



HAL
open science

From point processes to quantum optics and back

Rémi Bardenet, Alexandre Feller, Jérémie Bouttier, Pascal Degiovanni, Adrien Hardy, Adam Rançon, Benjamin Roussel, Grégory Schehr, Christoph I. Westbrook

► To cite this version:

Rémi Bardenet, Alexandre Feller, Jérémie Bouttier, Pascal Degiovanni, Adrien Hardy, et al.. From point processes to quantum optics and back. 2023. hal-03837697

HAL Id: hal-03837697

<https://hal.science/hal-03837697>

Preprint submitted on 25 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

From point processes to quantum optics and back

Rémi Bardenet^{*†1}, Alexandre Feller^{*‡1}, Jérémie Bouttier², Pascal Degiovanni³, Adrien Hardy⁴, Adam Rançon⁵, Benjamin Roussel⁶, Grégory Schehr⁷, and Christoph I. Westbrook⁸

¹ *Univ. Lille, CNRS, Centrale Lille, UMR 9189 – CRISTAL – Centre de Recherche en Informatique, Signal et Automatique, F-59000 Lille, France*

² *Université Paris-Saclay, CNRS, CEA, Institut de Physique Théorique, 91191, Gif-sur-Yvette, France*

³ *Univ Lyon, Ens de Lyon, Université Claude Bernard Lyon 1, CNRS, Laboratoire de Physique (UMR 5672), F-69342 Lyon, France*

⁴ *Qube Research and Technologies, 75008 Paris, France*

⁵ *Univ. Lille, CNRS, UMR 8523 – PhLAM – Laboratoire de Physique des Lasers, Atomes et Molécules, F-59000 Lille, France*

⁶ *Department of Applied Physics, Aalto University, 00076 Aalto, Finland*

⁷ *Sorbonne Université, Laboratoire de Physique Théorique et Hautes Energies, CNRS UMR 7589, 4 Place Jussieu, Tour 13, 5ème étage, 75252 Paris 05, France*

⁸ *Univ. Paris-Saclay, Institut d’Optique Graduate School, CNRS, UMR 8501 – Laboratoire Charles Fabry, F-91127 Palaiseau, France*

Abstract

[Disclaimer: this is Part I of a cross-disciplinary survey that is work in progress; All comments are welcome.] Some fifty years ago, in her seminal PhD thesis, Odile Macchi introduced permanental and determinantal point processes. Her initial motivation was to provide models for the set of detection times in fundamental bosonic or fermionic optical experiments, respectively. After two rather quiet decades, these point processes have quickly become standard examples of point processes with non-trivial, yet tractable, correlation structures. In particular, determinantal point processes have been since the 1990s a technical workhorse in random matrix theory and combinatorics, and a standard model for repulsive point patterns in machine learning and spatial statistics since the 2010s. Meanwhile, our ability to experimentally probe the correlations between detection events in bosonic and fermionic optics has progressed tremendously. In Part I of this survey, we provide a modern introduction to the concepts in Macchi’s thesis and their physical motivation, under the combined eye of mathematicians, physicists, and signal processors. Our objective is to provide a shared basis of knowledge for later cross-disciplinary work on point processes in quantum optics, and reconnect with the physical roots of permanental and determinantal point processes.

Contents

1	Introduction	2
2	Point processes and the semiclassical picture of HBT	5
2.1	The correlation functions of a point process	6
2.2	Poisson and Cox point processes	7
2.3	Permanental point processes	9
2.4	Determinantal point processes	11

*Joint first authorship and corresponding authors.

†remi.bardenet@univ-lille.fr

‡alexandre.feller@univ-lille.fr

3	Elements of quantum field theory	12
3.1	The mathematical framework of quantum theory	12
3.2	Two fundamental one-particle systems	15
3.2.1	The qubit	16
3.2.2	The harmonic oscillator	17
3.3	Modelling a finite number of particles	18
3.3.1	Subsystems, bosons and fermions	18
3.3.2	Example N -particle states	19
3.3.3	Occupation number representation	20
3.4	Modeling an indefinite number of particles	20
3.4.1	Fock spaces for bosons and fermions	21
3.4.2	Creation and annihilation operators	21
3.4.3	Modes	23
3.4.4	Field operators	23
3.4.5	Observables	24
3.5	Wick's theorem	25
3.5.1	Gaussian density matrices	25
3.5.2	Where permanents and determinants appear	26
3.6	Bosonic coherent states model classical fields	29
3.6.1	Definition and properties	29
3.6.2	The Husimi distribution and time-frequency analysis	31
4	Photodetection and bosonic coherences	32
4.1	Modeling photodetection events	32
4.1.1	Modeling the radiation	32
4.1.2	Modeling the detector: the first-order coherence function	33
4.1.3	Role of the detector structure function	37
4.2	Correlation between photodetection events	38
4.2.1	Second-order coherence	38
4.2.2	Higher-order coherences	40
4.3	Turning coherence functions into correlation functions	40
4.4	Single-photon sources can lead to anti-bunching	43
5	Electrodetection and fermionic coherences	45
5.1	Modeling electro-detection events	46
5.1.1	Modeling the fermionic field	46
5.1.2	First-order electronic coherence	47
5.2	Correlation between electro-detection events	49
5.3	Electron anti-bunching	52
5.4	Recovering a classical current is not as easy as for bosons	53
6	Wrapping up and open questions	54
6.1	From a point process to free particles	54
6.2	Going further: from simple to open questions	56
6.2.1	A quantum state is more than a point process	56
6.2.2	A DPP from bosons in a non-Gaussian state	58
6.2.3	Interacting field theories and point processes	58
6.2.4	Constructive arguments for point processes	59
6.3	A teaser for Part II	59

1 Introduction

The *photoelectric effect* is the release of individual electrons from a metal, when light falls onto that metal. The empirical observation that no electrons are released if the light frequency is beneath a certain threshold, along with an analogy with elastic collisions, prompted Einstein to posit the existence of light quanta, also known as *photons*. A century

later, the full explanation of the photoelectric effect is considered a success of the quantum theory of light.

In parallel, experimental devices have been designed to amplify the current resulting from a small number of released electrons, eventually giving detectors of light so sensitive that they are able to detect single photons. These detectors have led to investigations on the quantum coherence properties of light. Coherence here means the ability of sources of light to generate interference patterns, such as in Young’s celebrated double slit experiment [Mandel and Wolf, 1965, Sections 1-3]. One puzzling aspect of coherence was demonstrated by Hanbury Brown and Twiss [1958], and is called *photon bunching* or the HBT effect, after the initials of its discoverers. When placing a detector of single photons in the electric field created by a thermal light source, such as incandescent matter, Hanbury Brown and Twiss [1958] showed that the detector tends to produce clicks that are grouped in time. Everything happens as if the photons were tightly bunched together when arriving at the detector, hence the name of photon bunching. A few years later, with the invention of the laser, it was realized that this bunching effect disappeared for (so-called *coherent*) laser light.

At the turn of the 60s, there was intense research in finding the right mathematical framework to describe such coherence phenomena. A celebrated contribution is that of Glauber [1963], who introduced concepts like coherent states and coherence functions, which quickly became textbook material for quantum optics. Having met Glauber at Les Houches in 1964, French signal processing pioneer Bernard Picinbono launched a team in statistical optics in Orsay, determined to explore the transition between bunching and non-bunching measurements using the formalism of stochastic processes. Picinbono assembled a small and diverse group, including PhD student Odile Macchi, who had just obtained a degree in mathematics. The starting point of Macchi’s 1972 PhD thesis was to find the right stochastic object to describe the detection times in HBT and explain photon bunching. Her thesis turned out to be foundational in many respects. This justified a recent translation of Macchi’s thesis (originally in French) by Hans Zessin [Macchi, 2017], along with a large appendix written by Hans Zessin and Suren Poghosyan.

Odile Macchi introduced what we now call *correlation functions* to describe point processes, i.e., random configurations of points in a generic space. She showed how common models for physical sources and detection led to point processes with closed-form correlation functions. In particular, for a given model of the source used in the HBT experiment, she showed that the resulting point process of detection times is a permanent point process, for which bunching can be fully characterized and related to properties of the electric field. Teaming up with fellow PhD student in physics Christine Bénard, they used Glauber’s formalism to identify the point process describing the parallel situation of detection times of electrons. Unlike for photons, the latter point process naturally exhibits *anti-bunching*, with detection times being very regularly spaced. [Bénard and Macchi, 1973] is one the foundational stones¹ for what we call today *determinantal point processes* (DPPs).

Outside physics, determinantal point processes (DPPs) have since then become a cornerstone of the theory of random matrices [Anderson, Guionnet, and Zeitouni, 2010, Johansson, 2006], with applications in combinatorics [Borodin et al., 2000] and number theory [Rudnick and Sarnak, 1996]. DPPs are also a popular model for repulsive point pattern data in spatial statistics [Lavancier, Møller, and Rubak, 2014] and machine learning [Kulesza and Taskar, 2012]. Many DPP users in these fields² have little idea of the physical origin of DPPs. Given the late blooming of DPPs outside their original field, and some fifty years after Macchi’s first generic formalization of DPPs, we felt it a natural endeavour to reconnect DPPs with their physical roots, as models for non-interacting fermions and tools to probe generalizations of the HBT effect. This has been the purpose of two workshops so far.

¹Together with a 1974 conference communication, that appeared later as [Macchi, 1977], and earlier work by Jean Ginibre; see the preface to the recent re-print [Macchi, 2017] of Odile Macchi’s thesis for a broader history.

²including some of us prior to this work!

We started with a two-day event³ in 2019 in Lille, France, opened by Odile Macchi and featuring both theoretical and experimental physicists exposing their view of fermionic coherence to an audience of DPP users across mathematics, computer science, and signal processing. The workshop having been met with cross-disciplinary enthusiasm, we made two decisions. The first was to organize an ambitious, two-week follow-up to the workshop,⁴ which took place in 2022 in Lyon, France. The second decision was to unite forces and write a survey on the links between point processes and optics, to pin down a common ground for discussions. The current document is the first part of this survey. It is both a joint introduction to point processes and quantum optics, and organized notes from a modern reading of Odile Macchi’s thesis. The novelty of our document relies in its cross-disciplinary target audience of mathematicians, physicists, and signal processors, with a solid undergraduate background in probability and functional analysis. In particular, while tackling topics in modern physics, we assume little physics knowledge from the reader beyond undergraduate exposure to wave optics. One of our leitmotifs is to sketch the thought process behind some fundamental arguments in quantum optics. Indeed, in our experience, arguments thought as basic by physicists can be hard to grasp by people trained in mathematics, mostly because the implicitly assumed lore differs across communities. In a reverse movement towards physicists, and following the spirit of Macchi [1975], we motivate most mathematical concepts by their need as models in optics, including point processes.

We have striven to maintain a balance between mathematical rigour, clarity, and brevity, giving references whenever we had to take shortcuts. We expect that every reader will find some of the material basic and some more exotic, depending on their background, and we hope that all readers will eventually learn something useful. Our objective is to make the potential barrier for crossing from one discipline to another as low as possible, so that ideas can flow more easily. A second part of the document is in preparation, presenting selected advanced topics from the Lille workshop, including experimental measurements of the HBT effect, non-interacting trapped fermions in statistical physics [Dean et al., 2016, 2019], fermions in combinatorics, and electronic quantum signals. The two parts are ultimately to be bound in a single manuscript.

The rest of the document is organized as follows. In Section 2, we introduce key examples of point processes. Poisson, Cox, and permanent point processes are motivated there by the so-called *semi-classical* derivation of the HBT photon bunching effect, treating only the detector as a quantum object, not the electric field. Determinantal point processes are also introduced, but their physical motivation requires to go beyond the semi-classical picture, which justifies the next three sections. Section 3 is a crash course in quantum field theory, from the basics of quantum mechanics to Wick’s theorem on the average of products of ladder operators. Wick’s theorem yields two very different results depending on the commutation rules of the operators it applies to, which in turn derive from modeling either *bosonic* particles (like photons), or *fermionic* particles (like electrons). Ultimately, this dichotomy is at the origin of the appearance of permanent and determinantal point processes. The section concludes with a discussion of the coherent states of Glauber [1963] and their relation to time-frequency analysis in signal processing. Section 4 covers the modern view on photodetection, culminating in the full quantum justification of the HBT effect using permanent point processes as a model for the coincidence measurements of photons, as well as considerations on the role of the source in the bunching properties of the measurements. Following the lines of Section 4, Section 5 covers the detection of electrons, finally resulting in the appearance of determinantal point processes. The section concludes on the comparative difficulties of recovering non-quantum arguments from the quantum treatment in the case of fermions. Finally, Section 6 wraps up this first part, commenting on the generic derivation of a permanent or determinantal point process from a model of free fermions, and discussing open questions motivated by this construction.

A note on the style. Because of its cross-disciplinary objectives, the style of our document is hybrid. We mostly follow a style inspired by texts in mathematics, with definitions,

³<https://dpp-fermions.sciencesconf.org/>

⁴<https://indico.in2p3.fr/event/25182/>

theorems, assumptions, and examples, sometimes merged with the main text, sometimes fleshed out to draw the reader’s attention. Assumptions, in particular, often stand out. By *assumption* we mean a statement for the reader to accept in order to progress in a discussion or a computation. It can be, e.g., a modeling choice or a mathematical approximation.

Examples are often borrowed from physics; to make the text self-contained, some of these examples are relatively long. To help the reader to visually isolate examples from the rest of the text, we conclude each example with a \diamond symbol. Finally, footnotes abound, and are usually meant as side remarks to one of the targeted scientific communities, e.g. to discuss notational or minor conceptual differences between different domains.

2 Point processes and the semiclassical picture of HBT

In a physical experiment where a detector clicks when hit by a particle, the observation consists of a set of real numbers, the times at which the detector clicks. The natural probabilistic model for such a situation, where both the number and the location of the observed points are uncertain, is a *point process*, i.e, a random configuration of points. In Sections 2.1 to 2.4, we introduce some of the modern vocabulary of point process theory, along with three key families of point processes: Poisson, permanental and determinantal point processes.

While optical models that feature determinantal point processes will have to wait for the fully quantum treatment of electron detection in Section 5, we already motivate Poisson and permanental point processes in this section by the so-called *semi-classical* (as opposed to *fully quantum*) derivation of the photon bunching effect. Our running examples are based on the simple setup given here as Example 0.

Example 0 (A simple photodetection setup). As introduced in Section 1, photodetection is based on the photoelectric effect: when light falls onto a metal, electrons are released, generating a current that we can measure. Light is an electromagnetic field, so typically has an electric and magnetic component, but the physics of the photoelectric effect is essentially dependent on the electric component of the field.

The electric field at a point $\mathbf{r} \in \mathbb{R}^3$ in space and time t is modeled by a square-integrable function $\mathbf{E} : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3$ of space and time. The three components of $\mathbf{E}(\mathbf{r}, t)$ correspond to what is called in physics the *polarization* of light. Throughout this paper, for simplicity, we assume that the field is linearly polarized. In this section,⁵ linear polarization amounts to assuming the existence of a unit vector $\mathbf{u} \in \mathbb{R}^3$ such that, for all $(\mathbf{r}, t) \in \mathbb{R}^3 \times \mathbb{R}$, $\mathbf{E}(\mathbf{r}, t) = E(\mathbf{r}, t)\mathbf{u}$, with $E(\mathbf{r}, t) \in \mathbb{R}$. This simplification allows us to focus here on the scalar function $E : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}$. Moreover, we consider a photodetector placed at a fixed position $\mathbf{r} \in \mathbb{R}^3$. The effect of the field on the detector is assumed to only depend on the value $E(\mathbf{r}, \cdot)$ of the field at the detector position, and we thus further focus on the function $t \mapsto E(\mathbf{r}, t)$ in our examples, which we also denote⁶ by E in this section.

We are interested in the times at which our detector *clicks*, i.e., detects a single photon. Because the measured times vary from one run of the experiment to the next, we want to model them as a random set of (distinct) real numbers. \diamond

Finally, note that we only describe in this section an idealized version of the HBT experiment, following [Macchi, 1975, Section 4.2]. For physics arguments, we refer the reader with little prior exposition to physics to [Mandel and Wolf, 1995, Chapter 9], which is a recent textbook treatment of the survey [Mandel and Wolf, 1965] to which Macchi [1975] originally referred.

⁵In later sections, the field will be modeled by a collection of operators, and linear polarization will thus correspond to a different mathematical assumption.

⁶Overloading variable names is common in physics, and we shall follow this convention when possible without confusion.

2.1 The correlation functions of a point process

Let us fix a complete metric space \mathbb{X} , with μ a Borel measure on the Borel sets of \mathbb{X} . A majority of the examples in this paper will deal with $\mathbb{X} = \mathbb{R}$ equipped with the Lebesgue measure. As mentioned in the header of Section 2, this choice describes the ideal detection times of physical particles. In spatial statistics, we typically encounter $\mathbb{X} = \mathbb{R}^d$ with μ the Lebesgue measure [Baddeley et al., 2015]. It is also not uncommon to take \mathbb{X} to be a compact manifold M (like a sphere; [Beltrán and Hardy, 2019]) or a discrete space (like a large dataset in machine learning [Kulesza and Taskar, 2012]), in which case μ is typically taken to be the volume form on M or the counting measure of the discrete space, respectively.

A point process γ on \mathbb{X} is a random configuration⁷ of points in \mathbb{X} . In other words, γ is a random variable taking its values in the space of locally finite subsets of \mathbb{X} ,

$$\text{Conf}(\mathbb{X}) := \left\{ \gamma \subset \mathbb{X} : \#(\gamma \cap A) < \infty \text{ for all compact } A \subset \mathbb{X} \right\}.$$

In terms of modeling beams of physical particles, the *locally finite* assumption entails that for any given time interval, there is a finite number of particles that were detected in that interval.

For any $k \geq 1$, the k -point correlation function $\rho_k : \mathbb{X}^k \rightarrow [0, \infty]$ satisfies, when it exists,

$$\mathbb{E} \left[\sum_{\substack{x_1, \dots, x_k \in \gamma \\ x_i \neq x_j \text{ if } i \neq j}} f(x_1, \dots, x_k) \right] = \int_{\mathbb{X}^k} f(x_1, \dots, x_k) \rho_k(x_1, \dots, x_k) \prod_{i=1}^k d\mu(x_i) \quad (1)$$

for any bounded (or positive) and measurable test function f . Here

$$\mathbb{X}^k = \overbrace{\mathbb{X} \times \dots \times \mathbb{X}}^k$$

and the symbol \mathbb{E} in (1) stands for the expectation⁸ under the law of the random variable γ . Thus, the k -point correlation function ρ_k encodes the distribution of k -tuples of points from γ . Indeed, an informal rewriting of (1) reads

$$\rho^{(k)}(x_1, \dots, x_k) \mu(dx_1) \dots \mu(dx_k) = \mathbb{P} \left(\begin{array}{l} \text{There are at least } k \text{ points in } \gamma, \\ \text{one in each of the infinitesimal} \\ \text{balls } B(x_i, dx_i) \text{ for } i = 1, \dots, k \end{array} \right). \quad (2)$$

It is common to actually define a point process γ by a sequence (ρ_k) of *compatible* correlation functions; see [Daley and Vere-Jones, 2003]. By *compatible*, we mean that not every sequence (ρ_k) actually defines a point process, and that usually a mathematical argument for existence is necessary.

Of particular practical significance are the correlation functions for $k = 1$ and $k = 2$. The first correlation function ρ_1 describes the marginal distribution of particles, and is called the *intensity* in probability and statistics.⁹ The second correlation function ρ_2 describes pairwise correlations, and is often discussed in its normalized form

$$g(x, y) = \frac{\rho_2(x, y)}{\rho_1(x)\rho_1(y)}. \quad (3)$$

⁷*Stricto sensu*, we are defining here *simple* point processes, i.e., such that the samples never contain a given point of \mathbb{X} more than once. Since all point processes in this work are simple, we take this shortcut and identify a (simple) point process with a random configuration.

⁸Physicists speak of an ensemble average, and denote it using angle brackets $\langle \cdot \rangle$. We stick with the symbol \mathbb{E} for classical expectations, and we reserve brackets only for Hilbert space inner products and quantum averages.

⁹Physicists might prefer to call it *density of particles*, and reserve *intensity* for other physical quantities. To avoid confusion, we will write *first correlation function* in full.

Finally, when $\mathbb{X} = \mathbb{R}^d$ and the distribution of γ is translation- and rotation-invariant, g only depends on $r = \|x - y\|$, so that we write $g(x, y) = g_0(r)$. The function g_0 is called the *pair correlation function* of γ .

Figure 1 shows samples of three translation-invariant point processes in \mathbb{R} , along with their pair correlation functions. The three point processes share the same (constant) first-order correlation function. Note how the pair correlation function in Figure 1(f) is lower than 1 close to 0, indicating fewer small pairwise distances than the reference point process in Figure 1(d): this is a sign of a very regular, more grid-like distribution of the points, as seen in Figure 1(c). On the contrary, the pair correlation function of Figure 1(e) shows more small pairwise distances than the reference: points are lumped together, as confirmed by Figure 1(b). The rest of this section describes these three examples in more detail.

2.2 Poisson and Cox point processes

Let $\lambda : \mathbb{X} \rightarrow \mathbb{R}_+$ be locally integrable, that is, $\int_B \lambda(x) d\mu(x) < \infty$ for every bounded $B \subset \mathbb{X}$. Further assume for simplicity that the measure $\lambda d\mu$ has no atom.¹⁰ The point process with correlation functions

$$\rho_k(x_1, \dots, x_k) = \lambda(x_1) \dots \lambda(x_k), \quad k \geq 1, \quad (4)$$

always exists and is called the *Poisson point process* with parameter function¹¹ λ . The first correlation function of the process is thus $\rho_1 = \lambda$. Moreover, the separable form of (4) implies the lowest level of correlation among the points of the process. In particular, the pairwise correlation function $g \equiv 1$ is constant, so that no pairwise distances are preferred. Figure 1(a) shows a few Poisson samples where μ is the Lebesgue measure on $\mathbb{X} = \mathbb{R}$ and λ is constant.

Example 1 (Deterministic electric fields yield Poisson point processes). Consider the photodetection setup of Example 0. Let us further assume that the field is *quasi-monochromatic*, i.e., that the modulus $|\mathcal{F}E|$ of the Fourier transform of the function $E : \mathbb{R} \rightarrow \mathbb{R}$ concentrates around ω and $-\omega$ for a single value of the frequency $\omega > 0$. Note that the symmetry of $|\mathcal{F}E|$ is a consequence of E being real-valued. Denote by E^+ the *analytic signal* of the electromagnetic field E , i.e.,

$$E^+ = 2\mathcal{F}^{-1}(\mathcal{F}(E) \times 1_{(0, \infty)}). \quad (5)$$

Taking the analytic signal [Mandel and Wolf, 1995, Section 3.1] is a partial isometry of $L^2(\mathbb{R})$ commonly used in signal processing, which has several natural properties. For instance, E^+ removes the symmetry in $|\mathcal{F}E|$, so that each frequency is represented in an unambiguous, non-redundant manner. To a mathematician, the analytic signal is the boundary value of a particular analytic function of the upper-half plane, defined as the Cauchy transform of f ; see e.g. [Pugh, 1982, Section 2.1]. A more physical property is that when E is quasi-monochromatic, $|E^+|$ is a good approximation of the envelope of the signal E ; see Figure 2. This idea of an envelope, insensitive to rapid oscillations, along with intuition from classical electromagnetism relating a physical intensity to the square of the amplitude of a wave, is at the origin of the following modeling assumption, which we flesh out for future reference.

Assumption 1. *For a deterministic field E , the detection times follow a nonstationary Poisson process, with parameter function proportional to $t \mapsto |E^+(t)|^2$, where E^+ is given by (5).*

The full justification for the assumption can be obtained by a more precise model of light and matter interaction using quantum theory; see [Mandel and Wolf, 1995, Chapter 9] or our Section 4. Meanwhile, we see that Poisson processes naturally result from deterministic fields. \diamond

¹⁰Without this assumption, a Poisson point process would not necessarily be a *simple* point process.

¹¹Statisticians call this parameter function the *intensity* of the process. We refrain from using this naming convention, to avoid confusion with any physical intensity.

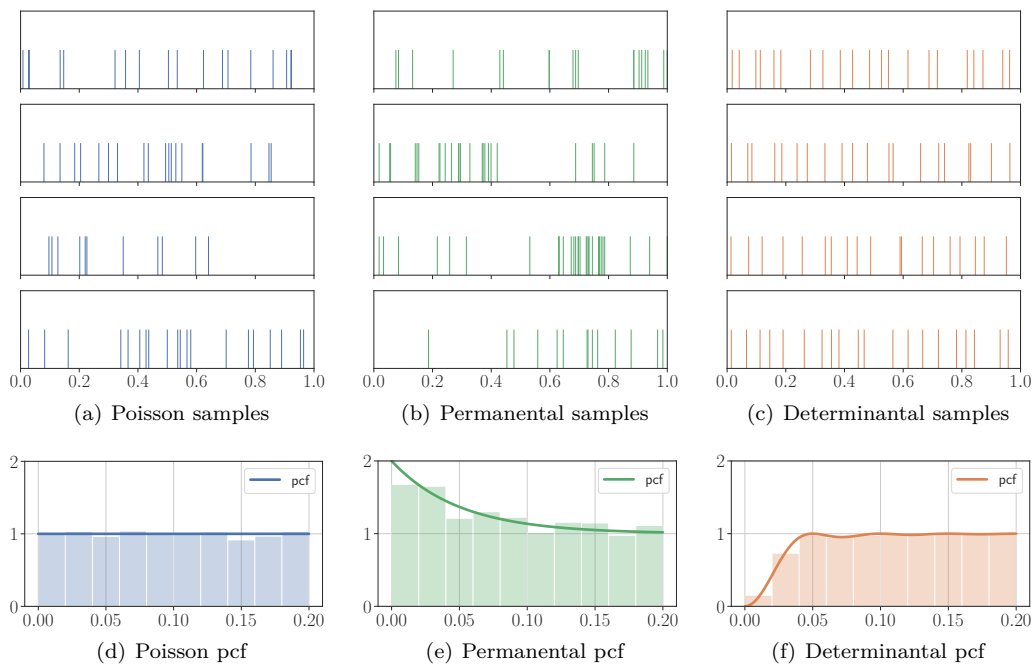


Figure 1: (a, b, c) The intersection with $W = [0, 1]$ of four samples of each of the three types of point processes introduced in Section 2. All three point processes are translation-invariant and are scaled to have the same expected number of points falling in W . (d, e, f) The pair correlation function (3) of the three point processes, and histograms of pairwise distances from 100 independent samples.

Poisson point processes are useful both as a reference point process and as a building block in statistical modeling. For instance, a Poisson point process with a random parameter function λ is known as a *Cox process* [Daley and Vere-Jones, 2003, Section 6.2]. A Cox process has correlation functions

$$\rho_k(x_1, \dots, x_k) = \mathbb{E}\lambda(x_1) \dots \lambda(x_k), \quad k \geq 1,$$

where the expectation is over λ . In particular, by Jensen's inequality,

$$\rho_2(x, x) = \mathbb{E}\lambda(x)^2 \geq (\mathbb{E}\lambda(x))^2 = \rho_1(x)^2. \quad (6)$$

If inequality is strict and ρ_2 is continuous, one can thus expect that for x, y close to each other, the probability that there is a point near x and a point near y in the same realization is larger for a Cox process than for a Poisson process with the same first correlation function. In other words, samples from a Cox process exhibit clusters, or *bunching*.

Example 2 (Random classical sources imply photon bunching). When the physical field E comes from a thermal source (i.e., incandescent matter or a gas discharge), experimentally observed detection events do not seem to form a Poisson process, but rather exhibit some form of *bunching*, or clustering. This is the HBT effect [Hanbury Brown and Twiss, 1958]. Mathematically, it can be seen as a consequence of Assumption 1, as soon as one represents the field resulting from the thermal source by a stochastic process making $x, y \mapsto \rho_2(x, y)$ smooth near the diagonal $x = y$. The point process of detection times is then a Cox process that favors clusters of points, compared to a Poisson process. \diamond

The archetypal Cox processes are the so-called permanental point processes, which we now introduce.

2.3 Permenal point processes

A point process γ is said to be *permenal* when there exists a so-called *correlation kernel* $K : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{C}$ such that the correlation functions of γ read

$$\rho_k(x_1, \dots, x_k) = \text{per} \left[K(x_i, x_j) \right]_{i,j=1}^k, \quad k \geq 1. \quad (7)$$

The permanent $\text{per} \mathbf{A}$ in (7) of a matrix $\mathbf{A} \in \mathbb{R}^{k \times k}$ is defined, analogously to the determinant, by a sum over permutations

$$\text{per} \mathbf{A} = \sum_{\sigma \in \mathfrak{S}_k} \prod_{i=1}^k a_{i\sigma(i)},$$

where \mathfrak{S}_k denotes the symmetric group. Assuming existence for a moment, the point process described by (7) has nontrivial correlation functions. Taking $k = 1$ and $k = 2$ in (1), we obtain for instance

$$\mathbb{E} \left[\sum_{x \in \gamma} f(x) \right] = \int f(x) K(x, x) \mu(dx), \quad (8)$$

$$\mathbb{E} \left[\sum_{\substack{x, y \in \gamma \\ x \neq y}} f(x, y) \right] = \int f(x, y) \left[K(x, x)K(y, y) + K(x, y)K(y, x) \right] \mu(dx) \mu(dy). \quad (9)$$

When the kernel K is Hermitian, that is, $K(x, y) = \overline{K(y, x)}$, the second correlation function in (9), once normalized, becomes

$$g(x, y) = \frac{\rho_2(x, y)}{\rho_1(x)\rho_1(y)} = 1 + \frac{|K(x, y)|^2}{K(x, x)K(y, y)} \geq 1. \quad (10)$$

In particular, (10) shows that permenal point processes are attractive: the larger $|K(x, y)|^2$, the more likely two particles at x and y are to co-occur. Moreover, they are always more likely to co-occur than if γ were a Poisson process with the same first correlation function.

The existence of a point process satisfying (7) requires conditions on K [Shirai and Takahashi, 2003]. A standard set of conditions comes from a representation of γ as a Cox process. This representation derives from the semiclassical treatment of the HBT effect for photons by Macchi [1975], which we present below as Example 3. To isolate the mathematical statement, let us simply mention that a Poisson point process with random first correlation function λ taken to be the squared modulus of a Gaussian process is a permenal point process; see e.g. [Hough et al., 2006, Proposition 35]. Making sure that the underlying Gaussian process exists in turn guarantees existence of the attached permenal point process.

Example 3 (Gaussian classical fields yield permenal point processes). Continuing Example 2 on thermal sources, quasimonochromatic thermal sources are actually represented by zero-mean stationary Gaussian processes, following the intuition that they result from the superposition of many zero-mean, independent random contributions from the source at roughly the same frequency [Macchi, 1975, Section 4.2]. Mathematically, a random function f is said to have for distribution a zero-mean Gaussian process with kernel K if, for any number n of observations and for any times $t_1, \dots, t_n \in \mathbb{R}$, we have

$$(f(t_1), \dots, f(t_n))^T \sim \mathcal{N} \left((0, \dots, 0)^T, ((K(t_i, t_j))_{1 \leq i, j \leq n}) \right), \quad (11)$$

where $\mathcal{N}(\mu, \Sigma)$ stands for the multivariate Gaussian with mean μ and covariance matrix Σ . Assuming E is such a zero-mean Gaussian process, and that its distribution is invariant

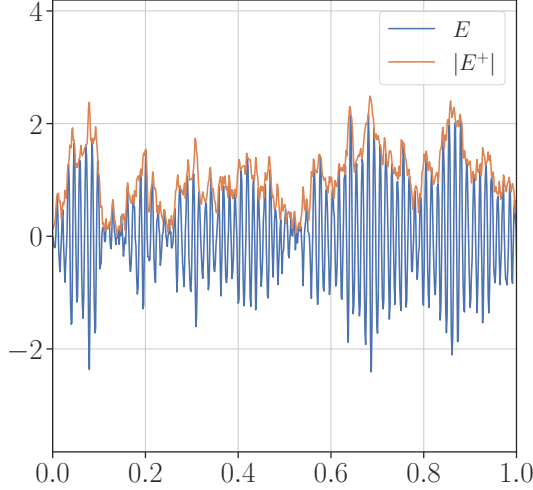


Figure 2: A sample E of the zero-mean Gaussian process with Lorentz kernel (12), and the modulus $|E^+|$ of the corresponding analytic signal.

under translations along the time axis, the linear transform E^+ in (5) is also a zero-mean, translation-invariant Gaussian process,¹² with the specific property that $\mathbb{E}E^+(t)E^+(s) = 0$.

As an example, the reader can think of E as having kernel

$$K_{\text{Lorentz}}(t, s) = \exp\left(-\frac{|t-s|}{\sigma}\right) \cos(\omega(t-s)) = K_{\text{Lorentz},0}(t-s), \quad (12)$$

for some $\omega \gg 1/\sigma$. K_{Lorentz} is the product of a slowly varying envelope, which we arbitrarily take to be exponential for concreteness, and a fast oscillating function. In particular, the Fourier spectrum of $K_{\text{Lorentz},0}$ is concentrated around ω , making E quasi-monochromatic in the sense of [Macchi, 1975]. In that case, E^+ has kernel

$$C_{\text{Lorentz}}(t, s) = 2K_{\text{Lorentz},0}^+(t-s) \approx 2\exp\left(-\frac{|t-s|}{\sigma}\right) e^{i\omega(t-s)}. \quad (13)$$

The first equality is a general property of analytic transforms of second-order stationary processes, and can be proved by direct computation. The approximation in (13) is a consequence of Bedrosian's theorem, stating that the slowly varying envelope is preserved when taking the analytic signal; see e.g. [Picinbono, 1997]. Finally, we note that the covariance (13) is a slight modification of the Lorentz kernel example given by [Macchi, 1975], where we have introduced a phase factor since, *stricto sensu*, analytic covariance kernels cannot take only real values. A sample E with kernel K_{Lorentz} is shown in Figure 2. The modulus of the analytic signal E^+ visibly plays the role of an envelope for E , averaging over local oscillations of E .

We now go back to a generic E following a zero-mean Gaussian process with kernel K . Following Assumption 1, the detection events now form a Poisson process with random parameter function, a.k.a. a Cox process; see Section 2.2. Its correlation functions read

$$\rho_k(t_1, \dots, t_k) = \mathbb{E}|E^+(t_1)|^2 \dots |E^+(t_k)|^2 = \mathbb{E}E^+(t_1) \dots E^+(t_k) \overline{E^+(t_1)} \dots \overline{E^+(t_k)}.$$

Now the expectation of a product of Gaussians can be expressed in terms of pairwise expectations using a theorem by Isserlis.¹³ Further recalling that $\mathbb{E}E^+(t)E^+(s) = 0$ for all $t \neq s$, we obtain

$$\rho_k(t_1, \dots, t_k) = \sum_{\sigma \in \mathfrak{S}_k} \prod_{i=1}^k \mathbb{E}E^+(t_i) \overline{E^+(t_{\sigma(i)})}. \quad (14)$$

¹²However, E^+ is complex-valued, i.e., the vector in (11) is a complex multivariate Gaussian vector.

¹³This theorem is also known to physicists as Wick's theorem, by analogy with a similar theorem for the quantum average of products of certain operators; see later in Section 3.5.

Recognizing a permanent in (14), we conclude that

$$\rho_k(t_1, \dots, t_k) = \text{per}((C(t_i, t_j))), \quad k \geq 1,$$

where $C(t, s) = \mathbb{E}E^+(t)\overline{E^+(s)}$ is the covariance kernel of the Gaussian process E^+ . The photon detection process is thus a permanental point process, see Section 2.3, with kernel the covariance kernel of the analytic signal of the electromagnetic field. In particular, for a Hermitian and translation-invariant kernel $C(t, s) = C_0(t - s) = \overline{C_0(s - t)}$, we obtain, for all s, t such that $t - s = r$,

$$g(r) = \frac{\rho_2(t, s)}{\rho(t)\rho(s)} = 1 + |C_0(r)|^2. \quad (15)$$

The probability of coincidence of a pair of detection times is thus larger than under a Poisson process with the same first correlation function. The pair correlation function g in Figure 1(e) is actually (15) with the Lorentz kernel in (13). The corresponding samples exhibit bunching when compared to the Poisson samples of Figure 1(a). \diamond

Equation (15) provides a remarkably simple mathematical derivation of photon bunching in terms of correlation functions, due to [Macchi, 1975, Chapter 4]. It is noteworthy that the representation of permanental point processes as Cox processes, which has become more of a side result for probabilists [Hough et al., 2006, Proposition 35], was actually the physical motivation for the introduction of permanental point processes by Macchi [1975].

2.4 Determinantal point processes

A point process γ is said to be *determinantal* when there exists a so-called *correlation kernel* $K : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{C}$ such that the correlation functions of γ read, for any $k \geq 1$,

$$\rho_k(x_1, \dots, x_k) = \det \left[K(x_i, x_j) \right]_{i,j=1}^k. \quad (16)$$

We write DPP(K, μ) for the point process with correlation functions (184) with respect to the reference measure μ . Note that, by definition of ρ_k , the kernel $K(x, y)$ has to be chosen so that the right hand side of (184) is nonnegative for any $k \geq 1$ and $x_1, \dots, x_k \in \mathbb{X}$; like for permanental point processes, not every kernel yields a well-defined DPP.

Assuming existence for a moment, we look at the first and second correlation functions. Taking $k = 1$ and $k = 2$ in (1), we obtain for instance

$$\mathbb{E} \left[\sum_{x \in \gamma} f(x) \right] = \int f(x) K(x, x) \mu(dx), \quad (17)$$

$$\mathbb{E} \left[\sum_{\substack{x, y \in \gamma \\ x \neq y}} f(x, y) \right] = \int f(x, y) \left[K(x, x) K(y, y) - K(x, y) K(y, x) \right] \mu(dx) \mu(dy). \quad (18)$$

When the kernel K is Hermitian, that is, $K(x, y) = \overline{K(y, x)}$, the second correlation function in (18), once normalized, becomes:

$$g(x, y) = \frac{\rho_2(x, y)}{\rho_1(x)\rho_1(y)} = 1 - \frac{|K(x, y)|^2}{K(x, x)K(y, y)} \leq 1. \quad (19)$$

In particular, (19) shows that DPPs with Hermitian kernels are repulsive: the larger $|K(x, y)|^2$, the less likely the two particles at x and y are to co-occur. Moreover, they are always less likely to co-occur than if γ were a Poisson process with the same first correlation function. In particular, if γ were to model the clicks of a detector as particles arrive, the arrival times would exhibit *antibunching*. We shall see examples of antibunching arrival

times in Section 5. Unlike Poisson and permanental point processes, whose appearance results from Assumption 1, to understand why DPPs appear in quantum optics, we shall need first to introduce elements of quantum field theory.

Necessary and sufficient conditions for existence of a DPP are known when the operator $\mathcal{K} : f \mapsto \int K(\cdot, y)f(y)d\mu(y)$ on $L^2(\mu)$ is Hermitian and locally trace-class. In particular, letting (λ_k) denote the eigenvalues of \mathcal{K} , $\text{DPP}(K, \mu)$ is well-defined if and only if $0 \leq \lambda_k \leq 1$ for any $k \in \mathbb{N}$, namely when K is a contraction operator. This is now known as the Macchi-Soshnikov theorem.

One constructive proof of this existence theorem relies on the following decomposition lemma [Hough et al., 2006]. Informally, a DPP with kernel

$$K(x, y) = \sum_{k=0}^{\infty} \lambda_k \varphi_k(x) \overline{\varphi_k(y)}, \quad (20)$$

is a statistical mixture of *projection DPPs*, i.e., DPPs with projection kernels. More precisely, if $(b_k)_{k \in \mathbb{N}}$ are independent Bernoulli random variables with $\mathbb{P}(b_k = 1) = 1 - \mathbb{P}(b_k = 0) = \lambda_k$ and $K^b(x, y) := \sum_k b_k \varphi_k(x) \overline{\varphi_k(y)}$, then $\text{DPP}(K^b, \mu) = \text{DPP}(K, \mu)$ in law; see again [Hough et al., 2006]. In particular, if we restrict ourselves to projection kernels, i.e., if $\lambda_k \in \{0, 1\}$ for every k , the cardinality of $\gamma \sim \text{DPP}(K, \mu)$ is thus $\text{Tr } K$ almost surely. If we further assume $\text{Tr } K = N < \infty$, the corresponding projection DPP generates exactly N particles, with joint probability distribution on \mathbb{X}^N given by

$$\frac{1}{N!} \left| \det \left[\varphi_{k-1}(x_j) \right]_{j,k=1}^N \right|^2 \prod_{j=1}^N \mu(dx_j). \quad (21)$$

3 Elements of quantum field theory

To deepen our understanding of the relationship between point processes and quantum physics, we now introduce a few elements of what is known as many-body quantum physics, quantum statistical physics or quantum field theory. We first introduce the basic quantum formalism to describe a single particle, its evolution in time and what information we can gather on it, mostly following [Folland, 2008]. Then, we introduce how to deal with systems made of many or an indefinite number of particles.

3.1 The mathematical framework of quantum theory

States. The state ψ of a quantum system is represented by a non-zero element of a complex Hilbert space $(\mathcal{H}, \langle \cdot | \cdot \rangle)$, the space of all possible states of the quantum system of interest. Two elements of \mathcal{H} that are equal up to multiplication by a (complex) scalar are understood to represent the same physical states,¹⁴ and we henceforth always assume that our states are normalized, i.e. $\|\psi\| = \langle \psi | \psi \rangle = 1$. Finally, we will often use the so-called bra-ket notation, denoting a vector by $|\psi\rangle = \psi \in \mathcal{H}$ and the associated linear form on \mathcal{H} by $\langle \psi | : h \mapsto \langle \psi | h \rangle$.

Observables. Any observable quantity is represented by a self-adjoint operator $A : \mathcal{H} \rightarrow \mathcal{H}$. To understand in what sense, take a self-adjoint operator $A : \mathcal{H} \rightarrow \mathcal{H}$ and a vector $|\psi\rangle \in \mathcal{H}$, and define an ordinary probability measure $\mathbb{P}_{A,\psi}$ on the spectrum $\sigma(A) \subset \mathbb{R}$ of A by

$$\mathbb{P}_{A,\psi} : E \mapsto \langle \psi | \mathbb{1}_E(A) | \psi \rangle, \quad E \subset \sigma(A). \quad (22)$$

The operator $\mathbb{1}_E(A)$ in (22) is defined by the *spectral functional calculus*; see [Reed and Simon, 1980, Chapter VII]. Informally, $\mathbb{1}_E(A)$ is the projection operator with the same eigenvectors as A , but replacing each eigenvalue λ by 1 if $\lambda \in E$, and 0 otherwise. One easy setting to understand the definition of $\mathbb{P}_{A,\psi}$ is that of a compact (self-adjoint) operator

¹⁴For brevity, we will stay informal on that point and ignore the subtleties coming from considering a projective Hilbert space rather than \mathcal{H} ; see [Folland, 2008, Chapter 3].

A , for in that case the support $\sigma(A) = \{\lambda_n\}$ of $\mathbb{P}_{A,\psi}$ is discrete, with no other accumulation point than 0, and there is an orthonormal basis $(|\varphi_n\rangle)$ of \mathcal{H} with $A|\varphi_n\rangle = \lambda_n|\varphi_n\rangle$; see e.g. [Reed and Simon, 1980, Theorem VI.16]. In particular, for $\lambda_n \neq 0$ with multiplicity one, $\mathbb{P}_{A,\psi}(\{\lambda_n\}) = |\langle\psi|\varphi_n\rangle|^2$, which indeed sums to $\|\psi\|^2 = 1$ in n . Without entering into details, we finally note that the definition (22) extends to unbounded operators by representing them as multiplication operators [Reed and Simon, 1980, Theorems VIII.4 and VIII.5]. Unbounded operators are necessary to represent physical observables that have an unbounded support, such as the position of a particle on the real line.

Interpreting a self-adjoint operator A as an observable means that the experimentally accessible information about the observable A is in $\mathbb{P}_{A,\psi}$. In particular, if we had a machine that could repeatedly prepare unrelated copies of the physical system in question in a state $|\psi\rangle$, we would model the measurements of observable A on such a sequence of states as independent draws from $\mathbb{P}_{A,\psi}$. Unless $|\psi\rangle$ is an eigenvector of A with eigenvalue λ , in which case $\mathbb{P}_{A,\psi} = \delta_\lambda$, measuring the physical quantity encoded by A thus becomes inherently probabilistic: one can only talk about the probability that a measured observable will lie in a given $E \subset \mathbb{R}$. Finally, we note that by the law of large numbers, the average of a large number of measurements of A is expected to be close to the expectation of a random variable with law $\mathbb{P}_{A,\psi}$, which is

$$\int \lambda d\mathbb{P}_{A,\psi}(\lambda) = \langle\psi|A|\psi\rangle. \quad (23)$$

Again, the derivation of (23) is easier for a compact operator $A = \sum_n \lambda_n |\varphi_n\rangle\langle\varphi_n|$, say even Hilbert-Schmidt, so that $\sum \lambda_n^2 < \infty$. In that case, both sides of (23) are equal to $\sum_i \lambda_n |\langle\psi|\varphi_n\rangle|^2 < \infty$. The general treatment of self-adjoint operators requires to carefully define the spectral measure of an operator, and we refer to [Reed and Simon, 1980, Chapters VII & VIII].

Mixed states and Born's rule. In practice, most experimental devices are not able to repeatedly produce a given state $|\psi\rangle \in \mathcal{H}$, but rather a noisy version of it. To model this noise, first associate to each state $|\psi\rangle \in \mathcal{H}$ the projector $|\psi\rangle\langle\psi| : |h\rangle \mapsto \langle\psi|h\rangle |\psi\rangle$, which we also abusively call a state. Now consider a linear operator on \mathcal{H} defined as

$$\rho = \mathbb{E}_{|\psi\rangle \sim p} |\psi\rangle\langle\psi|, \quad (24)$$

with p a probability measure on \mathcal{H} . The linear combination Equation (24) is typically interpreted as being the output of a noisy preparation device, which outputs the state $|\psi\rangle\langle\psi|$ with $|\psi\rangle$ drawn from p . When the agent runs its machine many times, they will describe the output by Equation (24).¹⁵ More abstractly, a *density matrix* ρ is a trace-class positive operator on \mathcal{H} with unit trace.¹⁶ A density matrix is often called a *mixed state*, as opposed to a *pure state*, which corresponds to a single element of \mathcal{H} . In other words, pure states correspond to density matrices that are also projectors.

Generally speaking, a density matrix ρ exhausts all the statistical content that an observer can predict about a system. For instance, for ρ defined by (24),

$$\text{Tr}[\rho \mathbf{1}_E(A)] = \mathbb{E} \text{Tr}[|\psi_s\rangle\langle\psi_s| \mathbf{1}_E(A)] = \mathbb{E} \langle\psi_s| \mathbf{1}_E(A) |\psi_s\rangle, \quad E \subset \mathbb{R}.$$

In particular, the map $\mathbb{P}_{A,\rho} : E \mapsto \text{Tr}(\rho \mathbf{1}_E(A))$ is a probability distribution on \mathbb{R} that corresponds to a statistical mixture of pure states, and simplifies to (22) for pure states. It gives the probability that a measurement of A will belong to E . By linearity, the expectation of the random variable¹⁷ describing the measurement of observable A when the system is in the state ρ , denoted by convention $\langle A \rangle_\rho$, is given by *Born's rule*

$$\langle A \rangle_\rho = \text{Tr}(\rho A). \quad (25)$$

¹⁵In other words, a mixed state can be used to describe epistemic uncertainty on the preparation process, while draws from $\mathbb{P}_{A,\psi}$ in (22) represent aleatoric uncertainty. Unlike epistemic uncertainty, aleatoric uncertainty cannot be reduced by better knowledge of the system.

¹⁶The use of the word *matrix* for an operator can be confusing at first glance, but it is standard here.

¹⁷Physicists talk of the *expectation value* of A in state ρ .

This formula is a cornerstone of the quantum formalism. It contains Equation (22) as a special case and thus describes both noisy and noiseless situations.

Conditioning on measurements. To define the joint distribution of the measurement of two observables A and B when the physical system is in a (possibly mixed) state ρ , one would like to use the operator AB in a spectral formula like (22). It turns out to be very natural to define a joint distribution using the spectral calculus when A and B commute, i.e. $AB = BA$. In particular, one can then talk of the conditional distribution of the measurement of B on ρ given that we observed A on ρ . Without entering into details, Bayes' formula yields that evaluating the conditional amounts to evaluating $\mathbb{P}_{B,\tilde{\rho}}$, where

$$\tilde{\rho} = C\rho C^\dagger, \quad (26)$$

and C is related to the measurement of A ; see [Bouten et al., 2007] for a precise statement that includes monitoring a state across time, i.e., filtering. The “sandwiched” updated state $\tilde{\rho}$ in (26) is interpreted as the state of the system immediately after the measurement of A .

Incompatible observables and Heisenberg's uncertainty principle. When $AB \neq BA$, it is not even guaranteed that AB is self-adjoint, let alone that A and B have common eigenspaces. Physicists associate this mathematical difficulty to the fact that non-commuting observables are *incompatible*: it is not possible to obtain a joint measurement of both observables using a *single* prepared copy of a state ρ .

Another hint that non-commuting observables are peculiar is Heisenberg's celebrated *uncertainty principle*; see [Folland, 2008, Section 3.3]. In a nutshell, for two non-commuting observables A and B and a state ρ , the product of the standard deviations $\sigma_{A,\rho}$ of $\mathbb{P}_{A,\rho}$ and $\sigma_{B,\rho}$ of $\mathbb{P}_{B,\rho}$ is lower-bounded,

$$\sigma_{A,\rho}\sigma_{B,\rho} \geq \frac{1}{2} |\text{Tr} [(AB - BA)\rho]|. \quad (27)$$

Since A and B cannot be measured simultaneously, this does not imply anything on measuring A and B on the same copy of ρ , but it is rather a property of the model of the overall experiment, i.e., of the physical system and the probes corresponding to A and B . The concrete consequences of the uncertainty relation (27) can be seen by repeatedly preparing a copy of the state ρ and measuring either A or B . If the measure represented by A yields measurements with a small empirical variance, then measuring B on similar repeated copies of ρ will lead to a comparatively large empirical variance.

Schrödinger's equation and Hamiltonians. To complete our description of a physical system, we need to model its evolution in time. In quantum physics, the evolution of the state of the system is given by a one-parameter group $\{U(t), t \in \mathbb{R}\}$ of unitary operators on \mathcal{H} , that is, $U(t)U(s) = U(t+s)$ and $U(t)^{-1} = U(t)^\dagger = U(-t)$. In particular $U(0) = \mathbb{1}$ is the identity operator. After time t , the state of a system that was in state $|\psi(0)\rangle$ at time $t = 0$ is considered to be

$$|\psi(t)\rangle = U(t) |\psi(0)\rangle. \quad (28)$$

Under weak assumptions on \mathcal{H} and $U(t)$, one can show that $U(t) = \exp(-\frac{it}{\hbar}H)$ in the sense of the spectral functional calculus again, for some (possibly unbounded) self-adjoint operator H on \mathcal{H} [Folland, 2008, Section 3.1]. Conversely, any choice of self-adjoint H gives a one-parameter unitary group. The operator H , thought as an observable, is called the *Hamiltonian* of the system. As the generator of the dynamics, H is often thought as the energy of the system. In particular, the group identity of U yields by differentiation that $U(t)^\dagger H U(t) = H$ for all t , so that the average energy $\langle \psi(t) | H | \psi(t) \rangle$ of the state $\psi(t)$ in (28) is constant over time.

Building a model in quantum physics usually boils down to choosing a state space \mathcal{H} , a set of self-adjoint operators as observables, and a Hamiltonian to describe the evolution

of the system in time. In particular, when $s \ll 1$, (28) informally yields

$$|\psi(t+s)\rangle = \left(\mathbb{1} - \frac{isH}{\hbar} + \mathcal{O}(s^2) \right) |\psi(t)\rangle,$$

so that

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle, \quad (29)$$

which the reader may recognize as the celebrated Schrödinger equation. This equation of motion can be generalized to mixed states (24) by linearity.

To finish, note that instead of letting $|\psi(t)\rangle$ depend on time through (28) and keeping observables constant (the so-called *Schrödinger picture*), we can obtain the same measurement probabilities (22) by keeping states constant and letting observables vary as $A(t) = U(t)^\dagger A(0)U(t)$ (the so-called *Heisenberg picture*). The two pictures are thus equivalent as to what observations they predict. There is a third equivalent convention that is commonly used, the *interaction picture*. The latter is tailored to problems where the Hamiltonian

$$H = H_0 + H_I$$

is the sum of a well-studied Hamiltonian H_0 , typically describing standard prior information on a subpart of the physical system (say, a *free field*), and an additional term H_I , called *interaction Hamiltonian*, typically encoding the interaction between parts of the system. In the interaction picture, one lets both states and observables evolve in time. Observables evolve according to $A(t) = U_0(t)^\dagger A(0)U_0(t)$, using the one-parameter unitary group U_0 associated to Hamiltonian H_0 . On the other hand, states evolve as $|\psi(t)\rangle = V(t)|\psi(0)\rangle$, where $V(t) = U_0(t)^\dagger U(t)$ and $U(t)$ is the group associated to the full Hamiltonian H . While the state evolution now differs from (29), the interaction picture yields again the same probabilities (22). Unless otherwise specified, we use the interaction picture in this work.

Perturbation theory. What is convenient with the interaction picture is that it singles out the role of H_I as follows. To compute $U(t)$, it is enough to compute V and then apply $U(t) = U_0(t)V(t)$, since $U_0(t)$ is assumed to be known. Now, by definition, V satisfies

$$\frac{d}{dt} V(t) = \frac{i}{\hbar} U_0(t)^\dagger H_0 U(t) - \frac{i}{\hbar} U_0(t)^\dagger H U(t) = \frac{1}{i\hbar} H_I(t) V(t). \quad (30)$$

Even in seemingly simple physical situations, as when modeling the detection of photons in Section 4.1, solutions to (30) have to be approximated. Writing (30) in integral form, and noting that $V(0) = \mathbb{1}$, we obtain

$$V(t) = V(0) + \int_0^t \frac{d}{d\tau} V(\tau) d\tau = \mathbb{1} + \frac{1}{i\hbar} \int_0^t H_I(\tau) V(\tau) d\tau. \quad (31)$$

Plugging (31) into itself yields

$$V(t) = \mathbb{1} + \frac{1}{i\hbar} \int_0^t H_I(\tau) d\tau - \frac{1}{\hbar^2} \int_0^t d\tau \int_0^\tau d\tau' H_I(\tau) H_I(\tau') V(\tau'). \quad (32)$$

Iteratively plugging (31) into (32) yields a series representation for V , the study of which is called *perturbation theory*, and is at the heart of quantum field theory and its current mathematical difficulties [Folland, 2008, Chapter 6]. In practice, physicists will often assume that stopping this iterative process early on yields a good approximation to V . For instance, we shall see in Section 4.1 a physical assumption that is mathematically interpreted as keeping only the first two terms in the right-hand side of (32).

3.2 Two fundamental one-particle systems

We saw in Section 3.1 that a quantum model consists in a Hilbert space \mathcal{H} , a Hamiltonian H specifying the dynamics, and self-adjoint operators that describe observables. To make things more concrete, we now describe two simple systems: the qubit and the harmonic oscillator.

3.2.1 The qubit

Consider a physical system whose state is described by one of two labels, say $\{e, g\}$, like whether a two-level atom is excited or in its ground state. One would then use $\mathcal{H} = \mathbb{C}^2$ with the usual inner product. Let now $(|e\rangle, |g\rangle)$ be any orthonormal basis of \mathcal{H} , and $|\psi\rangle = \alpha|e\rangle + \beta|g\rangle$ be a state. Normalization implies $|\alpha|^2 + |\beta|^2 = 1$.

Observables are described by self-adjoint operators of \mathcal{H} . For instance, the observable corresponding to checking whether the system is in state $|e\rangle$ is the projector $|e\rangle\langle e|$. Following (22), the probability to obtain the result e in the state $|\psi\rangle$ is given by

$$\mathbb{P}_{|e\rangle\langle e|, \psi}(\{e\}) = \langle e | (|\psi\rangle\langle\psi|) | e \rangle = |\langle e | \psi \rangle|^2 = |\alpha|^2 = 1 - \mathbb{P}_{|e\rangle\langle e|, \psi}(\{g\}) = 1 - |\beta|^2.$$

More generally, every self-adjoint operator on \mathbb{C}^2 can be expressed as a linear combination of four matrices. Expressed in the basis $(|e\rangle, |g\rangle)$, the four so-called *Pauli matrices* are

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (33)$$

All four matrices squared are the identity, so they are all diagonalizable with spectrum $\{\pm 1\}$. For instance, an orthonormal basis of eigenvectors of σ_x is given by

$$|\pm_x\rangle = 2^{-1/2}(|e\rangle \pm |g\rangle),$$

which is simply a rotation of the basis $(|e\rangle, |g\rangle)$. We can compute for instance the probability (22) of observing the outcome $+_x$, when the state is a generic $|\psi\rangle = \alpha|e\rangle + \beta|g\rangle$, namely

$$\begin{aligned} \mathbb{P}_{|+_x\rangle\langle+_x|, \psi}(\{+_x\}) &= \langle \psi | (|+_x\rangle\langle+_x|) | \psi \rangle \\ &= |\langle \psi |+_x\rangle|^2 \\ &= \frac{|\alpha + \beta|^2}{2} \leq 1. \end{aligned}$$

Similarly, the average value of the observable σ_x is given by

$$\langle \psi | \sigma_x | \psi \rangle = \frac{|\alpha + \beta|^2}{2} - \frac{|\alpha - \beta|^2}{2} = 2\Re(\alpha\beta^*).$$

Finally, we give the example of a rather common and simple Hamiltonian for the qubit, called *free evolution*. In the basis we have chosen, it reads

$$H_{\text{FQ}} = \frac{1}{2} \hbar \omega_{eg} \sigma_z, \quad (34)$$

where ω_{eg} is the transition frequency between the excited ($|e\rangle$) and ground ($|g\rangle$) states, and the label FQ stands for *free qubit*. The resulting evolution in time is easy to understand. The states $|g\rangle$ and $|e\rangle$ are eigenstates of H_{FQ} . Calling U_0 the one-parameter group corresponding to H_{FQ} , and starting e.g. at state $|g\rangle$ at $t = 0$, the probability of finding the evolved state $U_0(t)|g\rangle$ in state $|g\rangle$ is $\langle g | U_0(t) | g \rangle = 1$. The basis vectors $|e\rangle$ and $|g\rangle$ are thus stationary states. Finally, note that the form (34) of the free dynamics is mostly conventional; we could for instance add a term proportional to the identity to shift the energy reference without changing the physics. What matters is that the energy difference between the two states is $\hbar\omega_{eg}$.

Richer physical behaviors are modeled by adding interaction terms to the free Hamiltonian. There is no general prescription as to how to write interaction Hamiltonians and they mostly depend on the physical context we wish to describe. One simple but useful model is to consider a qubit interacting with an external field $\mathbf{B}(t)$. Note that we mean here a classical field, in the sense that $\mathbf{B}(t)$ is a vector $(B_x(t), B_y(t), B_z(t))^T$ of functions from \mathbb{R} to \mathbb{R} . Concretely, the model describes a small magnet or the spin of a particle in an external magnetic field. The basic ideas behind nuclear magnetic resonance, and magnetic resonance

imaging, come from this choice of interaction Hamiltonian [Cohen-Tannoudji et al., 2019a, Chapter IV, Appendix F]. Adding the free part, we obtain the total Hamiltonian

$$H = \frac{1}{2}\omega_{eg}\sigma_z + g\boldsymbol{\sigma} \cdot \mathbf{B}(t), \quad (35)$$

where $g \in \mathbb{R}$ is a coupling constant and $\boldsymbol{\sigma} \cdot \mathbf{B}(t)$ is a shorthand notation for the linear combination of matrices $\sigma_x B_x(t) + \sigma_y B_y(t) + \sigma_z B_z(t)$. When the field is time-independent, the dynamical problem can be interpreted geometrically. Indeed, the Hamiltonian is of the general form $H = \hbar\omega \mathbf{n} \cdot \boldsymbol{\sigma}$ with \mathbf{n} a unitary vector. The evolution operator is thus

$$U(t) = e^{-i\omega t \mathbf{n} \cdot \boldsymbol{\sigma}},$$

and we recognize one way of parametrizing an element of the rotation group $SU(2)$, where ωt is the angle of the rotation around the axis \mathbf{n} . The dynamics of a qubit in this model can be represented as a vector, representing the state $|\psi\rangle$, which precesses around an axis fixed by the external field; see [Degiovanni et al., 2021] for details.

3.2.2 The harmonic oscillator

To model the position of a particle living in \mathbb{R}^d , the state space is commonly taken to be $\mathcal{H} = L^2(\mathbb{R}^d)$. A natural observable is the position of the particle. We associate¹⁸ the measurement of the j th coordinate to the multiplication operator $X_j f : x \mapsto x_j f(x)$, defined for all functions f of $\mathcal{H} = L^2(\mathbb{R}^d)$ such that $x \mapsto x_j f(x) \in L^2(\mathbb{R}^d)$. Note that the only candidates to be common eigenvectors of all X_j are delta functions, and are thus not in L^2 . When performing computations, however, it is convenient to consider the tempered distributions

$$\langle x | := \delta_x, \quad x \in \mathbb{R}^d$$

as generalized states, and interpret $\int \delta_a(x)\delta_b(x)dx = \delta(a-b)$ and $f = \int f(x)\delta_x(\cdot)dx$ as making $|x\rangle$ a “generalized orthogonal basis”. In particular, for a smooth state $|\psi\rangle \in \mathcal{H}$, $\langle x|\psi\rangle = \psi(x)$, and the probability distribution (22) corresponding to X_j in a smooth state $|\psi\rangle$ becomes

$$\mathbb{P}_{X_j, \psi}(E_j) = \int_{\mathbb{R} \times \dots \times \mathbb{R} \times E_j \times \mathbb{R} \times \dots \times \mathbb{R}} |\psi(x)|^2 dx, \quad E_j \subset \mathbb{R}.$$

The reader may recognize here the common interpretation of the squared modulus $|\psi(x)|^2$ of the wave function as the probability density function for the position of the particle.

Similar considerations allow for considering the *momentum* operator P_j , defined for $1 \leq j \leq d$ through its Fourier transform

$$\mathcal{F}(P_j f)(k) = \hbar k_j \mathcal{F}(f)(k),$$

for all $f \in \mathcal{H}$ such that $k_j \mathcal{F}(f)(k) \in L^2(\mathbb{R}^d)$. The “generalized basis” of tempered distributions $\langle k | : f \mapsto \mathcal{F}(f)(k)$ allows talking about the *momentum representation* $\langle k|\psi\rangle$ of a smooth state $|\psi\rangle$. Furthermore, by the inverse Fourier transform, we can express the momentum operator P_j in the position basis as a derivative operator,

$$P_j = -i\hbar\partial_j.$$

A natural Hamiltonian in this setting is the so-called quantum harmonic oscillator, modelling the movement of a single particle in a quadratic potential. It writes

$$H = \sum_{j=1}^d \frac{1}{2m} P_j^2 + \frac{1}{2} m\omega^2 Q_j^2, \quad (36)$$

¹⁸Choosing what operator to associate to what classical quantity is a process known as quantization and is a whole research area; see [Folland, 2008] and references therein.

where $m > 0$ is called the mass and $\omega \in \mathbb{R}$ the angular frequency of the oscillator. For ψ a smooth (Schwartz) function, we get

$$\langle x | H | \psi \rangle = -\frac{\hbar^2}{2m} \Delta \psi(x) + \frac{1}{2} m \omega^2 \|x\|_2^2 \psi(x).$$

In particular, the eigenequation $\langle x | H | \psi \rangle = \lambda \psi(x)$ leads to a well-known differential equation. The solutions are tensor products of Hermite functions [Folland, 2008, Section 3.4], and form an orthonormal basis of $L^2(\mathbb{R}^d)$ of (smooth) eigenfunctions of H .

The harmonic oscillator can also be diagonalized algebraically. Define the so-called *ladder operators*

$$a_j := \sqrt{\frac{m\omega}{2\hbar}} \left(Q_j + \frac{i}{m\omega} P_j \right) \quad \text{and} \quad a_j^\dagger := \sqrt{\frac{m\omega}{2\hbar}} \left(Q_j - \frac{i}{m\omega} P_j \right).$$

They satisfy the commutation relation $a_i a_j^\dagger - a_j^\dagger a_i = \delta_{ij}$. The Hamiltonian can now be rewritten as

$$H = \hbar\omega \sum_{j=1}^d \left(a_j^\dagger a_j + \frac{1}{2} \right), \quad (37)$$

and its spectrum is given by that of the number operators $N_j = a_j^\dagger a_j$. Using that the number operator is positive, one can easily show that its eigenvalues are $n_j \in \mathbb{N}$, with corresponding eigenstates $|n_j\rangle$, i.e. $N_j |n_j\rangle = n_j |n_j\rangle$, called the Fock states [Cohen-Tannoudji et al., 2019a, Chapter V]. The corresponding wavefunctions $\langle x | n_j \rangle$ are the Hermite functions discussed above.

The ladder operators were introduced first by Dirac to solve the harmonic oscillator of a single quantum particle. But these operators have taken a life of their own, to describe many-particles quantum states, as we discuss now.

3.3 Modelling a finite number of particles

Section 3.1 focused on describing systems made of one particle. In this section, we show how to combine such systems to describe one made of a (known) finite number of particles.

3.3.1 Subsystems, bosons and fermions

Let us now consider a system of $N < \infty$ identical particles. Concatenating several physical systems is represented by tensor products, so that one is tempted to consider the Hilbert space $\mathcal{H}^{\otimes N} = \bigotimes_{i=1}^N \mathcal{H}$, where \mathcal{H} is the Hilbert space corresponding to a one-particle system. A generic state would thus be a linear combination of states of the form

$$|\psi\rangle = \bigotimes_{i=1}^N |\psi_i\rangle =: |\psi_1 \cdots \psi_N\rangle. \quad (38)$$

However, experiments rather suggest to use a state that encapsulates the property that the particles are *indistinguishable*, i.e., that permuting particles leaves the state invariant. Arguments from (projective) representation theory of the symmetric group [Degiovanni et al., 2021, Chapter 1.2] lead to two types of state spaces, at least when $d > 2$, bosons and fermions. *Bosons* have symmetric states, i.e., they are represented by states $|\psi\rangle \in \mathcal{H}^{\otimes N}$ such that for all permutations $\sigma \in \mathfrak{S}_N$

$$|\psi_1, \dots, \psi_N\rangle = U_\sigma |\psi_1, \dots, \psi_N\rangle := |\psi_{\sigma(1)}, \dots, \psi_{\sigma(N)}\rangle,$$

where we have implicitly defined the unitary representation U_σ of the permutation group on the Hilbert space $\mathcal{H}^{\otimes N}$. In other words, the Hilbert space for N bosons is the symmetric subspace of $\mathcal{H}^{\otimes N}$, namely the range of the orthogonal projector

$$S = \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} U_\sigma.$$

The other type of particles is called *fermion*, and corresponds to antisymmetric states

$$|\psi_1, \dots, \psi_N\rangle = \varepsilon(\sigma) U_\sigma |\psi_1, \dots, \psi_N\rangle, \quad \sigma \in \mathfrak{S}_N,$$

where $\varepsilon(\sigma)$ is the signature of σ . Systems of N fermions are thus represented by states in the antisymmetric subspace of $\mathcal{H}^{\otimes N}$, namely in the range of the orthogonal projector

$$\mathbf{A} = \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} \varepsilon(\sigma) U_\sigma.$$

Note that observables that are permutation-invariant leave both the ranges of \mathbf{S} and \mathbf{A} invariant. If the Hamiltonian is permutation-invariant, bosons thus stay bosons across time, and fermions stay fermions. One particular case is that of N *free* (a.k.a. *non-interacting*) particles. Letting H be the Hamiltonian for one particle, the dynamics of N free particles is given by the Hamiltonian

$$H^{(N)} = \sum_{n=1}^N \mathbf{1} \otimes \dots \otimes \mathbf{1} \otimes \underbrace{H}_{n\text{-th term}} \otimes \mathbf{1} \otimes \dots \otimes \mathbf{1}. \quad (39)$$

3.3.2 Example N -particle states

Consider the harmonic oscillator setting of Section 3.2.2, say with $d = 1$ for simplicity. Since the Hamiltonian is interpreted as measuring the energy of the system, and since the eigenstates of the Hamiltonian H in (36) form a basis of $L^2(\mathbb{R})$, the state with minimal energy is the eigenfunction of H with minimal eigenvalue. In position representation, it is the first Hermite function φ_0 , that is, a normalized Gaussian function.

Now, if we consider N free bosons, each coming with the same 1-dimensional harmonic oscillator Hamiltonian, the ground state of $H^{(N)}$ in (39) is easily seen to be, in position representation,

$$\psi_0^{\text{Sym}}(x_1, \dots, x_N) = \varphi_0(x_1) \dots \varphi_0(x_N),$$

so that, in this ground state, the probability density of having a configuration of N free bosons at $(x_1, \dots, x_N) \in \mathbb{R}^N$ is given by

$$|\psi_0^{\text{Sym}}(x_1, \dots, x_N)|^2 = |\varphi_0(x_1)|^2 \dots |\varphi_0(x_N)|^2.$$

This is the law of N independent random variables with identical distribution $|\varphi_0(x)|^2 dx$, i.e. N i.i.d. Gaussians.

Still within the harmonic oscillator setting, but this time for N free fermions, the situation is less straightforward. With similar eigendecomposition arguments, one can show that the ground state of $H^{(N)}$ on $\mathcal{A}\mathcal{H}^{\otimes N}$ is given by

$$\psi_0^{\text{Asym}}(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \det \begin{bmatrix} \varphi_0(x_1) & \dots & \varphi_0(x_N) \\ \vdots & & \vdots \\ \varphi_{N-1}(x_1) & \dots & \varphi_{N-1}(x_N) \end{bmatrix}, \quad (40)$$

where $\varphi_{k-1} \in L^2(\mathbb{R}^d)$ is the (unit-norm) eigenstate associated with the k^{th} smallest eigenvalue of H , i.e. here, the k^{th} Hermite function. The right hand side of (40) is known as a Slater determinant. In particular, in this ground state, the density probability of having a configuration at $(x_1, \dots, x_N) \in \mathbb{R}^{Nd}$ is given by

$$\begin{aligned} |\psi_0^{\text{Asym}}(x_1, \dots, x_N)|^2 &= \frac{1}{N!} \det \left[\begin{bmatrix} \varphi_0(x_1) & \dots & \varphi_0(x_N) \\ \vdots & & \vdots \\ \varphi_{N-1}(x_1) & \dots & \varphi_{N-1}(x_N) \end{bmatrix} \right]^2 \\ &= \frac{1}{N!} \det \left[K_N(x_i, x_j) \right]_{i,j=1}^N \end{aligned}$$

where

$$K_N(x, y) := \sum_{k=0}^{N-1} \varphi_k(x) \overline{\varphi_k(y)}.$$

Since the φ_k 's are orthonormal, K_N is a projection kernel. In other words, the N positions of the free fermions form a projection DPP on $\mathbb{X} = \mathbb{R}$, as introduced in Section 2.4. In our setting where $d = 1$ and the φ_k are Hermite functions, this DPP is known as the *Gaussian unitary ensemble* (GUE); see e.g. [Dean et al., 2019]. The GUE is a fundamental example in random matrix theory, as it arises as the eigenvalues of a $(A + A^*)/2$, where A is filled with i.i.d. complex Gaussian random variables [Anderson et al., 2010].

More generally, any projection DPP can be obtained as the position representation of the ground state of an (ad hoc and non-necessarily physically realizable) free Hamiltonian for fermions. This is our first encounter of DPPs as arising from a fermionic construction.¹⁹

3.3.3 Occupation number representation

Since the Hilbert space for modeling N bosons, respectively N fermions, is $\mathcal{S}\mathcal{H}^{\otimes N}$, respectively $\mathcal{A}\mathcal{H}^{\otimes N}$, it is useful to have an orthonormal basis of these spaces [Folland, 2008, Section 4.5]. A convenient way of writing states of indistinguishable particles is through the *occupation number basis*.

The rationale behind this basis is simply to count the number of particles that are in a specific basis state of the one-particle Hilbert space, chosen to make observables simple to express. Let (e_i) be a Hilbert basis of \mathcal{H} . Then an orthonormal basis of $\mathcal{S}\mathcal{H}^{\otimes N}$ is given by

$$|n_1, n_2, \dots\rangle := \frac{1}{Z} \mathcal{S} \left| \underbrace{e_1, \dots, e_1}_{n_1}, \underbrace{e_2, \dots, e_2}_{n_2}, \dots \right\rangle, \quad \text{where } \sum_{i \geq 1} n_i = N,$$

with Z the normalization constant. The label of the state $|n_1, n_2, \dots\rangle$ represents occupation numbers: n_i represents the number of particles in the state $|e_i\rangle$. By construction, the inner product between two states with different occupation numbers is zero. The normalization constant $Z > 0$ is fixed by requiring the state to be normalized, i.e.,

$$1 = \frac{1}{Z^2} \langle n_1, n_2, \dots | \mathcal{S}^\dagger \mathcal{S} | n_1, n_2, \dots \rangle = \frac{1}{Z^2} \frac{1}{N!} \prod_{i=1}^{\infty} n_i!.$$

The same construction applies for fermions. An orthonormal basis of $\mathcal{A}\mathcal{H}^{\otimes N}$ is given by

$$|n_1, n_2, \dots\rangle = \sqrt{N!} \mathcal{A} \left| \underbrace{e_1, \dots, e_1}_{n_1}, \underbrace{e_2, \dots, e_2}_{n_2}, \dots \right\rangle, \quad \text{where } \sum_{i \geq 1} n_i = N, \quad (41)$$

with the important difference that for fermions, $n_i \in \{0, 1\}$ for all i . Indeed, if there is some j such that $n_j \geq 2$, then the corresponding state of $\mathcal{H}^{\otimes N}$ is in the kernel of the antisymmetrization operator \mathcal{A} . Physicists identify this property with *Pauli's exclusion principle*: no two fermions of the same multi-particle state can be *in the same mode*, i.e., correspond to the same vector of the basis e_i of the single-particle space \mathcal{H} .

3.4 Modeling an indefinite number of particles

In order to form a state space that can accommodate any number of particles, and in particular to model uncertainty in that number, physicists introduce so-called *Fock* spaces.

¹⁹In Section 6, we shall come back to the idealized free fermion construction of this section, introducing a temperature parameter to obtain non-projection DPPs.

3.4.1 Fock spaces for bosons and fermions

Denote by $\mathcal{H}^{\otimes 0} \cong \mathbb{C}$ the Hilbert space consisting of all multiples of some unit norm state labeled as $|0\rangle$, called the *vacuum*, and representing the absence of any particle. Taking the convention $S h = A h = h$ for $h \in \mathcal{H}^{\otimes 0}$, we define two Hilbert spaces $\mathcal{H}_{\text{Bosons}}^{\infty}$ and $\mathcal{H}_{\text{Fermions}}^{\infty}$, respectively called the bosonic and fermionic Fock spaces, as orthogonal direct sums

$$\mathcal{H}_{\text{Bosons}}^{\infty} := \bigoplus_{k \geq 0} S \mathcal{H}^{\otimes k}, \quad \mathcal{H}_{\text{Fermions}}^{\infty} := \bigoplus_{k \geq 0} A \mathcal{H}^{\otimes k}. \quad (42)$$

Several remarks are in order. Note first that in (42), by an abuse of notation, we write the symmetrization and antisymmetrization operators in any dimension as S and A , respectively. Second, the orthogonal sum in (42) makes two vectors coming from different summands have inner product zero by definition. Third, we underline that the absence of particles is modeled by the vacuum state $|0\rangle$, not by the null element of either Hilbert space $\mathcal{H}_{\text{Bosons}}^{\infty}$ or $\mathcal{H}_{\text{Fermions}}^{\infty}$. In particular, unlike any null element, the vacuum state represents a physical state, and thus has norm 1 in the Fock space.

The Fock spaces in (42) are defined as orthogonal sums of Hilbert spaces, and are thus Hilbert spaces themselves. For instance, $\mathcal{H}_{\text{Bosons}}^{\infty}$ is the set of collections $h = (h_k)_{k \geq 0}$, where $h_k \in S \mathcal{H}^{\otimes k}$ and $\|h\| := \sum_k \|h_k\|^2 < \infty$. It is sometimes convenient to work instead with the *algebraic sum* of the same vector spaces, meaning that all but a finite set of coefficients are constrained to be zero. This is the case, for instance, when defining the number operator through

$$N |n_1, n_2, \dots\rangle := \left(\sum_i n_i \right) |n_1, n_2, \dots\rangle, \quad (43)$$

which counts the particles in a state, and is defined by the same formula for bosons and fermions. N is then uniquely extended to an unbounded operator on the Fock space [Folland, 2008, Section 4.5]. In this text, we will assume that this extension step can always be done unequivocally, and work directly with the (Hilbert) Fock spaces (42). Finally, a state of the form $|n_1, n_2, \dots\rangle$ with $\sum_{i \geq 1} n_i < \infty$ is called a *Fock state*, and together they form an orthonormal basis of the corresponding Fock space. Any state of a Fock space can thus be represented as a linear combination of Fock states.

3.4.2 Creation and annihilation operators

Besides the number operator, we now introduce two types of operators on Fock spaces that are respectively thought of as creating and destroying a particle in a given one-particle state. The choice of the symmetrization operator S or A has important consequences on the commutation relations between these operators, and we thus separate the treatment of bosons and fermions. We paraphrase here [Folland, 2008, Chapter 4.5].

Bosons. For $v \in \mathcal{H}$, define the operator $b(v)$ on the algebraic sum corresponding to (42), i.e., finite linear combinations of Fock states, by

$$b(v) S(u_1 \otimes \dots \otimes u_k) = \frac{1}{k} \sum_{j=1}^k \langle v | u_j \rangle S(u_1 \otimes \dots \otimes u_{j-1} \otimes u_{j+1} \otimes \dots \otimes u_k),$$

for any $k \geq 1$ and $u_1, \dots, u_k \in \mathcal{H}$. Now, for later ease of writing, we renormalize b and define the operator $a(v)$ on the algebraic sum corresponding to (42), i.e., finite linear combinations of Fock states, by $a(v)w = \sqrt{k}b(v)w$ whenever $w \in S \mathcal{H}^{\otimes k}$. Alternately, using the number operator introduced in Section 3.4.1, $a(v) = \sqrt{N+1}b(v)$. The operator $a(v)$ is called the annihilation operator in state v . Its adjoint $a^\dagger(v) := a(v)^\dagger$ on the same algebraic sum is defined as $a^\dagger(v) = S \sqrt{N}b(v)^\dagger$, and is called the creation operator in state v . Together with $a(v)$, they are called *ladder operators*.^{20 21}

²⁰Note that like the number operator later on, the ladder operators are defined only on the dense subset of finite linear combinations of Fock states; see Folland [2008] for discussions on their extension.

²¹We also note that these ladder operators were inspired by those introduced by Dirac to solve the harmonic oscillator; see Section 3.2.2. However, one should keep in mind that they do not act on the same

Writing $[u, v] = uv - vu$ for the commutator of two operators, we can check that the ladder operators satisfy the so-called *canonical commutation relations* (CCR),

$$[a(u), a(v)] = [a^\dagger(u), a^\dagger(v)] = 0, \quad [a(u), a^\dagger(v)] = \langle u|v \rangle I. \quad (44)$$

To see why the creation and annihilation operators bear such names, consider again a basis (e_i) of \mathcal{H} , and let $a_i = a(e_i)$ for $i \geq 1$. It can be checked that the ladder operators act on Fock states as

$$a_i^\dagger |n_1, \dots, n_i, \dots\rangle = \sqrt{n_i + 1} |n_1, \dots, n_i + 1, \dots\rangle, \quad (45a)$$

$$a_i |n_1, \dots, n_i, \dots\rangle = \sqrt{n_i} |n_1, \dots, n_i - 1, \dots\rangle, \quad (45b)$$

where, by convention, the right-hand side of (45b) is the null element of the Fock space whenever $n_i = 0$. Physically, a_i^\dagger thus models the creation of a particle in the state $|e_i\rangle$, while a_i removes a particle from the same state. Note also that $a_i |0\rangle = 0$ is the null element of the Fock space; we say that the annihilation operator annihilates the vacuum.

In terms of ladder operators, the number operator (43) can be rewritten

$$N = \sum_{i=1}^{\infty} a_i^\dagger a_i, \quad (46)$$

with $N|0\rangle = 0|0\rangle = 0$, the null element of the Fock space, as expected from an operator that counts particles. Note also that the CCRs imply $[N, a_i^\dagger] = a_i^\dagger$ and $[N, a_i] = -a_i$. Finally, any Fock state can be rewritten as the action of creation operators on the vacuum as

$$|n_1, \dots, n_i, \dots\rangle = \frac{1}{\sqrt{\prod_{i=1}^{\infty} n_i!}} \prod_{i=1}^{\infty} (a_i^\dagger)^{n_i} |0\rangle. \quad (47)$$

Note that by definition of Fock states, the products in (47) consist of a finite number of terms.

Fermions. Following the same lines, we can now define the creation and annihilation operators for fermions on the Fock space $\mathcal{H}_{\text{Fermions}}^\infty$. Formally, $a(v)$ is defined on finite linear combinations of Fock states²² by

$$a(v)A(u_1 \otimes \dots \otimes u_k) = \frac{1}{\sqrt{k}} \sum_{j=1}^k (-1)^j \langle v|u_j \rangle A(u_1 \otimes \dots \otimes u_{j-1} \otimes u_{j+1} \otimes \dots \otimes u_k),$$

while one can check that its adjoint $a^\dagger(v) := a(v)^\dagger$ satisfies

$$a^\dagger(v)A(u_1 \otimes \dots \otimes u_k) = \sqrt{k+1} A(u_1 \otimes \dots \otimes u_k \otimes v).$$

As for bosons, these definitions lead to particular commutation relations. Writing $\{u, v\} = uv + vu$ for the so-called *anti-commutator* of two operators, the fermionic annihilation and creation operators satisfy

$$\{a(u), a(v)\} = \{a^\dagger(u), a^\dagger(v)\} = 0 \quad \{a(u), a^\dagger(v)\} = \langle u|v \rangle I, \quad (48)$$

known in the literature as the *canonical anti-commutation relations* (CARs). Here again, the presence of an anti-commutator is interpreted by physicists as a manifestation of Pauli's exclusion principle. For instance, the identity $\{a^\dagger(u), a^\dagger(u)\} = 2a^\dagger(u)a^\dagger(u) = 0$ translates the fact that there cannot be two fermionic particles in the same quantum state.

spaces. In the rest of the paper, we only use the ladder operators acting on the Fock space.

²²and then extended to the whole Fock space.

Again, the ladder operators $a_i = a(e_i)$ in a given basis apply straightforwardly to the corresponding Fock states. For $i \geq 1$, and remembering that $n_i \in \{0, 1\}$, we obtain

$$a_i^\dagger |n_1, \dots, n_i, \dots\rangle = \delta_{n_i 0} \varepsilon(\sigma) |n_1, \dots, n_i + 1, \dots\rangle \quad (49a)$$

$$a_i |n_1, \dots, n_i, \dots\rangle = \delta_{n_i 1} \varepsilon(\sigma) |n_1, \dots, n_i - 1, \dots\rangle, \quad (49b)$$

where $n = \sum n_i < \infty$, σ is the permutation

$$\sigma = \begin{pmatrix} 1 & \cdots & S_i + 1 & S_i + 2 & \cdots & n + 1 \\ S_i + 1 & \cdots & S_i & S_i + 2 & \cdots & n + 1 \end{pmatrix}, \quad (50)$$

where $S_i = \sum_{j=1}^{i-1} n_j$ are the partial sums of the sequence of occupation numbers.

These fermionic ladder operators call for comments again. First, the presence of the delta symbols ensures that in order to create a fermion in a given mode, i.e. in a state described by one of the basis vectors, this mode must be empty; similarly, destroying a fermion in a mode requires that mode to be initially occupied by one particle. This is a natural implementation of the exclusion principle which requires that a mode can only be either empty or occupied by one and only one particle.

Finally, we note that the fermionic number operator can be written as $N = \sum_{i=1}^{\infty} a_i^\dagger a_i$, and that the following commutators are identical to their bosonic counterparts,

$$[N, a_i^\dagger] = a_i^\dagger \quad [N, a_i] = -a_i. \quad (51)$$

3.4.3 Modes

As seen in Section 3.4, a Fock space \mathcal{H}^∞ is built starting from a single-particle Hilbert state \mathcal{H} of arbitrary dimension, of which we single out a basis (e_k) . Basis vectors of the single-particle Hilbert space \mathcal{H} are typically called *modes* in physics. Fock states $|n_1, n_2, \dots\rangle$ form a basis of the Fock space, where the notation stands for the (anti-)symmetrization of the tensor product of n_1 times the basis vector (mode) e_1 , n_2 times the basis vector e_2 , etc.

There is a natural isomorphism of Hilbert spaces between the Fock space \mathcal{H}^∞ built on \mathcal{H} and the tensor product $\bigotimes_k (\mathbb{C}e_k)^\infty$ of the Fock spaces built on each of the one-dimensional (or “single-mode”) Hilbert spaces $\mathbb{C}e_k$. Indeed, taking bosons as an example, one simply needs to map the Fock state $|n_1, n_2, \dots\rangle \in \mathcal{H}^\infty$ to

$$\bigotimes_k \underbrace{S(e_k \otimes \cdots \otimes e_k)}_{n_k \text{ times}},$$

where S is the symmetrization operator, and should be replaced by A for fermions.

Thinking of the Fock space as a product of Fock spaces across modes is often implicit in physics texts.²³ It is, for instance, customary to introduce sophisticated states assuming a single mode, i.e., that \mathcal{H} is one-dimensional, and then to write tensor products of such states across modes to cover the case of an \mathcal{H} of arbitrary dimension. A concrete example will be bosonic coherent states in Section 3.6.

3.4.4 Field operators

There are often several natural bases for a physical situation, for instance a natural basis to describe a source of radiation, and a natural basis to describe a measurement. The linearity of the definition of ladder operators yields easy “change of basis formulas” connecting ladder operators in two different bases. Formally, let $(|u_i\rangle)_{i \in I}$ and $(|v_i\rangle)_{i \in I}$ be two bases of the single-particle Hilbert space \mathcal{H} . The creation and annihilation operators transform exactly as a regular change of basis, i.e.,

$$a^\dagger(v_k) = \sum_i a^\dagger(u_i) \langle u_i | v_k \rangle \quad \text{and} \quad a(v_k) = \sum_i a(u_i) \langle v_k | u_i \rangle. \quad (52)$$

²³Arguably, one already has this isomorphism in mind when writing $|n_1, n_2, \dots\rangle$ for a Fock state.

Note that we implicitly assumed that the two bases were actual Hilbert bases (i.e., countable), but the formula naturally extends in a weak sense to “generalized orthogonal bases”, such as the position basis for the harmonic oscillator in Section 3.2.2. The creation and annihilation *position field operators* are precisely the creation operator and the annihilation operator corresponding to the position basis, and are usually denoted by $\psi^\dagger(x)$ and $\psi(x)$, respectively. Their physical interpretation is that the operator $\psi^\dagger(x)$ creates a particle of a certain type at the position x . They are called (position) field operators because they are the building blocks for position-dependent operators, which we shall see are the mathematical description of quantum physical fields. As ladder operators, the field operators also satisfy the commutation relations corresponding to the particles being described, say for bosons

$$[\psi(x), \psi(y)] = 0, \quad [\psi^\dagger(x), \psi^\dagger(y)] = 0, \quad [\psi(x), \psi^\dagger(y)] = \delta(x - y). \quad (53)$$

To illustrate a generalized case of the change of basis in (52), consider the other fundamental basis for the harmonic oscillator, the momentum basis $|k\rangle$. Since, in an appropriate sense $\langle k|x\rangle = \frac{e^{-ik\cdot x}}{(2\pi)^d}$, Equation (52) then yields the Fourier transform relationship between position and momentum field operators,

$$\psi^\dagger(x) = \int a^\dagger(k) e^{-ik\cdot x} \frac{dk}{(2\pi)^d} \quad \text{and} \quad \psi(x) = \int a(k) e^{ik\cdot x} \frac{dk}{(2\pi)^d}, \quad (54)$$

with $a^\dagger(k)$ the creation operator corresponding to the momentum basis, creating a particle of momentum k . Once again, the important algebraic relations are the commutation relations written in the momentum basis, say for bosons,

$$[a(k), a(\ell)] = 0, \quad [a^\dagger(k), a^\dagger(\ell)] = 0, \quad [a(k), a^\dagger(\ell)] = (2\pi)^3 \delta_{k\ell}. \quad (55)$$

Note that the whole section translates to fermions, with the anticommutation relations (48) replacing the commutation relations of bosons.

Finally, we make a point on notation. While it is customary in physics to use $\psi(x)$ for the field operator of bosons and fermions, to avoid confusion, we write $a(x)$ for a position *bosonic* field operator, and reserve $\psi(x)$ for a position *fermionic* field operator.

3.4.5 Observables

Ladder operators allow one to define observables that model clusters of interacting particles. Depending on the number of particles interacting, we call the corresponding operators *one-particle* observables, *two-particle* observables, etc.

One-particle observables. An example of one-particle observable is the number operator N in (46).

A generic one-particle operator $O^{(1)}$ is an operator that takes as an input a one-particle state $|u_i\rangle$ and gives as an output $|u_j\rangle$, in a given basis (u_i) of the one-particle Hilbert space \mathcal{H} . This means that we destroy the particle in the state $|u_i\rangle$ and create one in the state $|u_j\rangle$. The operator $O^{(1)}$ can actually be rewritten [Cohen-Tannoudji et al., 2019b, Section XV.B] as

$$O^{(1)} = \sum_{ij} O_{ij}^{(1)} a^\dagger(u_j) a(u_i). \quad (56)$$

The number operator N is diagonal, in the sense that the only nonzero terms in the sum (56) are those for which $i = j$. A typical non-diagonal one-particle observable correspond to a hopping between states. For example, if i, j are labels for sites on a lattice, then $\sum_{ij} J_{ij} a_i^\dagger a_j$ with J_{ij} a symmetric matrix implements the hopping of particles from i to j with amplitude J_{ij} .

Two-particle observables. Two-particle observables are used to describe the interaction between two particles. The intuition is that an interaction takes two particles in a state $|u_k\rangle \otimes |u_l\rangle$ and yields a new two-particle state $|u_i\rangle \otimes |u_j\rangle$. Acting on the Fock space, a generic two-body operator $O^{(2)}$ can be written using ladder operators as

$$O^{(2)} = \sum_{ijkl} O_{ijkl}^{(2)} a^\dagger(u_i) a^\dagger(u_j) a(u_l) a(u_k); \quad (57)$$

see [Cohen-Tannoudji et al., 2019b, Chapter XV, Section C] for manipulations using multi-particle observables. A simple Hamiltonian featuring one- and two-particle observables is the Bose-Hubbard model [Fisher et al., 1989],

$$H = \sum_i \hbar\omega_i N_i + \sum_{i,j} J_{ij} a_i^\dagger a_j + \sum_i U_i N_i(N_i + 1), \quad (58)$$

where (ω_i, J_{ij}, U_i) are real numbers that parametrize the model, and $N_i = a_i^\dagger a_i$ counts the particles in mode i . The Hamiltonian (58) is composed of both one-particle and two-particle contributions. The two-particle term in $N_i(N_i + 1)$ models the extra energetic cost to pay to add a particle on a site i when some are already present.

3.5 Wick's theorem

Wick's theorem is a cornerstone of quantum field theory and the backbone of perturbation theory. It gives rise to the famous Feynman diagrams, and importantly for us, it yields the permanents and determinants in the correlation functions of fields, later to be turned into correlation functions of point processes when we introduce physical detection in Section 4 and Section 5.

In essence, Wick's theorem is a generalization of the calculation of the moments of multivariate Gaussian distributions to the case of Gaussian density matrices. We follow the derivation from [Cohen-Tannoudji et al., 2019b, Appendix C XVI].

3.5.1 Gaussian density matrices

A commonly used mixed state on the Fock space \mathcal{H}^∞ of either fermions or bosons corresponds to the so-called *Gaussian density matrices*, defined, when it exists, as the unit-trace, self-adjoint operator

$$\rho = \frac{1}{Z} e^{-\sum_{i,j} a_i^\dagger M_{ij} a_j}, \quad (59)$$

with $M_{ij} = \overline{M_{ji}}$. As in Section 3.1, the definition (59) relies on the spectral calculus: ρ is the operator with the same eigenvectors as the argument of the exponential, but with the exponential applied to the corresponding eigenvalues, followed by division by Z . The normalization constant

$$Z = \text{Tr} e^{-\sum_{i,j} a_i^\dagger M_{ij} a_j}$$

is called the *partition function* in statistical physics. The use of the word *Gaussian* in the name is sometimes confusing to non-physicists. It is motivated by the quadratic expression in the exponential, and the fact that, as we shall see below, the moments of products of creation and annihilation operator under these density matrices behave similarly to the moments of multidimensional Gaussian distributions. Finally, we note that there are conditions on the operator appearing in the exponential in (59) for ρ to be a proper mixed state, in particular to guarantee that $Z < \infty$. These conditions are best discussed for any particular M .

Example 4 (Grand canonical ensemble). Consider a system described by a Hamiltonian H acting on a Fock space of either bosons or fermions, with the Hamiltonian being quadratic in the ladder operators. In other words, we require that the Hamiltonian is a one-particle observable (56), like the Bose-Hubbard Hamiltonian (58) with $U_i = 0$. Physicists usually consider additional fluctuations in energy and number of particles, and describe the system

at thermal equilibrium by the so-called *grand canonical ensemble*. Formally, this amounts to considering the Gaussian density matrix

$$\rho_{\text{GC}} = \frac{e^{-\beta(H-\zeta N)}}{Z_{\text{GC}}} \quad (60)$$

where $N = \sum_i a_i^\dagger a_i$ is the number operator, $\beta > 0$ is the inverse temperature and $\zeta > 0$ is the chemical potential, which can be adjusted to change the average number of particles in the system. The normalization $Z_{\text{GC}} = \text{Tr} e^{-\beta(H-\zeta N)}$ is the so-called *grand-canonical partition function*. \diamond

As with any state ρ , a natural question is to compute the expectation $\langle A \rangle_\rho = \text{Tr}(\rho A)$ of an observable A ; see Section 3.1. It turns out that for Gaussian density matrices, this is a simple mechanical computation using canonical (anti)commutation relations, as soon as A is a product of linear combinations of annihilation and creation operators. The result of this computation is precisely Wick's theorem, a close parent to Isserlis' theorem in classical statistics on the computation of moments of a Gaussian distribution. Up to a change of basis, and because we only care about linear combinations of ladder operators, it is enough to treat the case of

$$\rho = \frac{1}{Z} e^{-\sum_i \nu_i a_i^\dagger a_i}. \quad (61)$$

In the rest of this section, we henceforth assume ρ to be in diagonal form (61).

3.5.2 Where permanents and determinants appear

For brevity, every average in this section is implicitly meant as under the Gaussian density ρ in (61), and we thus write $\langle \cdot \rangle$ instead of $\langle \cdot \rangle_\rho$. Additionally, to treat both bosons and fermions in a single theorem, we define the generalized commutator $[u, v]_\eta := uv - \eta vu$, where $\eta = \pm 1$. We note that both the CCR (44) and the CAR (48) can be rewritten as

$$[a_i, a_j]_\eta = 0, \quad [a_i^\dagger, a_j^\dagger]_\eta = 0, \quad \text{and} \quad [a_i, a_j^\dagger]_\eta = \delta_{ij}, \quad (62)$$

with $\eta = 1$ for CCR and $\eta = -1$ for CAR.

We note right away that the density matrix (61) conserves the number of particles, in the sense that $[\rho, N] = 0$, with $N = \sum_i a_i^\dagger a_i$. This leads to the following important lemma.

Lemma 2 (Creation and annihilation numbers must match). *Consider the Gaussian density matrix ρ given by (61), acting on either the bosonic or fermionic Fock space, and an operator O_{n_c, n_a} composed of the product, in any order, of n_c creation operators $a_{i_1}^\dagger, \dots, a_{i_{n_c}}^\dagger$ and n_a annihilation operators $a_{i_1}, \dots, a_{i_{n_a}}$. Then $\langle O_{n_c, n_a} \rangle = 0$ unless $n_c = n_a$.*

Note that this implies that if the total number of creation and annihilation operator is odd, the average necessarily vanishes.

Proof. The proof is simply based on the fact, to be shown below, that $[N, O_{n_c, n_a}] = (n_c - n_a)O_{n_c, n_a}$. Then it follows that $(n_c - n_a)\langle O_{n_c, n_a} \rangle = \langle [N, O_{n_c, n_a}] \rangle = \text{Tr}(\rho N O_{n_c, n_a} - \rho O_{n_c, n_a} N) = 0$ by the cyclicity of the trace and $[\rho, N] = 0$.

To prove $[N, O_{n_c, n_a}] = (n_c - n_a)O_{n_c, n_a}$, we proceed by recurrence. It is true for $\max(n_c, n_a) \leq 1$ by the commutation relations of the creation and annihilation operators with the number operator (51). Now assume that it is true up to some n_c, n_a , and write $O_{n_c, n_a+1} = O_{m_c, m_a} a_i O_{p_c, p_a}$ with $m_c + p_c = n_c$ and $m_a + p_a = n_a$. Then

$$\begin{aligned} [N, O_{n_c, n_a+1}] &= [N, O_{m_c, m_a}] a_i O_{p_c, p_a} + O_{m_c, m_a} [N, a_i] O_{p_c, p_a} + O_{m_c, m_a} a_i [N, O_{p_c, p_a}], \\ &= (m_c - m_a) O_{n_c, n_a+1} - O_{n_c, n_a+1} + (p_c - p_a) O_{n_c, n_a+1}, \\ &= (n_c - (n_a + 1)) O_{n_c, n_a+1}. \end{aligned} \quad (63)$$

The case O_{n_c+1, n_a} is similar. \square

Now, for an even integer $N = 2k$, define a *contraction* of order N as a permutation $\sigma \in \mathfrak{S}_N$ such that $\sigma(1) < \sigma(3) < \dots < \sigma(2k-1)$, and $\sigma(2i-1) < \sigma(2i)$ for $i = 1, \dots, k$. For instance, there are three contractions of order 4, namely

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}, \text{ and } \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}. \quad (64)$$

In words, contractions are built as follows. Starting from $Z = \{1, \dots, N\}$ and $i_1 = 1$, pair i_1 with an arbitrary $i_2 \in Z \setminus \{i_1\}$. Then select the smallest element i_3 of $Z \setminus \{i_1, i_2\}$ and pair it with any element of $Z \setminus \{i_1, i_2, i_3\}$. Repeat the procedure until all integers in Z have been paired. The corresponding permutation σ is the one such that $(\sigma(1), \sigma(2), \dots, \sigma(N)) = (i_1, i_2, \dots, i_N)$. This constructive definition shows that there are actually $(N-1)!! = 1 \times 3 \times 5 \dots \times (N-3) \times (N-1)$ contractions of order N .

Theorem 3 (Wick's theorem). *Consider the Gaussian density matrix ρ given by (61) and $N = 2k$ linear combinations b_1, \dots, b_N of creation and annihilation operators, in any basis. Then*

$$\langle b_1 \dots b_N \rangle = \sum_{\sigma \text{ contraction}} \eta^{\varepsilon(\sigma)} \langle b_{\sigma(1)} b_{\sigma(2)} \rangle \langle b_{\sigma(3)} b_{\sigma(4)} \rangle \dots \langle b_{\sigma(N-1)} b_{\sigma(N)} \rangle, \quad (65)$$

where $\eta = +1$ for bosons and $\eta = -1$ for fermions. The sum is over contractions of order N , with parity $\varepsilon(\sigma)$.

Before proving Wick's theorem for the sake of completeness, we give two illustrative examples. First, when $N = 4$, remembering the three contractions (64) of order 4, we have

$$\langle b_1 b_2 b_3 b_4 \rangle = \langle b_1 b_2 \rangle \langle b_3 b_4 \rangle + \eta \langle b_1 b_3 \rangle \langle b_2 b_4 \rangle + \langle b_1 b_4 \rangle \langle b_2 b_3 \rangle. \quad (66)$$

Second, Wick's theorem is the mathematical reason why permanents and determinants will appear when we consider point processes of detection times in Sections 4, 5, and more abstractly in Section 6. This application is so fundamental to our paper that we highlight it here.

Example 5 (Wick's theorem for coherence functions). We are interested in the expected number of k -uplets of particles simultaneously appearing in (distinct) modes 1 to k . Since the operator $N_i = a_i^\dagger a_i$ counts the number of particles in mode i , we aim to compute

$$\langle N_1 \dots N_k \rangle = \langle a_1^\dagger a_1 \dots a_k^\dagger a_k \rangle.$$

Note that all N_i s commute, so that the order is irrelevant. We now use the (anti-)commutation relations (62) to bring all creation operators to the front,

$$\langle N_1 \dots N_k \rangle = \eta^{1+2+\dots+(k-1)} \langle a_1^\dagger a_2^\dagger \dots a_k^\dagger a_1 a_2 \dots a_k \rangle.$$

By convention, we also put the annihilation operators in decreasing order, which removes the sign,

$$\langle N_1 \dots N_k \rangle = \langle a_1^\dagger a_2^\dagger \dots a_k^\dagger a_k a_{k-1} \dots a_1 \rangle.$$

Now, we are ready to apply Theorem 3. Upon noting, thanks to Lemma 2, that pairing two creation or two annihilation operators results in a zero average, the only non-zero terms in (65) result from permutations $\sigma \in \mathfrak{S}_{2k}$ of the form

$$\sigma = \begin{pmatrix} 1 & 2 & & 3 & 4 & & 5 & \dots & 2k-1 & 2k \\ 1 & \nu(k+1) & & 2 & \nu(k+2) & & 3 & \dots & k+1 & \nu(2k) \end{pmatrix},$$

where ν is a permutation of $\{k+1, \dots, 2k\}$. Composing σ with $2 \times (1 + 2 + \dots + (k-1))$ transpositions, we obtain

$$\begin{pmatrix} 1 & \dots & k & k+1 & \dots & 2k \\ 1 & \dots & k & \nu(k+1) & \dots & \nu(2k) \end{pmatrix},$$

so that $\varepsilon(\sigma) = \varepsilon(\nu)$. In particular, Wick's theorem yields

$$\langle N_1 \dots N_k \rangle = \sum_{\nu \in \mathfrak{S}_k} \eta^{\varepsilon(\nu)} \prod_{i=1}^k \langle a_i^\dagger a_{\nu(i)} \rangle,$$

where we recognize the permanent or determinant of the matrix $(\langle a_i^\dagger a_j \rangle)_{1 \leq i, j \leq k}$, depending on $\eta = \pm 1$. \diamond

We now prove Wick's theorem for the sake of completeness.

Proof. We want to compute the expectation $\text{Tr}(O_N \rho)$, where $O_N = b_1 b_2 \dots b_N$,

$$b_i = \sum_{\alpha} (A_{i,\alpha} a_{\alpha}^\dagger + B_{i,\alpha} a_{\alpha}), \quad \text{and} \quad \rho = e^{-\sum_{\alpha} \nu_{\alpha} a_{\alpha}^\dagger a_{\alpha}} / Z,$$

where we have expanded the b_i 's in the basis in which the density matrix is diagonal.

Let us start with two general remarks. First, by Lemma 2, we can restrict the number of b_i 's to be even, since otherwise the average vanishes. We henceforth assume that $N = 2k$ is even. We are going to repeatedly use the generalized commutation relations. In particular, we note that, for any η and i, j , $[b_i, b_j]_{\eta}$ is a multiple of the identity and thus commutes with all operators.

We are now ready to compute $\langle O_N \rangle = \langle b_1 \dots b_N \rangle = \text{Tr}(b_1 \dots b_N \rho)$. The general idea is (i) to push b_1 through all the other operators b_i using the (anti)commutation relations. Then, (ii) using the cyclicity of the trace, we bring b_1 back in front of the rest. And finally, (iii) using that ρ is Gaussian to recover the original average $\langle O_N \rangle$ up to a constant.

Starting this programme, we have

$$\begin{aligned} \langle O_N \rangle &= [b_1, b_2]_{\eta} \langle b_3 \dots b_N \rangle + \eta \langle b_2 b_1 \dots b_N \rangle, \\ &= [b_1, b_2]_{\eta} \langle b_3 \dots b_N \rangle + \eta [b_1, b_3]_{\eta} \langle b_2 b_4 \dots b_N \rangle + \eta^2 \langle b_2 b_3 b_1 \dots b_N \rangle. \end{aligned}$$

Iterating the same steps, we write

$$\langle O_N \rangle = \eta^{N-1} \langle b_2 b_3 \dots b_N b_1 \rangle + \sum_{j=2}^N [b_1, b_j]_{\eta} \eta^{j-2} \langle b_2 \dots b_{j-1} b_{j+1} \dots b_N \rangle. \quad (67)$$

Now, by the cyclicity of the trace, we have

$$\langle b_2 b_3 \dots b_N b_1 \rangle = \text{Tr}(\rho b_2 b_3 \dots b_N b_1), \quad (68)$$

$$= \text{Tr}(b_1 \rho b_2 b_3 \dots b_N). \quad (69)$$

Let us for a moment assume that b_1 is either a specific creation or annihilation operator, i.e. $b_1 = a_{\alpha}^{(\dagger)}$ for some specific α . Now, using the explicit Gaussian form of the density matrix, we have $b_1 \rho = \lambda \rho b_1$, with $\lambda = e^{-\nu_{\alpha}} / Z$ if $b_1 = a_{\alpha}$, and $\lambda = e^{\nu_{\alpha}} / Z$ if $b_1 = a_{\alpha}^{\dagger}$, as can be seen by checking the action of these operators on the basis of Fock states. We thus come to

$$\langle b_2 b_3 \dots b_N b_1 \rangle = \lambda \langle b_1 \dots b_N \rangle. \quad (70)$$

Since $\eta = \pm 1$, Equation (67) is simplified into:

$$\langle b_1 \dots b_N \rangle = \sum_{j=2}^N \eta^j \frac{[b_1, b_j]_{\eta}}{1 - \eta \lambda} \langle b_2 \dots b_{j-1} b_{j+1} \dots b_N \rangle. \quad (71)$$

In particular, note that (71) with $N = 2$ becomes

$$\langle b_1 b_2 \rangle = \frac{[b_1, b_2]_{\eta}}{1 - \eta \lambda}. \quad (72)$$

Plugging this into (71), we conclude that

$$\langle b_1 \dots b_N \rangle = \sum_{j=2}^N \eta^j \langle b_1 b_j \rangle \langle b_2 \dots b_{j-1} b_{j+1} \dots b_N \rangle. \quad (73)$$

Since this last expression is explicitly linear in b_1 , it is actually valid for any linear combination of creation and annihilation operators. In particular, (73) replaces the calculation of an average of a product of N operators by the weighted sum of $N - 1$ averages of products of $N - 2$ operators. Repeating the procedure for each product of $N - 2$ operators in (73), we obtain Wick’s theorem. \square

This is the simplest form of Wick’s theorem. It can be further generalized to Hamiltonians with in addition the “anomalous” quadratic terms $\sum_{nm} A_{nm} a_n a_m + \sum_{nm} \bar{A}_{nm} a_n^\dagger a_m^\dagger$, which arise in mean-field theories of interacting Bose gases and superconductors, and generate $\langle a_i a_j \rangle \neq 0$; or to the case of Hamiltonians with linear terms, which generate nonzero $\langle a_i \rangle$ [Berezin, 1966, Chapter 3].

Remark 4. The proof of Wick’s theorem also holds for the vacuum, and the vacuum is also usually said to be a Gaussian state. The vacuum can also be obtained as a limiting state when the temperature $1/\beta$ is taken to zero and $\zeta = 0$; see the computations in Section 6.

Finally, note that the definition of a Gaussian state is not universal and may be community-dependent. Quantum opticians, for instance, might define a Gaussian state as one that has a Gaussian Wigner transform; see Section 4.

3.6 Bosonic coherent states model classical fields

Many quantum systems have a classical (i.e., non-quantum) description, and physicists also sometimes start from a classical description to build a quantum theory, a procedure known as quantization. But in the end, a physical system is fundamentally quantum and we should try to understand how a classical behavior emerges from a quantum description. This phenomenon is known as *decoherence*, and is a vast modern research programme. An important part of the answer is to build quantum states that behave as closely as possible to classical ones. Glauber [1963] made seminal contributions in that regard, defining *coherent states* for systems of photons.²⁴ We follow the introduction of bosonic coherent states in [Mandel and Wolf, 1995, Chapter 11], before pointing out the strong connections with the subfield of signal processing called time-frequency analysis. We defer the more subtle discussion on fermionic coherent states to Section 5.4.

3.6.1 Definition and properties

As discussed around Equation (26), measuring an observable is modelled by sandwiching the state operator. After measurement, the quantum state is thus different from the initial state. This is very different from classical physics, where it is possible, in principle, to perform *passive* measurements, e.g. looking at a screen, or recording the intensity of the electromagnetic field. Consider henceforth the boson Fock space built using the modes of the harmonic oscillator from Section 3.2.2. If we are to build a state of this Fock space that behaves classically, we would intuitively like to take a state that remains unchanged when we measure it. Since most measurements involve absorbing bosons (think photons in optics; see Section 4.1), good candidates for classical-like states are the left eigenstates of the annihilation operator, or equivalently the right eigenstates of the creation operator.

To make things concrete, consider a single mode for simplicity, that is, a Fock space built on a one-dimensional $\mathcal{H} = \mathbb{C}e_1$; see Section 3.4.3. Denote the ladder operators by $a = a(e_1)$ and $a^\dagger = a^\dagger(e_1)$. Our candidate classical-like states are $|\alpha\rangle \in \mathcal{H}_{\text{Bosons}}^\infty$ such that

$$a |\alpha\rangle = \alpha |\alpha\rangle, \quad \alpha \in \mathbb{C}. \quad (74)$$

Informally,²⁵ the sandwiched state (26) that appears when modeling absorption mechanisms

²⁴The name coherent state comes from the concept of coherence in optics.

²⁵For details, see the Lindblad equation in [Bouten et al., 2007].

is then $a|\alpha\rangle\langle\alpha|a^\dagger = |\alpha|^2|\alpha\rangle\langle\alpha|$, so that coherent states are stable under absorption of a photon.

Solving the eigenstate equation (74) for $|\alpha\rangle = \sum_n c_n |n\rangle$ expressed in the Fock basis, we write

$$\sum_{n \geq 1} c_n \sqrt{n} |n-1\rangle = \sum_{n \geq 0} c_n |n\rangle,$$

so that $c_n = \alpha c_{n-1} / \sqrt{n}$ for $n \geq 1$. This leads to

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n \geq 0} \frac{\alpha^n}{\sqrt{n!}} |n\rangle, \quad (75)$$

up to a complex number of unit modulus, where we have determined the modulus of the normalizing constant by imposing $\langle\alpha|\alpha\rangle = 1$. The states (75), labeled by nonzero complex numbers α , are called *canonical coherent states*. By construction, they are mixed states formed as an infinite linear combination of Fock states. If we were to measure the number of photons in $|\alpha\rangle$, we would obtain n with probability $|\langle\alpha|n\rangle|^2 \propto |\alpha|^{2n}/n!$, and we recognize the Poisson distribution with mean $|\alpha|^2$. In particular, the Poisson distribution is in a sense the maximum entropy distribution for independent counts [Harremoës, 2001]: for a fixed mean $\lambda > 0$, the maximal entropy of a sum of k independent Bernoulli variables is increasing with k , and converges to the entropy of the Poisson variable with mean λ . In that limited sense, one can think of the coherent states as states with a maximally uncertain number of particles.

An equivalent point of view on coherent states, which sometimes serves as their definition in mathematical physics [Ali et al., 2000], is that they are naturally associated to a projective representation of the translation group in phase space,²⁶ the complex plane indexed by α . To see this, start from the description of Fock states $|n\rangle = (a^\dagger)^n / \sqrt{n!} |0\rangle$. The coherent state (75) is thus obtained from the vacuum by the action of an operator $D(\alpha)$ on the boson Fock space called a *displacement operator*,

$$|\alpha\rangle = D(\alpha) |0\rangle, \quad D(\alpha) = e^{\alpha a^\dagger - \alpha^* a}. \quad (76)$$

This formula is a consequence of the CCRs and the so-called Baker–Campbell–Hausdorff formula expanding the exponential of a sum of (non-commuting) operators. The same formula yields properties of D like

$$D^{-1}(\alpha) a D(\alpha) = a + \alpha \quad (77)$$

$$D(\alpha) D(\beta) = e^{i \operatorname{Im}\{\alpha \beta^*\}} D(\alpha + \beta). \quad (78)$$

These relations are central to the theory of coherent states. First, they show that D is a projective representation of the translation group in phase-space. Second, they naturally generalize to other groups, yielding coherent states for different systems than the harmonic oscillator; see e.g. spin coherent states. In fact, coherent states have been defined for any locally compact Lie group [Perelomov, 1972, Arecchi et al., 1972, Zhang et al., 1990].

Another way to justify that coherent states are almost classical is through Heisenberg’s uncertainty principle (27). Remember that for the harmonic oscillator, position and momentum operators X, P are incompatible observables, and that both can be expressed as sums of ladder operators; see the discussion around Equation (37). With the notation of Section 3.1, one can check that the product of $\sigma_{X,|\alpha\rangle}\langle\alpha| \sigma_{P,|\alpha\rangle}\langle\alpha|$ is minimal among states. One says that coherent states *saturate* Heisenberg’s uncertainty relations.

Another quasi-classical aspect of coherent states is their dynamical evolution. Since the Fock states are eigenvectors of the Hamiltonian of a harmonic oscillator $H = \hbar\omega a^\dagger a$, $U(t) = e^{-itH}$ applied to $|\alpha\rangle$ decomposed on the Fock basis (75) yields

$$|\alpha(t)\rangle = |ae^{-i\omega t}\rangle. \quad (79)$$

²⁶The name *phase space* comes from the connection through (37) to the classical phase space with position-momentum coordinates.

In phase space, i.e., the complex plane parametrized by α , the coherent state parameter $\alpha(t)$ simply rotates over time, with an angular frequency ω . This is precisely the phase-space behavior of a classical harmonic oscillator.

Finally, we mention two more properties of coherent states. First, they are an overcomplete family of the Fock space, with a reconstruction formula

$$\frac{1}{\pi} \int_{\mathbb{C}} |\alpha\rangle\langle\alpha| \, d\alpha = \mathbb{1}. \quad (80)$$

The fact that they are not an orthonormal basis is further seen from the overlap

$$\langle\beta|\alpha\rangle = \exp\left(\alpha^*\beta - \frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2\right), \quad \alpha, \beta \in \mathbb{C}. \quad (81)$$

A good reference for such properties of coherent states of the harmonic oscillator is [Folland, 1989], or, actually, books on harmonic analysis applied to time-frequency signal processing [Gröchenig, 2001], on which we say a few words in the next section.

Finally, we have built here bosonic coherent states assuming a single-mode Fock state. Multi-mode coherent states are naturally obtained as tensor products of single-mode coherent states, see Section 3.4.3, and one typically writes them

$$|\boldsymbol{\alpha}\rangle = |\alpha_1\rangle \otimes |\alpha_2\rangle \otimes \dots$$

3.6.2 The Husimi distribution and time-frequency analysis

Single-mode bosonic coherent states of the harmonic oscillator are intimately linked to the subfield of signal processing called time-frequency analysis; see e.g. [Flandrin, 1998, Gröchenig, 2001]. Consider e.g. the short-time Fourier transform $V_g : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}^2)$, defined as

$$V_g(f) : t, \omega \mapsto \int f(\tau) \overline{g(\tau - t)} e^{-2i\pi\omega\tau} \, d\tau,$$

where $g(t) = 2^{1/4}e^{-\pi t^2}$ is a unit-norm Gaussian window. Intuitively, if f is a signal, say representing an audio recording, then $|V_g(f)(t, \omega)|$ will be large whenever frequency ω is present at time t . In other words, $V_g(f)$ is akin to a musical score. Time-frequency transforms such as V_g are thoroughly used in signal processing, for tasks such as detection or estimation of signal corrupted with noise [Flandrin, 1998]. Mathematically, V_g is a unitary linear operator, which can be inverted [Gröchenig, 2001]. Furthermore, it is tightly linked to decomposing a state into coherent states. Indeed, up to a non-vanishing factor,

$$V_g(f)(t, -\omega) \propto \langle\alpha|\varphi\rangle, \quad (82)$$

where $\alpha = t + i\omega$ indexes a coherent state (75), and the state φ is described in the Fock basis by $\langle\varphi|n\rangle = \int f h_n^* dt$, where h_n is the n th Hermite function, a special basis of $L^2(\mathbb{R})$ that we already met in Section 3.2.2. Physicists call the squared modulus of the right-hand side of Equation (82) the Husimi distribution of φ , and think of it as a phase space representation of the state φ , in pretty much the same way signal processors think of the *spectrogram* $t, \omega \mapsto |V_g(f)(t, \omega)|^2$ as a time-frequency representation of the signal f .

As a final note and to introduce another related point process, one can prove that, in a suitable sense, the STFT V_g of white Gaussian noise is equal to the so-called *planar Gaussian analytic function*, up to a nonvanishing term [Bardenet et al., 2018, Bardenet and Hardy, 2019]. Similarly, if one had access to a white noise-like state $|\xi\rangle$ such that $\langle\xi|n\rangle \sim \mathcal{N}_{\mathbb{C}}(0, 1)$ are i.i.d. complex unit Gaussians, then its coherent-state “decomposition” $\alpha \mapsto \langle\xi|\alpha\rangle$ would be the so-called *planar Gaussian analytic function*, up to a nonvanishing term again. This planar Gaussian analytic function, along with the point process of its zeros, plays a role in the analysis of chaotic dynamical systems in statistical physics [Nonnenmacher, 2013], where it is also called the *chaotic analytic function* [Hannay, 1998]. The zeros of the planar Gaussian analytic function are not a DPP. Yet, being the zeros of a random *smooth* function, they have a repulsive behaviour. They actually share many properties with the Ginibre ensemble, itself a fundamental DPP [Hough et al., 2009].

4 Photodetection and bosonic coherences

To obtain signals from a physical situation, the picture to have in mind consists in the processes of *emission*, *propagation*, and *detection* of radiation. The first two processes are usually described by physicists using one or several sources, classical or quantum, and a field theory, classical or quantum. Questions that arise regard, for instance, the dynamics of the source(s), how they generate radiation, and the physical properties of this radiation. The third process, detection, lies at the interface of physics and signal processing. In this section, we discuss simple standard models for the sources, the field, and the detection for bosons, having in mind photodetection. Following the footsteps of Macchi [1975], our goal is to arrive at the description of the measurement of arrival times at a simple detector as a point process. We shall see how some of the point processes introduced in Section 2 naturally appear from Glauber's coherent state decompositions. As a guiding thread, we will comment on the HBT effect in its different guises.

In Section 4.1, we describe simple models for sources, fields, and detectors, and arrive at the marginal probability of detecting an event at a given time. In Section 4.2, we examine the correlation between more than one detection events. The central objects that encapsulate information about detection times are the coherence functions. In Section 4.3, we examine how to turn the coherence functions of a physical detection setup into the correlation functions of a point process. We discuss special particular cases, and recover the permanent point processes already announced using a semi-classical treatment in Section 2.3. Finally, in Section 4.4, we show that single-atom sources can yield anti-bunching detection events. This is a warning that bosons should not be identified with bunching particles in general, but that the properties of the source should be mentioned.

4.1 Modeling photodetection events

Photomultipliers are experimental devices that turn incoming radiation into a measurable electric current. The theory of photodetection aims at understanding what kind of signals can be observed as the output of a photomultiplier. As we shall see, these signals are directly related to the so-called coherence functions²⁷ of the field, quantum or classical.

4.1.1 Modeling the radiation

Before discussing how we introduce coherence functions, we need to specify how we describe the radiation we want to probe, using the objects of quantum field theory introduced in Section 3. Note that, while the framework below is enough to describe photodetection, a complete description would require to go into the details of the construction of (relativistic) quantum field theories and especially its dynamical content [Cohen-Tannoudji et al., 1998].

The electromagnetic field is described by a set of quantum fields satisfying some commutation relations. For photodetection, the dominant contribution comes from the *electric* field at the detection device. As a reminder, a field is a quantity that depends on both position and time. A *bosonic quantum field* is a collection of operators associated to an indefinite number of bosonic particles, called e.g. photons, on a Fock space $\mathcal{H}_{\text{Bosons}}^\infty$ (see Section 3.4). Formally, the state of the field is built by acting on the vacuum $|0\rangle$ with creation and annihilation operators $a_{\mathbf{k},\varepsilon}^\dagger$ and $a_{\mathbf{k},\varepsilon}$ [Mandel and Wolf, 1995, Chapter 10]. The quantities \mathbf{k} and ε index modes of the field (i.e., solutions to Maxwell's equations), and typically correspond to the momentum and the polarization of the photon.

For simplicity, and because it already contains the ingredients that relate the point processes of Section 2 to photodetection, we consider a fixed linear polarization, and henceforth drop the index ε . This corresponds to an assumption we made in the semi-classical treatment of Section 2.3. The electric field is then described by a single (instead of one per space coordinate) time- and space-dependent operator $E(\mathbf{r}, t)$, acting on the Fock space $\mathcal{H}_{\text{Bosons}}^\infty$.

²⁷Often called *correlation functions* in physics, but we refrain from using *correlation* here to avoid confusion with the concept of *correlation function* of a point process; see Section 2.

To obtain the explicit form of the electric field operator, the strategy is to consider the field as enclosed in a (large) box and perform the so-called *canonical quantization* of Maxwell's equations. We refer the reader to the literature for such details, e.g. [Mandel and Wolf, 1995, Chapter 10], and consider as given the resulting form of the field, namely

$$E(\mathbf{r}, t) = \sum_{\mathbf{k}} \mathcal{N}_{\mathbf{k}} \left[a_{\mathbf{k}} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega_{\mathbf{k}} t)} + a_{\mathbf{k}}^{\dagger} e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega_{\mathbf{k}} t)} \right], \quad (83)$$

where $\mathcal{N}_{\mathbf{k}}$ is a normalization coefficient, $\mathbf{r} \in \mathbb{R}^d$ indexes space and $t \in \mathbb{R}$ indexes time. The time dependence in (83) is a consequence of working in the interaction (or equivalently, so far, Heisenberg) picture. The explicit plane-wave form corresponds to the evolution implied by the free electromagnetic Hamiltonian

$$H_{\text{FF}} = \hbar \sum_{\mathbf{k}} \omega_{\mathbf{k}} a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}}. \quad (84)$$

4.1.2 Modeling the detector: the first-order coherence function

The usual image to have in mind for a detector is a two-level atom like the qubit described in Section 3.2.1, described quantum-mechanically with an eigenfrequency ω_{eg} . The atom interacts with the electric field $E(\mathbf{r}, t)$. In mathematical terms, the Hilbert space of the detector is finite-dimensional, isomorphic to \mathbb{C}^2 ; see Section 3.2.1. Since the Hilbert space of the electric field is $\mathcal{H}_{\text{Bosons}}^{\infty}$, the joint system of the detector and the field is $\mathbb{C}^2 \otimes \mathcal{H}_{\text{Bosons}}^{\infty}$.

To describe the time evolution of the system, we use the interaction picture and decompose the Hamiltonian as a free and an interaction part

$$H = H_0 + H_I, \quad (85)$$

see Section 3.1. The free part is taken to be $H_0 = H_{\text{FQ}} \otimes \mathbb{1} + \mathbb{1} \otimes H_{\text{FF}}$, a combination of the free Hamiltonians of a qubit and the field, respectively defined in (34) and (84). To describe the interaction Hamiltonian, let first $q \in \mathbb{R}$ model the modulus of the dipolar moment multiplied by the charge of an electron. Once again, Maxwell's equations hint that the interaction can be taken as [Cohen-Tannoudji et al., 1998]

$$H_I(t) = -q \sigma_x(t) \otimes E(\mathbf{r}, t), \quad (86)$$

where

$$\sigma_x(t) = U_{\text{FQ}}(t) \sigma_x(t) U_{\text{FQ}}(-t) = U_{\text{FQ}}(t) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} U_{\text{FQ}}(-t),$$

is the observable σ_x from Section 3.2.1, evolved through time using the group $U_{\text{FQ}}(t)$ corresponding to the free part of (85), as befits the interaction picture.²⁸

We want to compute the probability $p_{\omega_{eg}}(\mathbf{r}, t)$ of a detector at position \mathbf{r} to be in its excited state $|e\rangle$ after a time t , knowing that the initial state of the detector is $|g\rangle$ and the field is any arbitrary initial state, but irrespective of the final state of the field. We assume that the initially prepared state of the detector and the field is $|g\rangle \otimes |i\rangle = |g, i\rangle$, where $|g\rangle$ is the ground state of the detector and $|i\rangle$ is an arbitrary initial state of the field. The corresponding density matrix is thus the projector $\rho(0) = |g, i\rangle \langle g, i|$. In the interaction picture described in Section 3.1, the state evolves until time t through the action of the evolution operator $V(t) = U_0^{\dagger}(t) U(t)$ as $\rho(t) = V(t) \rho(0) V^{\dagger}(t)$. Measuring the detector in its excited state corresponds to the observable

$$|e\rangle \langle e| \otimes \mathbb{1} = |e\rangle \langle e| \otimes \sum_f |f\rangle \langle f| = \sum_f |e, f\rangle \langle e, f|, \quad (87)$$

where f indexes Fock states, and we decomposed the identity operator onto the Fock states, $\mathbb{1} = \sum_f |f\rangle \langle f|$. Now the observable (87), in the interaction picture, evolves as

$$U_0^{\dagger}(t) \sum_f |e, f\rangle \langle e, f| U_0(t).$$

²⁸Physicists will recognize here an instance of the more traditional form $H_I(t) = -\mathbf{d}(t) \cdot \mathbf{E}(\mathbf{r}, t)$, where we assumed the dipolar moment $\mathbf{d}(t)$ to be aligned with the polarization \mathbf{u} of the field $\mathbf{E}(\mathbf{r}, t) = E(\mathbf{r}, t) \mathbf{u}$.

We are now ready to compute the probability $p_{\omega_{eg}}(\mathbf{r}, t)$ of the detector being measured in its excited state at time t . By Born's rule (see Section 3.1), this probability is the average value of the evolved observable onto the evolved state $\rho(t)$, namely

$$p_{\omega_{eg}}(\mathbf{r}, t) = \text{Tr} \left(U_0^\dagger(t) \left[\sum_f |e, f\rangle\langle e, f| \right] U_0(t) V(t) |g, i\rangle\langle g, i| V^\dagger(t) \right) \quad (88a)$$

$$= \sum_f \text{Tr} (|e, f\rangle\langle e, f| U(t) |g, i\rangle\langle g, i| U(-t)) \quad (88b)$$

$$= \sum_f |\langle e, f| U(t) |g, i\rangle|^2 \quad (88c)$$

$$= \sum_f |\langle e, f| V(t) |g, i\rangle|^2, \quad (88d)$$

where the last line comes from the definition $V(t) = U_0^\dagger(t)U(t)$ and the fact that $|e, f\rangle$ is an eigenvector of $H_0(t)$, so that applying $U_0(t)$ to $|e, f\rangle$ simply multiplies by a complex number of modulus 1. Note also that the dependence in \mathbf{r} is hidden in $V(t)$.

Up until now, all the discussion is exact. We now introduce a physical assumption that allows for a simple expression of $V(t)$.

Assumption 5 (Weak Coupling). *The interaction between the detector and the field is assumed to be weak, allowing us to treat the dynamics using perturbation theory. Formally, this means that we assume the existence of a series expansion of the evolution operator in terms of the interaction Hamiltonian; see the discussion around (32). To first order, this amounts to*

$$V(t) \simeq \mathbb{1} - \frac{i}{\hbar} \int_0^t H_I(t') dt'. \quad (89)$$

Using (89) and $\langle e|\sigma_x(t)|g\rangle = e^{i\omega_{eg}t}$, we find

$$\langle e, f|V(t)|g, i\rangle = \frac{iq}{\hbar} \int_0^t e^{i\omega_{eg}t'} \langle f|E(\mathbf{r}, t')|i\rangle dt'. \quad (90)$$

Plugging this into (92), and remembering that the Fock states $|f\rangle$ form a basis of the Fock space, we obtain

$$p_{\omega_{eg}}(\mathbf{r}, t) = \left(\frac{q}{\hbar}\right)^2 \int_{[0, t]^2} e^{i\omega_{eg}(t'-t'')} \langle i|E(\mathbf{r}, t'')E(\mathbf{r}, t')|i\rangle dt' dt''. \quad (91)$$

This equation is straightforwardly generalized to the case where the initial state of the field is not a pure state but a mixed state, described by a density matrix $\rho = \sum_i p_i |i\rangle\langle i|$ in some basis ($|i\rangle$) of initial states; see Section 3.1. The probability of a detection event between 0 and t then becomes

$$p_{\omega_{eg}}(\mathbf{r}, t) = \left(\frac{q}{\hbar}\right)^2 \int_{[0, t]^2} e^{i\omega_{eg}(t'-t'')} \langle E(\mathbf{r}, t'')E(\mathbf{r}, t') \rangle_\rho dt' dt''. \quad (92)$$

Equation (92) underlies most of photodetection theory. It relates properties of the field, encoded in a *coherence function* $\langle E(\mathbf{r}, t'')E(\mathbf{r}, t') \rangle_\rho$, to the excitation probability of the detector.

The second important approximation usually made in photodetection theory is the *rotating wave approximation* (RWA), which translates the intuition that a proper photodetection event is an absorption of an excitation. To motivate and illustrate it, we consider first a simple example where the field is assumed to always be in a coherent state.

Example 6 (Monochromatic classical wave). We assume here that the field always remains in a coherent state as time evolves and is factorized with respect to the state of the

qubit; see Section 3.6. Since the free part of the Hamiltonian preserves coherent states up to a phase, what we are assuming is that the evolution under the interaction Hamiltonian approximately preserves coherent states. Physically, this assumption corresponds to, for instance, a macroscopic source emitting a coherent state or a statistical mixture thereof, like a common lightbulb: the fact that one observes the emitted light does not change its classical character. We also assume, for simplicity, that we have an incoming monochromatic coherent state at frequency ω . Formally, we reduce the sum in (83) to a single term \mathbf{k} such that $\omega_{\mathbf{k}} = \omega$, and denote by $\mathcal{E}(t) = \mathcal{E}(\mathbf{r}, t)$ the average value of this field at the fixed position \mathbf{r} of the detector. The terms in \mathbf{r} can be omitted by changing the time origin. More explicitly, we set $\mathcal{E}(t) = \mathcal{E} \sin \omega t$. We want a coherent state $|\alpha\rangle$ such that the average value of the electric field operator

$$\langle \alpha | E(t) | \alpha \rangle = \langle \alpha | a_{\mathbf{k}} e^{-i\omega t} + a_{\mathbf{k}}^\dagger e^{i\omega t} | \alpha \rangle = \mathcal{N}_{\mathbf{k}} (\alpha e^{-i\omega t} + \alpha^* e^{i\omega t})$$

is equal to $\mathcal{E}(t)$. We thus choose the coherent state parameter α to be $-\mathcal{E}/2i\mathcal{N}_{\mathbf{k}}$.²⁹ Furthermore, assume that we initially prepare the detector in its ground state $|g\rangle$. In the interaction picture, our approximation that the state of the field remains the same coherent state³⁰ and factorized allows to write

$$V(t) |g\rangle \otimes |\alpha\rangle \approx (V_{\text{sc}}(t) |g\rangle) \otimes |\alpha\rangle . \quad (93)$$

The operator $V_{\text{sc}}(t)$ in (93) is defined as the interaction picture evolution operator of the so-called *semi-classical* Hamiltonian $H_{\text{sc}}(t) = -q\mathcal{E}(t) \sigma_x(t)$. Using the weak coupling approximation to evaluate the time evolution, we obtain

$$p_{\omega_{eg}}(\mathbf{r}, t) = \text{Tr} \left(U_0^\dagger(t) [|e\rangle\langle e| \otimes \mathbf{1}] U_0(t) V(t) [|g\rangle\langle g| \otimes |\alpha\rangle\langle\alpha|] V^\dagger(t) \right) \quad (94a)$$

$$= \text{Tr} \left([|e\rangle\langle e| \otimes \mathbf{1}] V(t) [|g\rangle\langle g| \otimes |\alpha\rangle\langle\alpha|] V^\dagger(t) \right) \quad (94b)$$

$$= \text{Tr} \left(V_{\text{sc}}(t) |g\rangle\langle g| V_{\text{sc}}^\dagger(t) |e\rangle\langle e| \right) \quad (94c)$$

$$= \left| \frac{iq}{\hbar} \int_0^t \mathcal{E}(t') e^{i\omega_{eg}t'} dt' \right|^2 \quad (94d)$$

$$= \left| \frac{q\mathcal{E}}{2\hbar i} \left(\frac{1 - e^{i(\omega_{eg}-\omega)t}}{\omega_{eg}-\omega} - \frac{1 - e^{i(\omega_{eg}+\omega)t}}{\omega_{eg}+\omega} \right) \right|^2 , \quad (94e)$$

where the second line is obtained because the state $|e\rangle$ is an eigenvalue of the free evolution operator $U_0(t)$, the third follows from the semi-classical assumption that the state of the field always stays in a coherent state, the fourth from the weak coupling assumption and the form of H_{sc} , and the last from the monochromatic assumption.

At this point, we make an extra assumption called the *rotating wave approximation* (RWA). In our special case, the RWA amounts to neglecting the second term of the right-hand side of (94e). Because both ω and ω_{eg} are strictly positive, the denominator of the second term never vanishes, unlike the denominator of the first term. Around resonance $\omega \approx \omega_{eg}$, the first term will largely dominate the second. \diamond

We make the RWA from Example 6 a general principle that applies to arbitrary quantum fields. It echoes the use of the analytic signal in representing a real classical field; see Section 2.

Assumption 6 (Rotating wave approximation). *The photodetection response is a function of the positive frequency part of the quantum field only.*

With Assumption 6, the transition probability (92) now reads

$$p_{\omega_{eg}}(\mathbf{r}, t) = \left(\frac{eq}{\hbar} \right)^2 \int_0^t e^{i\omega_{eg}(t'-t'')} G_\rho^{(1)}(\mathbf{r}t', \mathbf{r}t'') dt' dt'' , \quad (95)$$

²⁹The normalization constant $\mathcal{N}_{\mathbf{k}}$ is reinserted into the coherent state parameter in this derivation to simplify the notation.

³⁰Strictly speaking, we should allow α to depend on time, but the argument below does not change.

where³¹ we introduced the *first-order coherence function*

$$G_\rho^{(1)}(\mathbf{r}' t', \mathbf{r} t) = \langle E^-(\mathbf{r}', t') E^+(\mathbf{r}, t) \rangle_\rho, \quad (96)$$

where E^+ is associated to the positive part of the spectrum (annihilation operators) and E^- is associated to the negative part (creation operators). The coherence function (96) was originally introduced by Glauber [1963].³²

At least informally, it is not difficult to generalize the computations leading to (92) to more sophisticated detectors. In experiments, for instance, photodetection is often performed through ionization: an electron of an atom is extracted by the incident light and this event is amplified and detected. This detector is best described by a system having, on top of its ground $|g\rangle$ state, a countable set of excited states $|e_i\rangle$, and a continuum of so-called *diffusive* states $|p\rangle, p \in \mathbb{R}$. This leads to non-compact Hamiltonian operators, having a spectrum that is partly discrete, partly continuous. Nonetheless, the computation of the transition probability $p(\mathbf{r}, t)$ follows from the same lines as above, with the new observation corresponding to projecting on all states that are not the ground state, that is, $(\sum_n |e_n\rangle\langle e_n| + \int |p\rangle\langle p| dp) \otimes \mathbb{1}$ instead of $|e\rangle\langle e| \otimes \mathbb{1}$. Because we are working with a first-order approximation in Assumption 5, this leads to a probability of transition that is an expectation of the single-state probability of transition (92), namely

$$p(\mathbf{r}, t) = \int p_{\omega_{eg}}(\mathbf{r}, t) d\xi(\omega_{eg}) = \int_{[0, t]^2} \kappa(t' - t'') G_\rho^{(1)}(\mathbf{r} t', \mathbf{r} t'') dt' dt'', \quad (97)$$

where ξ has a support that is the union of a discrete set, corresponding to the states $|e_n\rangle$, and an uncountable set, corresponding to $|p\rangle, p \in \mathbb{R}$, and the structure function is defined as

$$\kappa(\tau) := \left(\frac{e}{\hbar}\right)^2 \int e^{-i\omega\tau} d\xi(\omega).$$

We can think of either ξ or κ as being a characteristic of the detector, describing its time or frequency efficiency. For instance, if ξ is supported on a single point or closely around a single point, we only collect photons at a very specific energy, and call the detector *narrow-band*. If ξ is supported on a large interval, we collect all photons equally, and call the detector *broad-band*.

For completeness, we close this section by recovering the semi-classical treatment of Section 2.3 through our quantum treatment, using the coherent state formalism of Section 3.6.

Example 7 (General rotating wave). By linearity arguments, the monochromatic light case of Example 6 can be used to derive the excitation probability in the case of a more general field. Thinking of the field being enclosed in a large box, so as to be able to work with Fourier series instead of Fourier transforms for simplicity, let us decompose the electric field $\mathcal{E}_\alpha(\mathbf{r}, t)$ on a set of monochromatic plane waves

$$\mathcal{E}_\alpha(\mathbf{r}, t) = \sum_{\mathbf{k}} \mathcal{N}_{\mathbf{k}} \left[\alpha_{\mathbf{k}} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_{\mathbf{k}} t)} + \alpha_{\mathbf{k}}^* e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega_{\mathbf{k}} t)} \right], \quad (98)$$

with $\mathcal{N}_{\mathbf{k}}$ a normalization coefficient and $\alpha_{\mathbf{k}} \in \mathbb{C}$ the amplitude of the mode \mathbf{k} . Note that \mathcal{E} is a (real-valued) function, not an operator: we are still treating the field as a classical field. The first part of the sum corresponds to the positive frequency part of the field (the “analytic” field), while the second part corresponds to the negative frequency part \mathcal{E}^- .

An important observable of the electromagnetic field is its intensity. A general result of Maxwell’s equations [Mandel and Wolf, 1995] is that the energy carried by the field is given

³¹The missing commas in the arguments are a convention that will become handy when we consider higher-order coherence functions.

³²As a side remark, remember that this coherence function is defined after choosing a specific polarization of the field. In the general case, the electric field should be expanded along a basis of the three-dimensional polarization space and we should consider all combinations of coherence functions

$$G_{ij}^{(1)}(t', t) = \langle E_j^-(t') E_i^+(t) \rangle_\rho$$

where i, j label the basis elements.

by the squared modulus of the field. Under our rotating wave Assumption 6, the intensity $I(\mathbf{r}, t)$ of the field reads

$$I(\mathbf{r}, t) = \mathcal{E}_\alpha^-(\mathbf{r}, t)\mathcal{E}_\alpha^+(\mathbf{r}, t) = |\mathcal{E}_\alpha^+(\mathbf{r}, t)|^2. \quad (99)$$

Investigating this intensity, and especially its correlation structure in the case of random sources, has led to the major discoveries of photon bunching and in the end of the general theory of optical coherence in quantum theory [Mandel and Wolf, 1995]. More precisely, say the field (98) is not perfectly known; it is then reasonable to model it as a stochastic process, i.e., a random function. A very common assumption is to assume the process to be a stationary Gaussian process. One partial justification is that the field at a given spacetime point is the sum of the contributions of many independent sources, and is well approximated by a Gaussian. Formally, the field state becomes a statistical mixture of coherent states, a *mixed state* in the vocabulary of Section 3.1. The probability of excitation remains of the form (92), with first-order coherence function

$$G_\rho^{(1)}(\mathbf{r}t', \mathbf{r}t') = \mathbb{E}_\alpha \langle \alpha | E^-(\mathbf{r}', t') E^+(\mathbf{r}, t) | \alpha \rangle \quad (100)$$

$$= \mathbb{E}_\alpha \mathcal{E}_\alpha^-(\mathbf{r}', t') \mathcal{E}_\alpha^+(\mathbf{r}, t) \quad (101)$$

$$= \int p_{\text{classical}}(\alpha) \mathcal{E}_\alpha^-(\mathbf{r}', t') \mathcal{E}_\alpha^+(\mathbf{r}, t) d\alpha. \quad (102)$$

The expectation is over the stochastic process \mathcal{E} . Physicists prefer to denote it informally using a probability density $p_{\text{classical}}$ over the Fourier coefficients $\alpha = (\alpha_{\mathbf{k}})$ of the field in (98), and we use that notation here for future notational compatibility with Glauber-Sudarshan decompositions. In Section 4.2, we shall use Wick's theorem to derive higher-order coherence functions, and recognize the set of correlation functions of a permanent point process.

◇

4.1.3 Role of the detector structure function

In this section, we consider the influence of the detector response κ on the probability of excitation (97).

Broad-band detector. Consider first a broad-band detector, i.e. a detector with structure function that is sharply peaked in time or, equivalently, does not depend much on the frequency. Ideally, this corresponds to $\kappa(t - t') = \kappa_0 \delta(t - t')$ with κ_0 a real constant. The probability of excitation (97) becomes

$$p(\mathbf{r}, t) = \kappa_0 \int_0^t G_\rho^{(1)}(\mathbf{r}t', \mathbf{r}t') dt'. \quad (103)$$

Anticipating over Section 4.3, we would like to map this probability to a point process on the positive half-line \mathbb{R}_+ . One way to associate the two notions is to imagine $t > 0$ infinitely small in (103), and “re-setting” the detector in its ground state at each time multiple of t , keeping track of whether the detector was found in its excited state, say in the middle of each small time interval. Loosely speaking, in the limit $t \rightarrow 0$, these detection times would form a point process on the positive half of the real line, with first correlation function $G_\rho^{(1)}(\mathbf{r}t', \mathbf{r}t')$. To obtain a more formal correspondence, one needs to further model the detection process, for instance including de-excitation of the detector, see e.g. Bouten et al. [2007].

Narrow-band detector. A narrow-band detector is a detector that is well-resolved in frequency, like the two-level atom we used in our derivation of the photodetection signal. At the extreme, we think of the structure function as a delta distribution in Fourier space at

a frequency ω_0 , so that $\mathcal{F}\kappa(\omega) = \kappa_0\delta(\omega - \omega_0)$, and thus $\kappa(t) = \kappa_0/\sqrt{2\pi}e^{i\omega_0 t}$. The excitation probability (97) now reads

$$p(\mathbf{r}, t) = \frac{\kappa_0}{\sqrt{2\pi}} \int_{[0,t]^2} G_\rho^{(1)}(\mathbf{r} t', \mathbf{r} t'') e^{i\omega_0(t' - t'')} dt' dt'' \quad ., \quad (104)$$

The meaning of this equation is even more transparent when we switch to the frequency representation of the coherence function. Indeed, let us define the Fourier representation of the first-order coherence function through the Fourier transform(-like) equation

$$G_\rho^{(1)}(\mathbf{r}' t', \mathbf{r}'' t'') = \int_{\mathbb{R}^2} G_\rho^{(1)}(\mathbf{r}' \omega', \mathbf{r}'' \omega'') e^{-i\omega' t'} e^{i\omega'' t''} \frac{d\omega' d\omega''}{2\pi}, \quad (105)$$

where we follow the physicists' convention to denote both the function and its Fourier transform by the same notation and index the function by the name of its variable. Putting everything together, we arrive at

$$p(\mathbf{r}, t) = \frac{\kappa_0}{\sqrt{2\pi}} \int_{\mathbb{R}^2} G_\rho^{(1)}(\mathbf{r} \omega, \mathbf{r}' \omega') 2t \frac{\sin(t(\omega - \omega_0))}{(\omega - \omega_0)} \frac{d\omega d\omega'}{2\pi}. \quad (106)$$

In the long time limit³³, the integration window is approximately equal to $t\delta(\omega - \omega_0)$. We end up with the simple expression

$$p(\mathbf{r}, t) \approx \frac{\kappa_0 t}{\sqrt{2\pi}} G_\rho^{(1)}(\mathbf{r} \omega_0, \mathbf{r} \omega_0). \quad (107)$$

The interpretation of this expression is easier now. For a narrow-band detector in frequency space centered around a frequency ω_0 , the excitation probability is proportional to the diagonal ω_0 part of the first-order coherence function which is none other than the average intensity of the mode ω_0 .

4.2 Correlation between photodetection events

We now examine the joint probability of detecting two, and then more, photons.

4.2.1 Second-order coherence

Consider two detectors labelled 1 and 2 and placed at two distinct positions \mathbf{r}_1 and \mathbf{r}_2 . They are excited by an incident radiation, and we ask for the probability that detector 1 gets excited at time t_1 and detector 2 gets excited at time t_2 . Actually, the derivation of the joint probability of detection is quite similar to the derivation for the one photodetection event in Section 4.1.2, and we will only give the general principles of the method here again.

The dynamics are once again modeled through a Hamiltonian, whose interaction term is given by the following generalization of (86),

$$H_I(t) = q_1 \sigma_x^{(1)}(t) \otimes \mathbf{1} \otimes E(\mathbf{r}_1, t) + q_2 \mathbf{1} \otimes \sigma_x^{(2)}(t) \otimes E(\mathbf{r}_2, t). \quad (108)$$

Such a form is natural to physicists, in the sense that it models the local interaction of each of two detectors with the field at their respective position.

Computing the joint probability of excitation requires to compute first a matrix element of the evolution operator in the interaction picture, which is again done assuming weak coupling. We will see that in order to have a non-trivial result, we need to push the expansion in Assumption 5 up to second order.

Assumption 7 (Weak Coupling; second order). *The interaction between the detectors and the field is assumed to be weak, allowing to treat the dynamics using perturbation theory. To second order, this amounts to*

$$V(t) \simeq \mathbf{1} - \frac{i}{\hbar} \int_0^t H_I(t_1) dt_1 - \frac{1}{\hbar^2} \int_0^t \int_0^{t_1} H_I(t_1) H_I(t_2) dt_2 dt_1. \quad (109)$$

³³usually the coherence function varies over some timescale, so large time means t sufficiently large compared to that timescale

Being interested in the joint probability of excitation after a time t , we have to compute the matrix element $\langle e_1, e_2, f|V(t)|g_1, g_2, i\rangle$ where $(|g_1\rangle, |g_2\rangle)$ and $(|e_1\rangle, |e_2\rangle)$ are the ground and excited states of both detectors. Since they are all pairwise orthogonal, the $\mathbb{1}$ term in (109) will give a zero contribution. As for the linear term, and given the form of the interaction (108) as a sum of two contributions from each detector, it will split into two terms, each acting trivially on one of the detectors, again yielding a zero result. For instance

$$\langle e_1, e_2, f|\sigma_x^{(1)}(t) \otimes \mathbb{1} \otimes E(\mathbf{r}_1, t)|g_1, g_2, i\rangle = \langle e_1|\sigma_x^{(1)}(t)|g_1\rangle \langle e_2|g_2\rangle \langle f|E(\mathbf{r}_1, t)|i\rangle = 0$$

by orthogonality of $|e_2\rangle$ and $|g_2\rangle$. This is the reason why we have to go to the second order in perturbation theory to compute the joint probability distribution. In general, to compute the joint probability distribution of N detectors clicking, we will need to push the expansion of Assumption 7 to the order N , since none of the lower-order terms contributes. Going back to $N = 2$, the quadratic term also simplifies with only the cross terms $E(\mathbf{r}_2, t_2)E(\mathbf{r}_1, t_1)$ giving a non-zero contribution. Using again the RWA of Assumption 6, we end up with

$$\langle e_1, e_2, f|V(t)|g_1, g_2, i\rangle \quad (110)$$

$$= -\frac{1}{\hbar^2} \int_0^t \int_{t_1}^t e^{i\omega_{eg_1}t_1} e^{i\omega_{eg_2}t_2} \langle f|E^+(\mathbf{r}_2, t_2)E^+(\mathbf{r}_1, t_1)|i\rangle dt_2 dt_1 + 1 \leftrightarrow 2$$

$$= -\frac{1}{\hbar^2} \int_0^t \int_0^t e^{i\omega_{eg_1}t_1} e^{i\omega_{eg_2}t_2} \langle f|E^+(\mathbf{r}_2, t_2)E^+(\mathbf{r}_1, t_1)|i\rangle dt_2 dt_1. \quad (111)$$

The notation $1 \leftrightarrow 2$ is a shorthand for the same term with the two indices transposed. The second line is obtained by grouping the two terms into one and changing the integration range, noting that the involved operators commute. Just like we did in Section 4.1.2, we use Born's rule and sum over the final state of the field, since we are only interested in the state of the detectors, to obtain the joint excitation probability. Once again, we express it as the Fourier transform of a second-order coherence function. Assuming a fixed polarization for the field for simplicity, we have

$$G_\rho^{(2)}(\mathbf{r}'_1 t'_1, \mathbf{r}'_2 t'_2, \mathbf{r}_2 t_2, \mathbf{r}_1 t_1) := \langle E^-(\mathbf{r}'_1, t'_1)E^-(\mathbf{r}'_2, t'_2)E^+(\mathbf{r}_2, t_2)E^+(\mathbf{r}_1, t_1)\rangle_\rho. \quad (112)$$

By once again introducing structure functions κ_1 and κ_2 for both detectors, we obtain our final expression for the joint excitation probability

$$p(\mathbf{r}_2 t, \mathbf{r}_1 t) = \int_0^t \kappa_1(t'_1 - t''_1) \kappa_2(t'_2 - t''_2) G_\rho^{(2)}(\mathbf{r}_1 t'_1, \mathbf{r}_2 t'_2, \mathbf{r}_2 t''_2, \mathbf{r}_1 t''_1) dt'_1 dt''_1 dt'_2 dt''_2, \quad (113)$$

Intensity correlations. Of particular experimental relevance is the diagonal of the second-order coherence function, namely

$$G_\rho^{(2)}(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2, \mathbf{r}_2 t_2, \mathbf{r}_1 t_1) = \langle E^-(\mathbf{r}_1, t_1)E^-(\mathbf{r}_2, t_2)E^+(\mathbf{r}_2, t_2)E^+(\mathbf{r}_1, t_1)\rangle_\rho. \quad (114)$$

For reference, if we were treating a classical field (i.e., modeled as a real- or complex-valued stochastic process rather than a collection of noncommuting operators), all the fields would commute and from the definition of the field intensity, we would obtain

$$G_\rho^{(2)}(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2, \mathbf{r}_2 t_2, \mathbf{r}_1 t_1) = \mathbb{E}I(\mathbf{r}_2, t_2)I(\mathbf{r}_1, t_1). \quad (115)$$

Equation (114) is the non-commuting counterpart to (115), which gives us information about the intensity correlations of the field. The order of the operators in (114), called *normal order*, is crucial for the interpretation of the photodetection signal as the absorption of excitations by the detector, but also to take properly into account the fundamental quantum fluctuations that affect the detection signal.

Intensity correlations can be probed experimentally. For instance, assume a perfect temporal resolution for each detector, and that they are functioning in the time intervals $[0, t_i]$, $i = 1, 2$. This is modeled by choosing for structure functions

$$\kappa(t_i - t'_i) = \kappa_i \Theta_{[0, t_i]} \delta(t_i - t'_i), i = 1, 2,$$

with $\kappa_1, \kappa_2 > 0$. The probability of joint excitation is then given by

$$p(\mathbf{r}_2 t_2, \mathbf{r}_1 t_1) = \kappa_1 \kappa_2 \int_0^{t_2} \int_0^{t_1} G_\rho^{(2)}(\mathbf{r}_1 t', \mathbf{r}_2 t'', \mathbf{r}_2 t'', \mathbf{r}_1 t') dt' dt'' . \quad (116)$$

In particular, taking the derivative in t_1 and t_2 yields the so-called transition rate,

$$\Gamma(\mathbf{r}_2 t_2, \mathbf{r}_1 t_1) = \kappa_1 \kappa_2 G_\rho^{(2)}(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2, \mathbf{r}_2 t_2, \mathbf{r}_1 t_1) , \quad (117)$$

which is proportional to (114), and reduces to the intensity correlation function in the classical setting.

4.2.2 Higher-order coherences

The previous derivation can be generalized straightforwardly to obtain the joint probability of n detection events. Denoting the spacetime coordinates by $x_i = (\mathbf{r}_i, t_i)$, we define the n th order coherence function by

$$G_\rho^{(n)}(x_1, \dots, x_{2n}) = \langle E^-(x_1) \cdots E^-(x_n) E^+(x_{n+1}) \cdots E^+(x_{2n}) \rangle_\rho , \quad n \geq 1. \quad (118)$$

This definition generalizes the first- and second-order coherence functions. Upon noting that $\text{Tr}(\rho O^\dagger O) \geq 0$ for any operator O , we can derive general inequalities for the coherence functions (118). For instance, consider the operator $O = \sum_{k=1}^m \lambda_k E^+(x_k)$, for arbitrary λ_k 's. The nonnegativity of the quadratic form $\text{Tr}(\rho O^\dagger O)$ implies a constraint in the underlying determinant. For $m = 2$, we obtain

$$G_\rho^{(1)}(x, x) G_\rho^{(1)}(y, y) \geq |G_\rho^{(1)}(x, y)|^2 ,$$

which we interpret as a sign of bunching.

Finally, we note that for general states, higher-order coherence functions are typically hard to compute. One special case is for a Gaussian state (59), where Wick's Theorem 3 applies. The n th-order coherence function can then be expanded as

$$G_\rho^{(n)}(x_1, \dots, x_{2n}) = \sum_{\sigma \in \mathfrak{S}_n} \prod_{i=1}^n \langle E^-(x_i) E^+(x_{n+\sigma(i)}) \rangle_\rho . \quad (119)$$

We now have all the tools to identify the point processes behind the detection of photons.

4.3 Turning coherence functions into correlation functions

There is a specific set of states for which the coherence functions are particularly simple, namely the coherent states of Section 3.6. Specifically, take a multi-mode coherent state $|\alpha\rangle = \otimes_{\mathbf{k}} |\alpha_{\mathbf{k}}\rangle$. These states are eigenstates of the electric field (83) itself, with eigenvalue the classical field from (98), i.e.,

$$E^+(\mathbf{r}_1, t_1) |\alpha\rangle = \mathcal{E}_\alpha^+(\mathbf{r}_1, t_1) |\alpha\rangle = \left(\sum_{\mathbf{k}} \mathcal{E}_{\alpha_{\mathbf{k}}}^+(\mathbf{r}_1, t_1) \right) \otimes_{\mathbf{k}} |\alpha_{\mathbf{k}}\rangle . \quad (120)$$

Physically, coherent states correspond to ideal lasers. Now, by construction, the quantum coherence functions (118) are reduced to the classical coherence functions when the field is prepared in such a state, i.e., writing $x_i = (\mathbf{r}_i, t_i)$,

$$\begin{aligned} G_{|\alpha\rangle\langle\alpha|}^{(n)}(x_1, \dots, x_{2n}) &= \langle \alpha | E^-(x_1) \cdots E^-(x_n) E^+(x_{n+1}) \cdots E^+(x_{2n}) | \alpha \rangle \\ &= \prod_{i=1}^n \overline{\mathcal{E}_\alpha^+(x_i)} \mathcal{E}_\alpha^+(x_{n+i}) . \end{aligned} \quad (121)$$

This factorization across coherent states is what allows us recovering the correlation functions of a Poisson point process.

Example 8 (Ideal lasers yield Poisson processes). Take $x_i = x_{n+i} = (\mathbf{r}, t_i)$ for some fixed position $\mathbf{r} \in \mathbb{R}^3$ in (121), and assume a broad-band detector placed at \mathbf{r} , see Section 4.1.3. The joint detection probabilities yield measurement times that follow a point process with correlation functions

$$\rho_n(t_1, \dots, t_n) = |\mathcal{E}_\alpha^+(t_1) \cdots \mathcal{E}_\alpha^+(t_n)|^2.$$

Such a point process indeed exists: it is the Poisson point process of first correlation function $|\mathcal{E}_\alpha^+|^2$; see Section 2.2. \diamond

Now, consider a slightly more general state ρ that can be written as a statistical mixture $\rho = \mathbb{E}_{\alpha \sim p_{\text{classical}}} |\alpha\rangle\langle\alpha|$ of coherent states. Concretely, this state describes non-ideal laser light, e.g. where there is uncertainty on which light amplitude is emitted. For instance, a single-mode randomly phased laser falls into this category, where $p_{\text{classical}}$ is uniform over a circle in the complex plane: the mixture is over all possible phases with a fixed intensity $|\alpha|^2$. As for the coherence functions, by linearity, we obtain

$$G_\rho^{(n)}(x_1, \dots, x_{2n}) = \mathbb{E}_{\alpha \sim p_{\text{classical}}} \prod_{i=1}^n \overline{\mathcal{E}_\alpha^+(x_i)} \mathcal{E}_\alpha^+(x_{n+i}).$$

Example 9 (Mixture states in phase space yield Cox processes). Taking again $x_i = x_{n+i} = (\mathbf{r}, t_i)$ for some fixed position $\mathbf{r} \in \mathbb{R}^3$ in (121) and a broad-band detector, the detection times follow a point process with correlation functions

$$\rho_n(t_1, \dots, t_n) = \mathbb{E}_{\alpha \sim p_{\text{classical}}} |\mathcal{E}_\alpha^+(t_1) \cdots \mathcal{E}_\alpha^+(t_n)|^2.$$

The detection times thus form a Cox point process; see Section 2. In particular, arrival times tend to exhibit bunching: they are more clustered together than a Poisson process. Moreover, assuming that each α_k in (98) is Gaussian, \mathcal{E}_α is a Gaussian process, and the detection times then form a permanent point process. Permanent point processes are the archetype of a point process exhibiting bunching; see Section 2.3 and a few samples in Figure 1(b). The fact that we obtain a permanent point process validates *a posteriori* the semiclassical analysis carried out in Section 2. \diamond

Additionally, note that while we mention ideally resolved detectors in time for simplicity, non-broad-band detectors also yield standard point processes when the detector's characteristic functions can be normalized. Indeed, consider for simplicity the point process formed by the detection times at detectors that share the same characteristic function κ . The correlation functions then correspond to the point process formed by the marks in a marked point process with independent marks. In the vocabulary of [Daley and Vere-Jones, 2003, Chapter 6], the ground process is formed by the ideal detection times, and for each such time t_i , an actual detection time is drawn from the probability density function proportional to $\kappa(\cdot - t_i)$, independently from all other ideal detection times and marks.

Now we consider states that are not statistical mixtures of coherent states, the so-called *non-classical* or *genuinely quantum* states. As we saw in Section 3.6, coherent states form an overcomplete family of states. This suggests that even nonclassical states admit a decomposition as a linear combination of coherent states. The decomposition of any mixed state as a linear combination of projectors onto coherent states is known as the *Glauber-Sudarshan* decomposition. However, the linear decomposition is not a statistical mixture over a *bona fide* probability density, and needs to be interpreted in a weak sense. Informally, physicists write

$$\rho = \int P_\rho(\alpha) |\alpha\rangle\langle\alpha| d\alpha, \quad (122)$$

with $P_\rho(\alpha)$ a tempered distribution for each α ; see [Mandel and Wolf, 1995, Section 11.8] and references therein. For a general state, there is no reason for P_ρ to be a positive function, or even a function at all. We do not attempt to formalize the Glauber-Sudarshan decomposition here, but we study the simplest example of a nonclassical state: a Fock state.

Consider a single-mode Fock state

$$\rho = |k\rangle\langle k| \otimes I \otimes \dots$$

The associated Glauber-Sudarshan distribution, denoted for simplicity $P_{|k\rangle}$, is the “derivative of a Dirac delta”, namely

$$P_{|k\rangle}(\alpha) = \frac{e^{|\alpha|^2}}{k!} \frac{\partial^{2k}}{\partial^k \alpha \partial^k \bar{\alpha}} \delta(\alpha); \quad (123)$$

see [Mandel and Wolf, 1995, Section 11.8]. This representation of a Fock state is very singular and comes with no simple statistical interpretation. To get some intuition about Equation (123), let us fix two arbitrary states $|u\rangle$ and $|v\rangle$. We have in fact the matrix elements

$$\langle u|\rho|v\rangle = \int_{\mathbb{C}^2} \left(\frac{\partial^{2k}}{\partial^k \alpha \partial^k \bar{\alpha}} \frac{e^{|\alpha|^2}}{k!} \langle u|\alpha\rangle \langle \alpha|v\rangle \right) \delta(\alpha) d\alpha \quad (124)$$

$$= \left(\frac{\partial^{2k}}{\partial^k \alpha \partial^k \bar{\alpha}} \frac{e^{|\alpha|^2}}{k!} \langle u|\alpha\rangle \langle \alpha|v\rangle \right) \Big|_{\alpha=\bar{\alpha}=0}, \quad (125)$$

as long as $\alpha \mapsto \langle u|\alpha\rangle \langle \alpha|v\rangle$ is smooth enough. This indeed corresponds to the informal (122), when computing matrix elements and with P_ρ a specific tempered distribution. In particular, if $|m\rangle$ and $|n\rangle$ are two single-mode Fock states, and remembering from Section 3.6 that $\langle n|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \alpha^n / \sqrt{n!}$, we verify that (125) yields

$$\langle n|\rho|m\rangle = \left(\frac{\partial^{2k}}{\partial^k \alpha \partial^k \bar{\alpha}} \frac{e^{|\alpha|^2}}{k!} e^{-|\alpha|^2} \frac{\alpha^n \bar{\alpha}^m}{\sqrt{n!} \sqrt{m!}} \right) \Big|_{\alpha=\bar{\alpha}=0} = \delta_{kn} \delta_{km}$$

which is 1 if $n = m = k$ and otherwise 0, as expected.

As for the coherence functions themselves, we can use this representation to give a form very similar to the statistical mixture we discussed above. Indeed, by linearity, we can write

$$G_\rho^{(n)}(x_1, \dots, x_{2n}) = \int P_\rho(\alpha) \prod_{i=1}^n \overline{\mathcal{E}_\alpha^+(x_i)} \mathcal{E}_\alpha^+(x_{n+i}) d\alpha. \quad (126)$$

As an example, consider again a single-mode state in a Fock state with k photons. The coherence function can then be obtained directly from Equations (123) and (126)

$$G_{|k\rangle}^{(n)}(x_1, \dots, x_{2n}) = \frac{\partial^{2k}}{\partial^k \alpha \partial^k \bar{\alpha}} \left(\frac{e^{|\alpha|^2}}{k!} \prod_{i=1}^n \overline{\mathcal{E}_\alpha^+(x_i)} \mathcal{E}_\alpha^+(x_{n+i}) \right) \Big|_{\alpha=\bar{\alpha}=0}. \quad (127)$$

Example 10 (A candidate point process corresponding to a Fock state.). Taking again $x_i = x_{n+i}$ and a broad-band detector, we obtain candidate correlation functions to define a point process for a single-mode Fock state with k photons

$$\rho_n(t_1, \dots, t_n) = \frac{\partial^{2k}}{\partial^k \alpha \partial^k \bar{\alpha}} \left(\frac{e^{|\alpha|^2}}{k!} |\mathcal{E}_\alpha^+(t_1) \cdots \mathcal{E}_\alpha^+(t_n)|^2 \right) \Big|_{\alpha=\bar{\alpha}=0}. \quad (128)$$

Explicitly, we have

$$\mathcal{E}_\alpha^+(t) = \mathcal{N} \alpha e^{-i\omega t}. \quad (129)$$

Then (128) implies $\rho_n(t_1, \dots, t_n) = 0$ for $k < n$. For $k \geq n$, we obtain

$$\rho_n(t_1, \dots, t_n) = \mathcal{N}^n \frac{\partial^{2k}}{\partial^k \alpha \partial^k \bar{\alpha}} \frac{e^{|\alpha|^2}}{k!} \alpha^n \bar{\alpha}^n \Big|_{\alpha=\bar{\alpha}=0} \quad (130)$$

$$= \frac{\mathcal{N}^n}{k!} \frac{\partial^{2k}}{\partial^k \alpha \partial^k \bar{\alpha}} \frac{(\alpha \bar{\alpha})^k}{(n-k)!} \Big|_{\alpha=\bar{\alpha}=0} \quad (131)$$

$$= \mathcal{N}^n \frac{k!}{(k-n)!}. \quad (132)$$

We find an expression reminiscent of the correlation functions of k independent and uniformly distributed times [Johansson, 2006, Example 2.6]. However, these tentative correlation functions are not integrable, a problem linked with the fact that the plane waves in (83) or (98), while allowing simple formal derivations and a clear separation of positive and negative frequencies, are not actually in L^2 .

In fact, it is possible to build a quantum field theory of radiation replacing plane waves by any orthonormal set of L^2 solutions $\varphi_\ell(\mathbf{r}, t)$ of the classical equation of motion [Fabre and Treps, 2020]. The positive part of the field can then be decomposed on these modes as $\mathcal{E}_\alpha^+(\mathbf{r}, t) = \sum_\ell \alpha_\ell \varphi_\ell^+(\mathbf{r}, t)$, where the $+$ subscript still corresponds to the analytic signal 5 taken in the time variable. The whole framework introduced in Section 3 follows, with ladder operators and coherent states now relative to this new basis. Similarly, Equation (128) stills holds with α corresponding to the coherent state parameters in the chosen modes. However, in the current example of a single-mode Fock state with k photons, the electric field (129) becomes $\mathcal{E}_\alpha^+(t) = \mathcal{N}_\alpha \varphi^+(t)$, where we have dropped the dependence on \mathbf{r} as in (129) since the detector remains at a fixed position. We thus obtain, in lieu of (133),

$$\rho_n(t_1, \dots, t_n) = \mathcal{N}^n \frac{k!}{(k-n)!} |\varphi^+(t_1)|^2 \dots |\varphi^+(t_n)|^2, \quad n \geq 1. \quad (133)$$

Since $\varphi^+ \in L^2$, these are valid correlation functions, namely of the point process formed by k times t_1, \dots, t_k drawn independently with probability density function proportional to $t \mapsto |\varphi^+(t)|^2$. \diamond

4.4 Single-photon sources can lead to anti-bunching

The fact that photons emitted by a Gaussian classical field yield bunched detection times is known as the Hanbury-Brown & Twiss (HBT) effect, and was first evidenced by Hanbury Brown and Twiss [1958]. While we shall come back to HBT-type experiments in Part II, we pause here to insist on the fact that bunching is a consequence of considering multiple independent sources. It is perfectly possible to obtain point processes that anti-bunch with photons, in the sense that the resulting second correlation function $\rho_2(x, y)$ of the point process is small close to the diagonal $\{(x, y) : x = y\}$.

To gain intuition without the quantum overhead, we first recast the HBT effect as a consequence of interference between classical sources. Consider S independent³⁴ classical sources, represented by S independent, zero-mean stochastic processes $\mathcal{E}_s(\mathbf{r}, t)$. The total field is $\mathcal{E}(\mathbf{r}, t) = \sum_s \mathcal{E}_s(\mathbf{r}, t)$.

As derived in Section 4.1, the first-order coherence function is

$$G^{(1)}(\mathbf{r}' t', \mathbf{r} t) = \mathbb{E} \left(\sum_s \mathcal{E}_s^-(\mathbf{r}', t') \right) \left(\sum_s \mathcal{E}_s^+(\mathbf{r}, t) \right). \quad (134)$$

By assumption, all cross-terms in the product have expectation zero, so that

$$G^{(1)}(\mathbf{r}' t', \mathbf{r} t) = \sum_s \mathbb{E} \mathcal{E}_s^-(\mathbf{r}', t') \mathcal{E}_s^+(\mathbf{r}, t) := \sum_s G_s^{(1)}(\mathbf{r}' t', \mathbf{r} t). \quad (135)$$

The total first-order coherence is the sum of the first-order coherence of each source. Quite generically, the sum will tend to zero when $|\mathbf{r} - \mathbf{r}'|$ or $|t - t'|$ are large compared to the so-called *spatial coherence scale* and *correlation time*, respectively.

We can also compute physical quantities like the average field intensity, which, in this simple model, is the sum of the average field intensities of each source,

$$\mathbb{E} I(\mathbf{r}, t) = \mathbb{E} \mathcal{E}^*(\mathbf{r}, t) \mathcal{E}(\mathbf{r}, t) = G^{(1)}(\mathbf{r} t, \mathbf{r} t) = \sum_s \mathbb{E} (|\mathcal{E}_s(\mathbf{r}, t)|^2) = \sum_s \mathbb{E} I_s(\mathbf{r}, t). \quad (136)$$

³⁴For instance, we neglect interatomic coupling through the radiation for atomic sources.

To go further and be able to discuss bunching or anti-bunching, we need to consider the second-order coherence function, which characterizes the density of pairs of detection events; see (118). It writes

$$G^{(2)}(\mathbf{r}'_1 t'_1, \mathbf{r}'_2 t'_2, \mathbf{r}_2 t_2, \mathbf{r}_1 t_1) \quad (137)$$

$$= \mathbb{E} \left(\sum_s \mathcal{E}_s^-(\mathbf{r}'_1, t'_1) \right) \left(\sum_s \mathcal{E}_s^-(\mathbf{r}'_2, t'_2) \right) \left(\sum_s \mathcal{E}_s^+(\mathbf{r}_2, t_2) \right) \left(\sum_s \mathcal{E}_s^+(\mathbf{r}_1, t_1) \right). \quad (138)$$

Using our assumptions of independence and zero mean, this coherence function is a sum of one-source and two-source terms. We focus on two-source terms. Developing (138), the contribution of two-source terms to $\mathbb{E}I(\mathbf{r}, t)I(\mathbf{r}', t')$ is

$$\sum_{s \neq s'} \mathbb{E} (|\mathcal{E}_s^+(\mathbf{r}, t)\mathcal{E}_{s'}^+(\mathbf{r}', t') + \mathcal{E}_s^+(\mathbf{r}', t')\mathcal{E}_{s'}^+(\mathbf{r}, t)|^2). \quad (139)$$

At coincidence $(\mathbf{r}, t) = (\mathbf{r}', t')$, both amplitudes are the same and give a contribution $4|\mathcal{E}_s^+(\mathbf{r}, t)\mathcal{E}_{s'}^+(\mathbf{r}, t)|^2$. However, when (\mathbf{r}, t) and (\mathbf{r}', t') differ, a phase appears between the two contributions and reduces the modulus of the sum. We recover here the bunching effect of light as an interference effect between pairs of complex amplitudes. Note that $S = 2$ sources are enough to create this interference.

If we want to observe the opposite effect, a *reduction* in the rate at coincidence, we need a single-photon source such as a single two-level atom described by a Hilbert space \mathbb{C}^2 excited by a pumping laser field. A natural analogy is that of a gun (the atom) that needs to be reloaded (by the laser pump) before firing a new bullet (the photon). We shall see that the need to reload implies anti-bunching for the fired photons.

What we want to compute is once again the second-order coherence between two detections at the same photodetection location r , but delayed by a time τ , and abusively denoted by $G_\rho^{(2)}(r, r, \tau)$. The whole state space of the source, the field and the detector is then $\mathbb{C}^2 \otimes \mathcal{H}^\infty \otimes \mathbb{C}^2$. An accurate model of the situation would require a proper description of the interaction between the field, the source and the detector but we only give some elements of the general computation; see [Cohen-Tannoudji, 1979] for details. The new element here is the interaction of the field and the source and the main technical problem is to relate the field $E^+(t)$ emitted by the source to some of its physical properties. We ask the reader to admit that $E^+(t) \propto \sigma_-(t)$, where $\sigma_-(t) = U^\dagger(t)(\sigma_-(0) \otimes I \otimes I)U(t)$. Physically, this expression means that when the source is excited and then relaxes to its ground state, it emits light. The second-order coherence of the field can then be written as a function of the time evolution of a source operator as

$$\begin{aligned} G_\rho^{(2)}(r, r, \tau) &\propto \langle \sigma_+(0)\sigma_+(\tau)\sigma_-(\tau)\sigma_-(0) \rangle_\rho \\ &\propto \langle \sigma_+(0)U^\dagger(\tau)\sigma_+(0)U(\tau)U^\dagger(\tau)\sigma_-(0)U(\tau)\sigma_-(0) \rangle_\rho. \end{aligned} \quad (140)$$

Naturally, we assume that the initial state of the atom is excited and the field is in the vacuum so that $\rho = |e\rangle\langle e| \otimes |0\rangle\langle 0|$. With $\sigma_-(0) = |g\rangle\langle e|$ we end up with the very simple form:

$$\begin{aligned} G_\rho^{(2)}(r, r, \tau) &\propto \text{Tr} ((|g\rangle\langle g| \otimes |0\rangle\langle 0|) U^\dagger(\tau) |e\rangle\langle e| U(\tau)) \\ &\propto \langle e|U(\tau)(|g\rangle\langle g| \otimes |0\rangle\langle 0|)U^\dagger(\tau)|e\rangle. \end{aligned} \quad (141)$$

Two comments are in order. First, we could have also assumed a more general form for the atomic density matrix which would amount to multiply the above expression by the probability $p(e)$ for the atom to be initially excited. Its appearance is natural since the atom has to be excited initially. Second, the physical interpretation of the second-order coherence (141) is as follows. Starting from a factorized state in which the atom has already emitted a photon and is thus in its ground state, the system evolves until time τ where we compute the probability for the atom to be excited again. Equation (141) is the conditional probability that, knowing that the atom is in its ground state initially, it gets excited at time τ .

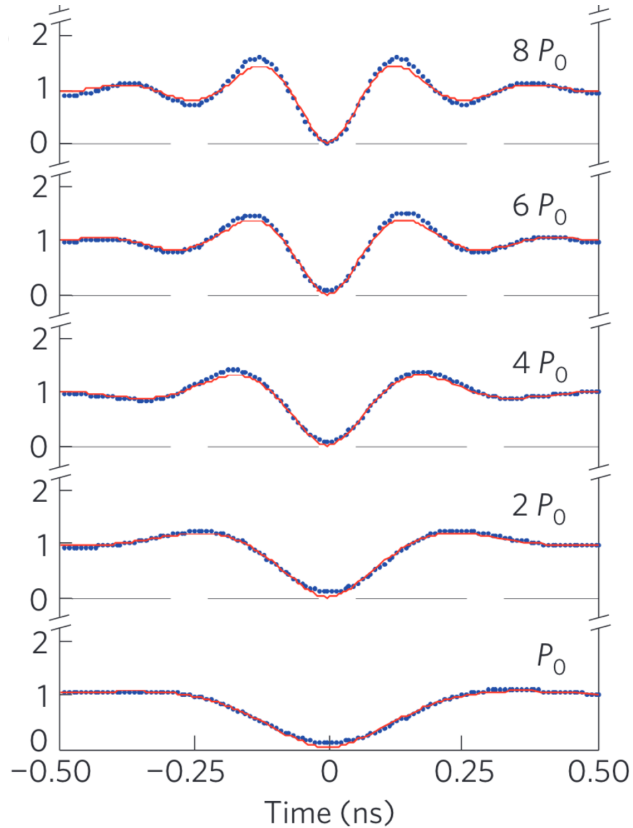


Figure 3: Experimental measurement of the second-order correlation signal of light coming from a single quantum dot showing anti-bunching at coincidence (zero time delay). The traces show different pump laser powers, all of which are multiples of the lowest power P_0 . The oscillations at larger delay correspond to so-called *Rabi oscillations* between the two energy levels; as the pump power is increased these oscillations increase in frequency. The solid line is a fit to a model. Figures taken from Flagg et al. [2009].

From there, predicting photon anti-bunching at coincidence is straightforward: near coincidence $\tau = 0$, the evolution operator is $U(0) = \mathbf{1}$. Since $|e\rangle$ and $|g\rangle$ are orthogonal, this implies

$$G_\rho^{(2)}(r, r, 0) = 0. \quad (142)$$

In other words, at coincidence, the probability to have a second emission is zero. This result is in fact physically quite intuitive if we think at the single-source level. Indeed, after the atom emits one photon, it is surely in its ground state and cannot emit a second photon right away. It has to be re-excited first by the laser source before emitting the second photon. A typical experimental signal is shown in (3) where we clearly see the coherence signal going to zero at coincidence. The oscillation pattern is also nicely understood as Rabi oscillations between the two atomic levels [Degiovanni et al., 2021]. Nowadays, this anti-bunching is used as signature of a good single photon source, along with a second major interference effect called Hong-Ou Mandel; see e.g. [Degiovanni et al., 2021].

5 Electro-detection and fermionic coherences

While photonic optics experiments have reached a high level of sophistication [Haroche and Raimond, 2006, Grynberg et al., 2010], it is only recently that electronic coherences

have started to be probed, thanks to experimental advances in the manipulation, control and measurement of small quantum systems in condensed matter and atomic physics. Formally, the main difference with bosons comes from the anti-commutation of the associated ladder operators, resulting in the Pauli exclusion principle and in determinants appearing in Wick's Theorem 3. Moreover, while we limit ourselves to free (i.e., non-interacting) fermions in this section, in realistic physical models fermions are usually subject to the effects of interaction. These two aspects drastically change the physics compared to photons. Photons are noninteracting objects with a truly empty reference state, the vacuum. Electrons, on the contrary, are subject to Coulomb interaction and their ground state, called the Fermi sea and representing a metal at equilibrium, is full of fermions: their vacuum is not empty.

This Section follows the lines of Section 4.2, explaining how similar computations are affected by the fermionic character of the particles. In Section 5.1, we remind what a fermionic field is and discuss how to model the detection of electrons. In Section 5.2, we discuss the coherence functions of arbitrary orders, and use the second-order one in Section 5.3 to justify the default anti-bunching character of fermions. In Section 5.4, we finally discuss the difficulties in recovering classical currents in the way coherent states do for bosonic fields.

5.1 Modeling electrodetetection events

We proceed as for photodetection by first modeling the field and then the detection, so as to be able to write correlation functions.

5.1.1 Modeling the fermionic field

A *fermionic quantum field* $\psi(\mathbf{r}, t)$ is a collection of operators indexed by space and time and acting on a Fock space $\mathcal{H}_{\text{Fermions}}^\infty$ of fermions; see Section 3.4. The field models an indefinite number of fermionic particles, like electrons. Similarly to a bosonic field, the state of a fermionic system is built by acting on a reference state with operators that create or annihilate fermions in a given state of a basis of the single-particle Hilbert space. When it is clear that we speak of fermions,³⁵ these ladder operators are usually denoted as c^\dagger and c , and indexed by the label of the single-particle state. Unlike for bosons, the creation and annihilation operators satisfy the canonical *anticommutation* relations (48).

In theory, the reference state can be taken to be the vacuum, meaning the complete absence of any particle, as we did for bosons. But when modeling electronic experiments, physicists often have to take another reference state, labeled as $|F\rangle$, and called the *Fermi sea*. The Fermi sea is meant to represent, for instance, the metal used in experiments, which is itself full of electrons. In the simplest case, corresponding to temperature zero, the Fermi sea is a state with one fermion in every energy level of the considered system, up to some reference level called the *Fermi level*. For instance, in a system built using a single-particle Hamiltonian with eigenpairs $(\varphi_n, \varepsilon_n)_{n \in \mathbb{N}}$, so that c_n^\dagger creates a particle in mode φ_n , the Fermi sea is

$$|F\rangle = \prod_{n: \varepsilon_n < \varepsilon_F} c_n^\dagger |0\rangle,$$

with the Fermi level ε_F defined by the experimental setting. Like the vacuum, the Fermi sea can also be built as the limit of a Gaussian state when $\beta \rightarrow \infty$ and the chemical potential ζ is fixed as a function of ε_F ; see the computations in Section 6.

Remark 8. Depending on the situation to model, more physical parameters can be introduced to describe the field in addition to the spacetime coordinates. For instance, the spin of the fermionic excitation could contribute, or more fermionic fields can be used to describe different channels, as in the description of the quantum Hall effect [Ezawa, 2008].

³⁵ a^\dagger and a are usually kept for photons, or general particles.

5.1.2 First-order electronic coherence

As we did for photons, we consider a two-level detector with Hilbert space \mathbb{C}^2 probing a fermionic field $\psi(\mathbf{r})$ living in the Hilbert space $\mathcal{H}_{\text{Fermions}}^\infty$.³⁶ The dynamics of the detector is described here by a fermionic operator d that satisfies the anticommutation rule $\{d, d^\dagger\} = 0$. Their interaction, again assumed to be weak, is modeled by Hamiltonian (in the interaction picture)

$$H_I(t) = d^\dagger(t)\psi(\mathbf{r}, t) + d(t)\psi^\dagger(\mathbf{r}, t), \quad (143)$$

Note that we place ourselves in the interaction picture, so that the time dependence in (143) results from the evolution of the free part of the Hamiltonian. The first term of the interaction Hamiltonian describes the absorption process of an electron that excites the detector while the second term describes an electron being emitted by the detector. With the same assumptions as in the photodetection problem, we can compute the excitation probability of the detector as a function of an electronic correlation function, naturally called in this context the *first-order electronic coherence function*, and defined as

$$G_\rho^{(1e)}(\mathbf{r}' t', \mathbf{r} t) = \langle \psi^\dagger(\mathbf{r}', t')\psi(\mathbf{r}, t) \rangle_\rho. \quad (144)$$

Remark 9. In concrete condensed matter systems, we argued above that the natural ground state is in fact a Fermi sea $|F\rangle$, contrary to the photonic case where the natural ground state is the true vacuum $|0\rangle$. In such a context, it is then also interesting to consider the de-excitation probability of the detector, sending back the electron into the system, or equivalently creating a *hole* in the Fermi sea. This transition probability is controlled by the *first-order hole coherence function*

$$G_\rho^{(1h)}(\mathbf{r}' t', \mathbf{r} t) = \langle \psi(\mathbf{r}', t')\psi^\dagger(\mathbf{r}, t) \rangle_\rho. \quad (145)$$

Note however that the two types of coherence functions are not independent. Indeed, thanks to the canonical anticommutation relations, they satisfy the equal time relation

$$G_\rho^{(1e)}(\mathbf{r}' t, \mathbf{r} t) + G_\rho^{(1h)}(\mathbf{r}' t, \mathbf{r} t) = \delta(\mathbf{r}' - \mathbf{r}).$$

In the following, we work with one of the two coherence functions, and we pick $G_\rho^{(1e)}$ by convention.

Example 11 (Free fermions at non-zero temperature T). The simplest example we can consider is a set of free fermions. In particular, we consider the free field Hamiltonian $H_{\text{FF}} = \sum_{\mathbf{k}} \nu_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}}$, and further assume that the energy level $\nu_{\mathbf{k}} = \nu_k$ only depends on $k = \|\mathbf{k}\|$. We prepare the state in the grand-canonical ensemble (60), which we denote here as $\rho = \rho_{\beta, \zeta}$. Note that ρ commutes with H_{FF} , so that Schrödinger's equation yields that ρ does not change in time: it is an equilibrium state. In particular, the coherence function $G_\rho^{(1e)}(\mathbf{r}' t', \mathbf{r} t)$ depends only on $t - t'$. We arbitrarily set $t = t'$, and write $G_\rho^{(1e)}(\mathbf{r}', \mathbf{r})$. Introducing the Fourier transform of the annihilation operator

$$\psi(\mathbf{r}) = \int_{\mathbb{R}^3} c_{\mathbf{k}} e^{i(\mathbf{k} \cdot \mathbf{r})} d\mathbf{k} / (2\pi)^3,$$

we can write the first electronic coherence function as

$$G_\rho^{(1e)}(\mathbf{r}', \mathbf{r}) = \langle \psi^\dagger(\mathbf{r}')\psi(\mathbf{r}) \rangle_\rho = \int_{\mathbb{R}^6} \langle c_{\mathbf{k}'}^\dagger c_{\mathbf{k}} \rangle_\rho e^{i(\mathbf{r} \cdot \mathbf{k} - \mathbf{r}' \cdot \mathbf{k}')} \frac{d\mathbf{k}}{(2\pi)^3} \frac{d\mathbf{k}'}{(2\pi)^3}. \quad (146)$$

Now note that³⁷

$$\langle c_{\mathbf{k}'}^\dagger c_{\mathbf{k}} \rangle_\rho = f_{\beta, \zeta}(k) \delta(\mathbf{k} - \mathbf{k}')$$

³⁶Actual experimental settings are more intricate to describe than for photons exciting a two-level atom; Part II will contain examples.

³⁷This computation is standard in physics. At this stage, we take it for granted; but see Equation (181) for an explicit derivation.

with $k = \|\mathbf{k}\|$ where $f_{\beta,\zeta}$ is the Fermi-Dirac distribution

$$f_{\beta,\zeta}(k) = \frac{1}{e^{\beta(\nu_k - \zeta)} + 1}. \quad (147)$$

We obtain

$$G_{\rho}^{(1e)}(\mathbf{r}', \mathbf{r}) = \int_{\mathbb{R}^3} f_{\beta,\zeta}(k) e^{i(\mathbf{r}-\mathbf{r}') \cdot \mathbf{k}} \frac{d\mathbf{k}}{(2\pi)^3}. \quad (148)$$

A practically relevant case is the zero-temperature limit, where the grand-canonical ensemble state is replaced by the Fermi sea itself $|F\rangle\langle F|$. The Fermi-Dirac distribution then simplifies to the indicator $\mathbb{1}_{B(k_F)}$ of the centered ball $B(k_F)$ of radius k_F known as the Fermi momentum, which is an input of the model³⁸. The first-order electronic coherence function is then

$$\begin{aligned} G_{|F\rangle}^{(1e)}(\mathbf{r}', \mathbf{r}) &= \frac{1}{\pi^2 \|\mathbf{r} - \mathbf{r}'\|} \int_0^{k_F} k \sin(k \|\mathbf{r} - \mathbf{r}'\|) dk \\ &= \frac{1}{\pi^2 \|\mathbf{r} - \mathbf{r}'\|} \left(\frac{\sin(k_F \|\mathbf{r} - \mathbf{r}'\|)}{\|\mathbf{r} - \mathbf{r}'\|^2} - \frac{k_F}{\|\mathbf{r} - \mathbf{r}'\|} \cos(k_F \|\mathbf{r} - \mathbf{r}'\|) \right). \end{aligned} \quad (149)$$

◇

Example 12 (Chiral free fermions at non-zero temperature T). A second example appears naturally when discussing electronic coherences. This time, we consider a simple model of a *chiral quantum wire*, that is, electrons moving in one dimension and only in a given direction (e.g. to the right). The Hilbert space is similar to Example 11, $\mathbb{C}^2 \times \mathcal{H}_{\text{Fermions}}^{\infty}$, with $\mathcal{H} = L^2(\mathbb{R})$ as a single-particle Hilbert space. Rather than describing the Hamiltonian, a typical shortcut to describe the model is to directly discuss the solutions to Schrödinger's equation, or, as we do here for the field operators, how operators evolve in the Heisenberg picture.

Compared to Example 11, we are not at equilibrium, and thus need to keep track of the time dependence of the field $\psi(x, t)$. The Fourier variable conjugated to the time t is denoted by ω , just like the momentum \mathbf{k} is conjugated to the position \mathbf{r} . Being in one dimension in this example, the vectors \mathbf{r} and \mathbf{k} are in fact scalar quantities that we will denote x and k respectively. To further constrain the model, we posit an equation of motion for the field, i.e., a relation between its time- and space derivatives³⁹. The non-zero solutions of this equation are described by a relation between all the Fourier variables, called a *dispersion relation*. In our quantum wire model, the dispersion relation is assumed to be linear, meaning that the eigenvalues of the free field Hamiltonian are given by $\nu_k = v_F |k|$, with v_F the so-called Fermi velocity of the electron in the quantum wire. Finally, we again assume that the electrons coming into the wire come from a reservoir of particles, whose state is the grand-canonical ensemble $\rho = \rho_{\beta,\zeta}$ at temperature $\beta = 1/k_B T$ and chemical potential ζ .

Now that we have all the additional elements for this model, we can proceed to compute the first-order coherence function. Assuming that the field lives in a large box, from the Fourier decomposition $\psi(x, t) = \sum_k c_k e^{i(kx - \nu_k t)}$, the first-order coherence expands as

$$G_{\rho}^{(1e)}(x' t', x t) = \langle \psi^{\dagger}(x', t') \psi(x, t) \rangle_{\rho} = \sum_{k, k'} \langle c_{k'}^{\dagger} c_k \rangle_{\rho} e^{i(kx - k'x' - (\nu_k t - \nu_{k'} t'))}. \quad (150)$$

The average value of $c_{k'}^{\dagger} c_k$ on the grand canonical state ρ can be shown to be $\langle c_{k'}^{\dagger} c_k \rangle = \delta(k - k') f_{\beta,\zeta}(k)$. In short, the first-order electronic coherence is simply a Fourier transform of the Fermi-Dirac distribution. Doing the change of variable to ω , we have

$$G_{\rho}^{(1e)}(x' t', x t) = \int_{\mathbb{R}} \frac{1}{e^{\beta(\hbar\omega - \zeta)} + 1} e^{i(x - \nu_F t - (x' - \nu_F t')) \frac{\omega}{v_F}} \frac{d\omega}{2\pi v_F}. \quad (151)$$

³⁸Equivalently, we could have specified the chemical potential ζ to a certain value ζ_F , called the Fermi chemical potential. Moreover, k_F is related to the Fermi velocity v_F or the Fermi energy ν_F by the free Hamiltonian so these too can be used as inputs of the model.

³⁹Again, this is a shortcut to avoid describing the Hamiltonian and Schrödinger's equation.

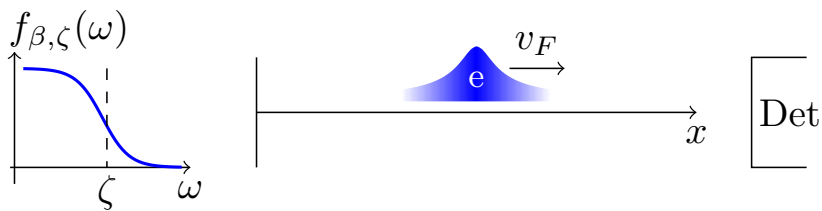


Figure 4: Representation of a simple model of a quantum wire where fermions are sent from a reservoir at equilibrium (grand-canonical ensemble) into a one-dimensional wire where they can travel and are probed by an electro-detection device. Detection probabilities are given by electronic coherence functions.

As it is traditionally done in this field, we will use a rescaled time variable $t - x/v_F$ as t . This simplification between time and space is very special to the one-dimensional case thanks to the ballistic propagation of the electron. We then write the simpler expression

$$G_\rho^{(1e)}(t', t) = \int_{\mathbb{R}} \frac{1}{e^{\beta(\hbar\omega - \zeta)} + 1} e^{i(t'-t)\omega} \frac{d\omega}{2\pi v_F}. \quad (152)$$

As such, the integral is divergent when $t = t'$ and has to be regularized. This divergence has a physical origin: the linear dispersion relation posits a Fermi sea with an infinite number of fermions with negative energy. In practice however, the system has a finite size and the Fermi sea has a finite bandwidth. The strategy is then to introduce a small parameter $\varepsilon > 0$, which corresponds to the inverse of that bandwidth, and compute the integral with the substitution $t - t' \rightarrow t - t' + i\varepsilon$. With these modifications, the residue theorem gives

$$G_F^{(1e)}(t', t) = \frac{i}{2\pi v_F \tau_{\text{th}}} \frac{e^{-i\frac{\zeta}{\hbar}(t-t')}}{\sinh\left(\frac{t-t'+i\varepsilon}{\tau_{\text{th}}}\right)}, \quad (153)$$

where $\tau_{\text{th}} = \hbar\beta/\pi$ is called the *thermal coherence time*. A particular case to have in mind is the zero-temperature limit of the coherence function. It can in fact be computed directly as the Fourier transform (in the sense of distributions) of the zero-temperature Fermi-Dirac energy density distribution $f_{0, \zeta}(\omega) = \Theta(\zeta - \omega)$, where Θ is the Heavyside function. It reads⁴⁰

$$G_F^{(1e)}(t', t) = \frac{i}{2\pi v_F} \frac{e^{-i\frac{\zeta}{\hbar}(t-t')}}{t - t' + i\varepsilon}. \quad (155)$$

◇

5.2 Correlation between electro-detection events

The first-order coherence gives information about the system at the one-particle level. If we want to probe two-particle or higher levels of information like the correlations between electro-detection events, we have to study the coherence functions of higher order. They are defined similarly as the photonic coherences of Section 4. For instance, the second-order electronic and hole coherence functions are defined as

$$G_\rho^{(2e)}(\mathbf{r}'_1 t'_1, \mathbf{r}'_2 t'_2, \mathbf{r}_2 t_2, \mathbf{r}_1 t_1) = \langle \psi^\dagger(\mathbf{r}'_1, t'_1) \psi^\dagger(\mathbf{r}'_2, t'_2) \psi(\mathbf{r}_2, t_2) \psi(\mathbf{r}_1, t_1) \rangle_\rho \quad (156a)$$

$$G_\rho^{(2h)}(\mathbf{r}'_1 t'_1, \mathbf{r}'_2 t'_2, \mathbf{r}_2 t_2, \mathbf{r}_1 t_1) = \langle \psi(\mathbf{r}'_1, t'_1) \psi(\mathbf{r}'_2, t'_2) \psi^\dagger(\mathbf{r}_2, t_2) \psi^\dagger(\mathbf{r}_1, t_1) \rangle_\rho. \quad (156b)$$

⁴⁰The $+i\varepsilon$ notation is a shorthand notation for the more rigorous distribution

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{x + i\varepsilon} = \text{P} \left(\frac{1}{x} \right) - i\pi \delta(x). \quad (154)$$

This straightforwardly generalizes to higher orders. Using the shorthand notation $x = (\mathbf{r}, t)$ for the spatio-temporal coordinates, the n -th-order electronic and hole coherences are

$$G_\rho^{(ne)}(x'_1, \dots, x'_n, x_n, \dots, x_1) = \langle \psi^\dagger(x'_1) \cdots \psi^\dagger(x'_n) \psi(x_n) \cdots \psi(x_1) \rangle_\rho \quad (157a)$$

$$G_\rho^{(nh)}(x'_1, \dots, x'_n, x_n, \dots, x_1) = \langle \psi(x'_1) \cdots \psi(x'_n) \psi^\dagger(x_n) \cdots \psi^\dagger(x_1) \rangle_\rho. \quad (157b)$$

Example 13 (Continuation of Example 11). The grand canonical ensemble $\rho = \rho_{\beta, \zeta}$ for free fermions is Gaussian, so that Theorem 3 (Wick's) applies and we obtain a 2×2 determinant

$$G_\rho^{(2e)}(\mathbf{r}'_1, \mathbf{r}'_2, \mathbf{r}_2, \mathbf{r}_2) = \langle \psi^\dagger(\mathbf{r}'_1) \psi^\dagger(\mathbf{r}'_2) \psi(\mathbf{r}_2) \psi(\mathbf{r}_2) \rangle_\rho \quad (158a)$$

$$\stackrel{\text{Wick}}{=} G_\rho^{(1e)}(\mathbf{r}'_1, \mathbf{r}_1) G_\rho^{(1e)}(\mathbf{r}'_2, \mathbf{r}_1) - G_\rho^{(1e)}(\mathbf{r}'_1, \mathbf{r}_2) G_\rho^{(1e)}(\mathbf{r}'_2, \mathbf{r}_1). \quad (158b)$$

The minus sign is of course a consequence of the fermionic character of the electronic field. In fact, we can already see at this level the repulsion between different fermions. Indeed, if we look for instance at the diagonal part $\mathbf{r}'_1 = \mathbf{r}_1$ and $\mathbf{r}'_2 = \mathbf{r}_2$, in the zero temperature limit, we have

$$G_{|F\rangle}^{(2e)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_2, \mathbf{r}_1) \propto \left(1 - \left| g_{|F\rangle}^{(1e)}(\|\mathbf{r}_1 - \mathbf{r}_2\|) \right|^2 \right), \quad (159)$$

where $g_{|F\rangle}^{(1e)}(\|\mathbf{r}_1 - \mathbf{r}_2\|)$ is the first-order electronic coherence function (149) normalized to unity. We clearly see that when $\mathbf{r}_1 = \mathbf{r}_2$, the second-order coherence function is exactly zero: two fermions cannot be at the same place at the same time. This is another avatar of the general anti-bunching effect of fermionic systems. \diamond

Example 14 (Second-order coherence of N excitations). To gain some intuitive understanding of the information content of the second-order coherence function, let's continue Example 12 of a quantum wire. Remember that in this simple one-dimensional model, time and space coordinates are identified thanks to the linear dispersion relation and that we are looking for the coherence functions at the fixed position of the detector (position that we omit in the arguments to simplify the notations). Consider a state containing two electrons above the true vacuum $|0\rangle$ (not the Fermi sea) in two orthogonal wavefunctions $|\varphi_1, \varphi_2\rangle = \psi^\dagger[\varphi_1] \psi^\dagger[\varphi_2] |0\rangle$, where $\psi[\varphi] = \int \psi(t) \varphi(t) dt$. Note again that wavefunctions, strictly speaking functions of x and t , are here functions of $x - v_F t$ only because of the linear dispersion relation. Since we further consider a fixed position here (that of the detector), we simply write $t \mapsto \psi(t)$ and $t \mapsto \varphi(t)$.

Since the true vacuum is a Gaussian state, we can apply Wick's Theorem 3 to the second-order coherence function

$$\begin{aligned} G_{|\varphi_1, \varphi_2\rangle}^{(2e)}(t'_1, t'_2, t_2, t_1) &= \int_{\mathbb{R}^4} \langle 0 | \psi(y'_1) \psi(y'_2) \psi^\dagger(t'_1) \psi^\dagger(t'_2) \psi(t_2) \psi(t_1) \psi^\dagger(y_2) \psi^\dagger(y_1) | 0 \rangle \\ &\quad \varphi_1^*(y'_1) \varphi_2^*(y'_2) \varphi_2(y_2) \varphi_1(y_1) dy_1 dy'_1 dy_2 dy'_2 \\ &= \varphi_1^*(t'_1) \varphi_2^*(t'_2) \varphi_1(t_1) \varphi_2(t_2) + \varphi_1^*(t'_2) \varphi_2^*(t'_1) \varphi_1(t_2) \varphi_2(t_1) \\ &\quad - \varphi_1^*(t'_2) \varphi_2^*(t'_1) \varphi_1(t_1) \varphi_2(t_2) - \varphi_1^*(t'_1) \varphi_2^*(t'_2) \varphi_1(t_2) \varphi_2(t_1). \end{aligned} \quad (160)$$

We see that the form of this coherence contains the expected anti-symmetries coming with fermionic statistics. It can be written even more explicitly by introducing the antisymmetrized wave-function

$$\Phi_{12}(t_1, t_2) = \det \begin{bmatrix} \varphi_1(t_1) & \varphi_2(t_1) \\ \varphi_1(t_2) & \varphi_2(t_2) \end{bmatrix}. \quad (161)$$

The second-order electronic coherence function then reads

$$G_{|\varphi_1, \varphi_2\rangle}^{(2e)}(t'_1, t'_2, t_2, t_1) = \Phi_{12}^*(t'_1, t'_2) \Phi_{12}(t_1, t_2). \quad (162)$$

	φ_{e_1}	\cdots	φ_{e_k}	\cdots	φ_{e_N}
$G^{(1e)}$	$\varphi_{e_1}^* \varphi_{e_1}$	\cdots	$\sum_{p=1}^k \varphi_{e_p}^* \varphi_{e_p}$	\cdots	$\sum_{p=1}^N \varphi_{e_p}^* \varphi_{e_p}$
\vdots	\vdots	\ddots			\vdots
$G^{(ke)}$	0		$\Phi_{1,\dots,k}^* \Phi_{1,\dots,k}$		$\sum_{i_1 < \dots < i_k} \Phi_{i_1,\dots,i_k}^* \Phi_{i_1,\dots,i_k}$
\vdots	\vdots			\ddots	\vdots
$G^{(Ne)}$	0	\cdots	0	\cdots	$\Phi_{1,\dots,N}^* \Phi_{1,\dots,N}$

Figure 5: Table showing the structure of the electronic coherence functions of order up to N in the presence of one up to N fermions in some given wavepackets.

The second-order coherence function is thus essentially the full many-body wavefunction of a two-particle state.

When more than two fermions are present, the second-order coherence function only keeps track of two-particle terms. To see this, consider a state containing N electrons above the vacuum in mutually orthogonal wavepackets $|\psi_N\rangle = \psi^\dagger[\varphi_1] \dots \psi^\dagger[\varphi_N] |0\rangle$. A straightforward, but cumbersome, application of Wick's theorem gives

$$G_{|\psi_N\rangle}^{(2e)}(t'_1, t'_2, t_2, t_1) = \sum_{k < l} \Phi_{kl}^*(t'_1, t'_2) \Phi_{kl}(t_1, t_2). \quad (163)$$

Informally, the second-order coherence function contains only the “two-particle physics”.

◇

Example 15 (k th-order coherence of N excitations). Continuing the quantum wire example, we can understand its k -particle physics by computing the k th-order coherence function. Consider again N fermions prepared in mutually orthogonal wavepackets $|\psi_N\rangle = \psi^\dagger[\varphi_1] \dots \psi^\dagger[\varphi_N] |0\rangle$. By Wick's theorem, the k th-order coherence function is

$$G_{|\psi_N\rangle}^{(ke)}(t'_1, \dots, t'_k, t_1, \dots, t_k) = \sum_{i_1 < \dots < i_k} \Phi_{i_1, \dots, i_k}^*(t'_1, \dots, t'_k) \Phi_{i_1, \dots, i_k}(t_1, \dots, t_k), \quad (164)$$

with

$$\Phi_{i_1, \dots, i_k}(t_1, \dots, t_k) = \det \begin{bmatrix} \varphi_{i_1}(t_1) & \cdots & \varphi_{i_k}(t_1) \\ \vdots & \ddots & \vdots \\ \varphi_{i_1}(t_k) & \cdots & \varphi_{i_k}(t_k) \end{bmatrix}. \quad (165)$$

In particular, the N -th coherence function is

$$G_{|\psi_N\rangle}^{(Ne)}(t'_1, \dots, t'_N, t_1, \dots, t_N) = \Phi_{1, \dots, N}^*(t'_1, \dots, t'_N) \Phi_{1, \dots, N}(t_1, \dots, t_N). \quad (166)$$

As expected, the N -th coherence function is essentially the full many-body wavefunction of a quantum state with N excitations in mutually orthogonal wavefunctions.

◇

To see the correlation functions of determinantal point processes appear, one can use an ideal broad-band detector as in Section 5. In particular, imagine that we could prepare N electrons in mutually orthogonal wavepackets $\varphi_1, \dots, \varphi_N$, and set up an ideal detector as we did for photons in Section 4, resulting in a point process with correlation functions given by the diagonal of $G_{|\psi_N\rangle}^{(ne)}$,

$$\rho_n(t_1, \dots, t_n) = G_{|\psi_N\rangle}^{(ne)}(t_1, \dots, t_n, t_1, \dots, t_n), \quad (167)$$

$$= |\Phi_{1, \dots, N}(t_1, \dots, t_n)|^2 \quad (168)$$

$$= \det [K(t_i, t_j)]_{i,j=1}^n, \quad (169)$$

where

$$K(t, s) = \sum_{i=1}^N \varphi_{e_i}(t) \varphi_{e_i}^*(s). \quad (170)$$

We recognize a DPP with (projection) kernel K , as introduced in Section 2.4.

Remark 10. To illustrate what reference computations different scientists may have in mind, note that, following a physics viewpoint, we computed in Examples 14 and 15 coherence functions applying Wick's theorem, one after the other, starting from the lower ones. Probabilists and statisticians are more used to derive correlation functions starting from the definition of the point process. Had we started by defining the DPP with kernel K in (170), and asked for the diagonal of, say, the second coherence function in (14), it would have been enough to notice, from the definition (184) of the correlation functions of a DPP, that

$$\rho_2(t_1, t_2) = \det \begin{bmatrix} K(t_1, t_1) & K(t_1, t_2) \\ K(t_2, t_1) & K(t_2, t_2) \end{bmatrix}. \quad (171)$$

By definition (170) of K ,

$$\rho_2(t_1, t_2) = \det \begin{bmatrix} \varphi_1(t_1) & \cdots & \varphi_N(t_1) \\ \varphi_1(t_2) & \cdots & \varphi_N(t_2) \end{bmatrix} \begin{bmatrix} \varphi_1^*(t_1) & \varphi_1^*(t_2) \\ \vdots & \vdots \\ \varphi_N^*(t_1) & \varphi_N^*(t_2) \end{bmatrix}. \quad (172)$$

The diagonal version of (163) then results from the Cauchy-Binet formula, which allows rewriting $\rho_2(t_1, t_2)$ as a sum of 2×2 determinants; see e.g. [Kulesza and Taskar, 2012].

5.3 Electron anti-bunching

For photons, we saw that HBT-type experiments can reveal both a bunching effect (for classical light beams) and an anti-bunching effect (for quantum beams), depending on the source.

A similar experimental setup as the one pictured for photons in Section 4 can be imagined for fermionic excitations. We defer the discussions on how to properly do this experimentally with fermionic atoms or single electronic excitations to Part II, but we can already foresee from either DPP constructions like the one with kernel (170), or directly from the coherence functions of Section 5.2 that fermions will exhibit anti-bunching detection times. Unlike bosons, this is a direct consequence of the anti-symmetric statistic of fermions, and does not qualitatively change when introducing source models.

If the arrival times follow a DPP with an Hermitian kernel like (170), we saw indeed in Section 2.4 that samples tend to spread regularly, and form less clusters than a Poisson point process. Alternately, this statistical anti-bunching effect can be directly seen from the coherence functions. Indeed, as for photons, the HBT signal is controlled by the second-order coherence function at the position \mathbf{r} of the detector,

$$G_\rho^{(2e)}(\mathbf{r}, t_2, t_1) = \langle \psi^\dagger(\mathbf{r}, t_1) \psi^\dagger(\mathbf{r}, t_2) \psi(\mathbf{r}, t_2) \psi(\mathbf{r}, t_1) \rangle_\rho. \quad (173)$$

At coincidence $\tau = t_2 - t_1 = 0$, because the CAR (48) impose $\psi(\mathbf{r}, t_1)^2 = 0$, we obtain

$$G_\rho^{(2e)}(\mathbf{r}, t, t) = 0. \quad (174)$$

If $G_\rho^{(2e)}$ is smooth, then we expect few coincidences at small time differences.

As a final remark, we insist that bunching or anti-bunching of quantum excitations is not a signature of the statistics of the elementary excitations. While it is a consequence of the statistics for fermions, bosons also anti-bunch when we add to the model a single-excitation source. Hence, it is better to think of anti-bunching as a signature of the quantum nature of the fundamental excitation.

5.4 Recovering a classical current is not as easy as for bosons

For photons, coherent states allowed for relating the quantum formalism to classical fields; see Section 4. Describing a classical electric current or, more generally, a classical theory of fermionic fields, is more difficult. In this section, we examine reasons why it is not possible to build a set of states having all the properties that make bosonic coherent states handy. First, even considering a single mode, bosonic coherent states are superpositions of arbitrary numbers of bosons, which allows for having very small fluctuations around the average boson number. Indeed the number of photons in the mode is a Poisson variable with arbitrarily large parameter, and thus arbitrarily small relative variance. Such small fluctuations are not achievable with fermions, because it is impossible to have more than one fermion per mode.

A second difference is that the eigenvalue of a tentative fermionic coherent state of the fermionic annihilation operator cannot be a complex number. Indeed, were $|\alpha\rangle$ an eigenstate of the fermionic annihilation operator a , then the CAR (48) would imply that its eigenvalue α satisfies $\alpha^2 = 0$. Moreover, in the many-mode case, a two-mode coherent state $|\alpha_1, \alpha_2\rangle$ would satisfy the anti-commutation relation $a_1 a_2 |\alpha_1, \alpha_2\rangle = \alpha_1 \alpha_2 |\alpha_1, \alpha_2\rangle = -a_2 a_1 |\alpha_1, \alpha_2\rangle = -\alpha_2 \alpha_1 |\alpha_1, \alpha_2\rangle$.⁴¹ The eigenvalues of a fermionic coherent state thus cannot be complex numbers.⁴²

A more subtle difference is that fermionic fields cannot have an observable average amplitude,⁴³ unlike bosonic fields prepared in a coherent state.⁴⁴ This impossibility is fundamental and is related to *superselection rules* [Wick et al., 1952]. For concreteness, consider a general fermionic state $|\varphi\rangle = \alpha |0\rangle + \beta |1\rangle$ with $|1\rangle = \psi^\dagger |0\rangle$, where ψ^\dagger is any fermionic creation operator. Now, physicists usually impose symmetries on their models, which implies that some states should yield the same measurements. In particular, one would like to model the fact that rotating the physical system around any axis by an angle of 2π should not modify the law of measurements. For fermions, the so-called *spin-statistics theorem* of relativistic quantum field theory tells us that this kind of rotations should be modelled by the action of the half-integer representations of the group $SU(2)$. For instance, in the representation of order $1/2$, a rotation of angle 2π corresponds to multiplying the state by $e^{i\pi N}$, where N is the number operator. Applied to $|\varphi\rangle = \alpha |0\rangle + \beta |1\rangle$, this yields $|\varphi'\rangle := e^{i\pi N} |\varphi\rangle = \alpha |0\rangle - \beta |1\rangle$, where the minus sign comes from the fact that 1 is an odd number. However, since such a rotation should not correspond to any change of the system, we impose that $|\varphi'\rangle = |\varphi\rangle$ up to a phase, and consequently either α or β is zero. Thus $|\varphi\rangle$ is in fact not a superposition, implying that $\langle\psi\rangle_{|\varphi\rangle} = 0$. More generally, the *parity superselection rule* imposes that a given state can only consist of superpositions of states with the same parity of the number of fermions. Therefore, a single fermionic annihilation operator can only have a vanishing average.⁴⁵

For all these reasons, defining a classical regime with fermionic fields is an open question. This actually raises deep conceptual questions when fermions are involved. For instance, it is not clear what is meant by a *classical electronic current* if we start from a pure quantum description of the electronic current.

⁴¹This result is independent on whether one assumes that α_i commutes or anti-commutes with a_j .

⁴²It is however possible to define fermionic coherent states in terms of anti-commuting variables (elements of an exterior algebra) by the action of the analogue of a displacement operator [Berezin, 1966, Chapter 1]. While they are very useful objects to define and use (especially to write path integrals), they do not possess the nice set of properties of bosonic coherent states that allow for deriving coherence functions.

⁴³Just like a for photons, ψ is not Hermitian, and thus not an observable, *stricto sensu*. Yet, for a , the average value corresponds to the intensity of the field, and is thus experimentally accessible as a statistical average. It thus makes sense to wonder whether the average value of ψ can be non-zero.

⁴⁴This property can actually be used as a definition of bosonic coherent states [Cohen-Tannoudji et al., 2019a].

⁴⁵This generalises beyond physically realistic fermions of half-integer spins, and applies to non-relativistic quantum mechanics, see for instance [Johansson, 2016, Szalay et al., 2021].

6 Wrapping up and open questions

To conclude this Part I, we first show how to use the formalism of Sections 2 to 5 to write a generic permanental or determinantal point process as the detection process of a system of free bosons or fermions, respectively. Second, we propose a list of questions, from simple ones to open problems, that are raised by the connections between point processes and physical measurements in quantum field theory. Third, we announce the sections that are to appear in Part II of this document.

6.1 From a point process to free particles

In Sections 4 and 5, we obtained some point processes, including permanental and determinantal point processes, by modeling the detection of physical particles. A natural question is whether all permanental and determinantal point processes arise in this form. We now show that, at least formally and under weak assumptions on the kernel, the answer is yes. More precisely, let μ be a Borel measure on a complete metric space \mathbb{X} , and (φ_i) an orthonormal family in $L^2(\mu)$. Consider the kernel

$$K(x, y) := \sum_i \lambda_i \overline{\varphi_i(x)} \varphi_i(y), \quad (175)$$

where, for all i , $\lambda_i \in [0, 1]$ if $\eta = -1$, and $\lambda_i \in \mathbb{R}_+$ if $\eta = +1$. We now build a Fock space, a quantum state, and a measurement model that together lead to a permanental or determinantal point process with K as its kernel. Let $\eta = \pm 1$ depending on whether one wants a permanental or a determinantal point process.

Building a Fock space. Build the Fock space like in Section 3.4, using the basis (φ_i) as single-particle basis and the symmetrization property dictated by η . Consider the operator

$$H = \sum_i \nu_i a_i^\dagger a_i,$$

where (ν_i) is left as a free parameter for now, and will later be chosen in relation to the spectrum of K . We think of H as a Hamiltonian.

A Gaussian density matrix. Let $\zeta > 0$, and define the so-called grand canonical ensemble as in (60) by the Gaussian density matrix

$$\rho = \frac{e^{-\beta(H - \zeta \sum_i a_i^\dagger a_i)}}{Z_\eta} = \frac{1}{Z_\eta} e^{-\beta \sum_i (\nu_i - \zeta) a_i^\dagger a_i}.$$

We explicitly write the dependence of Z to η , as bosons and fermions lead to different normalization constants. More precisely, since the Fock states $|\mathbf{n}\rangle = |n_1, n_2, \dots\rangle$ are eigenvectors of all the number operators $a_i^\dagger a_i$ in the exponential, the normalization constant is

$$Z_\eta = \text{Tr} e^{-\beta \sum_i (\nu_i - \zeta) a_i^\dagger a_i} = \sum_{\mathbf{n}} \langle \mathbf{n} | e^{-\beta \sum_i (\nu_i - \zeta) a_i^\dagger a_i} | \mathbf{n} \rangle = \sum_{\mathbf{n}} e^{-\beta \sum_i (\nu_i - \zeta) n_i}. \quad (176)$$

For bosons, we sum over all sequences of integers with only p non-zero components, and this for all $p \geq 1$. In particular,

$$Z_1 = \prod_{p \in \mathbb{N}} \sum_{n \in \mathbb{N}} e^{-\beta(\nu_p - \zeta)n} = \prod_{p \in \mathbb{N}} \frac{1}{1 - e^{-\beta(\nu_p - \zeta)}}, \quad (177)$$

where we implicitly assumed $\nu_p - \zeta > 0$ for all p , for the geometric sums in (177) to converge.⁴⁶ For Z_1 to be finite, we further need a condition on the (ν_p) , such as

$$\sum_p \log \left(\frac{1}{e^{\beta(\nu_p - \zeta)} - 1} \right) < \infty. \quad (178)$$

⁴⁶Note that for a given p , the limit $\nu_p \rightarrow \zeta$ corresponds to the onset of Bose-Einstein condensation, where a macroscopic number of bosons start to occupy the state p .

For fermions, we only sum in (176) over all sequences in $\{0, 1\}^{\mathbb{N}}$ that have p non-zeros components, for all sizes $p \geq 1$, so that

$$Z_{-1} = \prod_{p \in \mathbb{N}} \sum_{n \in \{0,1\}} e^{-\beta(\nu_p - \zeta)n} = \prod_{p \in \mathbb{N}} \left(1 + e^{-\beta(\nu_p - \zeta)}\right). \quad (179)$$

Guaranteeing convergence is easier than for bosons, e.g. assuming that $\sum_p(\nu_p - \zeta)$ converges.

Now that we have simple expressions for Z_η , $\eta = \pm 1$, we can use the partition function trick⁴⁷ to compute the expected number of particles in mode i ,

$$\langle a_i^\dagger a_i \rangle_\rho = \frac{1}{Z_\eta} \text{Tr } a_i^\dagger a_i \rho \quad (180a)$$

$$= \frac{\sum_n n_i e^{-\beta \sum_j (\nu_j - \zeta) n_j}}{\sum_n e^{-\beta \sum_j (\nu_j - \zeta) n_j}} \quad (180b)$$

$$= -\frac{1}{\beta} \frac{\partial}{\partial \nu_i} \log Z_\eta(\nu). \quad (180c)$$

Using (177) and (179) in turn, this yields

$$\langle a_i^\dagger a_i \rangle_\rho = \frac{1}{e^{\beta(\nu_i - \zeta)} - \eta}, \quad (181)$$

which is the so-called *Fermi-Dirac* distribution⁴⁸ for $\eta = -1$ and the *Bose-Einstein* distribution for $\eta = 1$. Note that along the lines of Lemma 2, one can show that $\langle a_i^\dagger a_j \rangle_\rho = 0$ if $i \neq j$.

A measurement basis. Now, consider the alternative basis $|x\rangle$, $x \in \mathbb{X}$, with annihilation operator

$$\psi(x) = \sum_i \langle x, \varphi_i \rangle a_i;$$

see (52). We think of $|x\rangle$ as the position basis, so that we write $\langle x, \varphi_i \rangle = \varphi_i(x)$, and we actually mean a generalized basis if necessary, e.g., if $X = \mathbb{R}^d$; see Section 3.1. In the basis $|x\rangle$, the state ρ rewrites

$$\rho = \frac{1}{Z} \exp \left(-\beta \sum_i (\nu_i - \zeta) \int \varphi_i(x) \overline{\varphi_i(y)} \psi^\dagger(x) \psi(y) d\mu(x) d\mu(y) \right).$$

Wick's theorem applies. By Theorem 3, see also Example 5, for all $n \geq 1$, we obtain the coherence functions

$$\langle \psi^\dagger(x_1) \psi(x_1) \dots \psi^\dagger(x_n) \psi(x_n) \rangle = f_\eta(\mathbf{K}),$$

where

$$\mathbf{K} = \left(\langle \psi(x_i)^\dagger \psi(x_j) \rangle_\rho \right)_{1 \leq i, j \leq n}$$

and f_η is the determinant if $\eta = -1$, and the permanent if $\eta = 1$. If the framework is consistent, an ideal detection experiment that measures all particles in ρ in the basis $|x\rangle$ should thus correspond to a permenantal/determinantal point process on \mathbb{X} , with kernel

$$\langle \psi^\dagger(x) \psi(y) \rangle_\rho = \sum_{i,j} \overline{\varphi_i(x)} \varphi_j(y) \langle a_i^\dagger a_j \rangle_\rho. \quad (182)$$

⁴⁷Statisticians might say the *score function* trick.

⁴⁸The word *distribution* is used here in a loose sense.

This is precisely K in (175), provided that we choose ν_i so that

$$\lambda_i = (e^{\beta(\nu_i - \zeta)} - \eta)^{-1}. \quad (183)$$

This is achieved by setting

$$\beta(\nu_i - \zeta) = \log\left(\frac{1 + \eta\lambda_i}{\lambda_i}\right).$$

Note that when $\eta = 1$, one can check that the corresponding permanental point process exists by an adaptation of Macchi’s Cox process construction of Section 2.3. When $\eta = -1$, the determinantal point process exists by the Macchi-Soshikov theorem; see Section 2.4.

Obtaining any kernel. We conclude with several comments. First, inverting (181) allows us to obtain any spectrum (λ_i) with $\lambda_i \in (0, 1)$ for fermions, and $\lambda_i > 0$ for bosons. With the right assumptions, projection kernels for DPPs can also be obtained, but as a limit as $\beta \rightarrow +\infty$, i.e. when the temperature $1/\beta$ goes to zero. The corresponding projection operator is onto the span of the φ_i s for which $\nu_i < \zeta$. In other words, the chemical potential controls the rank of the limiting projection operator. Note that for DPPs, uniform convergence of the kernel on compact subsets of \mathbb{X} implies convergence of the point process in a natural sense; see e.g. [Anderson et al., 2010, Section 4.2.8]. Another way to formally obtain projection DPPs is as a ground state; see Section 3.3.2.

Second, the construction in this section, with a state ρ and measuring correlators, is as close as one can hope for a Cox process-like decomposition of determinantal point processes. Loosely speaking, there is a Gaussian object above determinantal point processes with self-adjoint kernels, but it is a Gaussian density matrix, i.e., an operator, not a functional process.⁴⁹ Third, while we restricted here to Hermitian kernels, thus corresponding to an observable quantity, the Hermitian assumption does not play a major role in the mathematical framework. Furthermore, non-Hermitian free fermions may not be so removed from physics, see e.g. [Ashida et al., 2020] for a recent survey and [Guo et al., 2021] for a formal example.

On physical realizability. Finally, we have seen in Section 5 examples of experimentally realizable DPPs, in the sense that the corresponding idealized experimental setups can, in principle, be built in a physics laboratory. We will see in Part II concrete examples of DPP samples obtained from lab experiments. Yet, the generic construction given in Section 6.1 remains rather theoretical, as it is not clear how to build an experimental setup leading to *any* given DPP, that is, any choice of underlying space (\mathbb{X}, μ) and any valid kernel K . One might be able to treat $\mathbb{X} = \mathbb{R}^d$ for d up to 3 by associating mathematical coordinates to space-coordinates. Going slightly beyond 3 might be doable using either time or additional physical degrees of freedom (“pseudo-dimensions”) as the next dimensions, but it is unclear to physicists where fermions in arbitrary dimension would be needed as a model. Outside the problem of dimensionality, fixing the spectrum of the Hamiltonian to an arbitrary sequence can also be experimentally difficult.

6.2 Going further: from simple to open questions

Many natural questions come to mind when one looks back at the correspondence between permanental point processes and free bosons, or between DPPs and free fermions.

6.2.1 A quantum state is more than a point process

A point process is the result of observing a state ρ with a detector, but the state itself contains a lot more information. For instance, the correlation functions that we identified in

⁴⁹Alternatively, the formulation in terms of operators can be replaced by a “Gaussian functional integral” over complex commuting (anti-commuting) fields for permanents (determinants). The connection with Gaussian measures is then more direct, at the price of subtle mathematical difficulties. Furthermore, in the case of fermions, this implies generalising the notion of integration to variables that anti-commute, see for instance [Berezin, 1966, Chapter 1].

Sections 4 and 5 are only the diagonals of the corresponding coherence functions. Relatedly, depending on the measurement basis, a single state ρ can yield many point processes.

As a concrete example, consider for simplicity fermions living in $\mathbb{X} = \{1, \dots, N\}$, μ the counting measure on \mathbb{X} , and the Hamiltonian

$$H = \sum_{i=1}^N \nu_i a^\dagger(\varphi_i) a(\varphi_i)$$

acting on the corresponding Fock space. By (182), measuring in an orthonormal basis (v_i) of \mathbb{C}^N , with $\mathbf{V} = (\langle \varphi_i | v_j \rangle)_{i,j}$, we obtain the DPP with kernel matrix

$$\mathbf{K} = \left(\sum_k \lambda_k \langle \varphi_k | v_i \rangle \langle v_j | \varphi_k \rangle \right)_{1 \leq i, j \leq N} = \mathbf{V} \text{diag}(\lambda_k) \mathbf{V}^\dagger,$$

where (λ_i) is given by (183). All Hermitian kernels with spectrum (λ_k) can thus be obtained from the same Hamiltonian, just changing the measurement basis.

At one extreme, measuring in the original basis used to build the Hamiltonian, i.e. taking $\mathbf{V} = \mathbf{I}$, leads to a diagonal kernel matrix. This means that the corresponding point process is a set of independent Bernoulli samples for fermions, and independent geometric samples for bosons [Hough et al., 2006]. Correlation appears when the observation basis (v_k) and the creation basis (φ_k) do not match. For instance, let $v_k = \varphi_k$ for $k = 1, \dots, N-2$, but apply a special unitary transformation to the last two vectors. That is, for $\alpha, \beta \in \mathbb{C}$ such that $|\alpha|^2 + |\beta|^2 = 1$, complete the observation basis with

$$v_{N-1} = \alpha \varphi_{N-1} - \bar{\beta} \varphi_N \quad \text{and} \quad v_N = \beta \varphi_{N-1} + \bar{\alpha} \varphi_N.$$

The resulting permanental or determinantal point process, call it $\gamma^{\alpha, \beta}$, has a block diagonal kernel

$$\mathbf{K}^{\alpha, \beta} = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_{N-2} & & & \\ & & & \lambda_{N-1} |\alpha|^2 + \lambda_N |\beta|^2 & (\lambda_{N-1} - \lambda_N) \bar{\alpha} \beta & \\ & & & (\lambda_{N-1} - \lambda_N) \alpha \bar{\beta} & \lambda_{N-1} |\beta|^2 + \lambda_N |\alpha|^2 & \end{pmatrix}.$$

Note that since the diagonal of $\mathbf{K}^{\alpha, \beta}$ varies with (α, β) , the first correlation function of the point process is also altered. For instance, the marginal probability that item $N-1$ belongs to the point process is $\lambda_{N-1} |\alpha|^2 + \lambda_N |\beta|^2$. If $\lambda_{N-1} \geq \lambda_N$, then as soon as $\beta \neq 0$, the marginal probability of item $N-1$ occurring decreases. Note that for $\eta = -1$, though, the average number of points in $\gamma^{\alpha, \beta}$ remains constant, as the trace of $\mathbf{K}^{\alpha, \beta}$ does not depend on (α, β) . Meanwhile, the second correlation function $\rho_{2, \theta}(N-1, N)$ is the permanent/determinant of the trailing block of $\mathbf{K}^{\alpha, \beta}$, namely

$$\rho_2^{\alpha, \beta}(N-1, N) = |\alpha|^2 |\beta|^2 (\lambda_{N-1}^2 + \lambda_N^2 + \eta (\lambda_{N-1} - \lambda_N)^2) + (|\alpha|^4 + |\beta|^4) \lambda_{N-1} \lambda_N.$$

For the sake of illustration, we focus now on $\eta = -1$, in which case

$$\rho_2^{\alpha, \beta}(N-1, N) = \lambda_{N-1} \lambda_N (|\alpha|^2 + |\beta|^2)^2 = \lambda_{N-1} \lambda_N.$$

The probability of co-occurrence of items $N-1$ and N thus does not depend on α and β , while the product of their marginal probabilities of occurrence does change: we have thus introduced correlation between the events $\{N-1 \in \gamma^{\alpha, \beta}\}$ and $\{N \in \gamma^{\alpha, \beta}\}$. Because $\gamma^{\alpha, \beta}$ is a DPP with Hermitian kernel, this correlation is nonpositive.

To go further, an interesting question would be to relate the different ways to quantify repulsiveness in a DPP [Biscio and Lavancier, 2016, Møller and O'Reilly, 2021] to properties of the bases used in the construction of free fermions.

6.2.2 A DPP from bosons in a non-Gaussian state

Another example that illustrates the subtleties of associating correlation functions to states is that (non-Gaussian) density matrices corresponding to interacting (i.e., non-free) bosons can give rise to a DPP for some observables.

One example is the Tonks-Girardeau gas [Girardeau, 1960], which consists of impenetrable (statisticians would say *hardcore*) bosons in $\mathbb{X} = \mathbb{R}$. The very strong interactions prevent two bosons to be at the same position, which is reminiscent of the Pauli exclusion principle for fermions. In particular, one considers the usual bosonic field operators $a(x), a^\dagger(y)$, satisfying the canonical commutation relations as long as $x \neq y$, but the hardcore constraint is enforced by requiring that $a(x)^2 = a^\dagger(x)^2 = 0$ and $[a(x), a^\dagger(x)]_+ = 1$. These additional constraints are the results of adding a large penalty term to the Hamiltonian at coincidence.

Now, there actually is a transformation, called the *Jordan-Wigner transformation*, that maps these hardcore bosons onto free fermions [Lieb et al., 1961]. Concretely, one can write $a(x) = S(x)\psi(x)$ and $a^\dagger(x) = \psi^\dagger(x)S^\dagger(x)$, where $\psi(x), \psi^\dagger(x)$ obey the standard canonical anti-commutation relations, and $S(x)$ is a unitary operator called a *string operator*, which allows for preserving the commutation relations of the bosonic operators.⁵⁰

By construction, all observables that are built from number operators⁵¹ are the same as for free fermions, since $a^\dagger(x)a(x) = \psi^\dagger(x)\psi(x)$, and the correlation functions of an ideal detection experiment are those of a DPP with kernel $\langle \psi^\dagger(x)\psi(y) \rangle$. However, the coherence function $\langle a^\dagger(x)a(y) \rangle = \langle \psi^\dagger(x)S^\dagger(x)S(y)\psi(y) \rangle$ is very different from that of the free fermions. For instance, the largest eigenvalue of the operator with kernel $\langle a^\dagger(x)a(y) \rangle$ evaluated in a pure state typically scales like the square root of the number of particles; see e.g. Forrester et al. [2003] for investigations on the ground state properties of a Tonks gas in a harmonic trap.

6.2.3 Interacting field theories and point processes

The relationship between, on one side, permanental and determinantal point processes, and, on the other side, bosonic and fermionic quantum field theory, is based on free models, i.e., with a quadratic Hamiltonian in the fields. However, in physical models, Hamiltonians are rarely quadratic, but take into account interactions between the particles. For a given Hamiltonian, success is achieved when it is possible to approximate the coherence functions of the field, which correspond to concrete experimental measurements. The main issue is that Wick’s theorem does not apply.

There are interacting Hamiltonians that still yield closed-form correlation functions for the underlying ideal detection experiments. Such systems are usually called *integrable* by physicists. For instance, the Calogero-Sutherland model is a one-dimensional system of interacting fermions, for which ideal detection leads to well-known point processes, called β -ensembles in random matrix theory; see e.g. [Forrester, 2010, Chapter 11] as well as [Stéphan, 2019, Smith et al., 2021].

Yet, most interacting systems are not integrable. In Sections 4 and 5, we have used simple perturbation-theoretic arguments like Assumption 5 to work with the interaction Hamiltonian. There is now a significant and sophisticated toolbox to approximate coherence functions in the presence of interaction. Key tools include mean-field methods [Goldenfeld, 2018, Chapter 3], Feynman diagrams [Folland, 2008, Chapter 6] and renormalisation [Folland, 2008, Chapter 7]. At a high level, it would be interesting to investigate what these methods say about point processes. In words, one of the lessons of quantum field theory is that it is possible, even in the presence of interaction, to get “close” to a free situation in some regimes. Does this imply approximation results for the sophisticated point processes behind models with interaction? Conversely, can we use the mathematical technology of DPPs to perform non-trivial calculations for many-body problems?

Moreover, a central concept in quantum field theory is that of universality, i.e., the fact that many different interacting models behave, in a certain regime usually qualified

⁵⁰In particular, it satisfies $[S(x), S(y)] = 0$ and $S(x)\psi(y) + \text{sign}(x - y)\psi(y)S(x) = 0$.

⁵¹Usually rather called *density operators* when we use a generalized basis like here.

as the *low-energy regime*, essentially in the same way. Put differently, the detailed form of the interactions are irrelevant to understand this low-energy regime, only the dimension of the space and the symmetries of the model usually matter. The archetypal phenomenon illustrating this is phase transitions [Goldenfeld, 2018]. Does this universality connect with similar results for point processes, such as those obtained for the eigenvalues of random matrices [Anderson et al., 2010]?

6.2.4 Constructive arguments for point processes

The fermionic system “above” a DPP can help address fundamental questions on DPPs, about their invariance or their construction. For instance, a single DPP corresponds to many kernels. In particular, for any f , the kernel

$$x, y \mapsto \frac{f(x)}{f(y)} K(x, y)$$

yields the same correlation functions (184) as K . Without strong assumptions on the kernel, it has proven difficult to find all the transformations of a kernel that leaves a DPP invariant. It would be interesting if the fermionic framework helped us to understand these invariances; see e.g. recent partial results by Olshanski [2020].

As another example, generalizing permanental and determinantal point processes, one can define α -DPPs as having correlation functions

$$\rho_k(x_1, \dots, x_k) = \det_\alpha \left[K(x_i, x_j) \right]_{i,j=1}^k, \quad k \geq 1, \quad (184)$$

where, for an $n \times n$ matrix $\mathbf{A} = ((a_{ij}))$,

$$\det_\alpha(\mathbf{A}) := \sum_{\sigma \in \mathfrak{S}_n} \alpha^{n-\nu(\sigma)} \prod_{i=1}^n a_{i\sigma(i)}.$$

Shirai and Takahashi [2003] have studied the existence of α -DPPs for $\alpha \in [-1, 1]$, where $\alpha = -1$ corresponds to a DPP, and $+1$ to a permanental point process. Some α -DPPs appear naturally when marginalizing a projection DPP with a separable kernel over coordinates [Mazoyer et al., 2020]. It is a natural question whether α -DPPs correspond to any physical system of particles. As a partial affirmative answer, Cunden et al. [2019] give a limit procedure to construct certain α -DPPs out of fermionic processes.

In the same vein, can we build point processes from particles with more exotic commutation rules for their ladder operators? A natural physical example is anyons [Ezawa, 2008, Chapter 8], for which a phase factor appears for each transposition.

The links of fermions to Pfaffian point processes, another generalization of DPPs are also a promising research direction [Koshida, 2021], as well as the point processes behind quasi-free states [Bach et al., 1994, Lytvynov, 2002, Lytvynov and Mei, 2007, Olshanski, 2020], namely (possibly non-Gaussian) states to which Wick’s theorem still applies.

6.3 A teaser for Part II

This manuscript is intended to become Part I of a monograph. In Part II, we shall present selected topics at the intersection of point processes and quantum optics, using the vocabulary of Part I. We will describe landmark experimental measurements of HBT signals, with both photons and (bosonic and fermionic) atoms. We will discuss the application of determinantal point processes to the study of non-interacting trapped fermions in statistical physics. We will show how the formalism of quantum field theory can help to prove fundamental results on point processes appearing in combinatorics. Finally, we will discuss electronic quantum optics and its interactions with signal processing.

Acknowledgments

We thank all participants to the Lille and Lyon workshops, and all participants to the Lille workgroup on point processes and applications; the discussions held in these circles helped a lot in the construction of this document. We hope that the document, in return, participates to developing this cross-disciplinary research field. In particular, RB thanks all physicists with whom he has interacted over the last years for their enthusiasm and their patience, explaining basic physical concepts over and over until we converged to a description that spoke to all of us.

We acknowledge support from ERC grant BLACKJACK (ERC-2019-STG-851866), ANR AI chair BACCARAT (ANR-20-CHIA-0002), ANR grant c1SHOT RELOADED (ANR-14-CE32-0017), ANR grant QUSIG4QUSENSE (ANR-21-CE47-0012), the Joint Research Project SEQUOIA (17FUN04) within the European Metrology Programme for Innovation and Research (EMPIR), ANR grants QRITIC I-SITE ULNE and ANR-16-IDEX-0004 ULNE, ANR grant COSQUA (ANR-20-CE47-0001-01), ANR grant MENTA (ANR-22-QUA2-0008-01), ANR Grant RAMATRAF (ANR-17-CE30-0027-01), ANR grant DIMERS (ANR-18-CE40-0033), ANR grant COMBINÉ (ANR-19-CE48-0011), Labex PALM (ANR-10-LABX-0039-PALM), and Labex MILYON.

References

- S. T. Ali, J.-P. Antoine, and J.-P. Gazeau. *Coherent states, wavelets and their generalizations*. Springer, 2000.
- G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118. Cambridge university press, 2010.
- F. Arecchi, E. Courtens, R. Gilmore, and H. Thomas. Atomic coherent states in quantum optics. *Physical Review A*, 6(6):2211, 1972.
- Y. Ashida, Z. Gong, and M. Ueda. Non-hermitian physics. *Advances in Physics*, 69(3):249–435, 2020.
- V. Bach, E. H. Lieb, and J. P. Solovej. Generalized Hartree-Fock theory and the Hubbard model. *Journal of statistical physics*, 76(1):3–89, 1994.
- A. Baddeley, E. Rubak, and R. Turner. *Spatial point patterns: methodology and applications with R*. CRC press, 2015.
- R. Bardenet and A. Hardy. Time-frequency transforms of white noises and Gaussian analytic functions. *Applied and Computational Harmonic Analysis*, 2019.
- R. Bardenet, J. Flamant, and P. Chainais. On the zeros of the spectrogram of white noise. *Applied and Computational Harmonic Analysis*, 2018.
- C. Beltrán and A. Hardy. Energy of the Coulomb gas on the sphere at low temperature. *Archive for Rational Mechanics and Analysis*, 231(3):2007–2017, 2019.
- C. Bénard and O. Macchi. Detection and “emission” processes of quantum particles in a “chaotic” state. *Journal of mathematical physics*, 1973.
- F. Berezin. *The Method of Second Quantization*. Pure and applied physics : a series of monographs and textbooks. 24. Academic Press, 1966.
- C. A. N. Biscio and F. Lavancier. Quantifying repulsiveness of determinantal point processes. *Bernoulli*, 22(4):2001–2028, 2016.
- A. Borodin, A. Okounkov, and G. Olshanski. Asymptotics of Plancherel measures for symmetric groups. *Journal of the American Mathematical Society*, 13(3):481–515, 2000.

- L. Bouten, R. van Handel, and M. R. James. An introduction to quantum filtering. *SIAM Journal on Control and Optimization*, 46(6):2199–2241, 2007.
- C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum mechanics, Volume I*. Wiley-VCH, 2019a.
- C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum mechanics, Volume III*. Wiley-VCH, 2019b.
- C. Cohen-Tannoudji. Emission et détection de rayonnement : approches semi-classiques et approches quantiques, 1979.
- C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg. *Atom-photon interactions: basic processes and applications*. 1998.
- F. D. Cunden, S. N. Majumdar, and N. O’Connell. Free fermions and α -determinantal processes. *Journal of Physics A: Mathematical and Theoretical*, 52(16):165202, 2019.
- D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes*. Springer, 2nd edition, 2003.
- D. S. Dean, P. Le Doussal, S. N. Majumdar, and G. Schehr. Noninteracting fermions at finite temperature in a d-dimensional trap: Universal correlations. *Physical Review A*, 94(6):063622, 2016.
- D. S. Dean, P. Le Doussal, S. N. Majumdar, and G. Schehr. Noninteracting fermions in a trap and random matrix theory. *Journal of Physics A: Mathematical and Theoretical*, 52(14):144006, 2019.
- P. Degiovanni, N. Portier, C. Cabart, A. Feller, and B. Roussel. *Physique quantique, information et calcul: des concepts aux applications*. EDP sciences, 2021.
- Z. F. Ezawa. *Quantum Hall effects: Field theoretical approach and related topics*. World Scientific Publishing Company, 2008.
- C. Fabre and N. Treps. Modes and states in quantum optics. *Reviews of Modern Physics*, 92(3):035005, 2020.
- M. P. A. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher. Boson localization and the superfluid-insulator transition. *Phys. Rev. B*, 40:546–570, Jul 1989.
- E. B. Flagg, A. Muller, J. W. Robertson, S. Founta, D. G. Deppe, M. Xiao, W. Ma, G. J. Salamo, and C. K. Shih. Resonantly driven coherent oscillations in a solid-state quantum emitter. *Nature Physics*, 5(3):203–207, 2009.
- P. Flandrin. *Time-frequency/time-scale analysis*, volume 10. Academic press, 1998.
- G. B. Folland. *Harmonic analysis in phase space*. Number 122. Princeton university press, 1989.
- G. B. Folland. *Quantum Field Theory: A tourist guide for mathematicians*. Number 149. American Mathematical Society, 2008.
- P. J. Forrester. *Log-gases and random matrices*. Princeton University Press, 2010.
- P. J. Forrester, N. E. Frankel, T. M. Garoni, and N. S. Witte. Finite one-dimensional impenetrable bose systems: Occupation numbers. *Phys. Rev. A*, 67:043607, Apr 2003.
- M. Girardeau. Relationship between systems of impenetrable bosons and fermions in one dimension. *Journal of Mathematical Physics*, 1(6):516–523, 1960.
- R. J. Glauber. Coherent and incoherent states of the radiation field. *Physical Review*, 131(6):2766, 1963.

- N. Goldenfeld. *Lectures on phase transitions and the renormalization group*. CRC Press, 2018.
- K. Gröchenig. *Foundations of time-frequency analysis*. Birkhäuser, 2001.
- G. Grynberg, A. Aspect, and C. Fabre. *Introduction to quantum optics: from the semi-classical approach to quantized light*. Cambridge university press, 2010.
- Y.-B. Guo, Y.-C. Yu, R.-Z. Huang, L.-P. Yang, R.-Z. Chi, H.-J. Liao, and T. Xiang. Entanglement entropy of non-hermitian free fermions. *Journal of Physics: Condensed Matter*, 33(47):475502, sep 2021.
- R. Hanbury Brown and R. Q. Twiss. Interferometry of the intensity fluctuations in light. II. an experimental test of the theory for partially coherent light. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 243(1234):291–319, 1958.
- J. H. Hannay. The chaotic analytic function. *Journal of Physics A: Mathematical and General*, 31(49):L755, 1998.
- S. Haroche and J.-M. Raimond. *Exploring the quantum: atoms, cavities, and photons*. Oxford university press, 2006.
- P. Harremoës. Binomial and Poisson distributions as maximum entropy distributions. *IEEE Transactions on Information Theory*, 47(5):2039–2041, 2001.
- J. B. Hough, M. Krishnapur, Y. Peres, and B. Virág. Determinantal processes and independence. *Probability Surveys*, 3:206–229, 2006.
- J. B. Hough, M. Krishnapur, Y. Peres, and B. Virág. *Zeros of Gaussian analytic functions and determinantal point processes*, volume 51. American Mathematical Society, 2009.
- K. Johansson. *Random matrices and determinantal processes, Mathematical Statistical Physics*. Elsevier B.V. Amsterdam, 2006.
- M. Johansson. Comment on "reasonable fermionic quantum information theories require relativity". *arXiv preprint arXiv:1610.00539*, 2016.
- S. Koshida. Pfaffian point processes from free fermion algebras: Perfectness and conditional measures. *Symmetry, Integrability and Geometry: Methods and Applications*, 17:008, 2021.
- A. Kulesza and B. Taskar. Determinantal point processes for machine learning. *Foundations and Trends in Machine Learning*, 2012.
- F. Lavancier, J. Møller, and E. Rubak. Determinantal point process models and statistical inference: Extended version. *Preprint arXiv: 1205.4818*, 2014.
- E. Lieb, T. Schultz, and D. Mattis. Two soluble models of an antiferromagnetic chain. *Annals of Physics*, 16(3):407–466, 1961.
- E. Lytvynov. Fermion and boson random point processes as particle distributions of infinite free Fermi and Bose gases of finite density. *Reviews in Mathematical Physics*, 14(10): 1073–1098, 2002.
- E. Lytvynov and L. Mei. On the correlation measure of a family of commuting hermitian operators with applications to particle densities of the quasi-free representations of the car and ccr. *Journal of Functional Analysis*, 245(1):62–88, 2007.
- O. Macchi. The coincidence approach to stochastic point processes. *Advances in Applied Probability*, 7:83–122, 1975.

- O. Macchi. The fermion process – a model of stochastic point process with repulsive points. In *Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes and of the 1974 European Meeting of Statisticians*. Czechoslovak Academy of Sciences, 1977.
- O. Macchi. *Point processes and coincidences – Contributions to the theory, with applications to statistical optics and optical communication, augmented with a scholion by Suren Poghosyan and Hans Zessin*. Walter Warmuth Verlag, 2017.
- L. Mandel and E. Wolf. Coherence properties of optical fields. *Reviews of Modern Physics*, 37(2), 1965.
- L. Mandel and E. Wolf. *Optical coherence and quantum optics*. Cambridge university press, 1995.
- A. Mazoyer, J.-F. Coeurjolly, and P.-O. Amblard. Projections of determinantal point processes. *Spatial Statistics*, 38:100437, 2020.
- J. Møller and E. O’Reilly. Couplings for determinantal point processes and their reduced Palm distributions with a view to quantifying repulsiveness. *Journal of Applied Probability*, 58(2):469–483, 2021.
- S. Nonnenmacher. Anatomy of quantum chaotic eigenstates. In *Chaos*, pages 193–238. Springer, 2013.
- G. Olshanski. Determinantal point processes and fermion quasifree states. *Communications in Mathematical Physics*, 378(1):507–555, 2020.
- A. M. Perelomov. Coherent states for arbitrary lie group. *Communications in Mathematical Physics*, 26(3):222–236, 1972.
- B. Picinbono. On instantaneous amplitude and phase of signals. *IEEE Transactions on signal processing*, 45(3):552–560, 1997.
- E. L. Pugh. The generalized analytic signal. *Journal of Mathematical Analysis and Applications*, 89(2):674–699, 1982.
- M. Reed and B. Simon. *Methods of modern mathematical physics: Functional analysis*. Elsevier, 1980.
- Z. Rudnick and P. Sarnak. Zeros of principal L-functions and random matrix theory. *Duke Mathematical Journal*, 81(2):269–322, 1996.
- T. Shirai and Y. Takahashi. Random point fields associated with certain Fredholm determinants I: fermion, Poisson and boson point processes. *Journal of Functional Analysis*, 205(2):414–463, 2003.
- N. Smith, P. Le Doussal, S. Majumdar, and G. Schehr. Full counting statistics for interacting trapped fermions. *SciPost Physics*, 11(6):110, 2021.
- J.-M. Stéphan. Free fermions at the edge of interacting systems. *SciPost Physics*, 6(5):057, 2019.
- S. Szalay, Z. Zimborás, M. Máté, G. Barcza, C. Schilling, and Ö. Legeza. Fermionic systems for quantum information people. *Journal of Physics A: Mathematical and Theoretical*, 54(39):393001, 2021.
- G. C. Wick, A. S. Wightman, and E. P. Wigner. The intrinsic parity of elementary particles. *Phys. Rev.*, 88:101–105, Oct 1952.
- W.-M. Zhang, R. Gilmore, et al. Coherent states: theory and some applications. *Reviews of Modern Physics*, 62(4):867, 1990.