}/\mathrm{d}t = \mathbf{A}p_{n,N}$, where the matrix $\mathbf{A}$ summarizes the right-hand side of Eq. (3.1) acting on a suitably ordered vector $p_{n,N}$. Then, the eigenvector to the eigenvalue of $\mathbf{A}$, which is

closest to zero, gives the steady state (cf. Lettau, 2017).
(iii) Combining Eqs. (3.1) and (3.4), an update rule for a fixed-point iteration can be derived

$$p_{n,N} = \frac{\kappa(n+1)p_{n+1,N} + \beta n(N+1)p_{n-1,N+1} + (1-\beta)(N+1)p_{n,N+1} + Pp_{n,N-1}}{\kappa n + \beta nN + N + P}. \tag{3.5}$$

Again starting with the ground state $p_{0,0} = 1$, Eq. (3.5) is used to update the probability distribution until a fixed point is reached. For stronger pumping, starting with a Poissonian distribution leads to even faster convergence.
For the problems at hand, the fixed-point iteration method is found to be fastest and most versatile. Especially for high pump rates P and/or low β-factors the fixed-point iteration converges, while the other methods either need much longer or do not converge at all.

3.1.2 Photon and Emitter Statistics

From the full distribution $p_{n,N}$ the marginal distributions, i.e., the photon statistics p_n and the statistics of the excited emitters p_N are obtained by simply summing over all emitters or photons, respectively

$$p_n = \sum_{N=0}^{N_{\max}} p_{n,N} \quad \text{and} \quad p_N = \sum_{n=0}^{\infty} p_{n,N}. \tag{3.6}$$

As a first step, the steady-state photon statistics p_n that stems directly from the birth-death master equation (3.1) is compared to the maximum-entropy distributions $\hat{p}_n$ for various orders. For order k the first k photon moments $\left\langle n^k \right\rangle$ of the original probability distribution p_n are included (see Tab. 3.2). Accordingly, for the emitter statistics the moments $\left\langle N^k \right\rangle$ are used. For numerical implementation the maximum photon number needs to be restricted to a finite value $n_{\max}$ (cf. Secs. 2.2.3 and 2.2.4), which does not reflect the actual physical situation but is rather an artificial truncation. The value $n_{\max}$ is usually chosen such that no

Tab. 3.2.: Included input information for various orders of the maximum-entropy method.

Order	Input Information for Photon Statistics	Input Information for Emitter Statistics
1	$\langle n \rangle$	$\langle N \rangle$
2	$\langle n \rangle, \langle n^2 \rangle$	$\langle N \rangle, \langle N^2 \rangle$
3	$\langle n \rangle, \langle n^2 \rangle, \langle n^3 \rangle$	$\langle N \rangle, \langle N^2 \rangle, \langle N^3 \rangle$
4	$\langle n \rangle, \langle n^2 \rangle, \langle n^3 \rangle, \langle n^4 \rangle$	$\langle N \rangle, \langle N^2 \rangle, \langle N^3 \rangle, \langle N^4 \rangle$

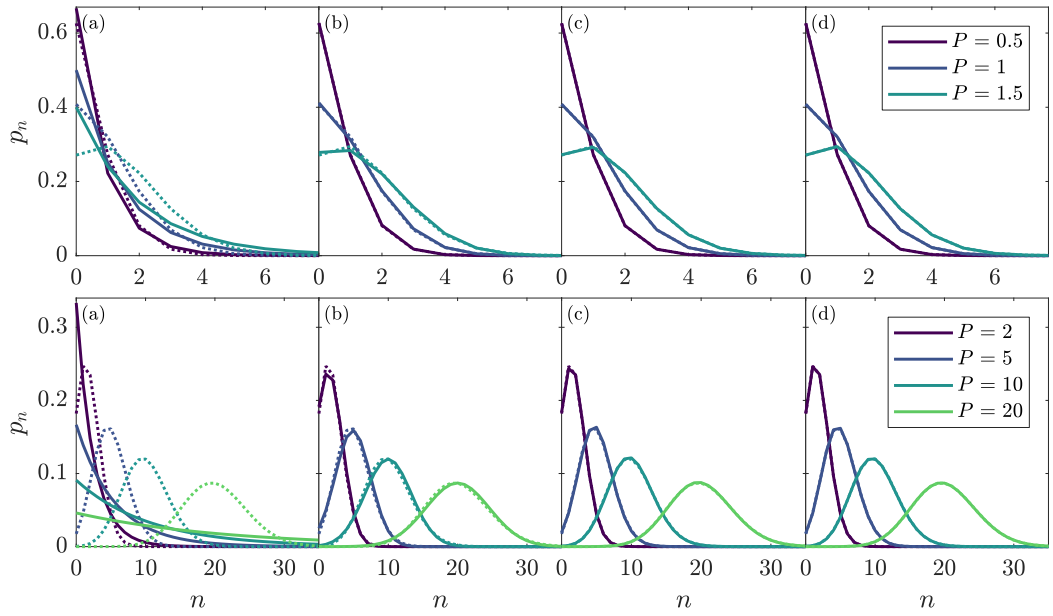


Fig. 3.1.: Comparison of the maximum-entropy method photon-statistics (solid curves) with the original distribution (dashed curves) for weaker (top row) and stronger pumping (bottom row). From left to right the order of the maximum-entropy method is increasing from first (a) to fourth order (d). Parameters are: $n_{\max} = 100$, $N_{\max} = 100$, $\kappa = 1$, $\beta = 1$.

relevant parts, i.e., parts with high probabilities are neglected. The maximum number of excited emitters $N_{\max} = 100$, on the other hand, is the reflection of a real physical situation and is set to a typical value of quantum dot laser (Lohof *et al.*, 2018). The parameters $\kappa = 1$ and $\beta = 1$ are chosen arbitrarily since the general results of the maximum-entropy method do not depend on the specific choice of their values. To numerically determine the maximum-entropy distribution the iterative Newton method is applied as described in Sec. 2.2.4. Varying the pump rate P while keeping all other parameters fixed results in the photon statistics shown in Fig. 3.1. In the top row the results of both approaches are shown for lower pump rates. Dotted lines indicate the original photon distribution from the master equation, whereas solid lines show the maximum-entropy distributions for the first four orders (from left to right). The bottom row depicts the results for stronger pumping. All first order maximum-entropy distributions show exponential behavior as expected from Eq. (2.15). In second order already, the maximum-entropy distributions resemble the original distributions, which indicates that the photon distribution has nearly the Gaussian form $\exp(-\lambda_0 - \lambda_1 n - \lambda_2 n^2)$. In third order the approximation improves further (although with a few caveats, cf. Sec. 3.1.3) and for the fourth-order maximum-entropy distributions the differences are no longer noticeable with the naked eye.

For the emitter statistics the results are quite similar (Fig. 3.2), although in general, a higher order is required to reproduce the original distribution. Note that the parameter $\beta = 0.2$

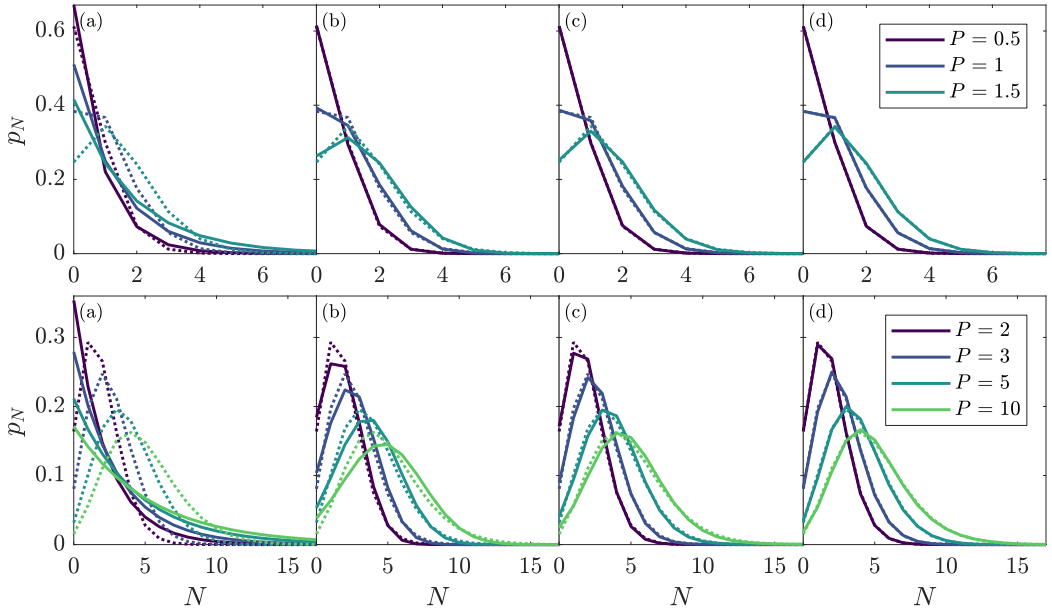


Fig. 3.2.: Comparison of the maximum-entropy method emitter-statistics (solid curves) with the original distribution (dashed curves) for weaker (top row) and stronger pumping (bottom row). From left to right the order of the maximum-entropy method is increasing from first (a) to fourth order (d). Parameters are: $n_{\max} = 100$, $N_{\max} = 100$, $\kappa = 1$, $\beta = 0.2$.

is chosen differently here for demonstration purposes. For a choice of $\beta = 1$ the emitter statistics look almost identically for different pump rates. With a suitable choice of input information that also includes mixed moments $\langle n^i N^j \rangle$ it is also possible to determine the full (photon and emitter) distribution $\hat{p}_{n,N}$ (cf. Gulyak *et al.,* 2018). Also an extension to more dimensions, for instance a bimodal birth-death model, is possible (cf. Gulyak, 2016).

As a measure of the approximation quality the relative entropy (or Kullback Leibler distance) (Vedral, 2002; Cover & Thomas, 2006)

$$D(p_n||\hat{p}_n) = \sum_{n=0}^{\infty} p_n \ln\left(\frac{p_n}{\hat{p}_n}\right) \tag{3.7}$$

is used here, which is always non-negative and exactly zero if and only if $p_n = \hat{p}_n$. It measures the informational inefficiency of choosing $\hat{p}_n$ instead of the original distribution p_n and is superior to, for instance, the summed absolute distance $\sum_i |p_n - \hat{p}_n|$ because it not only considers the deviation of the distributions but also takes the probability values of the original distribution into account. This means that for photon numbers n with a high probability deviations from the original distribution are punished more heavily than for low-probability photon numbers, which generally leads to better characterization of the approximation quality.

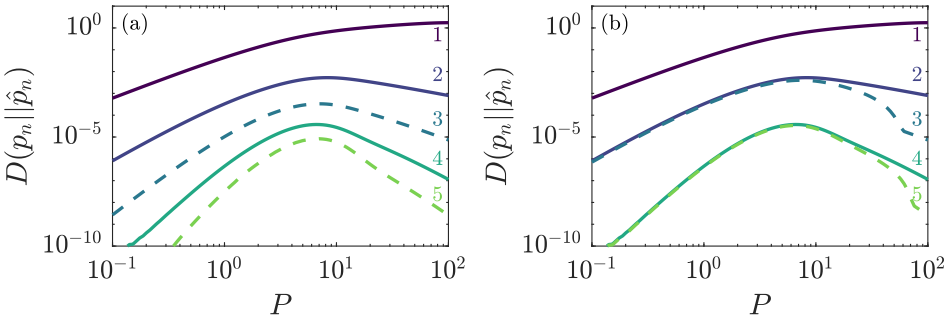


Fig. 3.3.: Relative entropy $D(p_n||\hat{p}_n)$ as a measure of the approximation quality of the maximum-entropy distribution $\hat{p}_n$ compared to the original distribution p_n for a range of pump rates P and maximum-entropy method orders. Dashed curves indicate odd orders, where the maximum-entropy distribution does not exist on the global range $\mathbb{N}_0$ (cf. Sec. 3.1.3). The parameters are: $\kappa = 1, \beta = 1, N_{\max} = 100$, and $n_{\max} = 200$ for panel (b). For panel (a) the maximum photon number $n_{\max}$was chosen according to App. A.3.

With higher order of the maximum-entropy method the relative entropy decreases, hence the approximation quality increases and the sequence of maximum-entropy distributions converges to the original distribution [Fig. 3.3 (a)]. This is true for a whole range of pump rates, although the approximation quality is best for low and high pump rates. The highest relative entropy lies in the transition region where the system changes from non-lasing to lasing because in this regime the photon distribution has the most complicated structure and is neither purely thermal nor Poissonian (cf. Fig. 3.1). However, this is only true for a careful choice of the maximum photon number $n_{\max}$. The behavior that is depicted Fig. 3.3 (b) will be explained in the next section.

3.1.3 Existence of the Distribution

While the maximum emitter number is always physically limited to a finite maximum value $N_{\max}$, the photon number n that in principle can take arbitrarily high values has to be artificially truncated to a finite value $n_{\max}$ for numerical calculations. Because it is assumed that the probability p_n of finding n photons in the cavity decreases for high photon numbers, the photon number is usually truncated at a high enough value $n_{\max}$ and numerical implementations are consequently restricted from the global range $\mathbb{N}_0 = \{0, 1, \ldots, \infty\}$ to an approximation space $\{0, 1, \ldots, n_{\max}\}$. Interestingly, this induces a substantial difference between the even and odd order maximum-entropy distributions (except the first order), which explains the relative entropy shown in Fig. 3.3 (b).

First, the behavior of the Lagrange multipliers is investigated for different orders of the maximum-entropy method and maximum photon numbers $n_{\max}$. For the birth-death model, the last Lagrange multiplier of each even (and first) order has a positive sign, whereas

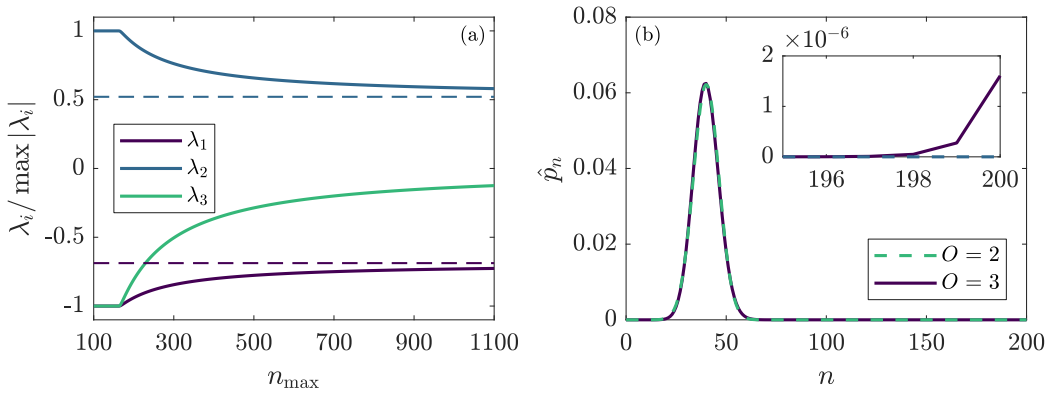


Fig. 3.4.: (a) Normalized Lagrange multipliers for the maximum-entropy method of second (dashed curves) and third order (solid curves). (b) Maximum-entropy distributions of second (dashed curve) and third order (solid curve) for $n_{\max} = 200$. The inset shows an enlarged view of the area $n \approx n_{\max}$. Other parameters are $N_{\max} = 100$, $P = 40$, $\kappa = 1$, $\beta = 1$.

the last Lagrange multiplier of odd orders has a negative sign. This holds true for all investigated parameters, throughout all pump rates and values of $n_{\max}$. In the even order case, the Lagrange multipliers are independent of the choice of $n_{\max}$. Contrariwise, in the odd order case the Lagrange multipliers are constant only for a range of values of $n_{\max}$. For higher choices of $n_{\max}$ though, the Lagrange multipliers start to tend towards the Lagrange multipliers of the previous order and the last Lagrange multiplier slowly approaches zero from below. This behavior is summarized in Fig. 3.4 (a), where the normalized Lagrange multipliers $\lambda_i/\max|\lambda_i|$ are depicted. This quantity gives better insight than the original λ_i because the Lagrange multipliers strongly differ in magnitude (roughly, λ_i is of the order 10^{i-1}). In that representation, the Lagrange multipliers keep their original sign but all lie in the range $[-1, 1]$. The dashed lines exemplarily demonstrate the $n_{\max}$-independence of the Lagrange multipliers for the second order. The solid curves illustrate the third-order Lagrange multipliers with a constant range for small $n_{\max}$, a kink, and then the Lagrange multipliers approaching the values of the second order. Consequently, because the Lagrange multipliers of the third order tend towards the Lagrange multipliers of the second order, also the third-order maximum-entropy distribution tends towards the maximum-entropy distribution of previous order for $n_{\max}$ approaching infinity.

Moreover, in the $n_{\max}$-dependent case the maximum-entropy distribution $\hat{p}_n$ behaves physically unreasonable. As expected, $\hat{p}_n$ decreases for n close to $n_{\max}$ in the $n_{\max}$-independent case (both, even order, and the constant region for odd orders). In the other case though, an increase of the probabilities p_n close to the maximum value $n \approx n_{\max}$ can be observed [Fig. 3.4 (b) inset]. Again, the dashed curve shows the second order maximum-entropy distribution whereas the solid curve demonstrates the behavior of the third-order maximum-entropy distribution close to $n_{\max}$, which is contrary to expectation. In the even-order case, the last Lagrange multiplier has positive sign and therefore the

photon statistics can be expanded from the implementation space $\{0, 1, \ldots, n_{\max}\}$ to the global range $\mathbb{N}_0$ because $\hat{p}_n$ is decreasing and approaching zero as the photon number is approaching infinity. In the odd-order case though, the last Lagrange multiplier has negative sign leading to probabilities exponentially approaching infinity for larger photon numbers. Hence, in the odd-order case, the maximum-entropy distribution cannot be expanded to the global range $\mathbb{N}_0$ because the distribution would no longer be normalizable. For a choice of $n_{\max}$ in the constant region [for instance $n_{\max} = 100$ in Fig. 3.4 (a)] $\hat{p}_n$ is small for large photon numbers. If one chooses another $n_{\max}$ (for instance $n_{\max} = 140$) that is still in the $n_{\max}$-independent range, the Lagrange multipliers and the maximum-entropy distribution are the same as for the previous choice of $n_{\max}$ except that there are more close-to-zero probabilities for large photon numbers. As the last Lagrange multiplier has negative sign, the photon statistics will exhibit large values exponentially approaching infinity for a large enough photon number n. This photon number roughly marks the point where the Lagrange multipliers exhibit a kink and try to compensate for the large probabilities close to $n_{\max}$. For a value of $n_{\max}$ in the $n_{\max}$-dependent regime [for instance $n_{\max} = 200$ as in Fig. 3.3 (b)] the Lagrange multipliers and the maximum-entropy distribution adjust themselves to avoid the overly large probabilities close to $n_{\max}$ (because that would result in a distribution with non-maximal entropy). Finally, with $n_{\max}$ approaching infinity the maximum-entropy distribution approaches the one of the previous (even) order. However, it is possible to derive a reasonable approximation in the odd-order case, if $n_{\max}$ is chosen in the constant range and probabilities for $n > n_{\max}$ are set to zero. A suitable strategy for the choice of $n_{\max}$ is presented in App. A.3. In general, if the maximum-entropy distribution of order k is known, existence of the maximum-entropy distribution of order $k + 1$ on the global range $\mathbb{N}_0$ can numerically be tested with criterion (2.26) [cf. (Gulyak *et al.,* 2018) for an example].

In conclusion, the even (and first) order maximum-entropy distributions are well-defined approximations of the photon statistics that also exist on the global range $\mathbb{N}_0$. For the odd-order maximum-entropy distributions though, a more careful choice of the maximum photon number $n_{\max}$ is necessary. Nevertheless, for all practical purposes, with a decent choice of $n_{\max}$, it is also possible to obtain physically reasonable maximum-entropy distributions in the odd-order case. This is summarized by Fig. 3.3. For the left panel the implementation space was carefully chosen with the algorithm in App. A.3 which leads to reasonable maximum-entropy distributions. For the right panel, on the other hand, a fixed maximum photon number $n_{\max} = 250$ was chosen that leads to the exact same results for the even (and first) order maximum-entropy distributions, but a different behavior in the odd-order case. For weaker pumping, the photon statistics is close to a thermal distribution with highest probabilities for low photon numbers n. There, the maximum photon number $n_{\max} = 250$ is so large that the third (fifth) order maximum-entropy distribution is the same

as the second (fourth) order maximum-entropy distribution. For strong pumping though, the photon statistics exhibits its maximum closer to the maximum photon number. In that case $n_{\max} = 250$ is a good choice and the same results are obtained as in Fig. 3.3 (a). The mid range reflects the transition from $n_{\max}$-dependent to $n_{\max}$-independent regime. So, even in the worst case of an incautious choice of $n_{\max}$ the odd order maximum-entropy distributions will still produce physically reasonable results, but lose the information of the highest order moment.

3.1.4 Threshold Definition Comparison

Already in the original paper of Rice and Carmichael there was put great effort into understanding and clarifying the "confusing picture"[3] of the laser threshold. While at that time typical β-factors of semiconductor lasers were in the order of 10^{-5} (Rice & Carmichael, 1994) and the laser threshold could be well defined by a kink in the input-output curve (or jump in a double-logarithmic plot), nowadays β-factors close to unity have been achieved both at low temperatures ($\beta = 0.85$, cf. Strauf *et al.,* 2006 and $\beta = 0.95$, cf. Khajavikhan *et al.,* 2012) and even at room temperature ($\beta = 0.85$, cf. Prieto *et al.,* 2015) making it impossible to clearly identify a threshold by a kink in the input-output curve [cf. Fig. 4 (a) in Khajavikhan *et al.,* 2012 and Fig. 2 (a) in Prieto *et al.,* 2015]. But besides the high lasing efficiency above the threshold also other features are characteristic. Amongst them a narrowing of the emission linewidth (Kozlov *et al.,* 1997; Samuel *et al.,* 2009; Khajavikhan *et al.,* 2012; Prieto *et al.,* 2015), a steep increase in the coherence time (Ulrich *et al.,* 2007; Wiersig, 2010), and a change in the fluctuations of the photon number (Rice & Carmichael, 1994; Strauf *et al.,* 2006; Ulrich *et al.,* 2007; Gies *et al.,* 2007; Leymann *et al.,* 2013b; Chow *et al.,* 2014; Schlottmann *et al.,* 2018; Lohof *et al.,* 2018). The latter is mostly taken into account by investigating the two-photon correlation function (at zero time-delay; the time argument is omitted throughout this thesis)

$$g^{(2)}(0) = g^{(2)} = \frac{\langle n^2 \rangle - \langle n \rangle}{\langle n \rangle^2}, \tag{3.8}$$

where a transition from the thermal value of $g^{(2)} = 2$, where spontaneous emission dominates, to $g^{(2)} = 1$, which is a characteristic fingerprint of coherent light emission, takes places (Loudon, 2000). Many attempts have been made to finally give a definite characterization of lasing (Kozlov *et al.,* 1997; Samuel *et al.,* 2009) which even lead to a laser checklist published by Nature Photonics in 2017 ("Scrutinizing lasers" 2017).

[3] Rice & Carmichael, 1994, p. 4321.

In this Section, however, a new lasing-threshold definition by a sign change of the first Lagrange multiplier is introduced that makes it possible to mark a distinct pump rate even in the case of thresholdless laser systems. This new definition is compared to common ones and linked to the features of the photon statistics. The focus lies on two main features: the input-output curve and the two-photon correlations captured in the value of $g^{(2)}$.

In terms of the photon statistics p_n the change from non-lasing to lasing can be assumed at the pump power where p_n changes from a monotonically decreasing function to a peaked function (Scully & Lamb, 1967) and therefore exhibits a local maximum. Consequently, for a second order maximum-entropy distribution the following condition has to be satisfied

$$\frac{\mathrm{d}}{\mathrm{d}n}\exp\Big(-\lambda_1 n - \lambda_2 n^2\Big) = 0 \quad \Rightarrow \quad \lambda_1 + 2\lambda_2 n = 0. \tag{3.9}$$

Because n and λ_2 are both non-negative, $\lambda_1 \leq 0$ has to be satisfied if the photon statistics is required to exhibit a local maximum. Consequently, a sign change of the first Lagrange multiplier marks a distinct threshold pump power that is linked to the photon-number distribution. This fact can be exploited to define a threshold for β-factors where a characterization with the above mentioned conditions fails.

In Fig. 3.5 the novel definition is compared to the threshold definitions by a jump in the laser intensity (which is proportional to the mean photon number $\langle n \rangle$) and a jump in the auto-correlation function $g^{(2)}$ for various β-factors, especially for those closer to the case of thresholdless lasing. In the Figure the fourth-order maximum-entropy-method results are depicted, but the sign change of the first Lagrange multiplier is found to be almost independent of the order. For smaller values β, both, the input-output curves [Fig. 3.5 (a)] and the two-photon correlation function $g^{(2)}$ [Fig. 3.5 (b)] exhibit a jump at the threshold pump power. For β close to unity, on the other hand, the jump smears out and it becomes harder to distinguish a distinct threshold pump rate. In contrast to that, even for β close to one a transition can be marked unambiguously by the sign change of the first Lagrange multiplier λ_1 [Fig. 3.5 (c)]. All three definitions are in perfect accordance (dashed line). Furthermore, the sign change of λ_1 can be linked to a value of $g^{(2)}$, which in contrast to λ_1 is directly accessible experimentally. This will be topic of Sec. 3.2.1.

As the full statistics and the entropy of the system is at hand, two alternative threshold definitions are possible. Firstly, the photon entropy S_n grows as the pump rate increases, yet, considerably slower in the region where stronger pumping only results in a shifting of the Poissonian distribution [Fig. 3.5 (e)]. Consequently, the transition from high to low slope can be interpreted as phase transition (Phoenix & Knight, 1988). But again, the kink disappears in the case of thresholdless lasing. Secondly, the mutual information

$$I = S_N + S_n - S_{n,N} \tag{3.10}$$

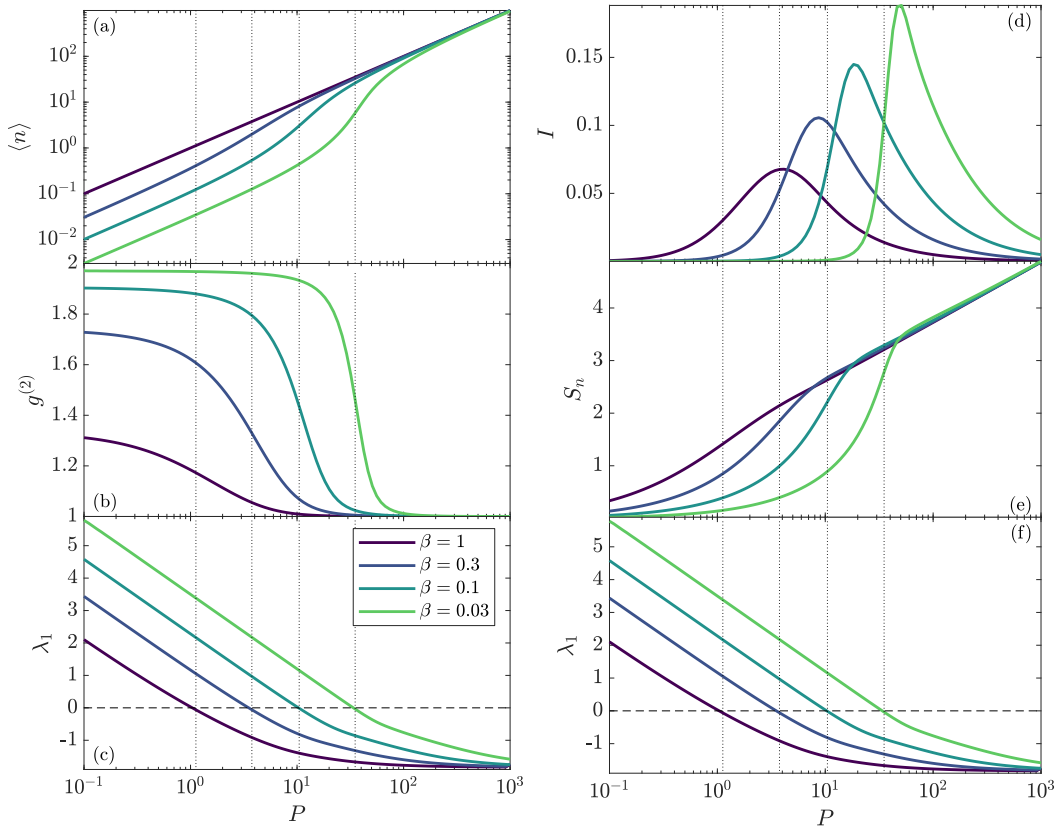


Fig. 3.5.: Threshold definition comparison. In the left column, common definitions by a jump in the input-output curve (a) and by a jump in the autocorrelation function (b) are compared to the sign change of the first Lagrange multiplier λ_1 (c) [and (e)]. The dashed horizontal line marks the zero line $\lambda_1 = 0$. Dashed vertical lines show the positions of the threshold by sign change of λ_1. The right column shows further possible threshold definitions by a kink in the photon entropy S_n (e) or a local maximum in the mutual information I (d), again compared with the sign change of λ_1 (f). The Lagrange multipliers stem from the fourth order maximum-entropy method. The parameters are: $n_{\max} = 1500$, $N_{\max} = 100$ and $\kappa = 1$.

where S_N and $S_{n,N}$ are the respective entropies of the emitter statistics p_N and the full statistics $p_{n,N}$. It measures the amount of information obtained about one random variable by measuring the other random variable and can also be used as a threshold definition because a local maximum can be observed even for $\beta = 1$ [Fig. 3.5 (d)]. In comparison to the traditional threshold definitions as well as the sign change of λ_1, the threshold is generally shifted to higher pump rates (cf. dashed lines in Fig. 3.5). Even though a threshold definition by entropy is possible, the definition by sign change is clearly superior.

In conclusion, the sign change of the first Lagrange multiplier λ_1 gives a distinct threshold definition that is also applicable in cases where the β-factor is close to unity and usual threshold definition do not further apply. Most importantly, the new definition is intimately connected to the features of the statistics of the light field. Moreover, there are cases where

the jump in the input-output curve occurs at a very different pump power than the jump in the auto-correlation function and a clear characterization becomes difficult. Such a system will be investigated more closely in Sec. 4.2.

3.1.5 Characterization of the Emitted Light by Entropy

It is common practice to characterize the emitted light by the moments of the photon distribution, i.e., the mean photon number $\langle n \rangle$ and the second order auto-correlation function $g^{(2)}$ that contains information about the second photon moment $\langle n^2 \rangle$. Sometimes also higher order correlation functions

$$g^{(k)} = \frac{\left\langle \left(b^\dagger\right)^k b^k \right\rangle}{\langle b^\dagger b \rangle^k} \tag{3.11}$$

are used (Rice & Carmichael, 1994; Jones *et al.,* 1999; Schneebeli *et al.,* 2008; Richter *et al.,* 2009; Wiersig *et al.,* 2009; Leymann *et al.,* 2013b; Chow *et al.,* 2014; Leymann *et al.,* 2014; Leymann *et al.,* 2015; Kazimierczuk *et al.,* 2015; Jahnke *et al.,* 2016; Fanaei *et al.,* 2016; Strauß *et al.,* 2016; Foerster *et al.,* 2017; Gies *et al.,* 2017; Schlottmann *et al.,* 2018; Lohof *et al.,* 2018; Lettau & Leymann, 2019). There, $b^\dagger$ and b are photonic creation and annihilation operators and since $b^\dagger b = n$, the auto-correlation function of order k contains the respective higher order moments $\left\langle n^k \right\rangle$. Here, another point of view is added by characterizing the emitted light in terms of entropy.

Usually, the photon distribution is assumed to be thermal below threshold and Poissonian above threshold (Gies *et al.,* 2007; Gies *et al.,* 2009; Gies *et al.,* 2017). In Sec. 3.1.2 it was pointed out that p_n closely resembles a Gaussian distribution. Therefore, the entropy of the photon distribution is compared with characteristic entropy curves that can be derived for thermal, Poissonian and Gaussian distributions. The considered distributions and their corresponding entropies are listed in Tab. 3.3, where the Poissonian entropy can be approximated for large $\langle n \rangle$ (Livesey & Skilling, 1985; Evans *et al.,* 1988) and the Gaussian entropy is derived for a continuous normal distribution on $(-\infty, \infty)$ (Lazo & Rathie, 1978). The entropy for the thermal distribution is derived in App. A.4. If one chooses the thermal

Tab. 3.3.: Thermal, Poissonian and Gaussian distributions and their respective entropies.

Distribution	p_n	Entropy
Thermal	$\exp(-\lambda_1 n)/Z(\lambda_1)$	$-\langle n \rangle \ln \langle n \rangle + \langle n+1 \rangle \ln \langle n+1 \rangle$
Poissonian	$\langle n \rangle^n \exp(-\langle n \rangle)/n!$	$\approx \frac{1}{2} \ln(2\pi e \langle n \rangle)$
Gaussian	$\exp(-\lambda_1 n - \lambda_2 n^2)/Z(\lambda_1, \lambda_2)$	$\approx \frac{1}{2} \ln[2\pi e \operatorname{var}(n)]$

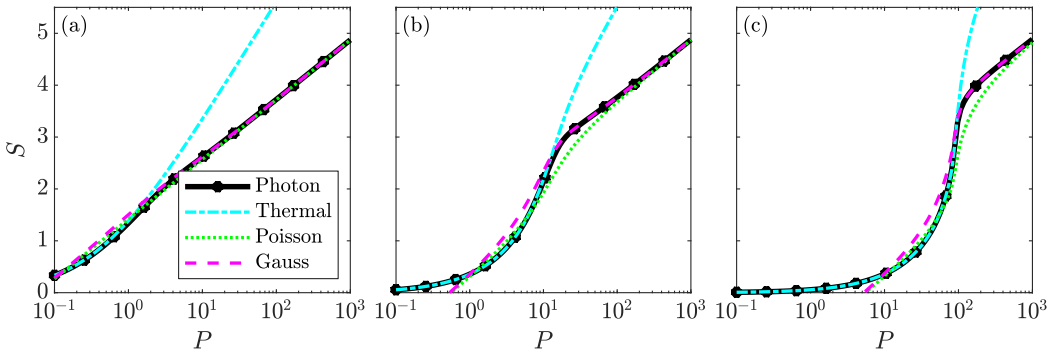


Fig. 3.6.: Entropy S of the original photon distribution (solid black curves) compared to the entropy curves of a purely thermal (dash-dotted), Poissonian (dotted) and Gaussian distribution (dashed) for different β-factors (a) $\beta = 1$, (b) $\beta = 0.1$, and (c) $\beta = 0.01$. Other parameters are: $\kappa = 1$, $n_{\max} = 1500$, $N_{\max} = 100$.

distribution such that the mean photon number $\langle n \rangle$ of the original distribution is reproduced, then the thermal distribution is equivalent to the first-order maximum-entropy distribution. Its entropy only depends on the first moment $\langle n \rangle$. For a Poissonian distribution, which is expected above threshold, the entropy value also only depends on the first moment $\langle n \rangle$. The entropy of a Gaussian distribution, on the other hand, also depends on the second moment $\langle n^2 \rangle$ which enters the variance as follows

$$\mathrm{var}(n) = \left\langle n^2 \right\rangle - \langle n \rangle^2 . \tag{3.12}$$

For large values of $\langle n \rangle$ the Poisson distribution is similar to a Gaussian distribution with variance $\mathrm{var}(n) = \langle n^2 \rangle - \langle n \rangle^2 = \langle n \rangle$.

A comparison of the entropy curves for different values of β is shown in Fig. 3.6. For low pump rates, as expected, the thermal distribution fits best. For higher pump rates though, surprisingly the Poisson distribution is not the best fit. Rather, the Gaussian entropy fits better, especially in the case of lower β-factors. Although unexpected, this finding is not completely counter-intuitive. While being a necessary condition, a value of $g^{(2)} = 1$ does not automatically imply a Poissonian distribution, because it only contains partial information about the distribution. Also, auto-correlations of higher order $g^{(k)}$ approaching one (cf. Fig. 8 in Gulyak *et al.*, 2018) do not necessarily imply a Poissonian distribution. In fact, the photon distribution is much broader than a Poissonian distribution with equal mean photon number [Fig. 3.7 (a)]. For a Poisson distribution the mean and variance coincide, hence

$$\mathrm{var}(n) - \langle n \rangle = \langle n \rangle^2 \left(g^{(2)} - 1 \right) = 0 \tag{3.13}$$

should hold. Although $g^{(2)} \to 1$ holds true, the left-hand side of the equation above grows for higher pump rates because $\langle n \rangle^2$ grows faster than $g^{(2)}$ approaches one. This is

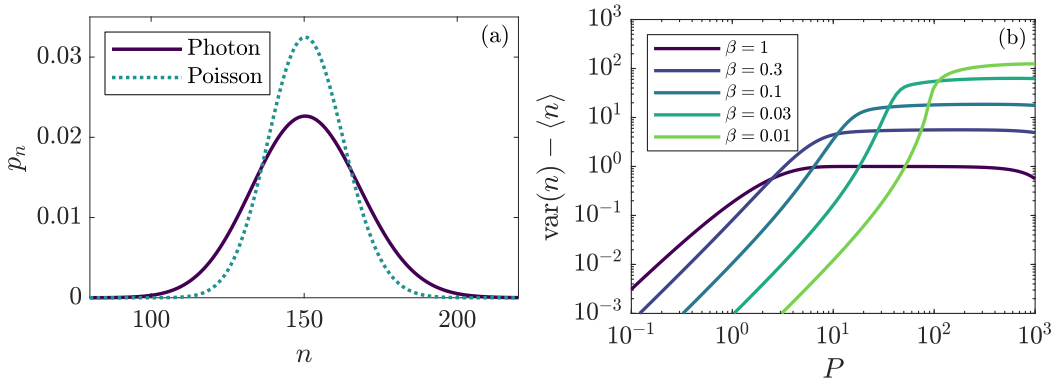


Fig. 3.7.: (a) Comparison of the original photon distribution (solid curve) with a Poissonian distribution with same mean photon number (dotted curve). Parameters: $\kappa = 1$, $\beta = 0.01$, $P = 250$, $n_{\max} = 1500$, $N_{\max} = 250$. (b) Deviation of the variance $\mathrm{var}(n)$ from the mean photon number. Parameters: $\kappa = 1$, $n_{\max} = 1500$, $N_{\max} = 100$.

illustrated in Fig. 3.7 (b) for different values of β. For a Poissonian distribution the quantity $\mathrm{var}(n) - \langle n \rangle$ should tend towards zero, which is clearly not the case here. Consequently, the photon distribution is much broader than a Poisson distribution with the same mean value would be. Moreover, it may well be the case that all $g^{(k)}$ are approaching unity, yet still the distribution is not Poissonian since $\langle n \rangle^k$ appears in the denominator of $g^{(k)}$ [cf. Eq. (3.11)] and it may be possible that $\langle n \rangle^k$ grows faster than $g^{(k)}$ approaches unity (cf. Gulyak *et al.,* 2018). Consequently, the conclusion, that the photon distribution is Poissonian just by observing that the second and/or higher order auto-correlation functions $g^{(k)}$ approach unity, might be deceptive. Or as Gies et al. accurately summarized: "While a distribution function can be accurately represented by all of its moments, the autocorrelation function $g^{(2)}$ [...] contains only partial information on the full photon statistics. As such, an interpretation of $g^{(2)}$ [...] requires at least some intuition about the underlying distribution function, otherwise results can be misleading."[4] This statement can be further generalized even to the case when higher order auto-correlation functions $g^{(k)}$ approach unity.

For this particular example the characterization by entropy is superior to the characterization of the emitted light by the value of $g^{(k)} \approx 1$. It thus can be recommended to also consider the photon distribution's entropy as a means to characterize the light field whenever it is possible. Instead of being Poissonian, the photon distribution above threshold for the birth-death model is rather Gaussian. Nevertheless, the auto-correlation function $g^{(2)}$ is still a valuable quantity to characterize the emitted light because it is also able to indicate effects that go beyond spontaneous and coherent emission, which will be topic of the next section.

[4] Gies *et al.,* 2017, p. 7.

3.2 Superthermal Photon Bunching

The value of the auto-correlation function has become a reliable measure when it comes to microcavity devices (Strauf *et al.,* 2006; Ulrich *et al.,* 2007; Chow *et al.,* 2014). Apart from being accessible relatively easy, both theoretically [for instance with low-order cluster-expansion method (Leymann *et al.,* 2014)] and experimentally [via Hanbury Brown and Twiss setup (Hanbury Brown & Twiss, 1956)], it is also capable of indicating effects that go beyond spontaneous and stimulated emission. While a value of $g^{(2)} = 1$ indicates completely uncorrelated emission events, higher values imply more pronounced bunching of the photons (Loudon, 2000). A value of $g^{(2)} = 2$ is characteristic for thermal light sources (Loudon, 2000), whereas for values $g^{(2)} > 2$ (superthermal light) the bunching of the photons increases steadily and hence indicates novel emission regimes. One of the possible mechanisms of superthermal light emission can be coupling of the emitters in the gain medium via the common light field (Dicke, 1954; Gross & Haroche, 1982; Scheibner *et al.,* 2007; Scully & Svidzinsky, 2009; Mlynek *et al.,* 2014; Leymann *et al.,* 2015; Jahnke *et al.,* 2016). Furthermore, in bimodal lasers where coupling between the two modes becomes relevant superthermal photon bunching in one of the modes has been observed (Leymann *et al.,* 2013a; Leymann *et al.,* 2013b; Fanaei *et al.,* 2016; Redlich *et al.,* 2016; Marconi *et al.,* 2016; Marconi *et al.,* 2018). Apart from being interesting from a theoretical point of view, superthermal photon bunching might also be beneficial for applications relying on nonlinear optical processes (Qu & Singh, 1992; Bennink *et al.,* 2002; Gatti *et al.,* 2004; Harder *et al.,* 2014; Spasibko *et al.,* 2017).

In this section, generic features are investigated that an unbiased photon distribution must have to produce superthermal values of $g^{(2)}$. It is shown that there cannot exist a single-mode photon distribution that is capable of producing superthermal values of $g^{(2)}$ without including additional information about the probability distribution (Sec. 3.2.1). Two anti-correlated modes, on the other hand, can produce superthermal light depending on the degree of anti-correlation between those modes (Sec. 3.2.2).

Parts of this section have been published in (Lettau *et al.,* 2018). Results concerning single-mode distributions were mainly developed by T. Lettau and H. A. M. Leymann. The thesis's author checked the results for plausibility. Results concerning two-mode distributions were mainly developed by T. Lettau, H. A. M. Leymann and the thesis's author. All contributing authors discussed the results. Parts of Sec. 3.2.1 have been published in (Gulyak *et al.,* 2018). Results concerning the connection of $g^{(2)}$ and sign change of the first Lagrange multiplier were mainly developed by B. Gulyak. All contributing authors discussed the results.

3.2.1 Superthermal Photon Bunching for Single-Mode Distributions

The existence of an unbiased single-mode photon distribution that exhibits superthermal values of the auto-correlation function $g^{(2)} > 2$ can be investigated by an application with the maximum-entropy method. As shown in the previous chapters, the procedure to generate an unbiased probability distribution is by finding the Lagrange multipliers of the maximum-entropy distribution that reproduces the given expectation values. Because the second order auto-correlation function contains the first two photon moments $\langle n \rangle$ and $\langle n^2 \rangle$, the maximally unbiased photon distribution is given by the second order maximum-entropy distribution

$$\hat{p}_n = \frac{1}{Z(\lambda)} \exp\Big(-\lambda_1 n - \lambda_2 n^2\Big), \quad Z(\lambda) = \sum_n \exp\Big(-\lambda_1 n - \lambda_2 n^2\Big). \tag{3.14}$$

As in the previous sections, for the maximum-entropy distribution to exist on the global range $\mathbb{N}_0$ the second Lagrange multiplier λ_2 has to be non-negative. Otherwise, the photon distribution cannot be normalized. For numerical reasons, on the other hand, the implementation space is always limited to a maximum photon number $n_{\max}$ and a maximum-entropy distribution might even be found when it does not exist on the global range $\mathbb{N}_0$ (cf. Secs. 2.2.3 and 3.1.3)

What might sound as a downside is actually exploited to decide whether or not a maximum-entropy distribution exists that is able to exhibit superthermal values of $g^{(2)}$. To do so, two different strategies are applied here. Firstly, the Lagrange multipliers in Eq. (3.14) are varied to generate a photon distribution $\hat{p}$ from scratch and to investigate the resulting mean photon numbers and the auto-correlation function values. As a second approach, the Lagrange multipliers are determined numerically for a wide range of input values of the mean photon number and the auto-correlation function to conclude which characteristic features a maximum-entropy distribution has.

For the first approach (Fig. 3.8), both Lagrange multipliers are varied. For the second Lagrange multiplier λ_2 though, only positive values are considered to keep the resulting photon distribution normalizable on the global range $\mathbb{N}_0$. In the left panel, the mean photon number is depicted as colorplot for different choices of λ_1 and λ_2. Larger positive values of λ_1 result in small mean photon numbers, whereas negative values lead to higher values of $\langle n \rangle$. This is because the maximum-entropy distribution exhibits a local maximum that shifts to higher photon numbers for lower values of λ_1 (cf. Sec. 3.1.4). Two further comments are added that are not clearly visible in the figure:
(i) Arbitrarily high mean photon numbers can be achieved by choosing a large negative value of λ_1 while simultaneously choosing a small positive value of λ_2 close to zero. This is not clearly visible in the colorplot, because all values of $\langle n \rangle > 10$ are shown in the same

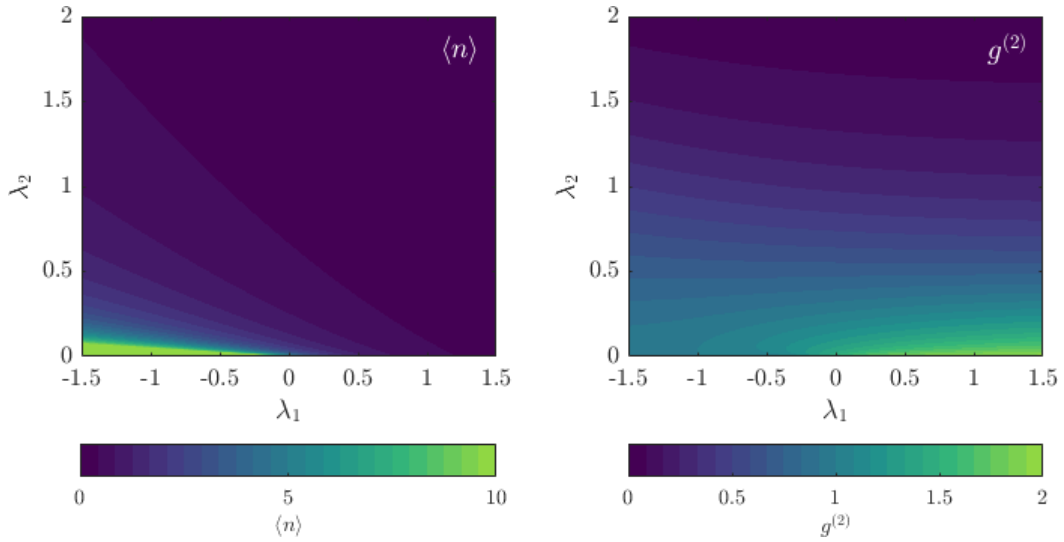


Fig. 3.8.: Mean photon number $\langle n \rangle$ (left) and auto-correlation function $g^{(2)}$(right) that result from a second order maximum-entropy distribution with given Lagrange multipliers λ_1 and λ_2. Note that, for the mean photon number (left panel) also values $\langle n \rangle > 10$ occur in the lower-left region but as the colormap is truncated at $\langle n \rangle = 10$ these are depicted in the same color. However, in the right panel the colormap is not truncated and the auto-correlation function value is limited to maximum value of $g^{(2)} = 2$. Parameter: $n_{\max} = 200$.

color.

(ii) For a choice of $\lambda_2 = 0$ the maximum-entropy distributions do only exist on the global range $\mathbb{N}_0$ for positive values of λ_1, again because for negative values, the photon distribution would not be normalizable.

The right panel of Fig. 3.8 depicts the possible $g^{(2)}$ values for different choices of the Lagrange multipliers. In contrast to the mean photon number, only a maximum value of $g^{(2)} = 2$ is possible here, which is a strong indication that superthermal photon bunching cannot be achieved in the single-mode case without including information beyond the second-order auto-correlation function.

For the second approach, the Lagrange multipliers are determined numerically for a wide range of the mean photon number $\langle n \rangle$ and the auto-correlation function $g^{(2)}$. Fig. 3.9 depicts the resulting Lagrange multipliers λ_1 (left panel) and λ_2 (right panel). As seen in the right panel, the second Lagrange multiplier has positive sign for values of $g^{(2)} \leq 2$ only, whereas its sign changes for superthermal values above $g^{(2)} > 2$ (dashed white line). Furthermore, for $g^{(2)} > 2$ the Lagrange multipliers become strongly dependent on the choice of $n_{\max}$. Referring to the arguments given in Sec. 3.1.3, this shows that there cannot exist an unbiased single-mode photon distribution that produces superthermal values of the auto-correlation

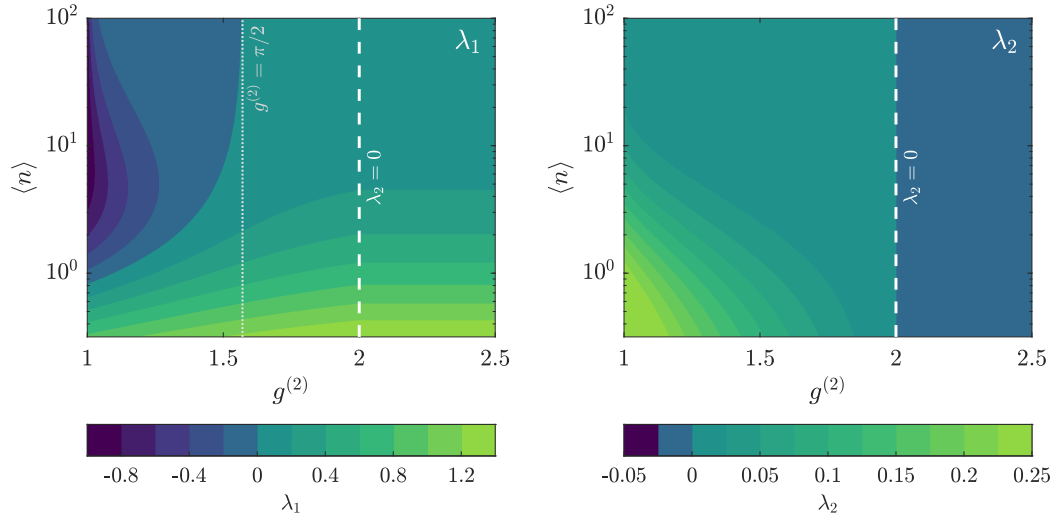


Fig. 3.9.: Numerically determined Lagrange multipliers λ_1 (left) and λ_2 (right) for a given mean photon number $\langle n \rangle$ and auto-correlation function value $g^{(2)}$. The regions where the second Lagrange multiplier changes its sign are separated by a dashed white line. The dotted gray line indicates $g^{(2)} = \frac{\pi}{2}$ which corresponds to the sign change of λ_1 for $\langle n \rangle$ approaching infinity. Parameter: $n_{\max} = 5000$.

function $g^{(2)} > 2$. Moreover, Eq. (2.26) can be used to analytically derive an upper bound on $\langle n^2 \rangle$ for a second-order maximum-entropy distribution (cf. App. A.1 and Sec. 2.2.3)

$$\left\langle n^2 \right\rangle \leq 2 \langle n \rangle^2 + \langle n \rangle . \tag{3.15}$$

This directly leads to a constraint for possible values of the auto-correlation function

$$g^{(2)} \leq \frac{2 \langle n \rangle^2 + \langle n \rangle - \langle n \rangle}{\langle n \rangle^2} = 2, \tag{3.16}$$

which confirms the numerical findings.

Additionally, a more general application of the results depicted in Fig. 3.9 is pointed out. The λ_1-landscape can be utilized to determine the change from thermal to Poisson-like distribution independently from kinks or jumps in the values of $\langle n \rangle$ and $g^{(2)}$ and without need for explicit implementation of the maximum-entropy method. To do so, for each given pair $\left(\langle n \rangle , g^{(2)} \right)$ the it can be used to determine whether or not the photon distribution has a local maximum by the sign of λ_1. More easily, for large $\langle n \rangle$ the sign change happens at a value of $g^{(2)} = \frac{\pi}{2}$ (Gulyak *et al.*, 2018), hence

$$g^{(2)}_{\text{thr}} = \frac{\pi}{2} \quad \text{if} \quad \langle n \rangle \gg 1 \tag{3.17}$$

can be used as threshold criterion. This approach can be especially handy when the kinks or jumps of $\langle n \rangle$ and $g^{(2)}$ occur at different pump rates. One example will be given in Sec. 4.2.

3.2.2 Anti-Correlated Two-Mode Distributions

It is pointed out that the findings of the previous section do not imply that a maximum-entropy distribution of third (or higher) order cannot produce values of $g^{(2)} > 2$. But for constructing a third-order maximum-entropy distribution it would be necessary to include information about third-order photon moments or the third-order auto-correlation function $g^{(3)}$. While introducing additional arbitrariness, $g^{(3)}$ is also more difficult to access experimentally so the simplest possible system to investigate beyond a single-mode second-order maximum-entropy distribution would rather be a second-order maximum-entropy distribution of a bimodal system. Also, superthermal photon bunching has previously been observed in such systems (Leymann *et al.,* 2013b; Hopfmann *et al.,* 2013; Redlich *et al.,* 2016; Fanaei *et al.,* 2016).

With photon numbers of the first and second mode n_1 and n_2 the second-order maximum-entropy distribution $\hat{p}_{n_1,n_2}$ reads

$$\begin{aligned} \hat{p}_{n_1,n_2} &= \frac{1}{Z(\lambda)} \exp\Big(-\lambda_{1,0} n_1 - \lambda_{0,1} n_2 - \lambda_{2,0} n_1^2 - \lambda_{0,2} n_2^2 - \lambda_{1,1} n_1 n_2\Big), \\ Z(\lambda) &= \sum_{n_1,n_2} \exp\Big(-\lambda_{1,0} n_1 - \lambda_{0,1} n_2 - \lambda_{2,0} n_1^2 - \lambda_{0,2} n_2^2 - \lambda_{1,1} n_1 n_2\Big), \end{aligned} \tag{3.18}$$

where the partition sum $Z(\lambda)$ ensures that the probability distribution is normalized. As input information, the mean photon numbers $\langle n_1 \rangle$ and $\langle n_2 \rangle$ as well as the second-order auto-correlation functions $g_1^{(2)}$ and $g_2^{(2)}$ of the individual modes ar required. Furthermore, the cross correlation between the two modes

$$g_{12}^{(2)} = \frac{\langle n_1 n_2 \rangle}{\langle n_1 \rangle \langle n_2 \rangle} \tag{3.19}$$

that also contains the mixed photon moment $\langle n_1 n_2 \rangle$ has to be included. Without the cross correlation (i.e. $\lambda_{1,1} = 0$) the maximum-entropy distribution (3.18) factorizes into a product of two single-mode maximum-entropy distributions, which were already discussed in Sec. 3.2.1.

Again, the highest-order Lagrange multipliers are investigated to conclude whether or not an unbiased photon distribution $\hat{p}_{n_1,n_2}$ exists on the global range $\mathbb{N}_0$. As an illustration, Fig. 3.10 (a) shows $\lambda_{2,0}$, $\lambda_{0,2}$, and $\lambda_{1,1}$ for a generic maximum-entropy distribution with fixed values $\langle n_1 \rangle = 7$, $\langle n_2 \rangle = 17$, $g_1^{(2)} = 2.5$, $g_2^{(2)} = 1.3$ that exhibits superthermal $g^{(2)}$

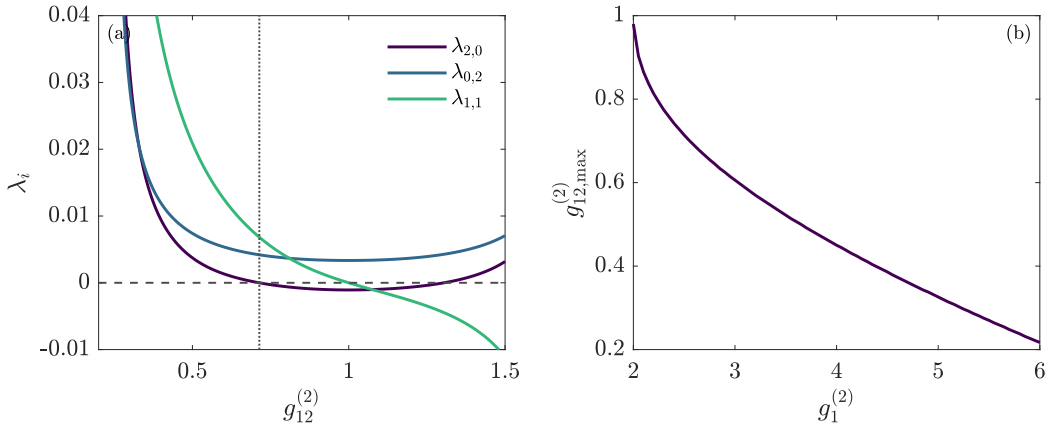


Fig. 3.10.: (a) Second-order Lagrange multipliers of a bimodal second-order maximum-entropy distribution where all constraints are fixed to $\langle n_1 \rangle = 7$, $\langle n_2 \rangle = 17$, $g_1^{(2)} = 2.5$, $g_2^{(2)} = 1.3$ except the cross-correlation $g_{12}^{(2)}$. The point where at least one of the three Lagrange multipliers gets negative (dashed and dotted lines) marks the highest possible value $g_{12,\max}^{(2)}$ to achieve a superthermal value $g_1^{(2)} = 2.5$ in the first mode. (b) Highest possible value $g_{12,\max}^{(2)}$ for different choices of $g_1^{(2)}$ with otherwise fixed constraints. Parameters: $n_{1,\max} = 80$, $n_{2,\max} = 80$.

values in the first mode. The cross correlation is varied and the resulting Lagrange multipliers are shown in the figure. For the maximum-entropy distribution to be normalizable, *all* second-order Lagrange multipliers $\lambda_{2,0}$, $\lambda_{0,2}$ and $\lambda_{1,1}$ must be non-negative. As indicated by the dashed and dotted lines, this condition is only met for values of the cross correlation of roughly $g_{12}^{(2)} \lessapprox 0.7$. Hence, for this generic choice of input information, to produce a superthermal $g^{(2)}$ value in one mode the two modes have to be sufficiently anti-correlated. To generalize the findings, $\langle n_1 \rangle$, $\langle n_2 \rangle$, and $g_2^{(2)}$ are chosen constant while the requested value of $g_1^{(2)}$ is altered within the superthermal regime. Then, the highest possible value of the cross correlation $g_{12,\max}^{(2)}$ is determined for each value of $g_1^{(2)}$. As depicted in Fig. 3.10 (b), the higher the superthermal photon bunching in one mode is, the stronger anti-correlated the two modes need to be. Or framed the other way around, stronger anti-correlation of the two modes allows for higher superthermal values in the first mode.

A typical resulting bimodal maximum-entropy distribution is shown in Fig. 3.11 (left panel) that exhibits the characteristic low probability along the $n_1 = n_2$ line. By summing over one mode the marginal distributions $\hat{p}_{n_1}$ and $\hat{p}_{n_2}$ can be obtained (Fig. 3.11, right panel). Both modes show a double-peaked structure which can be interpreted as "a mixture of a thermal-like and a coherent state"[5]. In fact, it is possible to generate superthermal values of $g^{(2)}$ by a linear combination of a thermal distribution with a low mean photon number and a normal distribution with a higher mean photon number (Lettau *et al.,* 2018).

[5]Leymann *et al.,* 2013b, p. 7.

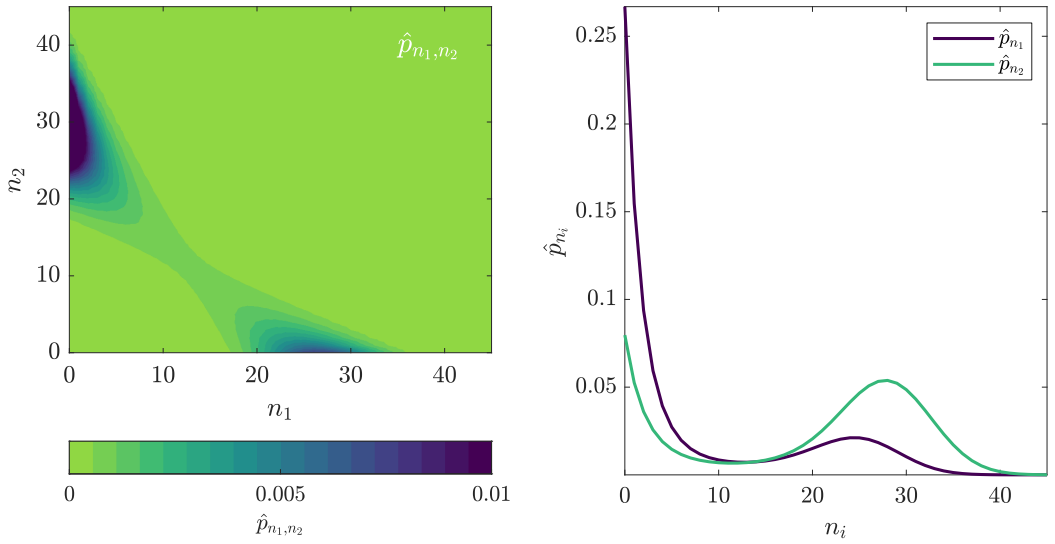


Fig. 3.11.: Full bimodal maximum-entropy distribution $\hat{p}_{n_1,n_2}$ (left panel) and marginal distributions, i.e., photon statistics of the individual modes $\hat{p}_{n_1}$ and $\hat{p}_{n_2}$ (right panel) for a highly anti-correlated system with superthermal $g^{(2)}$ values in the first mode. Parameters: $\langle n_1 \rangle = 8$, $\langle n_2 \rangle = 20$, $g_1^{(2)} = 2.5$, $g_2^{(2)} = 1.3$, $g_{12}^{(2)} = 0.3$, $n_{1,\max} = 45$, $n_{2,\max} = 45$.

In the sense of the maximum-entropy method the simplest system that is able to produce superthermal photon bunching when no information beyond the second-order correlation functions is included, is a bimodal system where superthermal photon bunching occurs in one of the modes as long as both modes are sufficiently anti-correlated. The higher the anti-correlation is, the higher the superthermal photon bunching in one mode can get. In the case of only one photon mode, the highest possible value of the auto-correlation function is $g^{(2)} = 2$, which can be shown both, analytically and numerically.

3.3 Chapter Conclusion

Basic features of the maximum-entropy method were investigated when the input information was given in form of moment constraints. First, the convergence properties of the maximum-entropy distribution were studied for a basic birth-death benchmark model, showing that the maximum-entropy distribution converges towards the original distribution for higher orders, that is for including higher order moments of the original distribution. For the investigated system only the even order (and first order) maximum-entropy distributions were well defined on the global range $\mathbb{N}_0$, whereas in the odd order case, the implementation space had to be chosen more carefully. Nevertheless, it was possible to give a reasonable maximum-entropy method approximation also in the odd-order case. With the resulting Lagrange multipliers a new lasing-threshold definition was given by a sign change of the

first Lagrange multiplier λ_1. It was applicable even in the case of high β-factors close to unity and perfectly matched conventional threshold definitions by kinks or jumps in the intensity or auto-correlation function values. The novel definition was connected to the photon statistics and marks the pump power at which the photon distribution exhibits a local maximum. By a comparison of the entropy of the original distribution with the respective entropies of a purely thermal, Poissonian, and Gaussian distribution it was revealed that, in contrast to expectation, the photon distribution above threshold was best approximated by a Gaussian distribution rather than a Poissonian. This was true although the second and higher order auto-correlation functions suggested the distribution to be Poissonian.

Finally, the existence properties of the maximum-entropy distributions were used to derive the simplest form of a photon distribution that is able to produce superthermal photon-bunching. It was shown that there cannot exist an unbiased maximum-entropy distribution with superthermal values of $g^{(2)} > 2$ in the case of a single-mode distribution. In the bimodal case, on the other hand, superthermal photon bunching can occur in one of the modes as long as the two modes are sufficiently anti-correlated. Additionally, a possibility to link given values of $\langle n \rangle$ and $g^{(2)}$ to the threshold definition by sign change of the first Lagrange multiplier was given that can be used without explicit application of the maximum-entropy method.

4 Combining Equation-of-Motion Techniques and the Maximum-Entropy Method

Since the convergence of the maximum-entropy distribution towards the original probability distribution was shown in the previous chapter, the maximum-entropy method can now be used as accompanying technique to cases where only the moments of a distribution are known but the knowledge of the full probability distribution is desirable. Apart from that, further insight about the investigated systems may be gained because the maximum-entropy method also yields the distribution's entropy as well as the Lagrange multipliers. In this chapter, the second-order maximum-entropy method is applied to two specific models where equations of motion for the first two photon moments are available. In Sec. 4.1 a semiconductor model for quantum-dot based microcavity lasers (Gies *et al.,* 2007) is investigated concerning the existence of a photon distribution, its entropy as well as the question of a threshold characterization via Lagrange multipliers. The second section (Sec. 4.2) of this chapter reviews the paper (Lohof *et al.,* 2018) where a strong separation of the laser threshold by a jump in the input-output curve and by a change of the auto-correlation function value was found for nanolasers with extended gain media. Here, the maximum-entropy method's perspective by the sign change of the first Lagrange multiplier is added. Lastly, the photon distribution's entropy is investigated.

4.1 Semiconductor Model

As at that time it became more popular to combine optical microcavities with quantum-dot emitters as active gain material, the paper (Gies *et al.,* 2007) can be seen as foundational. This work gave a concise theory that accounted for intrinsic semiconductor effects and predicted statistical features of the emitted light that go beyond the intensity. Since there, equations of motion for the mean photon number and the auto-correlation function (and other related quantities) are derived by considering the full semiconductor Hamiltonian,

semiconductor effects such as "a modified source term of spontaneous emission, Pauli-blocking effects of the occupied states, as well as many-body Coulomb effects"[1] are naturally included. Apart from being a helpful tool to interpret experimental data based on quantum-dot lasers more accurately, it also laid groundwork for numerous follow up investigations (Ulrich *et al.,* 2007; Ritter *et al.,* 2010; Florian *et al.,* 2012; Florian *et al.,* 2013; Leymann *et al.,* 2013b; Chow *et al.,* 2014; Leymann *et al.,* 2015; Fanaei *et al.,* 2016; Jahnke *et al.,* 2016; Kreinberg *et al.,* 2017; Moody *et al.,* 2018). In this section, after a brief introduction to the model, the output quantities (mean photon number $\langle n \rangle$ and auto-correlation function $g^{(2)}$) are used as input for the maximum-entropy method. Basic questions about the existence of a maximum-entropy distribution are answered and the resulting photon statistics are investigated concerning the question of a laser threshold and entropy.

Semiconductor model To derive a closed set of equations of motion for the quantities of interest the authors use the von Neumann equation in combination with the full semiconductor Hamiltonian [cf. Eq. (5) in the original paper (Gies *et al.,* 2007)]. Apart from the dipole interaction, the free parts of the charge carriers and the light field, also the two-particle Coulomb interaction Hamiltonian is included, accounting for above mentioned semiconductor-specific effects. Unlike more recent approaches where Lindblad terms are used (e.g. Florian *et al.,* 2013; Leymann *et al.,* 2015; Fanaei *et al.,* 2016), coupling to the environment, e.g., cavity losses, dephasing, are included phenomenologically. To truncate the arising infinite hierarchy of equations of motion that unfolds due to the many-particle contributions of the Hamiltonian, the cluster-expansion method is used to systematically factorize higher-order expectation values into lower-order expectation values plus correlation functions. Depending on the level of complexity, particular correlation functions are neglected to obtain a closed set of differential equations. In the case of what the authors call *extended laser equations*, a set of ten closed equations of motion remains, two of which give access to the mean photon number $\langle n \rangle$ and the auto-correlation function $g^{(2)}$. With a suitable choice of parameters the semiconductor model not only reproduces the laser rate-equations (3.3) but is also in accordance with the results of the birth-death model from Sec. 3.1. To generate input moments for the maximum-entropy method the extended laser equations, i.e., the ten closed equations of motion [Eqs. (8), (11), (16-17), (20-23) in the original paper (Gies *et al.,* 2007) with the parameters that lead to Fig. 3] are time-integrated until a steady state is reached. The values of the mean photon number $\langle n \rangle$ and auto-correlation function $g^{(2)}$ are then used as input to calculate a second-order maximum-entropy distribution.

Existence of a probability distribution Although the extended laser equations have proven to be physically reasonable, it is in general not clear beforehand, if the equation-of-

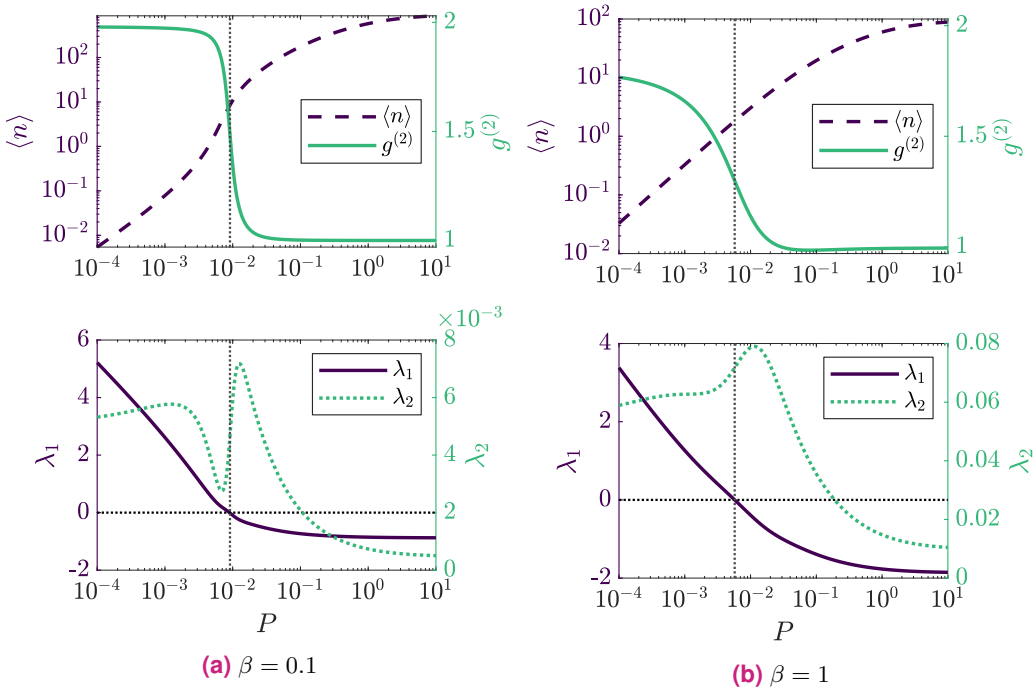
[1] Gies *et al.,* 2007, p. 1.

motion approach produces moment values for which a probability distribution can be found at all. Rather, it is known that this approach might as well produce physically unreasonable results, e.g., negative values for the auto-correlation function (Leymann *et al.,* 2014) and that the results should carefully be tested for plausibility. Since the underlying distribution for $\langle n \rangle$ and $g^{(2)}$ is a discrete probability distribution on $\mathbb{N}_0 = \{0, 1, \ldots, \infty\}$, the Stieltjes moment problem applies (cf. Sec. 2.2.3) and the existence of such a distribution can be checked with the conditions (2.20), which here indeed indicate that a distribution can be found for all investigated parameters. Nevertheless, these conditions do not answer the question of uniqueness. So the better inquiry is to check whether a maximum-entropy distribution can be found, that is compatible to the given values $\langle n \rangle$ and $g^{(2)}$. If so, it is automatically guaranteed that the distribution is unique and has least bias (cf. Sec. 2.2). Because only the first two moments are available, the existence condition (2.27) also applies here. Hence if

$$\left\langle n^2 \right\rangle \leq 2 \langle n \rangle^2 + \langle n \rangle \quad \text{or equivalently} \quad g^{(2)} \leq 2 \tag{4.1}$$

holds true, a unique second-order maximum-entropy distribution can be found that reproduces the given moment values. The authors give analytical results for $g^{(2)}$ [Eqs. (31) and (34) in the original paper (Gies *et al.,* 2007)] in the limiting cases of strong pumping, where $g^{(2)} = 1$ is obtained, and weak pumping, where in the bad cavity limit $g^{(2)} = 2 - \frac{2}{N}$ can be found with N being the number of quantum dots. Hence, in the limit of many quantum dots $g^{(2)} = 2$ is found, while $g^{(2)} = 0$ holds for a single quantum dot. This is a strong indication that the condition (4.1) is fulfilled. Looking at the numerically determined $g^{(2)}$ values in the original paper that all lie in the range between zero and two, it can be concluded that the equation-of-motion plus cluster-expansion method approach produces moment values for which a unique maximum-entropy distribution can be found. What might sound as technical minutiae is indeed an important criterion that can be applied to all equation-of-motion plus cluster-expansion method approaches where an important result would be to show existence and uniqueness of the underlying probability distribution and thus eliminating arbitrariness and refining the quality and physical reasonability of the method's results. Since for the semiconductor model a second-order maximum-entropy distribution can be found on $\mathbb{N}_0$, it is simultaneously ensured, that the second Lagrange multiplier λ_2 has positive sign.

Laser Threshold For two β values close to unity, the mean photon number $\langle n \rangle$ and auto-correlation function $g^{(2)}$ are shown with the Lagrange multipliers of the second-order maximum-entropy distribution in Fig. 4.1. Although available, the full photon statistics is not shown here, since it behaves exactly as expected (thermal below and Poisson-like above threshold, cf. Sec. 3.1.2). For the lower β-factor [Fig. 4.1 (a)] the threshold can still be defined by a jump of the $g^{(2)}$ value. In the $\beta = 1$ case [Fig. 4.1 (b)], on the other hand, the input-output curve has no jump at all and the auto-correlation function transitions



(a) $\beta = 0.1$ **(b)** $\beta = 1$

Fig. 4.1.: Multi figures for two different β-factors (a) $\beta = 0.1$ and (b) $\beta = 1$. In the respective upper subfigure, the mean photon number $\langle n \rangle$ (left y-axis) and the auto-correlation function $g^{(2)}$ (right y-axis) are shown as dashed and solid curves. The lower panel depicts the Lagrange multipliers λ_1 (left y-axis, solid curve) and λ_2 (right y-axis, dotted curve) as well as the threshold definition by the sign change of λ_1 (dotted horizontal and vertical lines).

rather smoothly to the value of $g^{(2)} = 1$. Nevertheless, with the sign change of the first Lagrange multiplier λ_1, the threshold pump power can be found unambiguously for both cases as indicated by the dashed lines in Fig. 4.1. In the $\beta = 0.1$ case, it perfectly matches the traditional definitions by a jump in the values of $\langle n \rangle$ and $g^{(2)}$, while for $\beta = 1$, the sign change happens in the center of the smooth transition area of the auto-correlation function.

Additionally, as pointed out in Sec. 3.2.1, with the aid of Fig. 3.9 the pump power at which the photon statistics exhibits a local maximum can be found without explicit application of the maximum-entropy method, but simply by the value of $g^{(2)}$. For the case of $\beta = 0.1$, it is justified to apply condition (3.17) and hence to use $g^{(2)}_{\mathrm{thr}} \approx 1.6$ as threshold criterion. This is also reflected in Fig. 4.1 (a), where the sign change of λ_1 is roughly at $g^{(2)}_{\mathrm{thr}} \approx 1.6$. This way, it is possible to link the pair of values $\left(\langle n \rangle, g^{(2)}\right)$ to the features of the photon distribution and find the threshold pump rate at which it first exhibits a local maximum without explicit application of the maximum-entropy method. For $\beta = 1$ one finds $\langle n \rangle \approx 1$, so $g^{(2)} = 1.1$ is the better choice here (cf. Fig. 3.9). In both cases, the criteria give a high quality approximation for the threshold pump power.

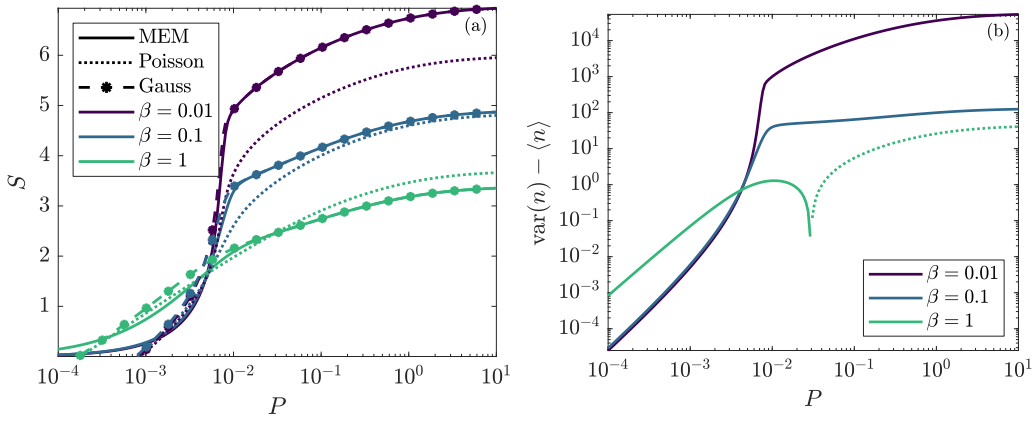


Fig. 4.2.: (a) Entropy curves for the extended laser equations. Different colors show the results for different β-factors. For each β-factor, the entropy of the second-order maximum-entropy distribution is shown as solid curve (MEM) whereas the entropies of a Poissonian and a Gaussian distribution are depicted as dotted and dashed (with a marker) curves, respectively. For higher pump rates the maximum-entropy distribution's and the Gaussian entropy curves lie on top of each other. (b) Difference between variance and mean photon number $\mathrm{var}(n) - \langle n \rangle$ for different β-factors in a double-logarithmic plot. The dotted curve indicates where instead $\langle n \rangle - \mathrm{var}(n)$ is shown to keep the plotted values positive.

Entropy As the maximum-entropy method also gives access to the entropy of the photon probability distribution, the light emission is characterized in terms of entropy (as in Sec. 3.1.5). To do so, the entropy of the maximum-entropy distribution is compared with the entropy values that a purely Poissonian or purely Gaussian distribution would have while reproducing the given values of $\langle n \rangle$ and $g^{(2)}$ (cf. Tab. 3.3). It is stressed that the maximum-entropy distribution is the unique probability distribution that is compatible with the given moment constraints without introducing additional arbitrariness, therefore, assuming any other photon probability distribution to be the origin of the known values of $\langle n \rangle$ and $g^{(2)}$ would be highly unreasonable. In Fig. 4.2 (a), the comparison is summarized for three different β-factors throughout the whole range of pump powers. Each color corresponds to a specific β-factor and the line styles indicate the different entropy approximations. In the weak pumping range, the curves of the purely Poissonian and purely Gaussian entropies deviate from the entropy of the maximum-entropy distribution which is as expected because there the photon statistics is thermal (not shown). Slightly above the transition region the entropy of the maximum-entropy distribution perfectly matches the entropy curve of a purely Gaussian distribution, while the entropy curves of the purely Poissonian distribution do not match at all. This holds true for all values of β. The same phenomenon was already observed for the birth-death model in Sec. 3.1.5 and is rather unexpected as one would assume Poissonian statistics above threshold. A closer look at the difference between variance and the mean photon number $\mathrm{var}(n) - \langle n \rangle$ [Fig. 4.2 (b)] reveals more details about the photon distribution itself. Since this quantity should approach zero for a Poissonian

distribution where mean value and variance coincide, a larger value indicates a much broader probability distribution than a purely Poissonian. This behavior can be seen for the two smaller β values, whereas for $\beta = 1$, the photon distribution is narrower than a Poissonian distribution (indicated by the dotted curve, where instead $\langle n \rangle - \mathrm{var}(n)$ is plotted). This can already be seen in the entropy curves in Fig. 4.2 (a), where the Gaussian entropy values are higher ($\beta = 0.01$ and $\beta = 0.1$) or lower ($\beta = 1$) than the Poissonian values. This again, as in the case of the birth-death model, underlines that a conclusion just from the knowledge of the $g^{(2)}$ value alone can be premature. Nevertheless, it should be pointed out that it is not surprising that the maximum-entropy method's entropy coincides with that of a Gaussian distribution since a second-order maximum-entropy distribution resembles a Gaussian distribution. However, this does not make the previous findings less true.

Concluding this section, it is emphasized that the maximum-entropy method can add valuable insight even for well established models by linking the features of the photon distribution to traditional ($\langle n \rangle$ and $g^{(2)}$) and novel (first Lagrange multiplier) threshold definitions, making it possible to mark a threshold pump power even in the case of $\beta = 1$ (and actually even without the need to perform the maximum-entropy method). Moreover, by checking the existence criteria, the output of all equation of motion plus cluster-expansion method approaches can easily be verified for plausibility since the existence of such a maximum-entropy distribution guarantees uniqueness and least possible arbitrariness.

4.2 Nanolasers with Extended Gain Media

Already in 2006, a bright future for nanolasers was predicted as they promised to be able to produce light extremely energy-efficiently while simultaneously having high integrability in optoelectronic devices (Noda, 2006). Back then, combining photonic nanostructures with quantum dots was predicted to be most promising since with this combination suppression of optical modes that lead to "undesired spontaneous emission [,] a single-cavity mode with a sufficiently high Q factor [...] and a small modal volume"[2] as well as suppression of non-radiative processes because of the three-dimensional carrier confinement in quantum dots was possible. With the rise of two-dimensional materials however, "[g]ain materials from atomically thin semiconductors have moved into the focus of investigations"[3]. Although two-dimensional materials were studied for decades (Brodie, 1859; Kosterlitz & Thouless, 1973; Thouless *et al.,* 1982; Novoselov *et al.,* 2004; Novoselov *et al.,* 2005; Bhimanapati *et al.,* 2015; Ren *et al.,* 2016), especially transition metal dichalcogenides are a strongly investigated field with growing applications in the whole field of nano- and optoelectronics

[2] Noda, 2006, p. 260.
[3] Lohof *et al.,* 2019, p. 210.

as well as spin- and valleytronics (Wang *et al.,* 2012; Xiao *et al.,* 2012; Bhimanapati *et al.,* 2015; Schaibley *et al.,* 2016; Ma & Oulton, 2019). Exceptionally strong Coulomb interaction between charge carriers as well as the weak dielectrical screening make the electrical and optical properties of these materials highly dependent on the environment as well as the presence of excited carriers, leading to a potpourri of different effects (Mak *et al.,* 2010; Butler *et al.,* 2013; Steinhoff *et al.,* 2018a; Steinhoff *et al.,* 2018b; Erben *et al.,* 2018; Lohof *et al.,* 2018; Lohof *et al.,* 2019; Steinhoff *et al.,* 2020). Many of these effects are in fact used to tailor desired properties by, for instance, combining and stacking layers of different materials on top of each other (Geim & Grigorieva, 2013; Xu *et al.,* 2014; Bhimanapati *et al.,* 2015; Chen *et al.,* 2019; Alexeev *et al.,* 2019; Tran *et al.,* 2019). Even quantum-dot-like states in molybdenum disulfide nanostructures due to surface wrinkling and strain have been observed recently (Carmesin *et al.,* 2019). Therefore, the combination of high-quality cavities with two-dimensional materials as gain media emerges as highly interesting field, making the prospects and limitations of this kind of systems a current topic of investigation (Lohof *et al.,* 2019). Since the carrier density in extended gain material strongly exceeds the carrier densities of quantum dots, new operational regimes are expected. Indeed, in the paper (Lohof *et al.,* 2018) that is considered in this section, a delayed transition to coherent light emission was found for nanolasers with extended gain material. As the maximum-entropy method turned out to be a valuable tool to characterize the transition to lasing, the threshold definition by sign change of the first Lagrange multiplier is added as a further perspective. Moreover, the photon statistics is compared in terms of entropy to investigate whether or not the presented model results in light emission with Poissonian statistics.

Model assumptions In (Lohof *et al.,* 2018) the authors derive rate equations for the mean photon number as well as an expression that links a unique value of the auto-correlation function to the mean photon number by considering "N independent two-level emmiters in resonant Jaynes-Cummings interaction with a single mode of an optical cavity"[4]. Essentially, this is the same approach as in (Gartner & Halati, 2016) that also delivers the standard laser rate-equations as well as the birth-death master-equation from Sec. 3.1. Using the von Neumann-Lindblad equation and a treatment of arising higher-order expectation values on the mean-field level, rate equations for the mean intracavity photon number $\langle n \rangle$ and excited-carrier population are derived [Eqs. (2a) and (2b) in the original paper]. In a steady-state situation the rate equations can be evaluated to define a pump value P_{thr} of what the authors call *intensity threshold*. It marks "the [emitter] population for which the

[4] Lohof *et al.,* 2018, p. 6.

gain (stimulated emission minus absorption) exactly compensates the cavity losses"[5] (cf. App. A in the original paper) and obeys

$$P_{\mathrm{thr}} - \gamma - (P_{\mathrm{thr}} - \gamma)\frac{\kappa}{RN} = 0, \tag{4.2}$$

where γ is the rate of radiative losses, κ the cavity loss rate and R is the spontaneous emission rate in the cavity mode. It further depends on the light-matter coupling-strength g and the dephasing processes $\Gamma = P + \gamma + \kappa$ where P is the pump rate:

$$R = \frac{4|g|^2}{\Gamma}. \tag{4.3}$$

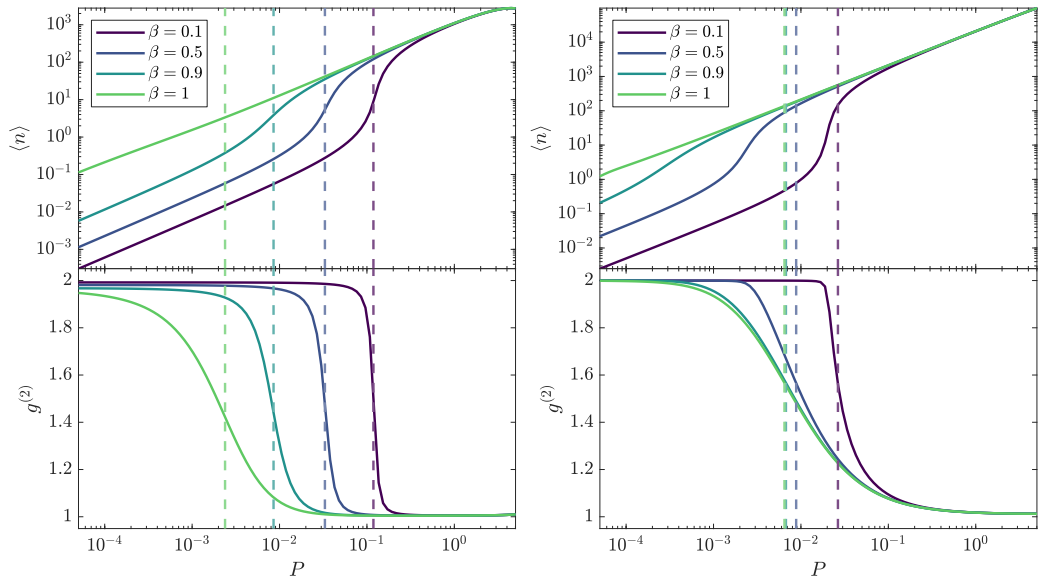
To obtain the two-photon correlation function $g^{(2)}$ the approach is extended to higher-order correlations that are truncated at the quadruplet-level of the cluster-expansion method (including photon-photon correlations). Combined with the stationary values of the rate equations, the authors arrive at a closed analytical expression for $g^{(2)}$ that depends on the mean photon number and the system's parameters only

$$g^{(2)} = 2 - \frac{R(2\langle n\rangle + 2) + \frac{3\kappa}{N}}{R(2\langle n\rangle + 1) + \kappa + \frac{\Gamma}{2\kappa}\frac{NR+\kappa}{2\langle n\rangle+1}}. \tag{4.4}$$

By using the steady-state results that stem from the rate equations and Eq. (4.4) the authors compare the intensity threshold with the *coherence threshold* P_{coh} that is chosen as the pump rate where $g^{(2)} = 1.1$ is achieved. On the one hand, a typical laser with high-quality (high-Q) cavity, i.e., low cavity loss rate κ, and quantum dots as gain material is studied [parameters from (Ulrich *et al.,* 2007)] where both thresholds are interlinked. On the other hand, extended gain media instead of individual emitters (e.g. thin layers of transition-metal dichalcogenides) are considered in a low-quality cavity with higher loss rate κ. There, the authors find that "the emission becomes coherent at much higher excitations, after the intensity threshold has been crossed"[6]. Therefore, they "refine the definition of the laser threshold [...] to take a combined look at the intensity threshold and the transition of the emission from thermal to coherent light reflected in $g^2(0)$"[7].

Threshold definition comparison and photon statistics For a comparison with the maximum-entropy method the results that lead to Figs. 1 and 2 in the original paper are reproduced by a time-integration of the rate equations and subsequent application of Eq. (4.4). The resulting input-output curves as well as the auto-correlation function $g^{(2)}$ are shown in Fig. 4.3. As can be found by application of condition (2.27), a unique probability

[5] Lohof *et al.,* 2018, p. 7.
[6] Lohof *et al.,* 2018, p. 2.
[7] Lohof *et al.,* 2018, p. 1.



(a) Few emitters in a high-Q cavity
$N = 100$, $\kappa = 0.04\,\mathrm{ps}^{-1}$

(b) Many emitters in a low-Q cavity
$N = 3 \cdot 10^4$, $\kappa = 0.7\,\mathrm{ps}^{-1}$

Fig. 4.3.: Mean photon number $\langle n \rangle$ (top) and auto-correlation function $g^{(2)}$ (bottom) for the few (a) and many (b) emitter cases for different values of β. The dashed lines show the position of the threshold by sign change of the first Lagrange multiplier λ_1. Parameters as in the original paper (Lohof *et al.,* 2018).

distribution that has maximum entropy can be found for all investigated parameters. In contrast to the original paper, here, the dashed lines indicate the threshold pump power by sign change of the first Lagrange multiplier of the maximum-entropy-method approach. In the case of few emitters [Fig. 4.3 (a)] this definition results in threshold pump rates that are located centrally in the jump region of the mean photon number (except for $\beta = 1$) and the auto-correlation function. Figure 4.4 (a) takes a combined look at all three different threshold pump rates for a range of different values of β. The maximum-entropy method's definition corresponds to the values obtained by evaluating Eq. (4.2) (intensity threshold) except that the maximum-entropy method's definition is also applicable to the case of $\beta \approx 1$. The coherence threshold definition gives slightly higher pump rate values throughout all examined β values. This is due to the cautious threshold choice of $g^{(2)} = 1.1$. Utilizing Fig. 3.9, a choice of $g^{(2)} \approx 1.5$ would be more reasonable since the mean photon number $\langle n \rangle$ is well above one but not large enough to apply the $g^{(2)} \approx \frac{\pi}{2}$ criterion. Interestingly, a look at the photon statistics at the respective threshold pump rates reveals slight inconsistencies for the intensity threshold definitions [Fig. 4.5 (a)]. While for the coherence threshold the photon statistics have Poisson-like shape for all β values, at intensity threshold the photon statistics have thermal shape for $\beta = 0.9$ while exhibiting a peak in the other two cases. From the probability distribution's point of view, the maximum-entropy method's threshold

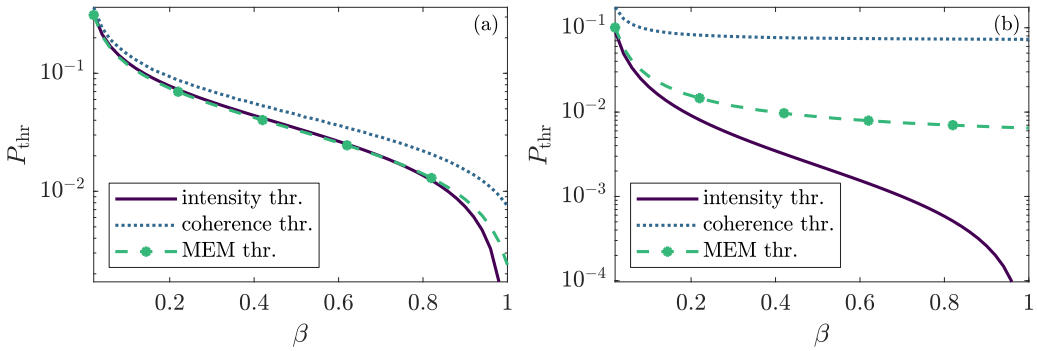


Fig. 4.4.: Threshold definition comparison for the case of few emitters (a) and many emitters (b). The threshold definitions by intensity (solid curve) and the auto-correlation function (dotted curve) from the original paper are compared to the threshold definition by sign change of the first Lagrange multiplier λ_1 (dashed curve with markers), cf. Fig. 4.3.

gives a more consistent picture, always marking the pump rate at which the photon statistics first exhibits a local maximum.

In the case of extended gain media a less consistent picture of the different threshold pump rates occurs [Fig. 4.3 (b)]. While the intensity threshold results in different threshold pump-rates for different β-factors, the coherence threshold condition is fulfilled almost independently of β. For high β-factors also the maximum-entropy method's thresholds are close to each other. Since here $\langle n \rangle \gg 1$ holds [cf. condition (3.17)], the sign change of the first Lagrange multiplier coincides with a value of $g^{(2)} = \frac{\pi}{2} \approx 1.6$ [cf. crossing of the dashed lines with the $g^{(2)}$ curves in Fig. 4.3 (b)]. A comparison of the threshold definitions for a range of β-factors [Fig. 4.4 (b)] makes the discrepancy of the intensity and coherence threshold apparent. While the first one gives highly varying threshold pump rates for different β-factors, the latter is almost independent of the fraction of spontaneous emission directed into the laser mode. This is why the authors conclude that, in the case of extended gain media, "an increase of the β factor has no benefit in terms of reducing the threshold power"[8] and beyond that, that "the pump rate at which coherent emission is achieved is rather determined by the cavity losses and can only be improved by increasing Q."[9] Compared to the coherence threshold, the maximum-entropy method's definition results in lower threshold pump rates, while still largely resembling its curve form. Although less pronounced, it supports the authors' statement of a nearly β-independent threshold pump rate. Looking at the photon distribution at the different threshold pump rates [Fig. 4.5 (b)] it becomes obvious that the intensity threshold is far away from the emitted light being coherent, since it clearly results in purely thermal photon statistics for all three β values. This underpins the authors' findings of a "[d]elayed [t]ransition to [c]oherent [e]mission

[8] Lohof *et al.*, 2018, p. 2.
[9] Lohof *et al.*, 2018, p. 6.

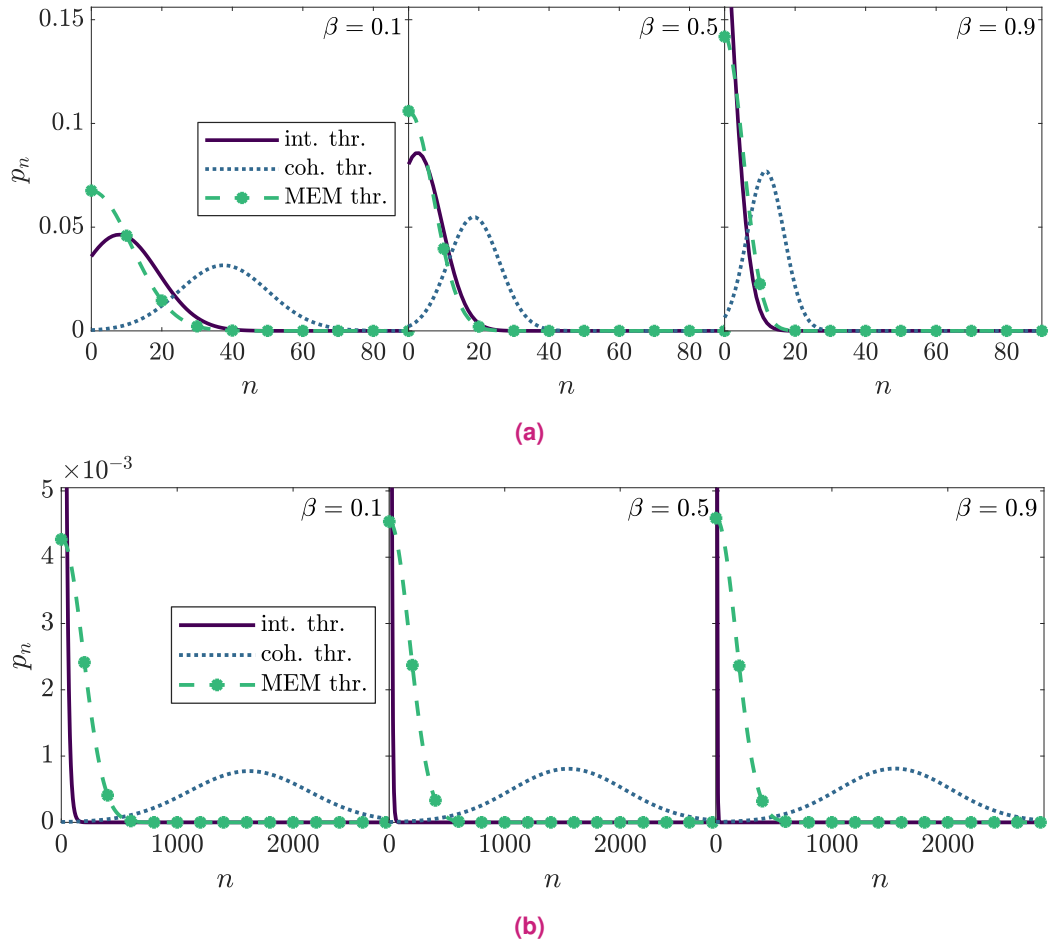


Fig. 4.5.: Photon statistics obtained by the maximum-entropy method at the different threshold pump powers for different β-factors in the few emitter (a) and many emitter (b) cases. Solid curves depict the photon statistics p_n at the intensity threshold, dotted curves at the coherence threshold and dashed (with markers) curves at the maximum-entropy method's threshold. For the many-emitter case, the photon statistics at intensity threshold are barely visible because they lie close to the y-axis.

in [n]anolasers with [e]xtended [g]ain [m]edia"[10]. The coherence threshold as well as the threshold by sign change of the first Lagrange multiplier both result in more reasonable probability distributions. In the first case, a Poisson-like distribution is obtained, while the maximum-entropy method's definition results in a probability distribution with a local maximum at the zero-photon state. Again, it can be argued that the sign change of the first Lagrange multiplier gives the more consistent definition, nonetheless reproducing the findings of the authors.

Finally, it is pointed out that a combination of the general maximum-entropy-method results that are reflected in Fig. 3.9 with Eq. (4.4) opens up the possibility to link results of simpler

[10] Lohof *et al.,* 2018, p. 1.

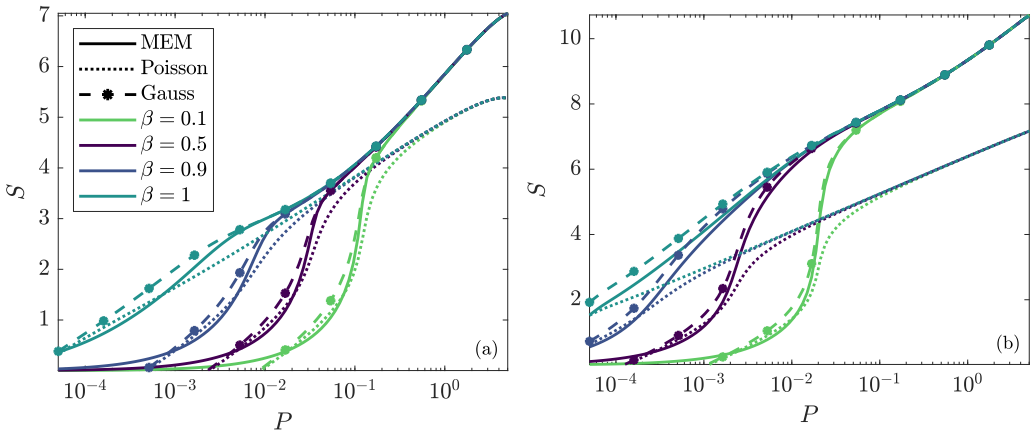


Fig. 4.6.: Entropy curves for the few (a) and many (b) emitter cases. Different colors show the results for different β-factors. For each β-factor, the entropy of the second-order maximum-entropy distribution is shown as solid curve whereas the entropies of a Poissonian and a Gaussian distribution are depicted as dotted and dashed (with a marker) curves, respectively. For higher pump rates the maximum-entropy distribution's and the Gaussian entropy curves lie on top of each other.

rate-equations, that only need to give access to the mean photon number, to statistical aspects of the photon distribution. This can be achieved by a subsequent application of Eq. (4.4) to determine $g^{(2)}$ out of the mean photon number and comparison of the values with the λ_1-landscape in Fig. 3.9. In the context of the growing interest in two-dimension materials as gain medium in nanolasers, this might be a valuable yet easy to use tool to investigate the question of threshold behavior and new operational regimes.

Entropy As for the birth-death model (Sec. 3.1.5) and the semiconductor model (Sec. 4.1), the photon distribution is characterized in terms of entropy to challenge the notion of Poissonian statistics above threshold. To do so, not unlike the other cases, the entropy of the maximum-entropy distribution is compared to characteristic entropy values of Gaussian and Poissonian statistics with the respective values of $\langle n \rangle$ and $g^{(2)}$. In Fig. 4.6 both cases of few (a) and many emitters (b) are shown. The solid curves represent the entropy of the maximum-entropy distribution, while dotted and dashed (with markers) curves give the respective entropy values of a Poissonian and Gaussian probability distribution. Different colors represent different values of the β-factor. In both cases and for all β values, the Gaussian entropy curve exactly matches the entropy of the maximum-entropy distribution above threshold. The Poissonian entropy gives lower values in all cases, indicating that the photon probability distribution has higher variance than a purely Poissonian probability distribution would have. This behavior is especially pronounced in the case of extended gain media, indicating much stronger fluctuations in the photon number than expected for a Poissonian distribution. It remains an open question though, if this reveals a deficiency

of the model or has real physical meaning. It is stressed again, that it is not surprising that the entropy values of the maximum-entropy distributions and purely Gaussian statistics coincide since the second-order maximum-entropy distribution resembles a Gaussian distribution. Nevertheless, the findings underline that premature conclusions about the nature of the underlying photon distributions solely based on the mean photon number and auto-correlation function should be avoided, since, as for this example, assuming that the photon statistics has Poissonian shape above threshold is only possible by a violation of the maximum-entropy principle.

4.3 Chapter Conclusion

In the two previous sections it was demonstrated how the maximum-entropy-method approach can be combined with equation-of-motion approaches by using the output moment-values of the equations of motion as input for the maximum-entropy method. As a first aspect, the maximum-entropy method gave a possibility to check the output of such approaches for plausibility. In the case of single-mode probability distributions and availability of only the first two moments, the criterion $g^{(2)} \leq 2$ ensured existence and uniqueness of the underlying probability distribution. The sign change of the first Lagrange multiplier was contrasted with common laser threshold definitions. Apart from perfectly matching usual definitions, it gave deeper insight in the case of extended gain material where the intensity and coherence threshold values strongly differed. A new possibility was pointed out to connect the sign-change criterion of the maximum-entropy method to pairs of values $\left(\langle n \rangle , g^{(2)}\right)$ without the explicit need to perform the maximum-entropy method. Moreover, with the aid of the results of (Lohof *et al.,* 2018), a possibility to link the novel threshold definition to rate-equation approaches that only give access to the mean photon number was suggested. Finally, an open question arose why the photon distribution above threshold, in contrast to common sense, deviated from a purely Poissonian distributions. While for the birth-death model, the deviation was rather small, for both equation-of-motion approaches of this chapter the discrepancies were more pronounced. Moreover, two different cases, a much broader and a narrower distribution were found, leading to the question, whether this finding is grounded in physical phenomena or rather reflects assumptions and approximations in the respective equation-of-motion approaches.

5 Maximum-Entropy Method for Quantum Systems

As outlined in Sec. 2.3, the maximum-entropy method is also applicable to the quantum case if one replaces the input of moments of a probability function with the input of expectation values of quantum mechanical operators. Then, the maximum-entropy method yields the least-biased density matrix that accounts for all given constraints but is otherwise maximally non-committal. Moreover, if the interest lies only in the system's steady state the need for explicit numerical values as input can be circumvented by inclusion of the equations of motion themselves as input information (cf. Sec. 2.3.3). In this sense, a four-level quantum dot laser system is considered in Sec. 5.1 where the maximum-entropy-method results are compared to the still available (due to the small system size) numerically exact steady-state solution of the von Neumann-Lindblad equation. Also, general strategies for a suitable choice of observation levels are presented and the results are compared in terms of entropy, mean photon number, auto-correlation functions as well as the full photon statistics. Furthermore, by considering two different excitation mechanisms, the application as trial-and-error method is exemplified. Finally, Sec. 5.2 gives an outlook on possible future research directions. There, two different extensions of the quantum-dot-laser model are suggested in slightly more detail.

Parts of this section have been published in (Melcher *et al.,* 2019). The general theory was developed in discussions by all contributing authors of the publication. All maximum-entropy method and von Neumann Lindblad calculations were done by the thesis's author. All authors discussed the results.

5.1 Benchmark Model

Open quantum systems gained considerable attention due to their ability to make the very heart of the quantum world explorable, where the interplay between driving mechanisms and dissipation processes strongly influences coherence properties of the probed system. There, numerous effects such as Bose-Einstein condensation of few photons (Walker *et al.,* 2018), Schrödinger cat states in photon resonators (Minganti *et al.,* 2016), superradiance of quantum dots (Jahnke *et al.,* 2016), cavity optomechanics (Aspelmeyer *et al.,* 2014) and the spin-boson model in superconducting quantum circuits (Magazzù *et al.,* 2018) have been examined. As sketched in Sec. 2.3.3, one common approach to describe those driven-dissipative quantum many-particle systems is by a combination of equation-of-motion techniques with approximation schemes such as the cluster-expansion method. In contrast to that, here, the maximum-entropy method is used as stand-alone approach. Consequently, the knowledge about the stationarity of several distinct observables is used to deduce the least-biased steady-state density matrix and thus make all relevant expectation values and correlation functions as well as the full statistics directly accessible. Importantly, a truncation of the many-particle hierarchy in terms of factorization schemes is completely avoided. To test the maximum-entropy method, a four-level single quantum-dot microcavity laser coupled to a Markovian environment is regarded, where the full density matrix is still available for comparison by numerically solving the von Neumann-Lindblad equation. In Sec. 5.1.1 the benchmark system is presented in detail. Then, in Sec. 5.1.2 the input information for the maximum-entropy method, i.e., the observation level, is concretized. Furthermore, general suggestions for the choice of a sufficient observation level are given. In Sec. 5.1.3 the numerical results of both approaches are compared in terms of entropy, mean photon number, auto-correlation functions as well as photon statistics. This is done for two different cases of pumping mechanisms where in one case the choice of input information is altered to exemplify how the maximum-entropy method can be applied as trial-and-error approach to find out which physical processes carry necessary information.

5.1.1 Four-Level Quantum-Dot Microcavity Laser

As a benchmark system, a four-level single quantum-dot (or single atom) microcavity laser (cf. Ralph & Savage, 1991; Mu & Savage, 1992; Grothe, 2014) is chosen, where two different pumping mechanisms are considered (cf. Fig. 5.1). Furthermore, the single electron approximation is used, hence the spin degree of freedom is neglected and only one electron is permitted in the system. The single charge carrier from the ground state $|1\rangle$ is either pumped incoherently with a pump rate P or coherently excited via an external field E into the highest level $|4\rangle$. E is a pumping rate (in units of ps^{-1}) proportional to the

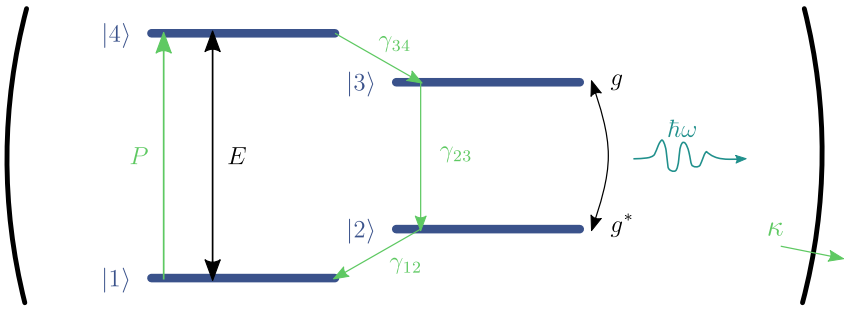


Fig. 5.1.: Four-level quantum dot in a microcavity. The electronic energy levels are denoted by $|1\rangle$, $|2\rangle$, $|3\rangle$, and $|4\rangle$. The inner levels are resonantly coupled to the cavity mode with frequency ω. The light-matter coupling strength is given by g. Interaction with the environment, namely electronic relaxation with rates γ_{ij}, cavity losses with rate κ and incoherent pumping with rate P are shown in lighter color. E represents an external field in the coherent excitation case. This figure has been published in (Melcher *et al.,* 2019).

field strength of the driving field. The charge carrier then relaxes non-radiatively via the inner levels $|3\rangle$ and $|2\rangle$ with decay rates γ_{34}, γ_{23}, and γ_{12}, back into the lowest level $|1\rangle$. The inner levels $|2\rangle$ and $|3\rangle$ are resonantly coupled to the cavity mode with frequency ω. Cavity losses occur at a rate κ and the light-matter coupling strength is denoted with g. In the interaction picture the interacting parts of the Hamiltonian read

$$\begin{aligned} H_{\mathrm{JC}} &= g^* a_2^\dagger a_3 b^\dagger + g a_3^\dagger a_2 b, \\ H_E &= \mathrm{i}\,\hbar E \left(a_1^\dagger a_4 - a_4^\dagger a_1 \right), \end{aligned} \tag{5.1}$$

where the Jaynes-Cummings Hamiltonian H_{JC} governs the light-matter interaction in electric dipole and rotating-wave approximation, while H_E represents the coherent pumping process via an external field, which is assumed to be in resonance with the transition of the outer electronic levels. In the equations above, $a_i^\dagger$ and a_i are fermionic creation and annihilation operators that create or annihilate an electron in state $|i\rangle$ and $b^\dagger$ and b are bosonic photon creation and annihilation operators, respectively. The external Markovian environment is including via Lindblad terms, hence the system is described by the von Neumann-Lindblad equation (2.49). With $H = H_{\mathrm{JC}} + H_E$, the resulting master equation for the full density matrix ρ reads

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\rho = &-\frac{\mathrm{i}}{\hbar}[H,\rho] \\ &+ \frac{P}{2}\left(a_4^\dagger a_1 \rho a_1^\dagger a_4 - a_1^\dagger a_4 a_4^\dagger a_1 \rho - \rho a_1^\dagger a_4 a_4^\dagger a_1 \right) \\ &+ \frac{\kappa}{2}\left(2 b \rho b^\dagger - b^\dagger b \rho - \rho b^\dagger b \right) \\ &+ \sum_{\substack{r,s \\ r=s-1}} \frac{\gamma_{rs}}{2}\left(2 a_r^\dagger a_s \rho a_s^\dagger a_r - a_s^\dagger a_r a_r^\dagger a_s \rho - \rho a_s^\dagger a_r a_r^\dagger a_s \right), \end{aligned} \tag{5.2}$$

where the commutator describes the unitary evolution, while the subsequent lines describe interaction with the environment, i.e., incoherent pumping with rate P, cavity losses with rate κ, and electronic relaxation with rates γ_{rs}. Only one of the pumping mechanisms is included at a time. Consequently, either $P \neq 0, E = 0$ and the system is assumed to be pumped incoherently, or $P = 0, E \neq 0$ for coherent excitation.

Note the fundamental differences between incoherent pumping via Lindblad term and the coherent pumping via H_E. While in the first case the charge carrier is pumped via coupling to an external bath, for instance, illumination with thermal light (Ralph & Savage, 1991; Mu & Savage, 1992) or via interaction with the wetting-layer and barrier material (Gies *et al.,* 2017), in the later case the system is pumped by the coherent light of another laser (Ralph & Savage, 1991; Mu & Savage, 1992). It will be demonstrated later, how for different pumping mechanisms different information turns out to be relevant as input for the maximum-entropy method. Another important fact is that for non-vanishing rates the system reaches a unique steady state (Schirmer & Wang, 2010). Therefore, the maximum-entropy method prediction will yield the unique steady state. For multiple accessible steady states (which usually is not the case), the maximum-entropy method gives an unbiased mixture of all possible outcomes (not shown here). Because of the small system size, here the von Neumann-Lindblad equation is still solvable numerically via time-integration (cf. Sec. 3.1.1) yielding the full density matrix ρ, which is in general not the case for larger systems. The solution ρ is assumed to be the numerically exact density matrix and used for comparison with the maximum-entropy-method results in the following sections. Note that also other approaches, such as a variational principle (Weimer, 2015) or computing the kernel of the Liouvillian (Navarrete-Benlloch, 2015), can be applied to determine the steady state. Also note that the equation-of-motion approach described in Sec. 2.3.3 would not be able to deliver the full density matrix.

5.1.2 Input Information and Observation Levels

The crucial part for the successful construction of a maximum-entropy density matrix according to Sec. 2.3 is a sufficient choice of a physically reasonable observation level $\{\alpha\}$. To do so, the procedure described in Sec. 2.3.3 is applied here. Accordingly, it is exploited that, in a steady state, any observable A_i is constant in time and

$$\langle \mathcal{L}(A_i) \rangle = \mu_i = 0, \quad i = 1, \ldots, k \tag{5.3}$$

is used as constraint. For the investigated benchmark system the master equation (5.2) results in

$$\begin{aligned}\mathcal{L}(A_i) =& \frac{\mathrm{i}}{\hbar}[H, A_i] \\ &+ \frac{P}{2}\Big(a_1^\dagger a_4 A_i a_4^\dagger a_1 - a_1^\dagger a_4 a_4^\dagger a_1 A_i - A_i a_1^\dagger a_4 a_4^\dagger a_1\Big) \\ &+ \frac{\kappa}{2}\Big(2b^\dagger A_i b - b^\dagger b A_i - A_i b^\dagger b\Big) \\ &+ \sum_{\substack{r,s \\ r=s-1}} \frac{\gamma_{rs}}{2}\Big(2a_s^\dagger a_r A_i a_r^\dagger a_s - a_s^\dagger a_r a_r^\dagger a_s A_i - A_i a_s^\dagger a_r a_r^\dagger a_s\Big).\end{aligned} \tag{5.4}$$

Note the subtle differences (sign change of the commutator and different order of the first summand of the Lindblad terms) between Eqs. (5.2) and (5.4) that stem from the derivation via $\langle A\rangle = \mathrm{Tr}(\rho A)$ (cf. Sec. 2.3.3). The art of applying the maximum-entropy method is to specify the observation level $\{\alpha\} = \{A_i : i = 1, \ldots, k\}$, and thus the input information, such that, on the one hand, as much physically relevant information as possible is included, while simultaneously the inclusion of irrelevant information is avoided.

Several strategies may be applied to find a suitable observation level. First, it is advisable to only choose combinations of operators that are selfadjoint (and hence represent physical observables) since that results in a selfadjoint maximum-entropy density matrix for real values of the Lagrange multipliers (cf. App. A.2). Although choices of A_i that are not selfadjoint are also allowed, they generally either lead to complex valued Lagrange multipliers, which drastically complicates the numerical search for them, or result in vanishing Lagrange multipliers. A good starting point is to consider operators that follow naturally from the observables of interest, the Hamiltonian (5.1), and the right-hand side of Eq. (5.4). Hence here, the linearly independent operators n, N_1, N_2, N_3, and N_4 where $n = b^\dagger b$ and $N_i = a_i^\dagger a_i$ are the occupation number operators, could be a first reasonable choice, since they correspond to the demand of a time-constant photon number and occupation of the electronic levels. The Hamiltonian (5.1) also suggest to include a selfadjoint combination of the photon-assisted polarisation $\mathrm{i}(a_2^\dagger a_3 b^\dagger - a_3^\dagger a_2 b)$ and congruously, to also take into account $a_1^\dagger a_4 + a_4^\dagger a_1$ for coherently pumped systems. What might seem as an obvious choice in hindsight, is the result of a rather fiddly trial-and-error process as described at the end of Sec. 2.3.3. Fortunately, the inclusion of irrelevant information leads to a vanishing corresponding Lagrange multiplier, so this gives a possibility to eliminate all unnecessary input. But on the other hand, it should also be mentioned that some information only becomes relevant as soon as other observables are also included. For instance, if one only includes stationarity of the polarization-like operator combination, this leads to a vanishing corresponding Lagrange multiplier. If on the other hand, stationarity of the photon number, as well as the electronic occupation numbers is included, the Lagrange multiplier of the polarization-like input is unequal zero and therefore carries relevant information only in

Tab. 5.1.: First four observation levels $\{1\}$, $\{2\}$, $\{3\}$, and $\{4\}$. For each shown operator combination A_i, $\mathcal{L}(A_i) = 0$ with Eq. (5.4) is used as constraint. For higher observation levels the constraints of all previous observation levels are also included. Operators in the rightmost column are only considered as constraint in the coherent excitation case.

$\{1\}$	n	N_i	$\mathrm{i}\left(a_2^\dagger a_3 b^\dagger - a_3^\dagger a_2 b\right)$	$a_1^\dagger a_4 + a_4^\dagger a_1$
$\{2\}$	n^2	nN_i	$\mathrm{i}\left(n a_2^\dagger a_3 b^\dagger - a_3^\dagger a_2 b n\right)$	$n\left(a_1^\dagger a_4 + a_4^\dagger a_1\right)$
$\{3\}$	n^3	n^2N_i	$\mathrm{i}\left(n^2 a_2^\dagger a_3 b^\dagger - a_3^\dagger a_2 b n^2\right)$	$n^2\left(a_1^\dagger a_4 + a_4^\dagger a_1\right)$
$\{4\}$	n^4	n^3N_i	$\mathrm{i}\left(n^3 a_2^\dagger a_3 b^\dagger - a_3^\dagger a_2 b n^3\right)$	$n^3\left(a_1^\dagger a_4 + a_4^\dagger a_1\right)$

Tab. 5.2.: Alternative choice of observation levels (cf. Tab. 5.1)

$\{1\}$	$b^\dagger b$	N_i	$\mathrm{i}\left(a_2^\dagger a_3 b^\dagger - a_3^\dagger a_2 b\right)$	$a_1^\dagger a_4 + a_4^\dagger a_1$
$\{2\}$	$(b^\dagger)^2 b^2$	$b^\dagger b N_i$	$\mathrm{i}\left(a_2^\dagger a_3 (b^\dagger)^2 b - a_3^\dagger a_2 b^\dagger b^2\right)$	$b^\dagger b\left(a_1^\dagger a_4 + a_4^\dagger a_1\right)$
$\{3\}$	$(b^\dagger)^3 b^3$	$(b^\dagger)^2 b^2 N_i$	$\mathrm{i}\left(a_2^\dagger a_3 (b^\dagger)^3 b^2 - a_3^\dagger a_2 (b^\dagger)^2 b^3\right)$	$(b^\dagger)^2 b^2\left(a_1^\dagger a_4 + a_4^\dagger a_1\right)$
$\{4\}$	$(b^\dagger)^4 b^4$	$(b^\dagger)^3 b^3 N_i$	$\mathrm{i}\left(a_2^\dagger a_3 (b^\dagger)^4 b^3 - a_3^\dagger a_2 (b^\dagger)^3 b^4\right)$	$(b^\dagger)^3 b^3\left(a_1^\dagger a_4 + a_4^\dagger a_1\right)$

combination with the other quantities. Consequently, it should be very carefully investigated which combination of input leads to a decrease of entropy.

In what is referred to as observation level $\{1\}$, only the above mentioned quantities are included (cf. Tab. 5.1). For higher observation levels though, to also include higher moments of the photon distribution, the operators from the previous observation level are successively multiplied with the photon number n and included together with all previous inputs. As it turns out, other possible inputs, e.g., combinations of the electronic occupation number operators, are not necessary and lead to vanishing Lagrange multipliers. This finally gives a set of 6 (7 in the case of coherent excitation) pieces of information and an equal number of Lagrange multipliers for the first observation level, 12 (14) for the second, 18 (21) for the third and 24 (28) for the fourth (cf. Tab. 5.1). It is emphasized that this choice is not unique. It is exactly in the spirit of the maximum-entropy method that the predicted maximum-entropy density matrix does not depend on the explicit way information is presented but rather only on the information itself (cf. Chap. 2). Another possibility that leads to the exact same predictions is given in Tab. 5.2 where instead of the photon number operator n the normal ordered creation and annihilation operators $b^\dagger$ and b are used. This may be less intuitive but allows for a better systematic inclusion of constraints, especially for more complex systems (cf. Sec. 5.2.1).

5.1.3 Numerical Results

Finally, the maximum-entropy-method results are benchmarked by comparison of the maximum-entropy density matrix $\hat{\rho}$ to the full steady-state density matrix ρ obtained by numerically solving the von Neumann-Lindblad equation (5.2) via time-integration. To get deeper insight into how the maximum-entropy method works in practice, three different cases are studied. First, the quantum-dot laser system is incoherently pumped via Lindblad term for a wide range of pump rates P. In that case only the smaller set of input information is included for the first four observation levels as shown in Tab. 5.1. As second and third case the coherent excitation with strength E is investigated. There, two different sets of input information are used to exemplify the behavior of the maximum-entropy-method predictions and to demonstrate a possible use as trial-and-error tool to find out which information is physically relevant. As a start, the same observation levels as for the incoherent pumping case are used to illustrate how results for an insufficient choice of observation levels manifest themselves. Then, the extended observation levels of Tab. 5.1 are used to underline that the additional constraints carry relevant information in one case but not the other. In each case, the investigations are made considering the entropy, the relative entropy, the mean photon number and auto-correlation functions as well as examples of the full photon statistics. For all results, the respective figures show the incoherent pumping case on the left hand side and the coherent excitation on the right hand side, whereas the insufficiently chosen observation level is illustrated as inset on the right hand side. For numerical implementation the photonic Hilbert space needs to be restricted to a finite value, which here is chosen to be $n_{\max} = 40$. For the mixed Hilbert space, since the electron can only occupy one of the four electronic levels, this leads to a sparse square 164×164 density matrix. It is stressed that no time evolution needs to be calculated for an application of the maximum-entropy method.

Entropy and relative entropy First, the entropy is calculated for the numerically exact density matrix ρ by evaluating the general entropy expression (2.35). Note that this might be a tricky task especially for larger matrices or when numerical errors become relevant such that the matrix logarithm loses unambiguousness. For the maximum-entropy method's results, on the other hand, the entropy can easily be obtained from the partition function via Eq. (2.55). The entropy values for incoherent and coherent pumping as well as the insufficient observation level choice in the latter case are depicted in Fig. 5.2. Providing more information, i.e., choosing a higher observation level, in general leads to a density matrix with less entropy and thus less uncertainty $S(\hat{\rho})$. The entropy decreases for higher observation levels and finally tends towards the numerically exact steady-state solution of the von Neumann-Lindblad equation. This is true for all pump rates for both, coherently and

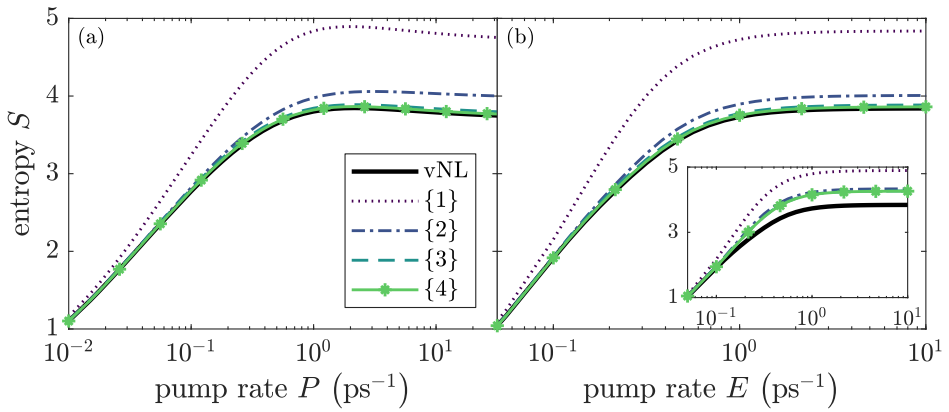


Fig. 5.2.: Entropy S for incoherent (a) and coherent pumping (b). Solid black curves show $S(\rho)$ for the steady-state solution of the von Neumann-Lindblad equation. Other curves show $S(\hat{\rho}_{\{\alpha\}})$ for the first four observation levels. For the inset in (b) input information of the type $a_1^\dagger a_4 + a_4^\dagger a_1$ was excluded. Throughout all calculations the following parameters are used: $\gamma_{12} = \gamma_{34} = 1\,\mathrm{ps}^{-1}$, $\gamma_{23} = 0.01\,\mathrm{ps}^{-1}$, $\kappa = 0.02\,\mathrm{ps}^{-1}$, and $g = 0.2\,\mathrm{meV}$. This figure has been published in (Melcher *et al.,* 2019).

incoherently pumped systems. For higher pump rates a higher observation level is required, whereas for low pump rates even the first observation level is sufficient.

To measure the quality of the maximum-entropy density matrix $\hat{\rho}$ compared to the numerically exact ρ the quantum relative entropy (Hiai *et al.,* 1981; Vedral, 2002)

$$S(\rho||\hat{\rho}) = \mathrm{Tr}\left[\rho\left(\ln\rho - \ln\hat{\rho}\right)\right] \tag{5.5}$$

is used (cf. Fig. 5.3). It measures how much information is lost when choosing the approximation $\hat{\rho}$ instead of the exact density matrix ρ. Note that the relative entropy is not symmetric. Alternatively, one could also use the trace distance (Hiai *et al.,* 1981)

$$T(\rho,\hat{\rho}) = \frac{1}{2}\,\mathrm{Tr}\left[\sqrt{(\rho-\hat{\rho})^2}\right] \tag{5.6}$$

which is a measure of the distinguishability between two states and gives very similar results. Both are connected via the quantum Pinsker inequality (Hiai *et al.,* 1981)

$$S(\rho||\hat{\rho}) \geq 2[T(\rho,\hat{\rho})]^2. \tag{5.7}$$

Higher observation levels lead to a decrease of relative entropy, hence the maximum entropy density matrix $\hat{\rho}$ tends towards the density matrix ρ of the steady-state solution of the von Neumann-Lindblad equation. Again, it is apparent that the maximum-entropy method delivers better results for weaker pumping.

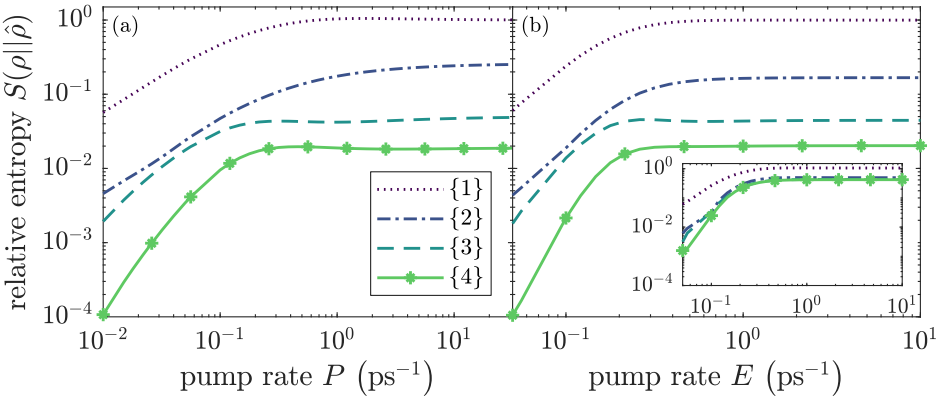


Fig. 5.3.: Relative entropy $S(\rho||\hat{\rho}_{\{\alpha\}})$ for incoherent (a) and coherent pumping (b) up to fourth observation level. For the inset in (b) input information of the type $a_1^\dagger a_4 + a_4^\dagger a_1$ was excluded. This figure has been published in (Melcher *et al.,* 2019).

The insets in Figs. 5.2 and 5.3 show that an appropriate choice of input information indeed is essential. There, the results are depicted when information about the stationarity of the coherent excitation process and its correlations is missing and input information of the form $a_1^\dagger a_4 + a_4^\dagger a_1$ is not included (cf. Tab. 5.1). Withholding it dramatically effects the results for the coherently pumped system, yielding less approximation quality, whereas for incoherent pumping this information is completely obsolete. Even explicitly adding it leads to vanishing of the corresponding Lagrange multipliers throughout the whole range of pump rates and for all observation levels. This corresponds to the notion that in the incoherent excitation regime averages of the form $\left\langle a_1^\dagger a_4 \right\rangle$ and $\left\langle a_4^\dagger a_1 \right\rangle$ vanish (Gies *et al.,* 2007; Gartner, 2019). In that sense the maximum-entropy method can be understood as a trial-and-error approach for learning and identifying the relevant processes within physical systems. This strategy can be applied, for instance, to decide which quantities can be neglected in equation-of-motion plus cluster-expansion-method approaches.

Mean photon number and correlation functions From the full density matrix, relevant physical observables, namely the mean photon number and autocorrelation functions

$$\langle n \rangle = \left\langle b^\dagger b \right\rangle, \quad g^{(k)} = \frac{\left\langle (b^\dagger)^k b^k \right\rangle}{\langle b^\dagger b \rangle^k} \tag{5.8}$$

are directly accessible (cf. Fig. 5.4). Contrary to naive expectation, for low pump rates the second-order auto-correlation function is not close to zero for this choice of parameters because of the high quality mode and the associated low loss rate $\kappa = 0.02\,\mathrm{ps}^{-1}$. For a higher loss rate though (e.g. $\kappa = 2\,\mathrm{ps}^{-1}$), $g^{(2)}$ is close to zero and the quantum dot microcavity system operates as single-photon source, which is nonetheless not shown here. Although the mean photon number already accords well within the first observation level

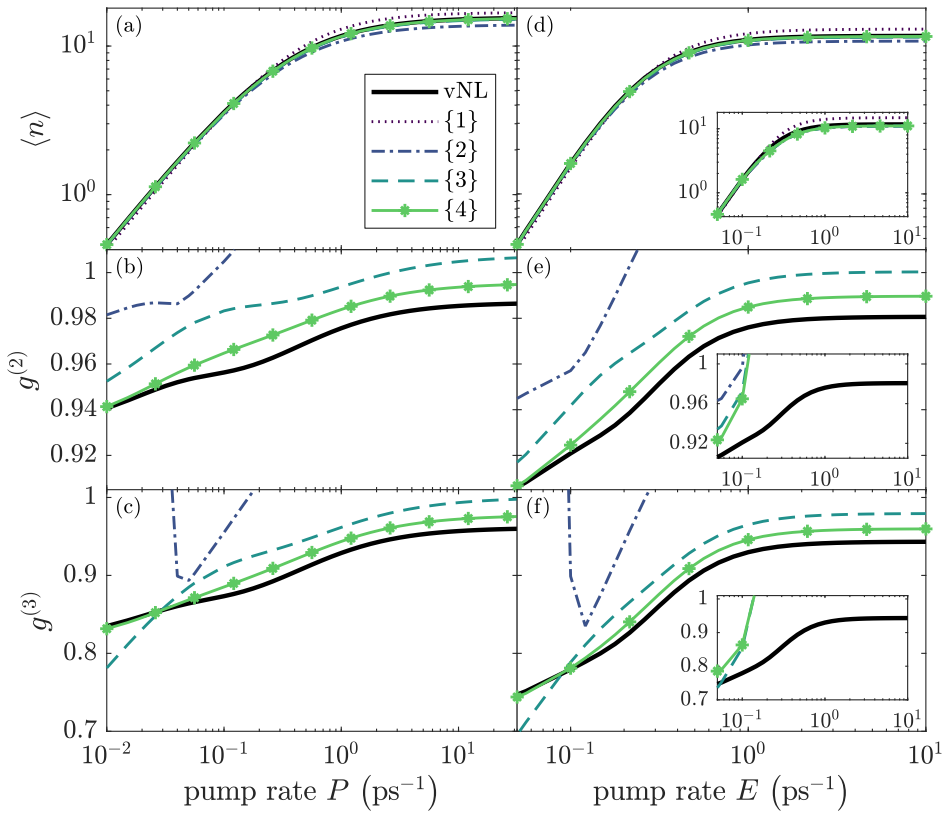


Fig. 5.4.: Mean photon number $\langle n \rangle$ [(a) and (d)] and second and third order autocorrelation function $g^{(2)}$ [(b) and (e)] and $g^{(3)}$ [(c) and (f)] for incoherent (a)–(c) and coherent pumping (d)–(f). Solid black curves show the results for the steady-state solution of the von Neumann-Lindblad equation. Other curves show the maximum-entropy-method results for the first four observation levels. For the insets in (d)–(f) input information of the type $a_1^\dagger a_4 + a_4^\dagger a_1$ was excluded. This figure has been published in (Melcher *et al.,* 2019).

[Fig. 5.4 (a) and (d)], to obtain reasonable values for the auto-correlation functions one needs to include higher-order moments. While observation level $\{\alpha\}$ contains information about the stationarity of operators of the form $n^{(\alpha-1)}O_i$ (cf. Tab. 5.1), the autocorrelation function of order k contains information about moments $\left\langle n^k \right\rangle$. Consequently, to get good accordance for $g^{(k)}$, observation level $\{k+1\}$ or higher is sufficient. For lower pump rates though, a rather odd behavior of the auto-correlation function is observed [cf. second observation level in Fig. 5.4 (c) and (f)]. This can be retraced to the restriction of the maximum photon number n to a finite value $n_{\max}$ that leads to an artificial increase of the probability values p_n close to $n_{\max}$. This was already found for the birth-death model (Sec. 3.1.3) but holds true here for all observation levels. This furthermore results in an overestimation of higher-order moments that is highly dependent on the choice of $n_{\max}$ and consequently explains the behavior of the auto-correlation function for weaker pumping. It is stressed that the Lagrange multipliers are system-size dependent throughout all observation levels, which is a strong indication that the numerically necessary restriction of the photonic

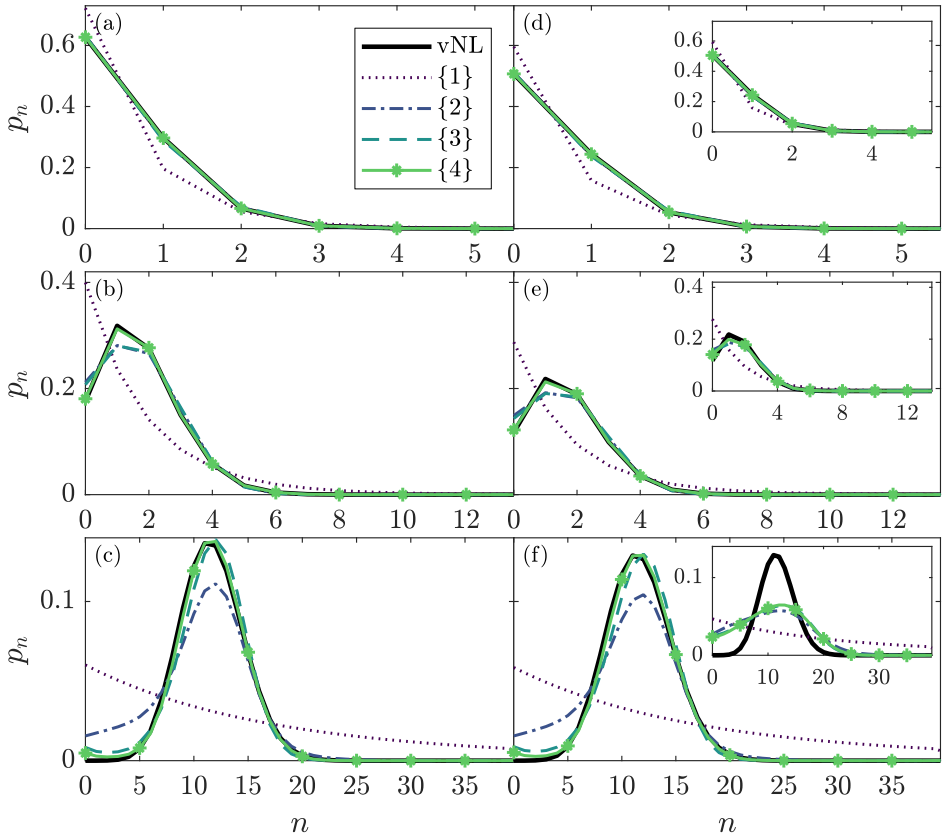


Fig. 5.5.: Photon statistics p_n for pump rates $P = 0.01\,\mathrm{ps}^{-1}$ (a), $P = 0.04\,\mathrm{ps}^{-1}$ (b), $P = 1\,\mathrm{ps}^{-1}$ (c) for incoherent pumping and $E = 0.05\,\mathrm{ps}^{-1}$ (d), $E = 0.1\,\mathrm{ps}^{-1}$ (e), $E = 10\,\mathrm{ps}^{-1}$ (f) for coherent pumping, respectively. Solid black curves show the photon statistics for the steady-state solution of the von Neumann-Lindblad equation. Other curves show p_n obtained by the maximum-entropy method for the first four observation levels. For the insets in (d)–(f) input information of the type $a_1^\dagger a_4 + a_4^\dagger a_1$ was excluded. In (a) and (d) the steady-state solution and the results for second, third, and fourth observation levels nearly lie on top of each other. This figure has been published in (Melcher *et al.,* 2019).

portion of the Hilbert space to a finite value $n_{\max}$ introduces an incorrect information that the maximum-entropy method accounts for but is not physically justified. Nevertheless, this can partly be fixed by a truncation closer to the relevant part of the photon distribution (e.g. $n_{\max} = 10$) that leads to better accordance for low pump rates as well. Fortunately, for intermediate and high pump rates the values for entropy, relative entropy, various expectation values as well as the photon statistics remain almost unchanged for different choices of $n_{\max}$. There, application of the maximum-entropy method is reasonably justified.

Photon statistics After tracing up the electronic Hilbert space, the photon statistics is easily obtained from the diagonal elements of the remaining photonic density matrix. For three pump rates each, a low, an intermediate and a high pump rate, the maximum-

entropy method's photon statistics are shown with the ones obtained form the numerically exact solution of the von Neumann-Lindblad equation in Fig. 5.5. For both, coherent and incoherent pumping, the maximum-entropy method photon statistics tend towards the steady-state solution of the von Neumann-Lindblad equation. Remarkably, already the second observation level resembles the photon statistics for low pump rates. For higher pump rates though, more information is required to reproduce the more complicated structure of the photon statistics; but still, already the third observation level's distribution is nearly identical to the steady-state solution. Only for low photon numbers n the maximum-entropy method tends to overestimate the probability p_n, which might be the consequence of the restriction of the maximum photon number to a finite value $n_{\max}$.

In the case of insufficient information (see insets) only low pump rate results are in agreement, whereas for higher pumping the maximum-entropy method solution is far away from the original one. This fact again can be interpreted as a hint that operators of the form $a_1^\dagger a_4 + a_4^\dagger a_1$ are of crucial importance especially for stronger coherent pumping while being less important for weak pumping and completely obsolete in the incoherent pumping case. As suggested in Chap. 2, this exemplifies one of the strengths of the maximum-entropy approach as trial-and-error method. Assume that the relevance of operators of the form $a_1^\dagger a_4 + a_4^\dagger a_1$ for coherently excited systems would not yet be known. For weak pumping, the maximum-entropy method would result in physically reasonable predictions of the photon statistics, as in that regime, stationarity of operators $a_1^\dagger a_4 + a_4^\dagger a_1$ carries little relevant information. For stronger pumping though, a much broader photon distribution [cf. Fig. 5.5 (f), inset] would be predicted by the maximum-entropy method. If compared to for instance the original (von Neumann-Lindblad equation) distribution, to experimental data or even to the physical intuition that the statistics should have Poisson-like shape, it would become apparent that a relevant physical process has been overlooked so far. Via trial-and-error this relevant process could then be revealed. Note that, even in the case of insufficient information, no unphysical predictions are made by the maximum-entropy method, since the mean photon number as well as the auto-correlation functions perfectly match the photon statistics that stem from the maximum-entropy density matrix. Quite the contrary, assuming any other prediction than the prediction given by the maximum-entropy method would introduce inconsistencies not represented by the given input information.

5.2 Outlook

In this section a brief outlook is given on possible future research directions. In Sec. 5.2.1 an expansion of the previous quantum-dot microcavity laser to a second laser mode is briefly outlined by presenting a simplified first attempt to derive suitable observation levels. Furthermore, the resulting photon statistics are compared to the numerically exact solution of the von Neumann-Lindblad equation for two different pump rates. In Sec. 5.2.2 a combination of both maximum-entropy-method approaches, input of moment values and demand of stationarity, is proposed to describe a quantum-dot laser system with one laser mode and a second light field that couples the outer electronic levels and serves as excitation mechanism. A general outlook is given in Sec. 5.2.3.

5.2.1 Bimodal Quantum-Dot Microcavity Laser

One specific extension of the model presented in Sec. 5.1 could be the inclusion of a second light field mode. In bimodal microcavity lasers a variety of different effects such as the competition for gain of the two modes as well as mode switching and mode coupling behavior have been previously examined (Leymann *et al.,* 2013b; Khanbekyan *et al.,* 2015; Fanaei *et al.,* 2016; Khanbekyan, 2018; Marconi *et al.,* 2018; Schlottmann *et al.,* 2018; Schmidt *et al.,* 2020). The maximum-entropy method's viewpoint could be an enriching addition to the previous findings, making the full photon statistics available and potentially shedding new light on the fundamental physical mechanisms. Here, a first step is presented that can be seen as a rough guideline to develop a physically justified model.

From a calculation and numerical viewpoint, the simplest extension of the single quantum-dot single-mode laser model of Sec. 5.1 can be achieved by simply including a second mode to the Jaynes-Cummings Hamiltonian

$$H_{\mathrm{JC}} = g_1^* a_2^\dagger a_3 b_1^\dagger + g_1 a_3^\dagger a_2 b_1 + g_2^* a_2^\dagger a_3 b_2^\dagger + g_2 a_3^\dagger a_2 b_2, \tag{5.9}$$

where $b_1^\dagger$, $b_2^\dagger$ and b_1, b_2 are the respective creation and annihilation operators of the first and second mode. Moreover, different light-matter coupling strengths $g_1 \neq g_2$ should be considered. Also, cavity losses are included for both modes and the Lindblad terms in Eqs. (5.2) and (5.4) are modified accordingly. It is pointed out that this extension is not yet physically justified but solely serves the purpose of artificially including two different modes in a implementationally least challenging way. Physically far more reasonable would be the inclusion of two modes that differ in energy $\hbar\omega_1 \neq \hbar\omega_2$ and thus at least one of them would be slightly off-resonant with the electronic transition, which should be the next step in a future investigation.

Tab. 5.3.: Possible choice of the first three observation levels $\{1\}$, $\{2\}$, and $\{3\}$ for the two-mode case. For each shown operator combination A_i, stationarity $\mathcal{L}(A_i) = 0$ is used as constraint. For higher observation levels the constraints of all previous observation levels are also included.

$\{1\}$
$b_1^\dagger b_1 \quad b_1^\dagger b_2 + b_2^\dagger b_1 \quad b_2^\dagger b_2$
N_i
$\mathrm{i}\left(a_2^\dagger a_3 b_1^\dagger - a_3^\dagger a_2 b_1\right) \quad \mathrm{i}\left(a_2^\dagger a_3 b_2^\dagger - a_3^\dagger a_2 b_2\right)$

$\{2\}$
$b_1^\dagger b_1^\dagger b_1 b_2 + b_2^\dagger b_1^\dagger b_1 b_1$
$b_1^\dagger b_1^\dagger b_1 b_1 \quad b_1^\dagger b_1^\dagger b_2 b_2 + b_2^\dagger b_2^\dagger b_1 b_1 \quad b_2^\dagger b_2^\dagger b_2 b_2$
$b_1^\dagger b_2^\dagger b_2 b_2 + b_2^\dagger b_2^\dagger b_2 b_1$
$b_1^\dagger b_1 N_i \quad \left(b_1^\dagger b_2 + b_2^\dagger b_1\right) N_i \quad b_2^\dagger b_2 N_i$
$\mathrm{i}\left(a_2^\dagger a_3 b_1^\dagger b_1^\dagger b_1 - a_3^\dagger a_2 b_1^\dagger b_1 b_1\right) \quad \mathrm{i}\left(a_2^\dagger a_3 b_2^\dagger b_2^\dagger b_2 - a_3^\dagger a_2 b_2^\dagger b_2 b_2\right)$
$\mathrm{i}\left(a_2^\dagger a_3 b_1^\dagger b_1^\dagger b_2 - a_3^\dagger a_2 b_2^\dagger b_1 b_1\right) \quad \mathrm{i}\left(a_2^\dagger a_3 b_2^\dagger b_2^\dagger b_1 - a_3^\dagger a_2 b_1^\dagger b_2 b_2\right)$
$\mathrm{i}\left(a_2^\dagger a_3 b_1^\dagger b_2^\dagger b_2 - a_3^\dagger a_2 b_2^\dagger b_2 b_1\right) \quad \mathrm{i}\left(a_2^\dagger a_3 b_2^\dagger b_1^\dagger b_1 - a_3^\dagger a_2 b_1^\dagger b_1 b_2\right)$

$\{3\}$
$b_1^\dagger b_1^\dagger b_1^\dagger b_1 b_1 b_2 + b_2^\dagger b_1^\dagger b_1^\dagger b_1 b_1 b_1$
$b_1^\dagger b_1^\dagger b_1^\dagger b_1 b_2 b_2 + b_2^\dagger b_2^\dagger b_1^\dagger b_1 b_1 b_1$
$b_1^\dagger b_1^\dagger b_1^\dagger b_1 b_1 b_1 \quad b_1^\dagger b_1^\dagger b_1^\dagger b_2 b_2 b_2 + b_2^\dagger b_2^\dagger b_2^\dagger b_1 b_1 b_1 \quad b_2^\dagger b_2^\dagger b_2^\dagger b_2 b_2 b_2$
$b_1^\dagger b_1^\dagger b_2^\dagger b_2 b_2 b_2 + b_2^\dagger b_2^\dagger b_2^\dagger b_2 b_1 b_1$
$b_1^\dagger b_2^\dagger b_2^\dagger b_2 b_2 b_2 + b_2^\dagger b_2^\dagger b_2^\dagger b_2 b_2 b_1$
$\left(b_1^\dagger b_1^\dagger b_1 b_2 + b_2^\dagger b_1^\dagger b_1 b_1\right) N_i$
$b_1^\dagger b_1^\dagger b_1 b_1 N_i \quad \left(b_1^\dagger b_1^\dagger b_2 b_2 + b_2^\dagger b_2^\dagger b_1 b_1\right) N_i \quad b_2^\dagger b_2^\dagger b_2 b_2 N_i$
$\left(b_1^\dagger b_2^\dagger b_2 b_2 + b_2^\dagger b_2^\dagger b_2 b_1\right) N_i$
$\mathrm{i}\left(a_2^\dagger a_3 b_1^\dagger b_1^\dagger b_1^\dagger b_1 b_1 - a_3^\dagger a_2 b_1^\dagger b_1^\dagger b_1 b_1 b_1\right) \quad \mathrm{i}\left(a_2^\dagger a_3 b_2^\dagger b_2^\dagger b_2^\dagger b_2 b_2 - a_3^\dagger a_2 b_2^\dagger b_2^\dagger b_2 b_2 b_2\right)$
$\mathrm{i}\left(a_2^\dagger a_3 b_1^\dagger b_1^\dagger b_1^\dagger b_1 b_2 - a_3^\dagger a_2 b_2^\dagger b_1^\dagger b_1 b_1 b_1\right) \quad \mathrm{i}\left(a_2^\dagger a_3 b_2^\dagger b_2^\dagger b_2^\dagger b_2 b_1 - a_3^\dagger a_2 b_1^\dagger b_2^\dagger b_2 b_2 b_2\right)$
$\mathrm{i}\left(a_2^\dagger a_3 b_1^\dagger b_1^\dagger b_1^\dagger b_2 b_2 - a_3^\dagger a_2 b_2^\dagger b_2^\dagger b_1 b_1 b_1\right) \quad \mathrm{i}\left(a_2^\dagger a_3 b_2^\dagger b_2^\dagger b_2^\dagger b_1 b_1 - a_3^\dagger a_2 b_1^\dagger b_2^\dagger b_2 b_2 b_2\right)$
$\mathrm{i}\left(a_2^\dagger a_3 b_1^\dagger b_1^\dagger b_2^\dagger b_2 b_2 - a_3^\dagger a_2 b_2^\dagger b_2^\dagger b_1 b_1 b_1\right) \quad \mathrm{i}\left(a_2^\dagger a_3 b_2^\dagger b_2^\dagger b_1^\dagger b_1 b_1 - a_3^\dagger a_2 b_1^\dagger b_1^\dagger b_1 b_2 b_2\right)$
$\mathrm{i}\left(a_2^\dagger a_3 b_1^\dagger b_2^\dagger b_2^\dagger b_2 b_2 - a_3^\dagger a_2 b_2^\dagger b_2^\dagger b_2 b_2 b_1\right) \quad \mathrm{i}\left(a_2^\dagger a_3 b_2^\dagger b_1^\dagger b_1^\dagger b_1 b_1 - a_3^\dagger a_2 b_1^\dagger b_1^\dagger b_1 b_1 b_2\right)$

Nevertheless, the assumptions at hand suffice to create two non-degenerate modes and demonstrate a possible choice of observation levels (Tab. 5.3). In contrast to the single-mode case, here, stationarity of operators of both modes need to be included for reasonable predictions. For the first observation level, apart from the already known quantities also transitions from one mode to the other need to be considered by inclusion of the selfadjoint combination $b_1^\dagger b_2 + b_2^\dagger b_1$. Interestingly, inclusion of the other possible selfadjoint combinations i $\left(b_1^\dagger b_2 - b_2^\dagger b_1\right)$ and $a_2^\dagger a_3 b_{1(2)}^\dagger + a_3^\dagger a_2 b_{1(2)}$ lead to vanishing corresponding Lagrange multipliers. This might be interpreted as a hint whether the intensity or phase information is the relevant one, but further investigations should be done to give a definite statement. As pointed out earlier, the choice of observation levels is not unique and working with normal ordering is especially handy in the two-mode case. For the second observation level operator combinations that only contain bosonic operators $b_{1(2)}^\dagger b_{1(2)}$ and $b_1^\dagger b_2 + b_2^\dagger b_1$ are multiplied with the electronic occupation number as well as the polarisation-like operators, which leads to $3 \cdot 4 = 12$ new inputs with electronic occupation numbers and $3 \cdot 2 = 6$ new inputs with polarization-like operator combinations. This is indicated by the block structure of Tab. 5.3. Furthermore, operator combinations of four bosonic operators are systematically included in the second observation level. This procedure is continued accordingly for the higher observation levels, leading to a set of 9 inputs for observation level $\{1\}$, 32 for observation level $\{2\}$, and 69 for observation level $\{3\}$. Congruously, the numerical search gets more demanding due to the high number of Lagrange multipliers. Moreover, since two photonic Hilberts subspaces are considered, the size of the density matrix increases. Nevertheless, convergence of the maximum-entropy method can still be achieved (especially with aid of the numerical strategies presented in Sec. 2.3.4) and also comparison with the numerically exact solution of the von Neumann-Lindblad equation is still possible for the two-mode case. It is again noted that, even though the choice of input information might seem obvious, actually, Tab. 5.3 is in itself an important result of the maximum-entropy method, since it summarizes the physically important processes in the system.

A restriction of the maximum photon numbers to $n_{1,\max} = 40$ and $n_{2,\max} = 10$ results in a density matrix of size 1804×1804. The first mode is assumed to be the strong mode and the second to be the weak mode ($g_1 > g_2$ and $\kappa_1 < \kappa_2$, cf. caption of Fig. 5.6). The maximum-entropy method's photon statistics of the individual modes in comparison to the solution of the von Neumann-Lindblad equation are shown for two different pump rates in Fig. 5.6. For the weak pumping case, the upper row shows a peaked structure of the first mode photon statistics and thermal shape for the second mode. Observation levels $\{1\}$ and $\{2\}$ give reasonable predictions where the latter suffices to give a decent approximation for both modes. Surprisingly, the photon statistics of observation level $\{3\}$ are of worse quality. One possible explanation could be that the numerical search for optimal Lagrange multipliers did not succeed even though the internal termination

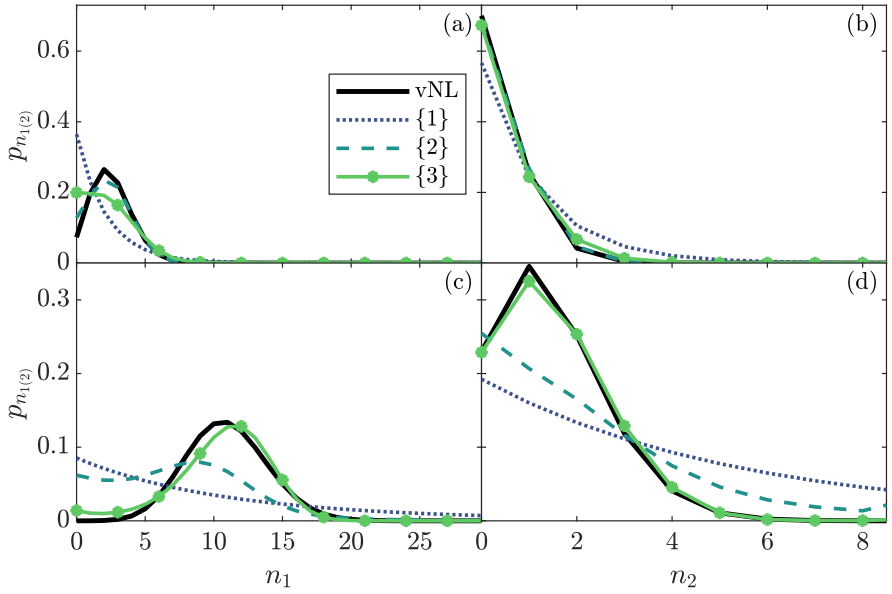


Fig. 5.6.: Photon statistics of first mode p_{n_1} (left column) and for second mode p_{n_2} (right column) for weak pumping ($P = 0.08\,\mathrm{ps}^{-1}$) (upper row) and strong pumping ($P = 1.67\,\mathrm{ps}^{-1}$) (lower row). Solid black curves show the photon statistics for the steady-state solution of the von Neumann-Lindblad equation. Other curves show the maximum-entropy-method results for observation levels $\{1\}$, $\{2\}$, and$\{3\}$. Note the differently scaled x-axes. Parameters: $\gamma_{12} = \gamma_{34} = 1\,\mathrm{ps}^{-1}$, $\gamma_{23} = 0.01\,\mathrm{ps}^{-1}$, $\kappa_1 = 0.02\,\mathrm{ps}^{-1}$, $\kappa_2 = 0.03\,\mathrm{ps}^{-1}$, $g_1 = 0.2\,\mathrm{meV}$, and $g_2 = 0.1\,\mathrm{meV}$

criteria of the Levenberg-Marquardt were fulfilled and hence the shown photon statistics is not the one with maximum entropy. Consequently, a next step should be the careful and systematic investigation of possible numerical effects on the search for the Lagrange multipliers. For the stronger pumping though (lower row), the statistics of both modes are adequately approximated by the maximum-entropy distributions and the consecutive observation levels lead to better quality approximations. Nevertheless, the first mode's statistics suffers from an overestimation of low photon number probabilities as was already found in the single-mode case (cf. Sec. 5.1.3). This is presumably due to the truncation of the maximum photon number to a finite value and is even more pronounced in the two-mode case since two Hilbert spaces are truncated.

However, this first example demonstrates that the maximum-entropy method can in principle also be applied to more complex systems. Although convergence is achieved pretty effortlessly for certain parameters for the lower observation levels, it should be mentioned that, in general, application of the maximum-entropy method is much more time consuming compared to the single-mode case and the von Neumann-Lindblad equation approach is clearly superior here. Nevertheless, the maximum-entropy-method approach might be improved by a more systematic optimization of the numerical search for the optimal Lagrange multipliers. Then, also more complex systems could be investigated by the maximum-entropy method.

An advantage of the maximum-entropy method is that the storage of the full density matrix becomes much more efficient since the system's whole information is contained within only few real-valued Lagrange multipliers. Also, the least-squares optimization of the maximum-entropy method is less memory demanding than time-integration of the von Neumann-Lindblad equation.

5.2.2 Excitation with a Second Light Field

A second possible extension of the single quantum-dot laser model could be the inclusion of an external light field mode that couples to the outer electronic levels $|1\rangle$ and $|4\rangle$ (cf. Fig. 5.7). In contrast to the coherent pumping of Sec. 5.1, one would explicitly include the second light field in the description and the system could be excited with a specific given photon-number distribution. It is known that excitation with different statistics leads to differences in the statistics of the emitted light (Grothe, 2014; Kazimierczuk *et al.,* 2015; Strauß *et al.,* 2016) and the maximum-entropy method could give valuable insights.

To investigate such systems the maximum-entropy method would be particularly well suited. For a time-integration of the von Neumann-Lindblad equation it is always required to choose an initial state ρ_0. Although it is possible to choose it such that the statistics of the exciting mode is fixed for the starting time $t = 0$, for later times, all photons of the exciting modes are finally lost through the dissipative processes (Grothe, 2014). To model continuous excitation though, von Neumann-Lindblad or equation-of-motion approaches would require artificial manipulation of the associated equations by hand. The maximum-entropy method does not suffer from this limitation and thus gives an intuitive approach to the problem at hand, since it only deals with information or constraints on the system and does not involve calculation of a time evolution. The idea would be to combine the approaches of Chaps. 3 and 4 with the requirement of stationarity of observables that was employed in Sec. 5.1. Consequently, one would use

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\langle A_i\rangle &= 0, \quad i = 1, \ldots, k \\ \langle n_2\rangle &= I, \quad g_2^{(2)} = G \end{aligned} \tag{5.10}$$

as constraints. The first line corresponds to the demand of stationarity. Moreover, the mean photon number of the second mode is fixed to the a known value I and the auto-correlation function $g^{(2)}$ is fixed to a value G (for instance $G = 1$ for coherent or $G = 2$ for thermal light). Inclusion of higher order moments would concretise the shape of the photon-number distribution (cf. Sec. 3.1). So the second line of (5.10) corresponds to the approaches of Chaps. 3 and 4 where moment values were used as constraint. This is a good example for the strengths of the maximum-entropy approach since it only deals with information.

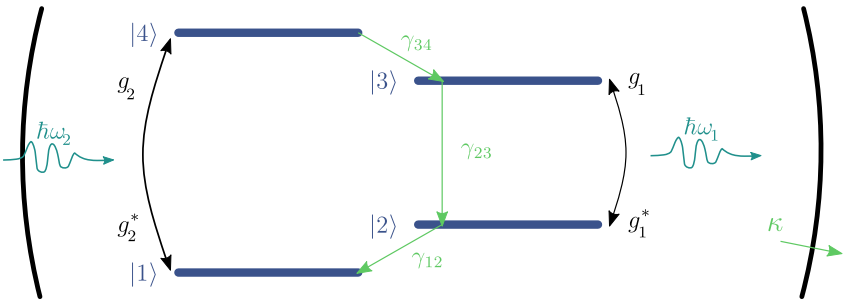


Fig. 5.7.: Four-level quantum dot in a microcavity excited by a second light field. The electronic energy levels are denoted by $|1\rangle$, $|2\rangle$, $|3\rangle$, and $|4\rangle$. The inner levels are resonantly coupled to the cavity mode with frequency ω_1 and the light-matter coupling strength is given by g_1. Interaction with the environment, namely electronic relaxation with rates γ_{ij} and cavity losses with rate κ are shown in lighter color. The system is excited via resonant interaction of the outer electronic levels $|1\rangle$ and $|4\rangle$ with a second external light field with frequency ω_2 and light-matter coupling strength g_2.

Application of the maximum-entropy method would translate to the question: "what is the least biased steady state of a system that is excited with a light field mode with mean photon number I and auto-correlation function value G?" An especially revealing first task would be to specify the set of operators A_i and thus find suitable observation levels that shed light on the relevant physical processes. A first reasonable choice would be the same as in Tab. 5.3 but it is not clear beforehand if stationarity of the mean photon number of the exciting mode needs to be included if the mean photon number itself is used as constraint. Furthermore, it should be carefully investigated how correlations of both modes need to be considered.

5.2.3 General Outlook

Apart from the more detailed propositions of the previous sections it would be highly desirable to investigate systems with several (non-identical) quantum dots (or atoms) as gain medium since there, the maximum-entropy method could help to gain deeper insight into the emerging collective effects (cf. e.g. Scheibner *et al.,* 2007; Scully & Svidzinsky, 2009; Mlynek *et al.,* 2014; Khanbekyan *et al.,* 2015; Leymann *et al.,* 2015; Jahnke *et al.,* 2016; Frisk Kockum *et al.,* 2019). Moreover, a more realistic modeling of the quantum dots, for instance, by considering many-particle configuration states that would also include excitonic effects (Leymann *et al.,* 2015; Lettau, 2017; Lettau *et al.,* 2017) and/or by using quantum-dot parameters that stem from more elaborate models, for instance, tight-binding calculations (Slater & Koster, 1954; Vogl *et al.,* 1983; Schulz, 2007; Zieliński *et al.,* 2010; Goldmann, 2014; Schumann, 2016; Melcher, 2016) would be beneficial. Also, it could be worthwhile to try to match maximum-entropy-method results to experimental data, since

there the full photon statistics is also available via measurements with a transition-edge sensor (Schlottmann *et al.,* 2018; Schmidt *et al.,* 2020). Moreover, although here only one specific model was investigated the maximum-entropy-method approach is not limited to quantum optics but can in principle be applied to any open quantum system that exhibits a steady-state situation. Consequently, the maximum-entropy method opens up the possibility to describe and investigate a wide range of different subfields of quantum mechanics. The possibility to treat non-Markovian situations (Pachón & Brumer, 2013a; Pachón *et al.,* 2013b; Pachón & Brumer, 2013; Triana *et al.,* 2016; de Vega & Alonso, 2017; Pachón *et al.,* 2019), where one might also include observables from the environment as input information could also be a possible direction that might lead to a better understanding of the role of the environment.

5.3 Chapter Conclusion

It was demonstrated how the maximum-entropy method can be used to derive the full density matrix of open quantum systems that are in a steady-state situation. This stand-alone approach was completely independent of equation-of-motion techniques and thus there was no need for any factorization-based truncation schemes concerning the many-particle hierarchy. Instead, a finite set of input information was used that lead to a self-consistent inference for the least biased full density matrix. As a benchmark the maximum-entropy-method results were compared to the steady-state solution for a single quantum-dot microcavity laser described by a von Neumann-Lindblad equation for both coherent and incoherent excitation. By comparing the entropy, relevant expectation values, and the full photon statistics of both independent approaches, remarkable accordance was demonstrated. Furthermore, by considering two different pumping mechanisms and different choice of input information, it was exemplified how the maximum-entropy method can be used as trial-and-error approach for identifying the relevant physical processes within the system. A possible outlook on future expansions of the model systems to two light field modes, either to a second laser mode or to excitation with an external light field with specific photon-number moments were given. For the former, a suitable choice of observation levels and first test calculations were presented, paving the way for more detailed investigations in the future. Although here, only few specific systems were examined, it was pointed out that the approach is neither limited to quantum optics nor to systems described by the von Neumann-Lindblad equation but can in principle be applied to all systems with stationary dynamics, i.e., in a steady state, but not necessarily in thermodynamic equilibrium. Beyond that, it should also be possible to treat non-Markovian situations, where one would also include observables from the environment as input information. Consequently, application of the maximum-entropy method opens up the possibility to circumvent numerous downsides of equation-of-motion techniques, gives an information theoretical perspective on the infinite many-particle hierarchy problem and provides the full density matrix making the full statistics and all relevant operators directly accessible. Nevertheless, the results of this chapter should merely be seen as a starting point that may ultimately lead to a possibility to apply the maximum-entropy method to physically more complex systems and use it as alternative or accompanying method to other techniques such as equation-of-motion approaches.

6 Conclusion

It was demonstrated how the maximum-entropy method can be applied in the field of micro- and nanolasers to gain access to the full statistical information, i.e., the full photon statistics or the full density matrix of the desired systems. To do so, the general form of the maximum-entropy distribution was derived out of the Shannonian entropy definition via a variational principle. For quantized descriptions, on the other hand, the von Neumann entropy was used to determine the general form of the maximum-entropy density matrix with. For both cases, suitable strategies were proposed to numerically determine the Lagrange multipliers.

Then, basic features of the maximum-entropy method were investigated when statistical moments were given as prior information. There, the maximum-entropy predictions tended towards the original probability distribution with increasing order as was validated for a birth-death benchmark model. Moreover, general existence criteria were exploited to demonstrate that there cannot exist an unbiased single-mode photon distribution that is capable of producing superthermal values of $g^{(2)}$ without assuming additional information about the probability distribution. In the bimodal case on the other hand, superthermal values were possible in one of the modes as long as the two modes were sufficiently anti-correlated.

Furthermore, the maximum-entropy method gave a possibility to define a laser threshold that was based on a sign change of the first Lagrange multiplier. Apart from perfect accordance with conventional threshold definitions, the threshold characterization by sign change was connected to the features of the photon probability distribution and occurred when the distribution first exhibited a peak. Consequently, the definition was also applicable for β-factors close to unity where conventional definitions failed. Moreover, the connection of the new threshold definition to values of the mean photon number and the auto-correlation function value were pointed out. By that, a valuable tool was established to unambiguously mark the change from thermal to a peaked distribution only by the knowledge of $\langle n \rangle$ and $g^{(2)}$; independently of any kinks or jumps in those quantities.

A combination of the maximum-entropy method with equation-of-motion approaches was successfully demonstrated for a semiconductor model for quantum-dot based microlasers and a model describing nanolasers with extended gain media. There, the output quantities of the equation-of-motion approaches were used as input for the maximum-entropy method to derive least biased estimations for the photon distribution. By considering the output

from a rate-equation model for extended gain media, the confusing picture of the threshold behavior in the few and many emitter cases was extended by the maximum-entropy method's perspective, indeed confirming that the laser threshold in the many-emitter case is almost independent of the β-factor. Aside from that, since the maximum-entropy method also delivered the entropy, a characterization of the light in terms of entropy revealed that the common argumentation that a value of $g^{(2)} \approx 1$ is a strong indication for a Poissonian photon distribution might often be premature and not necessarily reasonable. At least for the birth-death, semiconductor, and extended gain-media models, the unique consistent and unbiased estimations from the given photon moments, i.e., the distribution with maximum entropy, had rather Gaussian form with either smaller or larger variance that a Poissonian distribution.

Finally, a possibility to apply the maximum-entropy method as stand-alone approach completely independent of moment values from external sources was demonstrated for a single four-level quantum-dot microcavity laser. There, the equations of motion themselves were used as input. By that, the maximum-entropy method was applied to find the least biased steady-state density matrix. A sufficient choice of observation levels was presented and convergence of the maximum-entropy density matrix towards the results of the von Neumann-Lindblad-equation approach was demonstrated. In doing so, it was validated that requiring stationarity for several distinct observables as constraints indeed leads to a physically reasonable prediction of the density matrix. With this approach the many-particle hierarchy problem that commonly arises in equation-of-motion techniques and the need for a factorization or truncation were successfully circumvented. Furthermore, by considering two different pumping mechanisms and different choices of the input information the usage of the maximum-entropy method as trial-and-error approach to investigate the physically relevant processes was exemplified. Since the full density matrix was available, all relevant quantities, such as expectation values and (arbitrarily high order) auto-correlation function as well as the full photon statistics were easily available for comparison. Further research directions were pointed out, where inclusion of a second laser mode or a second light field mode as excitation mechanism were discussed in more detail. Still, for more complex systems, e.g., more (non-identical) quantum-dots, further improvements concerning the numerical stability should be pursued. Also, comparison with experimental data might be beneficial in future works.

Nonetheless, the usefulness of the maximum-entropy method, either as accompanying technique to deduce full probability distribution from given partial data or as stand-alone approach to derive the least biased steady-state density matrix was elucidated. The approach should hence be considered whenever knowledge of the full statistical features of the system seem advantageous.

Appendix

A

A.1 Upper Bound of the Second-Order Maximum-Entropy Distribution

A maximum-entropy distribution on the global range $\mathbb{N}_0$ does not exist for arbitrary moment constraints. But with condition (2.26) it is possible to derive an upper bound for the moments of the maximum-entropy distribution. For the second-order maximum-entropy method an upper bound for the second moment μ_2 can be calculated analytically. To do so, the Lagrange multipliers of a first-order maximum-entropy distribution are expressed in terms of the first moment. From the normalization condition (2.6) follows

$$\mu_0 = 1 = \sum_{n=0}^{\infty} \exp(-\lambda_0 - \lambda_1 n) \tag{A.1}$$

$$= e^{-\lambda_0} \sum_{n=0}^{\infty} \left(e^{-\lambda_1}\right)^n \tag{A.2}$$

$$= e^{-\lambda_0} \frac{1}{1 - e^{-\lambda_1}}. \tag{A.3}$$

In the last step, the properties of a geometric series were used. This leads to

$$\lambda_0 = -\ln\left(1 - e^{-\lambda_1}\right). \tag{A.4}$$

Then, the definition of the first moment

$$\mu_1 = \sum_{n=0}^{\infty} n \exp(-\lambda_0 - \lambda_1 n) \tag{A.5}$$

$$= -e^{-\lambda_0} \frac{\partial}{\partial \lambda_1} \sum_{n=0}^{\infty} \left(e^{-\lambda_1}\right)^n \tag{A.6}$$

$$= \frac{1}{e^{\lambda_1} - 1}, \tag{A.7}$$

where in the last step (A.4) was used, leads to an expression for the Lagrange multipliers in terms of the first moment

$$\lambda_1 = \ln\left(\frac{1 + \mu_1}{\mu_1}\right) \quad \text{and} \quad \lambda_0 = \ln(1 + \mu_1). \tag{A.8}$$

Because $\mu_1 > 0$ leads to a non-negative highest-order Lagrange multiplier λ_1, a first-order maximum-entropy distribution can always be found on the global range $\mathbb{N}_0$. Then, the maximum possible value of the second moment μ_2 for a second-order maximum-entropy distribution can be obtained by using Eq. (2.26)

$$\mu_2 \leq \sum_{n=0}^{\infty} n^2 \exp(-\lambda_0 - \lambda_1 n) \tag{A.9}$$

$$= e^{-\lambda_0} \frac{\partial^2}{\partial \lambda_1^2} \sum_{n=0}^{\infty} \left(e^{-\lambda_1}\right)^n \tag{A.10}$$

$$= \frac{e^{-\lambda_1}\left(e^{-\lambda_1}+1\right)}{\left(1-e^{\lambda_1}\right)^2}, \tag{A.11}$$

where again (A.4) was used. Inserting (A.8) into (A.11) finally leads to

$$\mu_2 \leq 2\mu_1^2 + \mu_1, \tag{A.12}$$

which gives an upper bound for the second-order moment. Note the subtle difference to the case of continuous maximum-entropy distributions, where the summation in Eq. (2.26) would have to be replaced by an integration (c.f. Einbu, 1977) leading to the more familiar Dowson-Wragg inequality (Dowson & Wragg, 1973)

$$\mu_2 \leq 2\mu_1^2. \tag{A.13}$$

A.2 Choice of Selfadjoint Operators as Input Information

A choice of selfadjoint operator combinations as input for a maximum-entropy method application in the sense of Sec. 2.3 is highly advisable, since it ensures selfadjointness of the maximum-entropy density matrix without need for complex-valued Lagrange multipliers. This can easily be verified for both cases, the direct input of operators and the demand for

stationarity of operators (cf. Sec. 2.3.3). Surely, since $\mathrm{Tr}\left(A^\dagger\right) = (\mathrm{Tr}\, A)^*$ and $\exp\left(A^\dagger\right) = (\exp A)^\dagger$,

$$Z^*(\lambda) = \left\{ \mathrm{Tr}\left[\exp\left(-\sum_{i=1}^{k} \lambda_i A_i \right) \right] \right\}^* \tag{A.14}$$

$$= \mathrm{Tr}\left[\exp\left(-\sum_{i=1}^{k} \lambda_i^* A_i^\dagger \right) \right] \tag{A.15}$$

$$= Z(\lambda), \tag{A.16}$$

if the Lagrange multipliers $\lambda_i \in \mathbb{R}$ are assumed to be real and operators $A_i^\dagger = A_i$ are chosen selfadjoint. Furthermore, for the maximum-entropy density matrix

$$\hat{\rho}^\dagger = \left[\frac{1}{Z(\lambda)} \exp\left(-\sum_{i=1}^{k} \lambda_i A_i \right) \right]^\dagger \tag{A.17}$$

$$= \frac{1}{Z^*(\lambda)} \exp\left(-\sum_{i=1}^{k} \lambda_i^* A_i^\dagger \right) \tag{A.18}$$

$$= \hat{\rho} \tag{A.19}$$

holds, which demonstrates that a choice of selfadjoint input operators A_i ensures self-adjointness of the maximum-entropy density matrix with real-valued Lagrange multipliers. Note that this does not simultaneously imply that all inputs need to be selfadjoint. Nevertheless, a non-selfadjoint choice will either lead to vanishing or complex-valued Lagrange multipliers to ensure that the maximum-entropy density matrix stays selfadjoint. In practice, inclusion of non-selfadjoint operators, e.g., $b^\dagger$ and b leads to either vanishing of both respective Lagrange multipliers or to a pair of complex-valued Lagrange multipliers whose imaginary parts cancel out each other.

The above prove directly translates to the demand of stationarity of operator combinations in the sense of Sec. 2.3.3 if one replaces operators A_i with the respective equations of motion $\mathcal{L}(A_i)$ in the exponential function. It remains to show that $\mathcal{L}^\dagger(A) = \mathcal{L}(A)$ holds.

$$\mathcal{L}^\dagger(A) = \left[\frac{\mathrm{i}}{\hbar}[H, A] + \sum_{\nu,\nu'} \vartheta_{\nu,\nu'} \left(2L_\nu^\dagger A L_{\nu'} - L_\nu^\dagger L_{\nu'} A - A L_\nu^\dagger L_{\nu'} \right) \right]^\dagger \tag{A.20}$$

$$= \frac{\mathrm{i}}{\hbar}\left[H, A^\dagger\right] + \sum_{\nu,\nu'} \vartheta_{\nu,\nu'}^* \left(2L_{\nu'}^\dagger A^\dagger L_\nu - L_{\nu'}^\dagger L_\nu A^\dagger - A^\dagger L_{\nu'}^\dagger L_\nu \right) \tag{A.21}$$

$$= \frac{\mathrm{i}}{\hbar}\left[H, A^\dagger\right] + \sum_{\nu,\nu'} \vartheta_{\nu',\nu}^* \left(2L_\nu^\dagger A^\dagger L_{\nu'} - L_\nu^\dagger L_{\nu'} A^\dagger - A^\dagger L_\nu^\dagger L_{\nu'} \right) \tag{A.22}$$

where in the first step, self-adjointness of the Hamiltonian $H^\dagger = H$ and $(ABC)^\dagger = C^\dagger B^\dagger A^\dagger$ were used (note the slightly different indices). In the second step the indices were renamed: $\nu \to \nu'$ and $\nu' \to \nu$. The last line equals $\mathcal{L}(A)$ if A is selfadjoint and if

$$\vartheta_{\nu,\nu'} = \vartheta^*_{\nu',\nu} \tag{A.23}$$

holds. The quantity $\vartheta = (\vartheta_{\nu,\nu'})$ is referred to as damping matrix and contains the rates of the dissipative processes. In most cases, the damping matrix contains real values only and is chosen diagonal, hence the condition (A.23) is satisfied (cf. Fanaei *et al.,* 2016; Fanaei, 2019 and references therein). The effect of a non-diagonal choice on the competing behavior of the two modes of a quantum dot microcavity laser has been studied in (Fanaei *et al.,* 2016; Fanaei, 2019), where, nevertheless, condition (A.23) is satisfied since the damping matrix is chosen symmetrically. Consequently, in most cases (including all cases in this thesis), a selfadjoint choice of operators A_i ensures that $\mathcal{L}^\dagger(A_i) = \mathcal{L}(A_i)$ and consequently that the maximum-entropy density matrix is selfadjoint with real-valued Lagrange multipliers. But again, this does not automatically imply that only selfadjoint A_i are allowed. For a non-selfadjoint choice however, it is expected that the respective Lagrange multipliers either vanish or exhibit complex values such that, in combination with all other inputs, the imaginary parts are canceled out. It would be an interesting topic though, to study systems where condition (A.23) is not satisfied since there, stationarity of non-selfadjoint operators is expected to be a necessary input for a reasonable maximum-entropy prediction.

A.3 Choice of the Maximum Photon Number for Numerical Calculations

As shown in Sec. 2.2.3, existence of a maximum-entropy distribution on the finite range $\{0, 1, \ldots, n_{\max}\}$ does not guarantee existence of a maximum-entropy distribution on the global range $\mathbb{N}_0 = \{0, 1, \ldots, \infty\}$. Nevertheless, with a suitable choice of $n_{\max}$, it is possible to artificially extend the maximum-entropy distribution from the finite to the global range. This can be done by setting the probability values p_n of the probability distribution to zero for $n > n_{\max}$, i.e., by cutting off the inaccurate tail of the maximum-entropy distribution that exhibits large probability values for $n \approx n_{\max}$ but is expected to decline rapidly for the probability distribution on the infinite range. This procedure enables to predict usable maximum-entropy distributions also in the case of negative last Lagrange multiplier (cf. Sec. 3.1.3). The crucial part is a suitable choice of $n_{\max}$, for which one possibility is presented here.

The general idea is to start with a first rough estimation $\tilde{n}_{\max}$ and to calculate the Lagrange multipliers $\tilde{\lambda}_i$ and corresponding probability distribution $\tilde{p}_n$ on the implementation space $\{0, 1, \ldots, \tilde{n}_{\max}\}$. Then, the Lagrange multipliers $\tilde{\lambda}_i$ can be used to calculate the corresponding probability distribution on the global range $\mathbb{N}_0$. It will exhibit a tail that grows to infinity for large n [cf. Fig. 3.4 (b) inset]. Consequently, a local minimum can be found in between the real maximum (peak of the thermal or poisson-like distribution) for small values of n and the numerical peak for large n. Taking $n_{\max}$ close to this minimum is usually a good choice that lies in the $n_{\max}$-independent range [cf. Fig. 3.4 (a)] and results in a probability distribution with decreasing probability values for n approaching the edge of the implementation space.

For the practical implementation however, more details need to be clarified. First, an initial guess for the maximum photon number can be estimated from the given moment values. For the problems at hand, a decent choice is given by

$$\tilde{n}_{\max} \approx \langle n \rangle + 3\sqrt{\langle n^2 \rangle}. \tag{A.24}$$

This may seem somewhat arbitrary but in most cases suffices to include all relevant parts of the distribution. Note that also other choices are possible and may even be necessary for different problems with differently shaped probability distributions. Furthermore, one should ensure that $\tilde{n}_{\max}$ is not to small and for instance require that at least $\tilde{n}_{\max} \geq 10$.

The second problem is, as with the corresponding Lagrange multipliers $\tilde{\lambda}_i$ one wishes to determine $\tilde{p}_n$ on the global range, it is again numerically necessary to limit the global range to a finite one $\{0, 1, \ldots, n_{\inf}\}$. This second truncation however, is only needed temporarily and omitted afterwards and should not be confused with the search for $n_{\max}$. Nevertheless, several points need to be considered for an adequate choice of $n_{\inf}$. On one hand, it should be chosen large enough such that the tail of the distribution that exhibits the large probabilities for large n is considered. On the other hand, if $n_{\inf}$ is chosen to large, this leads to values of the exponential function [cf. Eq. (2.15)] that are ∞ because they exceed the range of numerically representable values. This further leads to NaN (not a number) entries in $\tilde{p}_n$ that are caused by the division of ∞ (exponential function) by ∞ (partition sum) in Eq. (2.15). One strategy to avoid this behavior, is to start with a rather large $n_{\inf}$ (for instance $n_{\inf} = 10000$) and to successively reduce it until no NaN entries occur in $\tilde{p}_n$. If finally $\tilde{p}_n$ is found on $\{0, 1, \ldots, n_{\inf}\}$, the local minimum between both peaks, the peak of the thermal or poisson-like maximum and the artificial peak for large n, can be determined numerically. This minimum gives the estimation for $n_{\max}$ and the final maximum-entropy distribution can be found with a second search for optimal Lagrange multipliers λ_i. All probabilities for $n > n_{\max}$ are then simply assumed to be zero. Although this procedure might seem costly, it gives a possibility to also calculate maximum-entropy

method predictions for odd orders (cf. Fig. 3.3) with the only time-consuming task being the additional second search for the Lagrange multipliers.

A.4 Entropy of a Thermal Distribution

The entropy of a thermal distribution

$$p_n = \frac{1}{Z(\lambda)} \exp(-\lambda n) \quad \text{with} \quad Z(\lambda) = \sum_{n=0}^{\infty} \exp(-\lambda n) \tag{A.25}$$

can be expressed in terms of the mean photon number $\langle n \rangle$ by using Eq. (2.13)

$$S = \lambda \langle n \rangle + \ln Z(\lambda). \tag{A.26}$$

First, rewriting

$$Z(\lambda) = \sum_{n=0}^{\infty} \left(e^{-\lambda}\right)^n = \frac{1}{1 - e^{-\lambda}} \tag{A.27}$$

leads to

$$\ln Z(\lambda) = -\ln[1 - \exp(-\lambda)]. \tag{A.28}$$

And second, since

$$\langle n \rangle = -\frac{\partial}{\partial \lambda} \ln Z(\lambda) \tag{A.29}$$

the Lagrange multiplier λ can be expressed in terms of the mean photon number $\langle n \rangle$

$$\lambda = \ln\left(\frac{\langle n \rangle + 1}{\langle n \rangle}\right). \tag{A.30}$$

Inserting (A.30) into (A.28) then results in

$$\ln Z(\lambda) = \ln(\langle n \rangle + 1). \tag{A.31}$$

Finally, Eq. (A.26) can be evaluated by inserting (A.30) and (A.31) leading to an expression for the entropy of a thermal distribution in terms of the mean photon number

$$S = -\langle n \rangle \ln \langle n \rangle + (\langle n \rangle + 1) \ln(\langle n \rangle + 1). \tag{A.32}$$

Bibliography

Abramov, R. V. A practical computational framework for the multidimensional moment-constrained maximum entropy principle. *J. Comput. Phys.* **211,** 198–209 (2006) (cit. on p. 23).

Abramov, R. V. An improved algorithm for the multidimensional moment-constrained maximum entropy problem. *J. Comput. Phys.* **226,** 621–644 (2007) (cit. on p. 23).

Abramov, R. V. The multidimensional moment-constrained maximum entropy problem: A BFGS algorithm with constraint scaling. *J. Comput. Phys.* **228,** 96–108 (2009) (cit. on p. 23).

Abramov, R. V. The multidimensional maximum entropy moment problem: a review of numerical methods. *Commun. Math. Sci.* **8,** 377–392 (2010) (cit. on pp. 23–25).

Alexeev, E. M., Ruiz-Tijerina, D. A., Danovich, M., Hamer, M. J., Terry, D. J., Nayak, P. K., Ahn, S., Pak, S., *et al.* Resonantly hybridized excitons in moiré superlattices in van der Waals heterostructures. *Nature* **567,** 81–86 (2019) (cit. on pp. 2, 67).

Aspelmeyer, M., Kippenberg, T. J. & Marquardt, F. Cavity optomechanics. *Rev. Mod. Phys.* **86,** 1391–1452 (2014) (cit. on p. 76).

Auffèves, A., Gerace, D., Gérard, J.-M., Santos, M. F., Andreani, L. C. & Poizat, J.-P. Controlling the dynamics of a coupled atom-cavity system by pure dephasing. *Phys. Rev. B* **81,** 245419 (2010) (cit. on p. 37).

Baer, N., Gies, C., Wiersig, J. & Jahnke, F. Luminescence of a semiconductor quantum dot system. *Eur. Phys. Jour. B* **50,** 411–418 (2006) (cit. on pp. 3, 31).

Bandyopadhyay, K., Bhattacharya, A. K., Biswas, P. & Drabold, D. A. Maximum entropy and the problem of moments: A stable algorithm. *Phys. Rev. E* **71,** 057701 (2005) (cit. on p. 23).

Bandyopadhyay, K., Bhattacharyya, K. & Bhattacharya, A. K. Maximum-entropy principle with moment recursion relations as constraints. *Phys. Rev. A* **63,** 064101 (2001) (cit. on p. 23).

Batou, A. & Soize, C. Calculation of Lagrange Multipliers in the Construction of Maximum Entropy Distributions in High Stochastic Dimension. *SIAM/ASA J. Uncertain. Quantif.* **1,** 431–451 (2013) (cit. on p. 23).

Bennink, R. S., Bentley, S. J. & Boyd, R. W. "Two-Photon" Coincidence Imaging with a Classical Source. *Phys. Rev. Lett.* **89,** 113601 (2002) (cit. on p. 52).

Bhimanapati, G. R., Lin, Z., Meunier, V., Jung, Y., Cha, J., Das, S., Xiao, D., Son, Y., *et al.* Recent Advances in Two-Dimensional Materials beyond Graphene. *ACS Nano* **9,** 11509–11539 (2015) (cit. on pp. 2, 66, 67).

Boyd, S. & Vandenberghe, L. *Convex Optimization* (Cambridge University Press, 2004) (cit. on p. 23).

Brillouin, L. Maxwell's Demon Cannot Operate: Information and Entropy. I. *J. Appl. Phys.* **22,** 334–337 (1951) (cit. on pp. 1, 8).

Brodie, B. C. XIII. On the atomic weight of graphite. *Philos. Trans. R. Soc. London* **149,** 249–259 (1859) (cit. on p. 66).

Brukner, C. & Zeilinger, A. Conceptual inadequacy of the Shannon information in quantum measurements. *Phys. Rev. A* **63,** 022113 (2001) (cit. on p. 11).

Butler, S. Z., Hollen, S. M., Cao, L., Cui, Y., Gupta, J. A., Gutiérrez, H. R., Heinz, T. F., Hong, S. S., *et al.* Progress, Challenges, and Opportunities in Two-Dimensional Materials Beyond Graphene. *ACS Nano* **7,** 2898–2926 (2013) (cit. on pp. 2, 67).

Carmesin, C., Lorke, M., Florian, M., Erben, D., Schulz, A., Wehling, T. O. & Jahnke, F. Quantum-Dot-Like States in Molybdenum Disulfide Nanostructures Due to the Interplay of Local Surface Wrinkling, Strain, and Dielectric Confinement. *Nano Lett.* **19,** 3182–3186 (2019) (cit. on p. 67).

Carmichael, H. J. in *Statistical Methods in Quantum Optics 1: Master Equations and Fokker-Planck Equations* 1–28 (Springer, 1999) (cit. on p. 30).

Carnot, S. *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance* (Bachelier Libraire Paris, 1824) (cit. on pp. 1, 5, 6).

Carnot, S. in *Reflections on the Motive Power of Fire: And Other Papers on the Second Law of Thermodynamics* (ed Mendoza, E.) (Dover Publications, 2012) (cit. on pp. 1, 5, 6).

Chanda, P., Costa, E., Hu, J., Sukumar, S., Van Hemert, J. & Walia, R. Information Theory in Computational Biology: Where We Stand Today. *Entropy* **22,** 627 (2020) (cit. on pp. 1, 14).

Chehade, S. S. & Vershynina, A. Quantum entropies. *Scholarpedia* **14,** 53131 (2019) (cit. on pp. 1, 7).

Chen, W., Sun, Z., Wang, Z., Gu, L., Xu, X., Wu, S. & Gao, C. Direct observation of van der Waals stacking–dependent interlayer magnetism. *Science* **366,** 983–987 (2019) (cit. on pp. 2, 67).

Chow, W., Jahnke, F. & Gies, C. Emission properties of nanolasers during the transition to lasing. *Light Sci. Appl.* **3,** e201 (2014) (cit. on pp. 46, 49, 52, 62).

Cisco. Cisco Visual Networking Index: Forecast and Trends, 2017–2022 (2019) (cit. on p. 1).

Clausius, R. Ueber die bewegende Kraft der Wärme und die Gesetze, welche sich daraus für die Wärmelehre selbst ableiten lassen. *Ann. Phys.* **155,** 368–397 (1850a) (cit. on pp. 1, 6).

Clausius, R. Ueber die bewegende Kraft der Wärme und die Gesetze, welche sich daraus für die Wärmelehre selbst ableiten lassen. *Ann. Phys.* **155,** 500–524 (1850b) (cit. on pp. 1, 6).

Clausius, R. Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie. *Ann. Phys.* **201,** 353–400 (1865) (cit. on pp. 1, 6).

Clausius, R. *Über den zweiten Hauptsatz der mechanischen Wärmetheorie* (Friedrich Vieweg und Sohn, 1867) (cit. on p. 6).

Cover, T. M. & Thomas, J. A. *Elements of Information Theory* Second Edition (Wiley, 2006) (cit. on pp. 17, 18, 42).

Csiszar, I. Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inference for Linear Inverse Problems. *Ann. Statist.* **19,** 2032–2066 (1991) (cit. on p. 11).

De Vega, I. & Alonso, D. Dynamics of non-Markovian open quantum systems. *Rev. Mod. Phys.* **89,** 015001 (2017) (cit. on pp. 30, 93).

Dicke, R. H. Coherence in Spontaneous Radiation Processes. *Phys. Rev.* **93,** 99–110 (1954) (cit. on p. 52).

Dovzhenko, D. S., Ryabchuk, S. V., Rakovich, Y. P. & Nabiev, I. R. Light–matter interaction in the strong coupling regime: configurations, conditions, and applications. *Nanoscale* **10,** 3589–3605 (2018) (cit. on p. 2).

Downarowicz, T. Entropy. *Scholarpedia* **2,** 3901 (2007) (cit. on pp. 1, 7).

Dowson, D. & Wragg, A. Maximum-entropy distributions having prescribed first and second moments (Corresp.) *IEEE Trans. Inf. Theory* **19,** 689–693 (1973) (cit. on pp. 22, 98).

Dubin, F., Russo, C., Barros, H. G., Stute, A., Becher, C., Schmidt, P. O. & Blatt, R. Quantum to classical transition in a single-ion laser. *Nat. Phys.* **6,** 350–353 (2010) (cit. on p. 37).

Ehrenberg, W. Maxwell's Demon. *Sci. Am.* **217,** 103–111 (1967) (cit. on pp. 1, 6).

Einbu, J. On the existence of a class of maximum-entropy probability density functions (Corresp.) *IEEE Trans. Inf. Theory* **23,** 772–775 (1977) (cit. on pp. 22, 98).

Erben, D., Steinhoff, A., Gies, C., Schönhoff, G., Wehling, T. O. & Jahnke, F. Excitation-induced transition to indirect band gaps in atomically thin transition-metal dichalcogenide semiconductors. *Phys. Rev. B* **98,** 035434 (2018) (cit. on pp. 2, 67).

Evans, R. J., Boersma, J., Blachman, N. M. & Jagers, A. A. The Entropy of a Poisson Distribution: Problem 87-6. *SIAM Rev.* **30,** 314–317 (1988) (cit. on p. 49).

Fanaei, M. *Effect of the second mode on the optical properties of quantum-dot microcavity lasers* PhD thesis (Otto-von-Guericke-Universität Magdeburg, 2019) (cit. on pp. 3, 100).

Fanaei, M., Foerster, A., Leymann, H. A. M. & Wiersig, J. Effect of direct dissipative coupling of two competing modes on intensity fluctuations in a quantum-dot-microcavity laser. *Phys. Rev. A* **94,** 043814 (2016) (cit. on pp. 3, 32, 49, 52, 56, 62, 87, 100).

Feldtmann, T., Schneebeli, L., Kira, M. & Koch, S. W. Quantum theory of light emission from a semiconductor quantum dot. *Phys. Rev. B* **73,** 155319 (2006) (cit. on pp. 3, 31).

Fick, E. & Sauermann, G. *The Quantum Statistics of Dynamic Processes* (Springer, 1990) (cit. on pp. 26–31).

Florian, M., Gies, C., Gartner, P. & Jahnke, F. Improved antibunching by using high-excitation pulses from a single semiconductor quantum dot—a theoretical study. *J. Opt. Soc. Am. B* **29,** A31–A35 (2012) (cit. on pp. 3, 62).

Florian, M., Gies, C., Jahnke, F., Leymann, H. A. M. & Wiersig, J. Equation-of-motion technique for finite-size quantum-dot systems: Cluster expansion method. *Phys. Rev. B* **87,** 165306 (2013) (cit. on pp. 3, 31, 62).

Foerster, A. *Theory of Semiconductor Quantum-Dot Microcavity Lasers* PhD thesis (Otto-von-Guericke-Universität Magdeburg, 2017) (cit. on pp. 3, 30, 31).

Foerster, A., Leymann, H. A. M. & Wiersig, J. Computer-aided cluster expansion: An efficient algebraic approach for open quantum many-particle systems. *Comput. Phys. Commun.* **212,** 210–219 (2017) (cit. on p. 49).

Fricke, J. Transport Equations Including Many-Particle Correlations for an Arbitrary Quantum System: A General Formalism. *Ann. Phys.* **252,** 479 –498 (1996) (cit. on pp. 3, 31).

Frisk Kockum, A., Miranowicz, A., De Liberato, S., Savasta, S. & Nori, F. Ultrastrong coupling between light and matter. *Nat. Rev. Phys.* **1,** 19–40 (2019) (cit. on pp. 2, 92).

Frontini, M. & Tagliani, A. Maximum entropy in the finite Stieltjes and Hamburger moment problem. *J. Math. Phys.* **35,** 6748–6756 (1994) (cit. on pp. 20, 21).

Frontini, M. & Tagliani, A. Entropy-convergence in Stieltjes and Hamburger moment problem. *Appl. Math. Comput.* **88,** 39–51 (1997) (cit. on p. 22).

Garber, E. W. Clausius and Maxwell's Kinetic Theory of Gases. *Chymia* **2,** 299–319 (1970) (cit. on pp. 1, 6).

Gartner, P. Spontaneous symmetry breaking in the laser transition. *Phys. Rev. B* **99,** 115313 (2019) (cit. on pp. 37, 83).

Gartner, P. & Halati, C. M. Laser transition in the thermodynamic limit for identical emitters in a cavity. *Phys. Rev. A* **93,** 013817 (2016) (cit. on pp. 38, 67).

Gatti, A., Brambilla, E., Bache, M. & Lugiato, L. A. Ghost Imaging with Thermal Light: Comparing Entanglement and Classical Correlation. *Phys. Rev. Lett.* **93,** 093602 (2004) (cit. on p. 52).

Geim, A. K. & Grigorieva, I. V. Van der Waals heterostructures. *Nature* **499,** 419–425 (2013) (cit. on pp. 2, 67).

Gibbs, J. W. *Elementary principles in statistical mechanics* (Scribner's sons, 1902) (cit. on pp. 1, 7).

Gies, C., Florian, M., Steinhoff, A. & Jahnke, F. in *Quantum Dots for Quantum Information Technologies* 3–40 (Springer, 2017) (cit. on pp. 49, 51, 78).

Gies, C., Wiersig, J. & Jahnke, F. in *Single Semiconductor Quantum Dots* 1–31 (2009) (cit. on pp. 39, 49).

Gies, C., Wiersig, J., Lorke, M. & Jahnke, F. Semiconductor model for quantum-dot-based microcavity lasers. *Phys. Rev. A* **75,** 013803 (2007) (cit. on pp. 3, 31, 32, 37, 46, 49, 61–63, 83).

Global e-Sustainability Initiative. *GeSI SMARTer 2020: The Role of ICT in Driving a Sustainable Future* (2012) (cit. on p. 2).

Global e-Sustainability Initiative. *SMARTer2030 – ICT Solutions for 21st Century Challenges* (2015) (cit. on p. 2).

Goldmann, E. *From Structure to Spectra: Tight-Binding Theory of InGaAs Quantum Dots* PhD thesis (Universität Bremen, 2014) (cit. on pp. 2, 92).

Griffin, A. *Maximum Entropy: The Universal Method for Inference* PhD thesis (State University of New York, 2008) (cit. on pp. 1, 7–9).

Gross, M. & Haroche, S. Superradiance: An essay on the theory of collective spontaneous emission. *Phys. Rep.* **93,** 301–396 (1982) (cit. on p. 52).

Grothe, I. H. *Theoretisches Modell der Quantenspektroskopie an Einzelatomen* Bachelor's thesis (Universität Bremen, 2014) (cit. on pp. 76, 91).

Grothe, I. H. *Effects of Radiative Emitter Coupling in Coupled Cavity Arrays* Master's thesis (Universität Bremen, 2018) (cit. on pp. 31, 32).

Gulyak, B. *Maximum-Entropie-Methoden in der Vielteilchenphysik* Diploma thesis (Otto-von-Guericke-Universität Magdeburg, 2016) (cit. on pp. 17, 18, 26–29, 32, 33, 39, 42).

Gulyak, B., Melcher, B. & Wiersig, J. Determination of the full statistics of quantum observables using the maximum-entropy method. *Phys. Rev. A* **98,** 053857 (2018) (cit. on pp. 14, 23, 38, 42, 45, 50–52, 55).

Gzyl, H. *The Method of Maximum Entropy* (World Scientific, 1995) (cit. on p. 17).

Hanbury Brown, R. & Twiss, R. Q. A Test of a New Type of Stellar Interferometer on Sirius. *Nature* **178,** 1046–1048 (1956) (cit. on p. 52).

Hao, W. & Harlim, J. An equation-by-equation method for solving the multidimensional moment constrained maximum entropy problem. *Comm. App. Math. Comp. Sci.* **13,** 189–214 (2018) (cit. on p. 23).

Harder, G., Mogilevtsev, D., Korolkova, N. & Silberhorn, C. Tomography by Noise. *Phys. Rev. Lett.* **113,** 070403 (2014) (cit. on p. 52).

He, J. & Kolovos, A. Bayesian maximum entropy approach and its applications: a review. *Stoch. Environ. Res. Risk Assess.* **32,** 859–877 (2018) (cit. on pp. 1, 14).

He, L., Özdemir, c. K. & Yang, L. Whispering gallery microcavity lasers. *Laser Photonics Rev.* **7,** 60–82 (2013) (cit. on p. 2).

Hiai, F., Ohya, M. & Tsukada, M. Sufficiency, KMS condition and relative entropy in von Neumann algebras. en. *Pacific J. Math.* **96,** 99–109 (1981) (cit. on p. 82).

Hopfmann, C., Albert, F., Schneider, C., Höfling, S., Kamp, M., Forchel, A., Kanter, I. & Reitzenstein, S. Nonlinear emission characteristics of quantum dot–micropillar lasers in the presence of polarized optical feedback. *New J. Phys.* **15,** 025030 (2013) (cit. on p. 56).

Hoyer, W., Kira, M. & Koch, S. W. Influence of Coulomb and phonon interaction on the exciton formation dynamics in semiconductor heterostructures. *Phys. Rev. B* **67,** 155113 (2003) (cit. on pp. 3, 31).

Huo, Y. H., Křápek, V., Rastelli, A. & Schmidt, O. G. Volume dependence of excitonic fine structure splitting in geometrically similar quantum dots. *Phys. Rev. B* **90,** 041304 (2014a) (cit. on p. 2).

Huo, Y. H., Rastelli, A. & Schmidt, O. G. Ultra-small excitonic fine structure splitting in highly symmetric quantum dots on GaAs (001) substrate. *Appl. Phys. Lett.* **102** (2013) (cit. on p. 2).

Huo, Y. H., Witek, B. J., Kumar, S., Cardenas, J. R., Zhang, J. X., Akopian, N., Singh, R., Zallo, E., *et al.* A light-hole exciton in a quantum dot. *Nat. Phys.* **10,** 46–51 (2014b) (cit. on p. 2).

Jagsch, S. T., Triviño, N. V., Lohof, F., Callsen, G., Kalinowski, S., Rousseau, I. M., Barzel, R., Carlin, J.-F., *et al.* A quantum optical study of thresholdless lasing features in high-β nitride nanobeam cavities. *Nat. Comm.* **9,** 564 (2018) (cit. on p. 2).

Jahnke, F., Gies, C., Aßmann, M., Bayer, M., Leymann, H. A. M., Foerster, A., Wiersig, J., Schneider, C., *et al.* Giant photon bunching, superradiant pulse emission and excitation trapping in quantum-dot nanolasers. *Nat. Comm.* **7,** 11540 (2016) (cit. on pp. 3, 32, 49, 52, 62, 76, 92).

Jariwala, D., Sangwan, V. K., Lauhon, L. J., Marks, T. J. & Hersam, M. C. Emerging Device Applications for Semiconducting Two-Dimensional Transition Metal Dichalcogenides. *ACS Nano* **8,** 1102–1120 (2014) (cit. on p. 2).

Jaynes, E. T. Information Theory and Statistical Mechanics. *Phys. Rev.* **106,** 620–630 (1957) (cit. on pp. 1, 8, 9, 19, 26, 28, 29, 33).

Jaynes, E. T. Information Theory and Statistical Mechanics. II. *Phys. Rev.* **108,** 171–190 (1957) (cit. on pp. 1, 8, 9, 19, 26, 28, 29, 33).

Jaynes, E. T. Gibbs vs Boltzmann Entropies. *Am. J. Phys.* **33,** 391–398 (1965) (cit. on pp. 1, 7, 9).

Jaynes, E. T. The minimum entropy production principle. *Ann. Rev. Phys. Chem.* (1980) (cit. on p. 9).

Jaynes, E. T. *Probability Theory* (Cambridge University Press, 2003) (cit. on pp. 1, 7–9, 12, 13, 17–19, 29, 33).

Jiang, X., Shao, L., Zhang, S.-X., Yi, X., Wiersig, J., Wang, L., Gong, Q., Lončar, M., *et al.* Chaos-assisted broadband momentum transformation in optical microresonators. *Science* **358,** 344 (2017) (cit. on p. 3).

Jizba, P. & Korbel, J. Maximum Entropy Principle in Statistical Inference: Case for Non-Shannonian Entropies. *Phys. Rev. Lett.* **122,** 120601 (2019) (cit. on pp. 1, 11–16).

Jizba, P. & Korbel, J. When Shannon and Khinchin meet Shore and Johnson: Equivalence of information theory and statistical inference axiomatics. *Phys. Rev. E* **101,** 042126 (2020) (cit. on pp. 1, 12–17).

Jones, B., Ghose, S., Clemens, J. P., Rice, P. R. & Pedrotti, L. M. Photon statistics of a single atom laser. *Phys. Rev. A* **60,** 3267–3275 (1999) (cit. on p. 49).

Kabuss, J., Carmele, A., Brandes, T. & Knorr, A. Optically Driven Quantum Dots as Source of Coherent Cavity Phonons: A Proposal for a Phonon Laser Scheme. *Phys. Rev. Lett.* **109,** 054301 (2012) (cit. on pp. 3, 31).

Kapetanakis, M. D. & Perakis, I. E. Spin Dynamics in (III,Mn)V Ferromagnetic Semiconductors: The Role of Correlations. *Phys. Rev. Lett.* **101,** 097201 (2008) (cit. on pp. 3, 31).

Kazimierczuk, T., Schmutzler, J., Aßmann, M., Schneider, C., Kamp, M., Höfling, S. & Bayer, M. Photon-Statistics Excitation Spectroscopy of a Quantum-Dot Micropillar Laser. *Phys. Rev. Lett.* **115,** 027401 (2015) (cit. on pp. 3, 49, 91).

Khajavikhan, M., Simic, A., Katz, M., Lee, J. H., Slutsky, B., Mizrahi, A., Lomakin, V. & Fainman, Y. Thresholdless nanoscale coaxial lasers. *Nature* **482,** 204–207 (2012) (cit. on pp. 2, 37, 46).

Khanbekyan, M. Three-dimensional cavity-assisted spontaneous emission as a single-photon source: Two cavity modes and Rabi resonance. *Phys. Rev. A* **97,** 023809 (2018) (cit. on p. 87).

Khanbekyan, M., Leymann, H. A. M., Hopfmann, C., Foerster, A., Schneider, C., Höfling, S., Kamp, M., Wiersig, J., *et al.* Unconventional collective normal-mode coupling in quantum-dot-based bimodal microlasers. *Phys. Rev. A* **91,** 043840 (2015) (cit. on pp. 3, 32, 87, 92).

Kira, M., Jahnke, F. & Koch, S. W. Microscopic Theory of Excitonic Signatures in Semiconductor Photoluminescence. *Phys. Rev. Lett.* **81,** 3263–3266 (1998) (cit. on pp. 3, 31).

Köhler, T. & Burnett, K. Microscopic quantum dynamics approach to the dilute condensed Bose gas. *Phys. Rev. A* **65,** 033601 (2002) (cit. on pp. 3, 31).

Kosterlitz, J. M. & Thouless, D. J. Ordering, metastability and phase transitions in two-dimensional systems. *J. Phys. C* **6,** 1181–1203 (1973) (cit. on p. 66).

Kozlov, V. G., Bulović, V., Burrows, P. E. & Forrest, S. R. Laser action in organic semiconductor waveguide and double-heterostructure devices. *Nature* **389,** 362–364 (1997) (cit. on p. 46).

Kreinberg, S., Chow, W. W., Wolters, J., Schneider, C., Gies, C., Jahnke, F., Höfling, S., Kamp, M., *et al.* Emission from quantum-dot high-β microcavities: transition from spontaneous emission to lasing and the effects of superradiant emitter coupling. *Light Sci. Appl.* **6,** e17030–e17030 (2017) (cit. on pp. 3, 62).

Lazo, A. V. & Rathie, P. On the entropy of continuous probability distributions (Corresp.) *IEEE Trans. Inf. Theory* **24,** 120–122 (1978) (cit. on p. 49).

Lettau, T. *Pumpleistungsgetriebener Modenwechsel in einem bimodalen Mikrolaser und Fallstricke in der Beschreibung der Ladungsträger in Halbleiterquantenpunkten* Diploma thesis (Otto-von-Guericke-Universität Magdeburg, 2017) (cit. on pp. 40, 92).

Lettau, T. & Leymann, H. A. M. Differences and similarities between lasing and multiple-photon subtracted states. *Phys. Rev. A* **99,** 023815 (2019) (cit. on pp. 37, 49).

Lettau, T., Leymann, H. A. M., Melcher, B. & Wiersig, J. Superthermal photon bunching in terms of simple probability distributions. *Phys. Rev. A* **97,** 053835 (2018) (cit. on pp. 14, 22, 52, 57).

Lettau, T., Leymann, H. A. M. & Wiersig, J. Pitfalls in the theory of carrier dynamics in semiconductor quantum dots: Single-particle basis versus the many-particle configuration basis. *Phys. Rev. B* **95,** 085314 (2017) (cit. on p. 92).

Leymann, H. A. M. *Theory of many-particle correlations and optical properties of semiconductor quantum dots* PhD thesis (Otto-von-Guericke-Universität Magdeburg, 2015) (cit. on pp. 3, 30, 31).

Leymann, H. A. M., Foerster, A., Jahnke, F., Wiersig, J. & Gies, C. Sub- and Superradiance in Nanolasers. *Phys. Rev. Appl.* **4,** 044018 (2015) (cit. on pp. 3, 32, 49, 52, 62, 92).

Leymann, H. A. M., Foerster, A. & Wiersig, J. Expectation value based cluster expansion. *Phys. Status Solidi C* **10,** 1242–1245 (2013a) (cit. on p. 52).

Leymann, H. A. M., Foerster, A. & Wiersig, J. Expectation value based equation-of-motion approach for open quantum systems: A general formalism. *Phys. Rev. B* **89,** 085308 (2014) (cit. on pp. 3, 31, 32, 49, 52, 63).

Leymann, H. A. M., Hopfmann, C., Albert, F., Foerster, A., Khanbekyan, M., Schneider, C., Höfling, S., Forchel, A., *et al.* Intensity fluctuations in bimodal micropillar lasers enhanced by quantum-dot gain competition. *Phys. Rev. A* **87,** 053819 (2013b) (cit. on pp. 3, 32, 37, 46, 49, 52, 56, 57, 62, 87).

Lieb, E. H. & Yngvason, J. The physics and mathematics of the second law of thermodynamics. *Phys. Rep.* **310,** 1 –96 (1999) (cit. on pp. 1, 6).

Lima, J. A. S. & Deppman, A. Tsallis meets Boltzmann: q-index for a finite ideal gas and its thermodynamic limit. *Phys. Rev. E* **101,** 040102 (2020) (cit. on p. 11).

Livesey, A. K. & Skilling, J. Maximum entropy theory. *Acta Crystallogr. A* **41,** 113–122 (1985) (cit. on p. 49).

Lohof, F., Barzel, R., Gartner, P. & Gies, C. Delayed Transition to Coherent Emission in Nanolasers with Extended Gain Media. *Phys. Rev. Appl.* **10,** 054055 (2018) (cit. on pp. 2, 41, 46, 49, 61, 67–71, 73).

Lohof, F., Steinhoff, A., Florian, M., Lorke, M., Erben, D., Jahnke, F. & Gies, C. Prospects and Limitations of Transition Metal Dichalcogenide Laser Gain Materials. *Nano Lett.* **19,** 210–217 (2019) (cit. on pp. 2, 3, 66, 67).

Loudon, R. *The Quantum Theory of Light* Third Edition (Oxford Science Publications, 2000) (cit. on pp. 46, 52).

Ma, R.-M. & Oulton, R. F. Applications of nanolasers. *Nat. Nanotechnol.* **14,** 12–22 (2019) (cit. on pp. 1, 2, 67).

Magazzù, L., Forn-Díaz, P., Belyansky, R., Orgiazzi, J.-L., Yurtalan, M. A., Otto, M. R., Lupascu, A., Wilson, C. M., *et al.* Probing the strongly driven spin-boson model in a superconducting quantum circuit. *Nat. Comm.* **9,** 1403 (2018) (cit. on p. 76).

Mak, K. F., Lee, C., Hone, J., Shan, J. & Heinz, T. F. Atomically Thin MoS_2: A New Direct-Gap Semiconductor. *Phys. Rev. Lett.* **105,** 136805 (2010) (cit. on pp. 2, 67).

Malmodin, J. & Lundén, D. The Energy and Carbon Footprint of the Global ICT and E&M Sectors 2010–2015. *Sustainability* **10** (2018) (cit. on p. 1).

Marconi, M., Javaloyes, J., Hamel, P., Raineri, F., Levenson, A. & Yacomotti, A. M. Far-from-Equilibrium Route to Superthermal Light in Bimodal Nanolasers. *Phys. Rev. X* **8,** 011013 (2018) (cit. on pp. 52, 87).

Marconi, M., Javaloyes, J., Raineri, F., Levenson, J. A. & Yacomotti, A. M. Asymmetric mode scattering in strongly coupled photonic crystal nanolasers. *Opt. Lett.* **41,** 5628–5631 (2016) (cit. on p. 52).

Martínez, J. M. Practical quasi-Newton methods for solving nonlinear systems. *J. Comput. Appl. Math.* **124,** 97 –121 (2000) (cit. on p. 23).

Meadow, C. T. *Making Connections – Communication through the Ages* (Scarecrow Press, Inc., 2002) (cit. on p. 1).

Melcher, B. *Tight-Binding-Beschreibung von GaAs/AlGaAs Quantenpunkten* Master's thesis (Universtität Bremen, 2016) (cit. on pp. 2, 92).

Melcher, B., Gulyak, B. & Wiersig, J. Information-theoretical approach to the many-particle hierarchy problem. *Phys. Rev. A* **100,** 013854 (2019) (cit. on pp. 26, 75, 77, 82–85).

Milev, M. & Tagliani, A. Entropy convergence of finite moment approximations in Hamburger and Stieltjes problems. *Stat. Probab. Lett.* **120,** 114 –117 (2017) (cit. on p. 22).

Minganti, F., Bartolo, N., Lolli, J., Casteels, W. & Ciuti, C. Exact results for Schrödinger cats in driven-dissipative systems and their feedback control. *Sci. Rep.* **6,** 26987 (2016) (cit. on p. 76).

Mlynek, J. A., Abdumalikov, A. A., Eichler, C. & Wallraff, A. Observation of Dicke superradiance for two artificial atoms in a cavity with high decay rate. *Nat. Comm.* **5,** 5186 (2014) (cit. on pp. 52, 92).

Moody, G., Segnon, M., Sagnes, I., Braive, R., Beveratos, A., Robert-Philip, I., Belabas, N., Jahnke, F., *et al.* Delayed formation of coherence in the emission dynamics of high-Q nanolasers. *Optica* **5,** 395–401 (2018) (cit. on pp. 3, 62).

Mu, Y. & Savage, C. M. One-atom lasers. *Phys. Rev. A* **46,** 5944–5954 (1992) (cit. on pp. 76, 78).

Navarrete-Benlloch, C. Open systems dynamics: Simulating master equations in the computer. *arXiv e-prints* (2015) (cit. on p. 78).

Nocedal, J. & Wright, S. J. in *Numerical Optimization* 245–269 (Springer, 2006a) (cit. on p. 34).

Nocedal, J. & Wright, S. J. in *Numerical Optimization* 30–65 (Springer, 2006b) (cit. on p. 25).

Noda, S. Seeking the Ultimate Nanolaser. *Science* **314,** 260–261 (2006) (cit. on pp. 2, 66).

Nolting, W. *Grundkurs Theoretische Physik 5/1* 8. Auflage (Springer, 2013) (cit. on p. 26).

Norman, J. C., Jung, D., Zhang, Z., Wan, Y., Liu, S., Shang, C., Herrick, R. W., Chow, W. W., *et al.* A Review of High-Performance Quantum Dot Lasers on Silicon. *IEEE J. Quantum Electron.* **55,** 1–11 (2019) (cit. on p. 2).

Novoselov, K. S., Geim, A. K., Morozov, S. V., Jiang, D., Zhang, Y., Dubonos, S. V., Grigorieva, I. V. & Firsov, A. A. Electric Field Effect in Atomically Thin Carbon Films. *Science* **306,** 666–669 (2004) (cit. on p. 66).

Novoselov, K. S., Jiang, D., Schedin, F., Booth, T. J., Khotkevich, V. V., Morozov, S. V. & Geim, A. K. Two-dimensional atomic crystals. *Proc Natl Acad Sci* **102,** 10451–10453 (2005) (cit. on p. 66).

Ota, Y., Kakuda, M., Watanabe, K., Iwamoto, S. & Arakawa, Y. Thresholdless quantum dot nanolaser. *Opt. Express* **25,** 19981–19994 (2017) (cit. on p. 2).

Pachón, L. A. & Brumer, P. Mechanisms in environmentally assisted one-photon phase control. *J. Chem. Phys.* **139,** 164123 (2013) (cit. on p. 93).

Pachón, L. A. & Brumer, P. Incoherent excitation of thermally equilibrated open quantum systems. *Phys. Rev. A* **87,** 022106 (2013a) (cit. on p. 93).

Pachón, L. A., Triana, J. F., Zueco, D. & Brumer, P. Influence of non-Markovian dynamics in equilibrium uncertainty-relations. *J. Chem. Phys.* **150,** 034105 (2019) (cit. on p. 93).

Pachón, L. A., Yu, L. & Brumer, P. Coherent one-photon phase control in closed and open quantum systems: A general master equation approach. *Faraday Discuss.* **163,** 485–495 (2013b) (cit. on p. 93).

Phoenix, S. J. D. & Knight, P. L. Fluctuations and entropy in models of quantum optical resonance. *Ann. Phys.* **186,** 381–407 (1988) (cit. on p. 47).

Pressé, S., Ghosh, K., Lee, J. & Dill, K. A. Principles of maximum entropy and maximum caliber in statistical physics. *Rev. Mod. Phys.* **85,** 1115–1141 (2013) (cit. on pp. 1, 6–11, 14).

Pressé, S., Ghosh, K., Lee, J. & Dill, K. Reply to C. Tsallis' "Conceptual Inadequacy of the Shore and Johnson Axioms for Wide Classes of Complex Systems". *Entropy* **17,** 5043–5046 (2015) (cit. on p. 11).

Prieto, I., Llorens, J. M., Muñoz-Camúñez, L. E., Taboada, A. G., Canet-Ferrer, J., Ripalda, J. M., Robles, C., Muñoz-Matutano, G., *et al.* Near thresholdless laser operation at room temperature. *Optica* **2,** 66–69 (2015) (cit. on pp. 2, 37, 46).

Qu, Y. & Singh, S. Photon correlation effects in second harmonic generation. *Opt. Commun.* **90,** 111–114 (1992) (cit. on p. 52).

Ralph, T. C. & Savage, C. M. Squeezed light from a coherently pumped four-level laser. *Phys. Rev. A* **44,** 7809–7814 (1991) (cit. on pp. 76, 78).

Redlich, C., Lingnau, B., Holzinger, S., Schlottmann, E., Kreinberg, S., Schneider, C., Kamp, M., Höfling, S., *et al.* Mode-switching induced super-thermal bunching in quantum-dot microlasers. *New J. Phys.* **18,** 063011 (2016) (cit. on pp. 52, 56).

Reitzenstein, S. Semiconductor Quantum Dot–Microcavities for Quantum Optics in Solid State. *IEEE J. Sel. Top. Quantum Electron.* **18,** 1733–1746 (2012) (cit. on p. 2).

Reitzenstein, S. & Forchel, A. Quantum dot micropillars. *J. Phys. D Appl. Phys.* **43,** 033001 (2010) (cit. on p. 2).

Ren, Y., Qiao, Z. & Niu, Q. Topological phases in two-dimensional materials: a review. *Rep. Prog. Phys.* **79,** 066501 (2016) (cit. on p. 66).

Rice, P. R. & Carmichael, H. J. Photon statistics of a cavity-QED laser: A comment on the laser-phase-transition analogy. *Phys. Rev. A* **50,** 4318–4329 (1994) (cit. on pp. 37–39, 46, 49).

Richter, M., Carmele, A., Sitek, A. & Knorr, A. Few-Photon Model of the Optical Emission of Semiconductor Quantum Dots. *Phys. Rev. Lett.* **103,** 087407 (2009) (cit. on p. 49).

Ritter, S., Gartner, P., Gies, C. & Jahnke, F. Emission properties and photon statistics of a single quantum dot laser. *Opt. Express* **18,** 9909–9921 (2010) (cit. on pp. 3, 32, 62).

Samuel, I. D. W., Namdas, E. B. & Turnbull, G. A. How to recognize lasing. *Nat. Photonics* **3,** 546–549 (2009) (cit. on p. 46).

Schaibley, J. R., Yu, H., Clark, G., Rivera, P., Ross, J. S., Seyler, K. L., Yao, W. & Xu, X. Valleytronics in 2D materials. *Nat. Rev. Mater.* **1,** 16055 (2016) (cit. on pp. 2, 67).

Scheibner, M., Schmidt, T., Worschech, L., Forchel, A., Bacher, G., Passow, T. & Hommel, D. Superradiance of quantum dots. *Nat. Phys.* **3,** 106–110 (2007) (cit. on pp. 52, 92).

Schirmer, S. G. & Wang, X. Stabilizing open quantum systems by Markovian reservoir engineering. *Phys. Rev. A* **81,** 062306 (2010) (cit. on p. 78).

Schlottmann, E., von Helversen, M., Leymann, H. A. M., Lettau, T., Krüger, F., Schmidt, M., Schneider, C., Kamp, M., *et al.* Exploring the Photon-Number Distribution of Bimodal Microlasers with a Transition Edge Sensor. *Phys. Rev. Appl.* **9,** 064030 (2018) (cit. on pp. 3, 32, 46, 49, 87, 93).

Schmidt, M., Grothe, I. H., Neumeier, S., Bremer, L., von Helversen, M., Zent, W., Melcher, B., Beyer, J., *et al.* Investigation of the bimodal behavior of microlasers with a two-channel photon-number-resolving transition-edge sensor system. *Phys. Rev. Res. [submitted]* (2020) (cit. on pp. 3, 87, 93).

Schneebeli, L., Kira, M. & Koch, S. W. Characterization of Strong Light-Matter Coupling in Semiconductor Quantum-Dot Microcavities via Photon-Statistics Spectroscopy. *Phys. Rev. Lett.* **101,** 097401 (2008) (cit. on pp. 3, 49).

Schulz, S. *Electronic and Optical Properties of Quantum Dots: A Tight-Binding Approach* PhD thesis (Universität Bremen, 2007) (cit. on p. 92).

Schumann, J. *Semi-empirisches Tight-Binding für InGaAs-Quantenpunkte* Master's thesis (Universität Bremen, 2016) (cit. on pp. 2, 92).

Scrutinizing lasers. *Nat. Photonics* **11,** 139–139 (2017) (cit. on p. 46).

Scully, M. O. & Lamb, W. E. Quantum Theory of an Optical Maser. I. General Theory. *Phys. Rev.* **159,** 208–226 (1967) (cit. on p. 47).

Scully, M. O. & Svidzinsky, A. A. The Super of Superradiance. *Science* **325,** 1510–1511 (2009) (cit. on pp. 52, 92).

Shannon, C. E. A mathematical theory of communication. *Bell Syst. tech. J.* **27,** 379–423 (1948) (cit. on pp. 1, 7).

Shohat, A. & Tamarkin, J. *The Problem of Moments* (American Mathematical Society, 1970) (cit. on pp. 20, 21).

Shore, J. & Johnson, R. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inf. Theory* **26,** 26–37 (1980) (cit. on pp. 10, 16).

Shore, J. & Johnson, R. Properties of cross-entropy minimization. *IEEE Trans. Inf. Theory* **27,** 472–482 (1981) (cit. on pp. 10, 16).

Skajaa, A. *Limited Memory BFGS for Nonsmooth Optimization* Master's thesis (New York University, 2010) (cit. on p. 25).

Skilling, J. Data analysis: The maximum entropy method. *Nature* **309,** 748–749 (1984) (cit. on p. 10).

Slater, J. C. & Koster, G. F. Simplified LCAO Method for the Periodic Potential Problem. *Phys. Rev.* **94,** 1498–1524 (1954) (cit. on p. 92).

Sobczyk, K. & Trębicki, J. Maximum entropy principle in stochastic dynamics. *Probabilistic Eng. Mech.* **5,** 102 –110 (1990) (cit. on p. 32).

Sobczyk, K. & Trębicki, J. Maximum entropy principle and nonlinear stochastic oscillators. *Physica A* **193,** 448 –468 (1993) (cit. on p. 32).

Sobczyk, K. & Trębicki, J. Approximate probability distributions for stochastic systems: maximum entropy method. *Comput. Methods Appl. Mech. Eng.* **168,** 91 –111 (1999) (cit. on pp. 32, 33).

Spasibko, K. Y., Kopylov, D. A., Krutyanskiy, V. L., Murzina, T. V., Leuchs, G. & Chekhova, M. V. Multiphoton Effects Enhanced due to Ultrafast Photon-Number Fluctuations. *Phys. Rev. Lett.* **119,** 223603 (2017) (cit. on p. 52).

Steinhoff, A., Florian, M. & Jahnke, F. Dynamical screening effects of substrate phonons on two-dimensional excitons. *Phys. Rev. B* **101,** 045411 (2020) (cit. on pp. 2, 67).

Steinhoff, A., Florian, M., Singh, A., Tran, K., Kolarczik, M., Helmrich, S., Achtstein, A. W., Woggon, U., *et al.* Biexciton fine structure in monolayer transition metal dichalcogenides. *Nat. Phys.* **14,** 1199–1204 (2018a) (cit. on pp. 2, 67).

Steinhoff, A., Wehling, T. O. & Rösner, M. Frequency-dependent substrate screening of excitons in atomically thin transition metal dichalcogenide semiconductors. *Phys. Rev. B* **98,** 045304 (2018b) (cit. on pp. 2, 67).

Strauf, S., Hennessy, K., Rakher, M. T., Choi, Y.-S., Badolato, A., Andreani, L. C., Hu, E. L., Petroff, P. M., *et al.* Self-Tuned Quantum Dot Gain in Photonic Crystal Lasers. *Phys. Rev. Lett.* **96,** 127404 (2006) (cit. on pp. 2, 46, 52).

Strauß, M., Placke, M., Kreinberg, S., Schneider, C., Kamp, M., Höfling, S., Wolters, J. & Reitzenstein, S. Photon-statistics excitation spectroscopy of a single two-level system. *Phys. Rev. B* **93,** 241306 (2016) (cit. on pp. 3, 49, 91).

Tagliani, A. On the existence of maximum entropy distributions with four and more assigned moments. *Probabilistic Eng. Mech.* **5,** 167 –170 (1990) (cit. on p. 22).

Tagliani, A. Existence and stability of a discrete probability distribution by maximum entropy approach. *Appl. Math. Comput.* **110,** 105 –114 (2000) (cit. on pp. 21–23).

Takemura, N., Takiguchi, M., Kuramochi, E., Shinya, A., Sato, T., Takeda, K., Matsuo, S. & Notomi, M. Lasing thresholds and photon statistics in high-β buried multiple quantum well photonic crystal nanocavity lasers. *Phys. Rev. A* **99,** 053820 (2019) (cit. on p. 37).

Templeman, A. B. & Xingsi, L. Entropy Duals. *Eng. Optim.* **9,** 107–119 (1985) (cit. on p. 23).

Thouless, D. J., Kohmoto, M., Nightingale, M. P. & den Nijs, M. Quantized Hall Conductance in a Two-Dimensional Periodic Potential. *Phys. Rev. Lett.* **49,** 405–408 (1982) (cit. on p. 66).

Thyrrestrup, H., Sapienza, L. & Lodahl, P. Extraction of the β-factor for single quantum dots coupled to a photonic crystal waveguide. *Appl. Phys. Lett.* **96,** 231106 (2010) (cit. on p. 2).

Tikochinsky, Y., Tishby, N. Z. & Levine, R. D. Alternative approach to maximum-entropy inference. *Phys. Rev. A* **30,** 2638–2644 (1984a) (cit. on p. 10).

Tikochinsky, Y., Tishby, N. Z. & Levine, R. D. Consistent Inference of Probabilities for Reproducible Experiments. *Phys. Rev. Lett.* **52,** 1357–1360 (1984b) (cit. on p. 10).

Tran, K., Moody, G., Wu, F., Lu, X., Choi, J., Kim, K., Rai, A., Sanchez, D. A., *et al.* Evidence for moiré excitons in van der Waals heterostructures. *Nature* **567,** 71–75 (2019) (cit. on pp. 3, 67).

Trębicki, J. & Sobczyk, K. Maximum entropy principle and non-stationary distributions of stochastic systems. *Probabilistic Eng. Mech.* **11,** 169–178 (1996) (cit. on pp. 32, 33).

Triana, J. F., Estrada, A. F. & Pachón, L. A. Ultrafast Optimal Sideband Cooling under Non-Markovian Evolution. *Phys. Rev. Lett.* **116,** 183602 (2016) (cit. on p. 93).

Tribus, M. & McIrvine, E. C. Energy and Information. *Sci. Am.* **225,** 179–190 (1971) (cit. on pp. 1, 7, 8).

Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **52,** 479–487 (1988) (cit. on pp. 11, 12).

Tsallis, C. Conceptual Inadequacy of the Shore and Johnson Axioms for Wide Classes of Complex Systems. *Entropy* **17,** 2853–2861 (2015) (cit. on pp. 11, 12).

Tsallis, C. Beyond Boltzmann–Gibbs–Shannon in Physics and Elsewhere. *Entropy* **21,** 696 (2019) (cit. on pp. 1, 11, 12, 14).

Uffink, J. Can the maximum entropy principle be explained as a consistency requirement? *Stud. Hist. Philos. Sci. B* **26,** 223–261 (1995) (cit. on pp. 10, 11, 16).

Uffink, J. The constraint rule of the maximum entropy principle. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **27,** 47 –79 (1996) (cit. on p. 10).

Uffink, J. in *The Stanford Encyclopedia of Philosophy* (ed Zalta, E. N.) Spring 2017 (Metaphysics Research Lab, Stanford University, 2017) (cit. on pp. 1, 7).

Ulrich, S. M., Gies, C., Ates, S., Wiersig, J., Reitzenstein, S., Hofmann, C., Löffler, A., Forchel, A., *et al.* Photon Statistics of Semiconductor Microcavity Lasers. *Phys. Rev. Lett.* **98,** 043906 (2007) (cit. on pp. 37, 46, 52, 62, 68).

Vahala, K. J. Optical microcavities. *Nature* **424,** 839 (2003) (cit. on pp. 2, 37).

Van Fraassen, B. C. A Problem for Relative Information Minimizers in Probability Kinematics. *Br. J. Philos. Sci.* **32,** 375–379 (1981) (cit. on p. 11).

Van Fraassen, B. C., Hughes, R. I. G. & Harman, G. A Problem for Relative Information Minimizers, Continued. *The British Journal for the Philosophy of Science* **37,** 453–463 (1986) (cit. on p. 11).

Vedral, V. The role of relative entropy in quantum information theory. *Rev. Mod. Phys.* **74,** 197–234 (2002) (cit. on pp. 17, 26, 27, 42, 82).

Vogl, P., Hjalmarson, H. P. & Dow, J. D. A Semi-empirical tight-binding theory of the electronic structure of semiconductors. *J. Phys. Chem. Solids* **44,** 365 –378 (1983) (cit. on p. 92).

Volkenstein, M. V. *Entropy and Information* (Birkhäuser, 2009) (cit. on pp. 1, 5, 8).

Von Neumann, J. *Mathematische Grundlagen der Quantenmechanik* (Springer, 1996) (cit. on pp. 1, 7).

Walker, B. T., Flatten, L. C., Hesten, H. J., Mintert, F., Hunger, D., Trichet, A. A. P., Smith, J. M. & Nyman, R. A. Driven-dissipative non-equilibrium Bose-Einstein condensation of less than ten photons. *Nat. Phys.* (2018) (cit. on p. 76).

Wang, Q. H., Kalantar-Zadeh, K., Kis, A., Coleman, J. N. & Strano, M. S. Electronics and optoelectronics of two-dimensional transition metal dichalcogenides. *Nat. Nanotechnol.* **7,** 699–712 (2012) (cit. on pp. 2, 67).

Wang, T., Deng, Z. L., Sun, J. C., Wang, X. H., Puccioni, G. P., Wang, G. F. & Lippi, G. L. Photon statistics and dynamics of nanolasers subject to intensity feedback. *Phys. Rev. A* **101,** 023803 (2020) (cit. on p. 37).

Weimer, H. Variational Principle for Steady States of Dissipative Quantum Many-Body Systems. *Phys. Rev. Lett.* **114,** 040402 (2015) (cit. on p. 78).

Wiersig, J. Microscopic theory of first-order coherence in microcavity lasers based on semiconductor quantum dots. *Phys. Rev. B* **82,** 155320 (2010) (cit. on p. 46).

Wiersig, J., Gies, C., Jahnke, F., Aßmann, M., Berstermann, T., Bayer, M., Kistner, C., Reitzenstein, S., *et al.* Direct observation of correlations between individual photon emission events of a microcavity laser. *Nature* **460,** 245 (2009) (cit. on pp. 37, 49).

Wolfe, P. Convergence Conditions for Ascent Methods. *SIAM Rev.* **11,** 226–235 (1969) (cit. on p. 25).

Wolfe, P. Convergence Conditions for Ascent Methods. II: Some Corrections. *SIAM Rev.* **13,** 185–188 (1971) (cit. on p. 25).

Wu, Z., Phillips, G. N., Tapia, R. & Zhang, Y. A Fast Newton Algorithm for Entropy Maximization in Phase Determination. *SIAM Rev.* **43,** 623–642 (2001) (cit. on p. 23).

Xiao, D., Liu, G.-B., Feng, W., Xu, X. & Yao, W. Coupled Spin and Valley Physics in Monolayers of MoS_2 and Other Group-VI Dichalcogenides. *Phys. Rev. Lett.* **108,** 196802 (2012) (cit. on pp. 2, 67).

Xu, X., Yao, W., Xiao, D. & Heinz, T. F. Spin and pseudospins in layered transition metal dichalcogenides. *Nat. Phys.* **10,** 343–350 (2014) (cit. on pp. 2, 67).

Yokoyama, H. & Brorson, S. D. Rate equation analysis of microcavity lasers. *J. Appl. Phys.* **66,** 4801–4805 (1989) (cit. on p. 39).

Zheng, W., Jiang, Y., Hu, X., Li, H., Zeng, Z., Wang, X. & Pan, A. Light Emission Properties of 2D Transition Metal Dichalcogenides: Fundamentals and Applications. *Adv. Opt. Mater.* **6,** 1800420 (2018) (cit. on p. 2).

Zhou, R., Cai, R. & Tong, G. Applications of Entropy in Finance: A Review. *Entropy* **15,** 4909–4931 (2013) (cit. on pp. 1, 14).

Zieliński, M., Korkusiński, M. & Hawrylak, P. Atomistic tight-binding theory of multiexciton complexes in a self-assembled InAs quantum dot. *Phys. Rev. B* **81,** 085301 (2010) (cit. on p. 92).

Acknowledgements

First of all, I wholeheartedly thank everybody who reads more than just the acknowledgments (although only reading the acknowledgments is perfectly fine). Countless hours were invested in this work and I am glad for everyone who gets any value out of this thesis.

However, this work would not have been possible without the help and support of many people. First, I want to thank J. Wiersig for the opportunity to work in his group, for the many valuable insights, advises and comments on the topic and the manuscript, and for the gentle introduction into the world of scientific research. Furthermore, I am very grateful that I got the opportunity to attend three superb conferences in Berlin (NOEKS), Regensburg (DPG spring meeting), and Banff (FOPS, with awesome Canadian landscapes and wildlife). Support from the DFG within the project "Maxium-Entropie-Methode angewandt auf das Vielteilchenhierarchie-Problem in Quantenpunkt-Mikroresonator-Systemen" is gratefully acknowledged.

Additionally, B. Gulyak deserves special thanks for the decision to rather work on a mathematical subject of which I probability wouldn't understand a word. His decision and a fortunate coincidence gave me the chance to work on this inspiring topic and in J. Wiersig's group. Aside from that, I thank him for the uncomplicated collaboration that lead to the two main publications of this thesis, his thorough preliminary work on the topic, and comments from a mathematical viewpoint that often made me marvel at the beautiful profundity of this discipline. Also, I would like to thank T. Lettau and H. A. M. Leymann for the collaboration and for valuable insights on how to prepare a manuscript for publication in a scientific journal.

Moreover, I want to thank all colleagues in the group, especially S. Neumeier and M. Badel with whom I was pleased to share an office room. Thank you for suffering through my constant need for better lighting and for the successful rearrangement of the office furniture. Apart from them, I thank I. H. Grothe, J. Kullig, M. Khanbekyan, C.-H. Yi, J. Guckel, C. Vogel, H. Maasland, G. Kasner, and D. Vogt for contributing to a calm and pleasant working atmosphere and for help with any technical, administrative, physics-related or non-physics-related problem.

I would like to thank G. Kasner for the opportunity to host the "Computer & Software" seminar and I thank all participating students for reminding me of the enthusiasm which I also had when I was writing my first programs during my time in Bremen. Also, thanks to

G. Kasner for sharing his first-hand experiences of the GDR-era during the coffee breaks. I learned a lot.

I explicitly would like to thank D. Vogt for being in the team. Your presence, help, and kindness is greatly appreciated and makes many previous problems less of an issue.

Special credit goes to all proofreaders of the thesis, I. H. Grothe, M. Badel, J. Kullig, and S. Neumeier. After countless hours of staring at the screen, even the most obvious fallacies and typos seem to magically disappear from the field of vision. Your effort helped to greatly improve the manuscript. Thank goes to G. Hollens for proofreading the (few) German parts of the manuscript and administrative documents.

Finally, a warm thanks goes to my family for their support and for the beautiful holidays we had together. Thank you for helping me to stay positive and optimistic. A special thanks goes to Чебурашка for helping me to get through the very darkest hours.

Lastly, thank you, Isa. For everything.

Colophon

This thesis was typeset with LaTeX 2_ε. It uses the *Clean Thesis* style developed by Ricardo Langner. The design of the *Clean Thesis* style is inspired by user guide documents from Apple Inc.

Download the *Clean Thesis* style at `http://cleanthesis.der-ric.de/`.

Declaration of Honor

I hereby declare that I prepared this thesis without impermissible help of third parties and that none other than the indicated tools have been used; all sources of information are clearly marked, including my own publications.

In particular I have not consciously:

- Fabricated data or rejected undesired results
- Misused statistical methods with the aim of drawing other conclusions than those warranted by the available data
- Plagiarized external data or publications
- Presented the results of other researchers in a distorted way

I am aware that violations of copyright may lead to injunction and damage claims of the author and also to prosecution by the law enforcement authorities.

I hereby agree that the thesis may be reviewed for plagiarism by mean of electronic data processing.

This work has not yet been submitted as a doctoral thesis in the same or a similar form in Germany or in any other country. It has not yet been published as a whole.

Magdeburg, 15.12.2020

M. Sc. Boris Melcher