UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master Degree in Physics

Final Dissertation

Light Scattering from Dust Contamination in Gravitational-Wave Interferometric Detectors

Thesis supervisor

Prof. Giacomo Ciani

Thesis co-supervisor

Dr. Livia Conti

Candidate

Andrea Moscatello

Academic Year 2021/2022

Contents

Introduction

The GW community is putting a large effort towards increasing the sensitivity of the current interferometric detectors at the lowest frequencies. The observation of low-frequency signals, such as the earliest stages of the inspiral phase of merging binaries, would allow for a better reconstruction of the source parameters. In addition, signals from merging of black holes with higher masses can be detected. Stray light manifests as an excess noise in the low-frequency range of the interferometers. While stray light generated by optical surfaces roughness is usually studied, improvements in the interferometers sensitivity and in the quality of polishing can cause the dust particles contribution to be significant too, even in the clean environments in which the detectors are built. It is therefore crucial to characterize and possibly mitigate this noise contribution, both for the present LIGO-Virgo interferometers and for the next generation detectors, such as Einstein Telescopes, which is designed to have a higher sensitivity in the low-frequency region.

The aim of this thesis is to start a systematic study of the stray light contribution from dust particles in the environments of the Virgo gravitational-wave (GW) Interferometer. In order to apply scattering models to estimate the stray-light contribution from dust particles, it is crucial to correctly characterize the number and size of the dust particles that can deposit on the optics. To do so, clean Si wafers are exposed as witness samples in the Quantum Noise Reduction environments of the Advanced Virgo interferometer, next to optical components that are both in air or in vacuum. The exposed witness samples are then photographed and the images analyzed with an algorithm to detect, count and measure dust particles present.

For the analysis, we use a code developed by LIGO, and, as part of this thesis, we upgrade and improve the code, minimizing the need for user input and including the capability of using adaptive thresholds. This results in higher detection efficiency and better reconstruction of particle size at various luminosity levels. We then analyze a variety of witness samples that have been exposed in different conditions in the Quantum Noise Reduction environments of the Advanced Virgo interferometer. From the number and size of the particles detected, we estimate the effective cleanliness of the environments and project dust accumulation on optics. We finally use this data to estimate the amount of scattered light contributed by the measured dust deposit. This information will be of key importance to keep dust contribution to stray light under control by setting requirements on the cleanliness of the environment and defining adequate cleaning protocols.

After an overview on Gravitational Waves Physics and the current GW interferometric detectors (Chapter 1), we will discuss how stray light from dust particulates can be studied and modelled in practice (Chapter 2). In Chapter 3, we describe how the dust monitoring campaign in Virgo is performed. We will describe the algorithm used to analyze the witness samples, presenting the improvements that we have added and its performances, in Chapter 4. Finally, in Chapter 5, we will discuss the analysis of the witness samples collected in the first stage of this monitoring campaign, and present the preliminary results obtained.

Chapter 1

Gravitational Waves Physics

In 1915, Albert Einstein gave the world a new theory of gravity: the Theory of General Relativity. This new theory is based on the simple and beautiful idea that gravity is a manifestation of warps and curves in the fabric of space-time, and not a long-range attraction force as described by Newton. Barely one year after, Einstein realized that if space-time is flexible enough to warp, it should also be able to ripple and undulate. The existence of gravitational waves (GWs) is predicted by the theory of General Relativity (GR), by Albert Einstein. They emerge as the simple idea that, as light travelling is a propagation of an Electro-Magnetic field (from Maxwell equations), a new kind of signal must correspond to the propagation of a gravitational field perturbation. Indeed, in 1916, for the first time Einstein mentioned GW in his works. Although the basic idea is surprisingly simple, the math behind is less, and Einstein himself doubted his whole life whether of not those waves are a real physical phenomena or just a mathematical artifact.

Even on the experimental side, the reputation of GWs in the scientific community does not get better when analysing the expected effect on matter. In fact, if one consider two test masses separated by a distance L, the effect of GWs is to change this distance by a relative quantity h called strain

$$h = \frac{\Delta L}{L}.$$

By doing all the math and the theory behind, scientists realized that this effect, even in the case of a cataclysmic event, the strain would be extremely small. In the case of a coalescing binary, the strain amplitude for a GW is [1]:

$$h \sim \frac{4G^2}{c^4}\frac{m_1 m_2}{rR},$$

where $m_{1,2}$ are the masses of the two bodies, r is the binary-observer distance and R is the distance between the two objects. G and c are the gravitational constant and the speed of light constant respectively. The distance R changes during the evolution of the binary, so it is not the most convenient variable to use. We can use instead the frequency of rotation of the two objects and assume it to be constant (as is the case far from the merging), which satisfies the equation

$$\omega_s^2 = \frac{GM}{R^3},$$

where $M = m_1 + m_2$ is the total mass of the binary. By substituting the R from the above equation inside the expression for h, one finds

$$h \sim \frac{4G^{\frac{5}{2}}}{c^4}\frac{\omega_s^{\frac{2}{3}}\mu M^{\frac{2}{3}}}{r},$$

where μ is defined as $\mu = m_1 m_2/M$. Let us now consider the collision of two Black Holes (BHs). BHs are among the heaviest objects present in the Universe: let's assume, as an average mass scale, they have masses $m_1 \sim m_2 = 30M_{\odot} \approx 6 \cdot 10^{31}$kg. Moreover, assume they are orbiting with a frequency of

$f_s = 250\text{Hz}$, and at a luminosity distance $r \sim 400\text{Mpc} \approx 1.3 \cdot 10^{25}\text{m}$ from Earth. This cataclysmic event would only generate a strain in the order of

$$h \approx 10^{-21},$$

which means that if the test masses are separated by 1km, their relative distance changes by a factor 10^{-18}m. This is 0.001 times the diameter of a single proton[1].

Eventually, the scientific community came to the idea of using interferometers to detect GWs: by using a laser beam to measure the length of the arm (the distance between a beam-splitter and a test mass mirror) we would be able to measure distance changes on a very small scale, and small enough to detect GWs. This idea will see the light in the USA and Italy, with the construction of the Virgo interferometer in Cascina (IT) and the two LIGO interferometers in Livingston and Handford (USA).

Finally, on September 14th, 2015, for the first time in history a Gravitational Wave (GW) signal was detected [2]. The detection was made in coincidence by the two LIGO interferometers. In 2017, the Virgo interferometer joined the LIGO detectors in an observing run, and on August the 14th, 2017, for the first time a GW signal was detected by both LIGO and Virgo [3].

The advent of Gravitational Waves astronomy provided us with a new tool to observe the Universe. Before that, the only information that we collected came from either photons, neutrinos and cosmic rays in general. GWs carry many useful informations on our Universe and its components: they allow us to observe Black-Holes (BH) and Neutron Stars (NS) merging, explore the physics and generations of such populations and probe the limits of physics knowledge that we are not able to reproduce on Earth. Citing the words of Brian Green, following the announce of the first GW detection: *"[...] for millennia we explored the Universe using waves of light, this discovery marks the opening of a new era in which we will explore the the Universe using waves of gravity"* [2].

1.1 Crash course in General Relativity

1.1.1 General Relativity: the Theory

In General Relativity, spacetime is described as a Rienmannian manifold, and the local geometry, around a spacetime point x, is governed by the metric tensor $g_{\mu\nu}(x)$. Moreover, spacetime curvature is affected by the presence of a energy-momentum, which is quantified by the Energy-Momentum Tensor $T_{\mu\nu}$. The idea is then simple: spacetime tells a mass how to move, and the mass tells the spacetime how to curve. This non-linear interaction between spacetime geometry (described by the metric tensor $g_{\mu\nu}$) and the energy-momentum (described by the energy-momentum tensor $T_{\mu\nu}$) is engraved in the famous Einstein's Field Equations (EFE) [4]:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{1.1}$$

where G is the universal gravitational constant and c is the speed of light. The quantities $R_{\mu\nu}$ and R are the Ricci Tensor and Ricci Scalar respectively, and they are defined as

$$R_{\mu\nu} = R^{\alpha}_{\mu\alpha\nu} \tag{1.2}$$

$$R = g^{\mu\nu} R_{\mu\nu}, \tag{1.3}$$

where $R^{\mu}_{\nu\rho\sigma}$ is the Riemann Tensor, defined as

$$R^{\mu}_{\nu\rho\sigma} = \partial_\sigma \Gamma^{\mu}_{\nu\rho} + \Gamma^{\mu}_{\alpha\rho}\Gamma^{\alpha}_{\nu\sigma} - \Gamma^{\mu}_{\alpha\sigma}\Gamma^{\alpha}_{\nu\rho}. \tag{1.4}$$

Finally, the Christoffel Symbols $\Gamma^{\rho}_{\mu\nu}$ are defined as

$$\Gamma^{\rho}_{\mu\nu} = \frac{1}{2} g^{\rho\sigma}(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}). \tag{1.5}$$

---

[1] Proton's diameter is approximately $\approx 10^{-15}$m

[2] "Brian Greene Explains The Discovery Of Gravitational Waves": `https://youtu.be/s06_jRK939I`

One can see that the left-hand side of the EFE (eq. 1.1) is completely dependent on the metric tensor $g_{\mu\nu}$, so that it is completely (and only) dependent of spacetime geometry. Indeed, the left-hand represents how the spacetime curves and the right-hand side is the source of this curvature, and the two quantities are related by the simple equation 1.1.

Linearized GR

The mathematical complexity behind this set of non-linear differential equations is very high, and in practice it is easier and helpful to set ourselves in a particular regime with a series of simplifications. The most useful approximation that is often used when working in GR and GWs physics is the *linearized approximation*: this allows us to rewrite the metric tensor $g_{\mu\nu}$ in the form

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{1.6}$$

where we now use the flat-space metric tensor $\eta_{\mu\nu} = diag(-1, 1, 1, 1)$ and the perturbation term is small: $|h_{\mu\nu}| \ll 1$. Physically we are now in a frame in which the spacetime is *almost* flat, with a very small perturbation term. It is easy to understand why the linear approximation can be considered good in many physical situations, if one notes that the coupling factor in the EFE is very small:

$$\frac{8\pi G}{c^4} \approx 10^{-43}\mathrm{s}^2\mathrm{m}^{-1}\mathrm{kg}^{-1}.$$

This very small value indicates that spacetime is very stiff and hard to deform, and as a result the GWs amplitude will be very small in general.

Once we introduce the metric tensor in this form, one can linearise all the tensors and quantities defined in the general theory by expanding them in powers of h and truncating the expansion at first order. One obtains:

$$\Gamma^\sigma_{\mu\nu} = \frac{1}{2}(\partial_\mu h^\sigma_\nu + \partial_\nu h^\sigma_\mu - \partial^\sigma_{\mu\nu})$$

$$R^\sigma_{\mu\nu\rho} = \frac{1}{2}(\partial_\nu\partial_\mu h^\sigma_\rho + \partial_\rho\partial_\sigma h_{\mu\nu} - \partial_\nu\partial_\sigma h^\sigma_\mu - \partial_\sigma\partial_\mu h^\sigma_\nu)$$

$$R_{\mu\nu} = \frac{1}{2}(\partial_\nu\partial_\mu h + \Box h_{\mu\nu} - \partial_\nu\partial_\sigma h^\sigma_\mu - \partial_\sigma\partial_\mu h^\sigma_\nu)$$

$$R = \Box h - \partial_\mu\partial_\sigma h^{\mu\sigma}$$

where $h = h^\alpha_\alpha$ is the trace of the linear term of the metric tensor and we also introduce the D'Alembertian operator

$$\Box = \partial_\sigma\partial^\sigma = \left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}\right).$$

By substituting the non-linear quantities in the EFE with their respective linearised expression, one obtains the linearised version of the EFE:

$$\Box\bar{h}_{\mu\nu} + \eta_{\mu\nu}\partial^\rho\partial^\sigma\bar{h}_{\rho\sigma} - \partial^\rho\partial_\sigma\bar{h}_{\mu\rho} - \partial^\rho\partial_\mu\bar{h}_{\nu\rho} = \frac{16\pi G}{c^4}T_{\mu\nu}, \tag{1.7}$$

where we also introduced the quantity

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h.$$

An important simplification in the expression of equation 1.7 comes from the freedom of performing a gauge transformation which leaves the equations covariant. In this case, the wisest choice is the Lorenz Gauge, which reads

$$\partial^\nu h_{\mu\nu} = 0. \tag{1.8}$$

Eventually, by applying the Lorenz Gauge to the linear EFE, we can rewrite such equations as

$$\Box \bar{h}_{\mu\nu} = \frac{16\pi G}{c^4} T_{\mu\nu}. \tag{1.9}$$

Remarkably, this equation has the exact form of a *wave equation*. This is the first hint that gravitational field perturbations propagates through waves, and in particular with plane waves propagating at the speed of light c, as can be seen from the c acting as the propagation speed inside the D'Alembertian operator definition.

1.1.2 Gravitational Waves

As shown by the linearised EFE (eq. 1.9), GWs seem to come out of the theory as a consequence of the wave-like propagation of the gravitational field. However, we still lack any sort of informations on the actual wave and the effect that it has on matter.

In order to dig in the study of GWs, we analyse eq. 1.9 in vacuum so that the energy-momentum tensor is null, and the wave equation simplifies to

$$\Box \bar{h}_{\mu\nu} = 0. \tag{1.10}$$

A solution of this equation is given as:

$$\bar{h}_{\mu\nu} = A_{\mu\nu} e^{ik_\rho x^\rho}. \tag{1.11}$$

By inserting this solution inside the wave equation two constraints emerges:

$$A^{\mu\nu} k_\mu = 0,$$

and

$$k^\alpha k_\alpha = 0.$$

The first condition implies that GWs are transverse waves, while from the second condition it follows that GWs group and phase velocities are equal to c. In general, those two condition allows us to lower the degrees of freedom of our $\bar{h}_{\mu\nu}$ tensor from 10 (being a 4x4 symmetric tensor) to 6.

Another useful operation that we can do to reduce even further the degrees of freedom of our tensor is to impose the *TT Gauge* (Transverse-Traceless gauge). This gauge choice reads

$$\bar{h}_{TT} = 0. \tag{1.12}$$

In this new TT gauge frame of reference, the perturbation tensor for a GWs-like perturbation is rather simple and can be expressed as

$$\bar{h}^{TT}_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} e^{i\omega(t - z/c)}, \tag{1.13}$$

where the wave is assumed to travel along the z direction with frequency ω. From this expression, it is manifest that the GW has two different polarizations with amplitude respectively h_+ (plus-polarization) and $h_\times$ (cross-polarization).

1.1.3 GW's Effect on Test Masses

To quantify the effect of GWs on test masses is not a trivial task, since whether GWs have any physical effect or they were just an artifact that could be set to 0 with an appropriate choice of gauge it was debated for decades.

Let's take two close-by test particles initially at rest, whose geodesics reads $x(\tau)$ and $x(\tau) + \xi(\tau)$, where τ is the proper time. The equation for the separation terms between the geodesics reads

$$\frac{d^2\xi^\sigma}{d\tau^2} + 2\Gamma^\sigma_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{d\xi^\nu}{d\tau} + \xi^\gamma(\partial_\gamma\Gamma^\sigma_{\mu\nu})\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0,$$

and if we restrict our derivation to the spatial coordinates we obtain

$$\left[\frac{d^2\xi^i}{d\tau^2}\right]_{\tau=0} = -\left[c\partial_0 h^i_j\frac{d\xi^j}{d\tau}\right]_{\tau=0}.$$

If the geodesics are parallel the derivative $d\xi^j/d\tau = 0$ is vanishing, hence $[d^2\xi^i/d\tau^2]_{\tau=0}$ is null too. This implies that parallel geodesics remain parallel under the passage of GWs. The explanation to this apparent issue is that we are working on coordinate distances, while we should be looking at proper distances instead. In fact, TT coordinates stretch themselves to follow the GW and keep the test masses coordinates fixed.

Luckily for us, coordinates systems are arbitrary and we can safely move to the proper detector frame in which we measure things by using rigid rules. Assume a GW propagating along the z direction and consider a ring of particles placed in the xy plane. By placing ourselves in a free falling frame, the equation governing the difference in geodesics finally reads

$$\ddot{\xi}^i = \frac{1}{2}\ddot{h}^{TT}_{ij}\xi^j.$$

The second derivative of the displacement is now non vanishing. If we looks closely, we can see that the above equation looks like to same degree to the Newtonian Force equation

$$F_i = \frac{m}{2}\ddot{h}^{TT}_{ij}\xi^j.$$

Physically, this pictorially suggests that the particles accelerate as if under the effect of a fictitious Newtonian force. Note, in fact, that this force is proportional to the mass and the distance from the expansion point.

By recalling the expression for $\bar{h}^{TT}_{\mu\nu}$ (equation 1.13), we can see right away that there is no effect on the direction of propagation:

$$\ddot{\xi}^3 = \frac{1}{2}\ddot{h}^{TT}_{3j}\xi^j = 0.$$

This was also expected since we derived that GWs are transverse waves (and we assumed that the test masses are initially at rest).

We can now compute the effect of the plus and cross polarizations. Firstly, we rewrite the displacement as $\xi_a(t) = (x_0 + \delta x(t), y_0 + \delta y(t))$. For the plus polarization ($h_\times = 0$):

$$\ddot{\xi}^1_a = \delta\ddot{x} = -\frac{1}{2}h_+ x_0\omega^2\cos(\omega t) \tag{1.14}$$

$$\ddot{\xi}^2_a = \delta\ddot{y} = \frac{1}{2}h_+ y_0\omega^2\cos(\omega t) \tag{1.15}$$

which lead to the solutions

$$\delta x(t) = \frac{h_+}{2}x_0\cos(\omega t) \tag{1.16}$$

$$\delta y(t) = -\frac{h_+}{2}y_0\cos(\omega t). \tag{1.17}$$

In the same fashion, for the cross polarization ($h_+ = 0$) we find the solutions:

$$\delta x(t) = -\frac{h_\times}{2} y_0 \cos(\omega t) \tag{1.18}$$

$$\delta y(t) = \frac{h_\times}{2} x_0 \cos(\omega t). \tag{1.19}$$

The effect of GWs on a ring of test masses results in a periodic modulation in their displacement with amplitude proportional to the diameter of the ring, and the overall deformation depends on the two different polarization amplitudes. Visually, the effect of the two polarizations on a ring of test masses is displayed in 1.1 and 1.2.

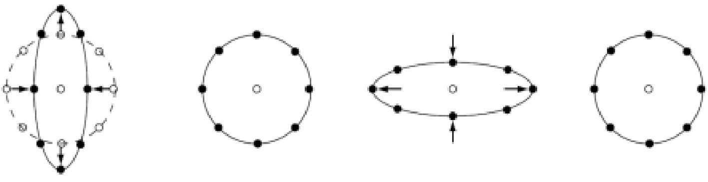

Figure 1.1: + polarization. The wave is incident in the figure plane, and time is passing from left to right.

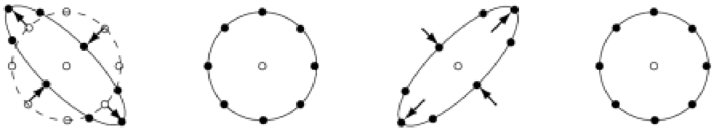

Figure 1.2: × polarization. The wave is incident in the figure plane, and time is passing from left to right.

1.1.4 GW Signal

Up to now, the idea of gravitational waves was only considered as an effect arising from the theory, and no information on what this signal actually looks like was provided. The shape of the gravitational waves signal is described by its *waveform* and it can be derived from the theory by solving the gravitational wave equation according to the stress energy tensor of our system. Even in the most familiar case (the binary merger) the complexity of the equations is very high, and analytical solutions are available under certain assumptions, requiring numerical models and simulations to compute an accurate version of the waveform. As for a fully detailed discussion on General Relativity, the full derivation on the waveform parameters is out of the scope of this thesis and can be found in [1].

Nevertheless, as will appear clearer in the future sections, one particular parameter of the waveform is important to us, as is worth spending some time on it: the frequency[3]. In fact, as will be shown in Section 1.2, the detection of a gravitational wave is made in the frequency domain, and the sensitivity range of the GW detectors is most of the time expressed as a frequency sensitivity range. A GW with a too low or too high frequency will not be detected.

In practice, how does a gravitational wave look like? Experimentally, an example of a gravitational wave signal is reported in Fig.1.3: it is GW150914, the first gravitational wave signal ever detected by human kind (more details will be provided in Sec.1.3).

For what concerns an analytical expression for the frequency of the signal, let us consider the case of a Binary Black Hole (BBH) merger, as is the case of GW150914. The frequency of the wave will depend on the mass of those objects, and on time. In fact, pictorially, as the two BHs are spiralling closer

---

[3] Actually, one can argue that the most important parameter would be the amplitude, since a signal with an amplitude too small will never be detected. This is true. However, we can safely (since we actually detected them) assume that a signal is loud enough to be detected, and move on with the treatment.

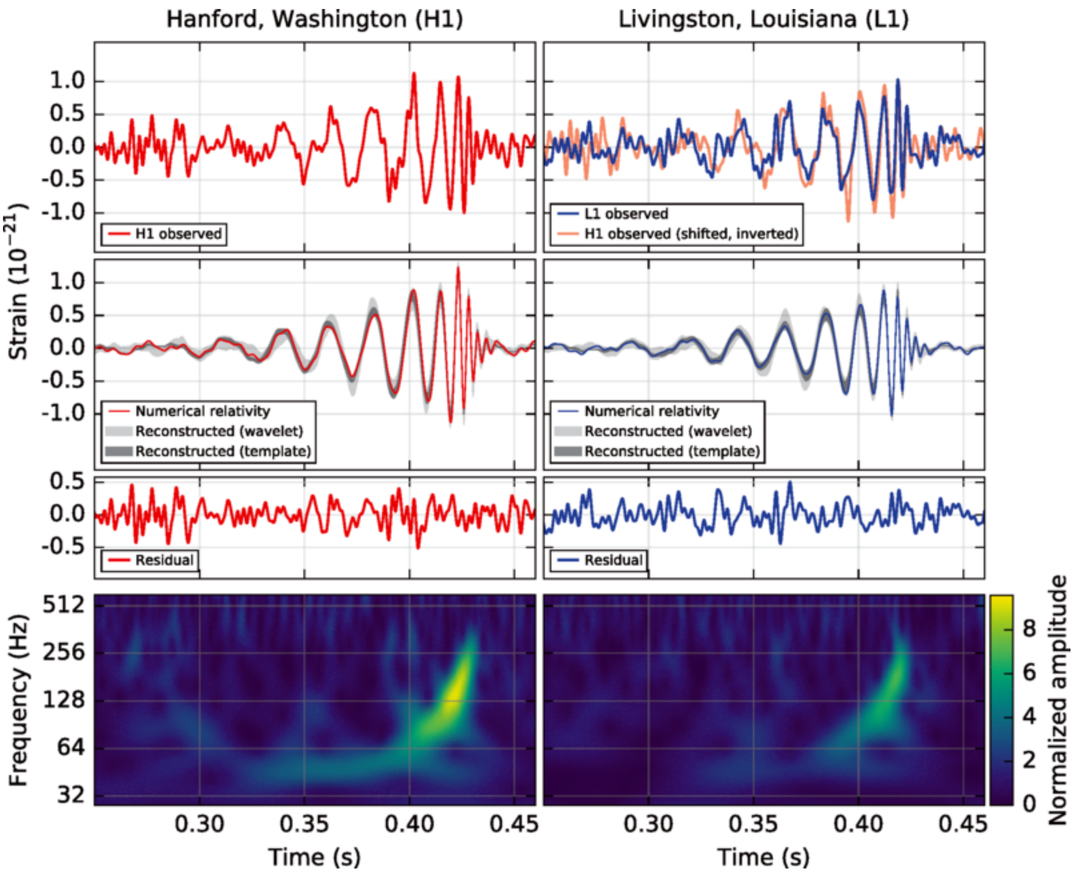

Figure 1.3: Waveform of the GW150914 event in the Hanford (H1, left) and Livingston (L1, right) LIGO detectors. The first row represents the observed (filtered) waveform, while in the second row it is reported the numerical relativity reconstruction of the signal, along with the residual from the experimental data. In the bottom row, the PSD of the signal is reported. Image from [2].

and closer to each other, they are also moving faster, hence rotating with an increasing frequency. In the same (pictorial) fashion, the frequency of the wave increases with time. This can be seen from the experimental data in figure 1.3 (bottom panel): in the frequency domain the signal's frequency can be seen to increase as we are approaching the coalescence. Analytically, the time dependence of the GW frequency, for a merging binary is [1]

$$f_{gw}(\tau) = \frac{1}{\pi}\left(\frac{c^3}{GM_c}\right)^{5/8}\left(\frac{5}{256\tau}\right)^{3/8}, \tag{1.20}$$

where $\tau = t - t_{coal}$ is the time-to-coalescence parameter (t_{coal} is the time when the two body collide at the end of the inspiral phase) and M_c is the chirp mass of the system, defined as

$$M_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}. \tag{1.21}$$

In general it can be seen that the chirp mass increases for increasing component masses, if we assume that the two masses are comparable and we do not have a large gap between them. From the $1/\tau$ dependence in equation 1.20, it may seem that at coalescence ($\tau = 0$), f_{gw} diverges to infinite. However, it should be noted that the previous equations are valid under the quasi-circular motion approximation, i.e. the radius of the orbit does not change substantially in a period, so they correctly apply on the inspiral phase of the binary only. When the binary is close to merging, this assumption is no longer valid and more complex models has to be considered. Close to the coalescence, much more complex models and numerical simulations are required to provide a more accurate waveform.

To provide an order of magnitude estimate for this frequency range, let us compute the f_{gw}, following Eq.1.20, in the case of GW150914. With $m_1 = 36M_\odot$ and $m_2 = 29M_\odot$, the chirp mass for the

GW150914 binary is $M_c = 28M_\odot$. As a time frame, let us assume $\tau \approx 0.2$s to avoid the late inspiral phase of the binary. By plugging the number in the above equation we find

$$f_{gw} \approx 30\text{Hz}.$$

This approximate prediction actually matches very well the detected frequency in figure 1.3.

The other physically important quantity entering in f_{gw} is the chirp mass. As shown in equation 1.20, the f_{gw} is also inversely dependent on the chirp mass M_c of the system, meaning that a system with an higher chirp mass would result in signal emitted with a lower frequency. As a comparison to GW150914, let us now consider the heaviest binary merger observed: GW190521, which was generated from the coalesce of the BHs with masses $m_1 = 85M_\odot$ and $m_2 = 66M_\odot$ [5]. In this case, the chirp mass is $M_c = 65M_\odot$ and in the same fashion as before (keeping the time $\tau = 0.2$s fixed) we obtain a frequency

$$f_{gw} \approx 20\text{Hz}.$$

We can see that, compared to the lighter BBHs signal GW150914, the frequency is indeed lower in this case. This was expected from the $1/M_c$ dependence inside f_{gw}.

1.2 GW Interferometers

The effect of gravitational waves is to create ripples in spacetime. Although this may lead to the possibility to detect such a (small) effect, how to detect them is not trivial at all. What we can exploit, is that light travels at the constant speed c. If the space between two objects gets stretched light would take longer to travel and, in the same fashion, if space gets squeezed light would take less time to travel. This is the idea behind using interferometers (ITF) in the detection of GWs.

1.2.1 Michelson Interferometer

The basic scheme behind the current GWs ITF detectors is the Michelson Interferometer (MI) [1]. A MI is a configuration of two mirrors and a beam splitter, along with a monochromatic light source, which is able to sense small variations in the length of the arms by exploiting the interference of the recombined light coming from the two arms. The basic scheme of a MI can be seen in figure 1.4.

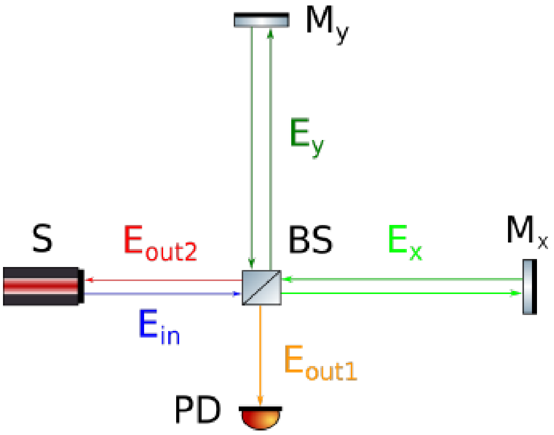

Figure 1.4: Scheme of a Michelson Interferometer. $M_{x,y}$ are the two mirrors, BS is the beam splitter and S is the source of the laser light. The E_i indicated the various amplitudes of the electromagnetic field travelling inside each section of the ITF. The interference is read at the photo detector (PD)

The electromagnetic field entering the ITF can be expressed in complex notation as

$$E_{in} = E_0 e^{-i(\omega t - k\cdot x)},$$

where $k = \omega/c$ is the wavelength of the laser and ω is the frequency of the laser. Out of the BS we have the fields:

$$E_x = \frac{E_0}{2} e^{i(kx-\omega t+\phi_x)} \tag{1.22}$$

$$E_y = \frac{E_0}{2} e^{i(ky-\omega t+\phi_y)}, \tag{1.23}$$

so that the total amplitude outside the BS, after the light travelled inside the arms and bounced back, is

$$E_{out} = E_x + E_y = \frac{E_0}{2} \left(e^{i(kx-\omega t+\phi_x)} + e^{i(ky-\omega t+\phi_y)} \right),$$

and the intensity is

$$I_{out} = |E_{out}|^2 \tag{1.24}$$

$$= \frac{E_0^2}{2} \left(1 + \cos(k(x-y) + (\phi_x - \phi_y))\right). \tag{1.25}$$

Now, we must differentiate between the two different outputs present in the ITF, $E_{out,1}$ and $E_{out,2}$. In fact, depending of which direction we consider as output, we would have different reflection/transmissions additional phases from the BS. For $E_{out,1}$ we will have $\phi_{x,1} = \pi$ and $\phi_{y,1} = 2\pi$, while for the second output $E_{out,2}$ we will have $\phi_{x,2} = \pi$ and $\phi_{y,2} = 3\pi$. The different intensities that we will eventually read are

$$I_{out,1} = E_0^2 \sin^2\left(\frac{k(x-y)}{2}\right) \tag{1.26}$$

$$I_{out,2} = E_0^2 \sin^2\left(\frac{k(x-y)}{2} + \frac{\pi}{2}\right). \tag{1.27}$$

$$\tag{1.28}$$

Physically, what we are interested in is the modulation in intensities due to the total phase difference $k(x-y)$. For example, we can design our simple interferometer to have

$$x = y + n\lambda/2,$$

where λ is the laser wavelength and n is an integer number. In this way we will measure $I_{out,2} = 0$ at the photo detector (*dark fringe configuration*) and if somehow we start to record a non vanishing signal at the PD, $I_{out,2} \neq 0$, we deduce that the total pathlength of the light has changed, meaning that the length of the arms has changed: $x \neq y$. This simple idea is actually the basic working principle used in the current (and much more complex) gravitational waves detectors: if we measure a change in the intensity of the interference of the beams, it means that the length of the arms changed[4].

GW Effect on the Interferometer

To quantify the GW-ITF interaction, it is convenient to move in the TT gauge frame. In this particular frame, we can assume the mirrors to be free-falling[5] so they do not move in the TT coordinates. Light propagation in a spacetime deformed by a +-polarized GW is described by the equation

$$ds^2 = -c^2dt^2 + (1 + h_+(t))dx^2 + (1 - h_+(t))dy^2 + dz^2 = 0,$$

which considering only the x-arm becomes

$$dx = \sqrt{\frac{c^2dt^2}{1 + h_+(t)}} \approx \pm c dt \left(1 - \frac{1}{2} h_+(t)\right),$$

---

[4] This statement is true in the proper detector frame (PDF), where we can interpret the effect of GW as a force that moves the TM and changes the length of the ITF arms. However, the PDF is a local approximation, and because of this higher order are neglected. A deeper analysis of the GW-ITF interaction will be presented in Sec.1.2.1.

[5] In the x-y plane and at frequencies above the resonance of the suspensions.

where the ± sign depends on the direction of propagation. We can now integrate the above equation in a round trip BS-TM-BS, by defining the time t_0 as the instant at which the photon starts from the BS, t_1 as the time at which the photon reaches the x-arm end mirror and t_2 as the time when the photon arrives back to the BS. The integrals of the two trips read:

$$\begin{aligned}\int_0^{L_x} dx &= +c\int_0^{t_1} dt\left(1-\frac{1}{2}h_+(t)\right)\\ \int_{L_x}^{0} dx &= +c\int_{t_1}^{t_2} dt\left(1-\frac{1}{2}h_+(t)\right)\end{aligned}$$

whose solutions are:

$$\begin{aligned}L_x &= c(t_1-t_0)-\frac{c}{2}\int_{t_0}^{t_1} dth_+(t)\\ L_x &= c(t_2-t_1)-\frac{c}{2}\int_{t_1}^{t_2} dth_+(t)\end{aligned}$$

The total travel time is obtained by summing the above quantities:

$$t_2-t_0=\frac{2L_x}{c}+\frac{1}{2}\int_{t_0}^{t_2} dth_+(t).$$

The integrand is already of order $O(h_0)$, so we can approximate the upper limit of the integral to the flat metric roundtrip time $t_2 = t_0 + 2L_x/c$. Remembering that $h_+(t) = h_0\cos(\omega_{gw}t)$, where ω_{gw} is the frequency of the gravitational wave, we can solve the integral, and eventually find

$$t_2-t_0=\frac{2L_x}{c}+\frac{L_x}{c}h_+\left(t_0+\frac{L_x}{c}\right)\frac{\sin(\omega_{gwL_x}/c)}{\omega_{gwL_x}/c}.$$

The analogous calculations can be made for the y-axis arm. For the light to arrive at the BS at the time $t = t_2$, we can invert the above equation and use $t_0 \approx t - \frac{2L}{c}$ in h. Once we obtain t_0^x and t_0^y the phase difference can be quantified as

$$\begin{aligned}\Delta\phi &= \omega_l(t_0^x-t_0^y) && (1.29)\\ &\approx \omega_l\left(2\frac{L_x-L_y}{c}+\frac{2L}{c}\text{sinc}\left(\frac{\omega_{gw}L}{c}\right)h_0\cos(\omega_{gw}t+\alpha)\right) && (1.30)\\ &= \Delta\phi_0+\Delta\phi_{gw}. && (1.31)\end{aligned}$$

The cardinal sine function, $\text{sinc}(\alpha) = \sin(\alpha)/\alpha$ is introduced. The physical phase difference induced by GW is proportional to the amplitude h_0 and the sinc, with $\omega_{gw}L$ as an argument (we have approximated $L_x \approx L_y \approx L$, which is the usual case). A plot of the sinc() function is reported in figure 1.5.

As can be seen from the plot, for $\omega_{gw}L \to +\infty$ (where the + sign is due to both quantities being positive) the sinc function, hence the phase effect of the GW, goes to zero. This may happen in two different scenarios:

- For $\frac{L_x}{c} \ll T_{gw} = \frac{2\pi}{\omega_{gw}}$: the frequency of the GW is too low for our system detector. In this case, the GW perturbation is "static" during the light's travel time, and we measure no effect on the mirrors.
- For $\frac{L_x}{c} \gg T_{gw}$: the GW frequency is now too high and during the light's travel time h_+ oscillates very fast. Those fast oscillations cancel out in a full period and we measure no effect of the GW.

Lastly, the intensity of the recombined beam in a Michelson ITF is modulated by the passage of a GW as

$$I_{out} = E_0^2\sin^2(\Delta\phi_0+\Delta\phi_{gw}).$$

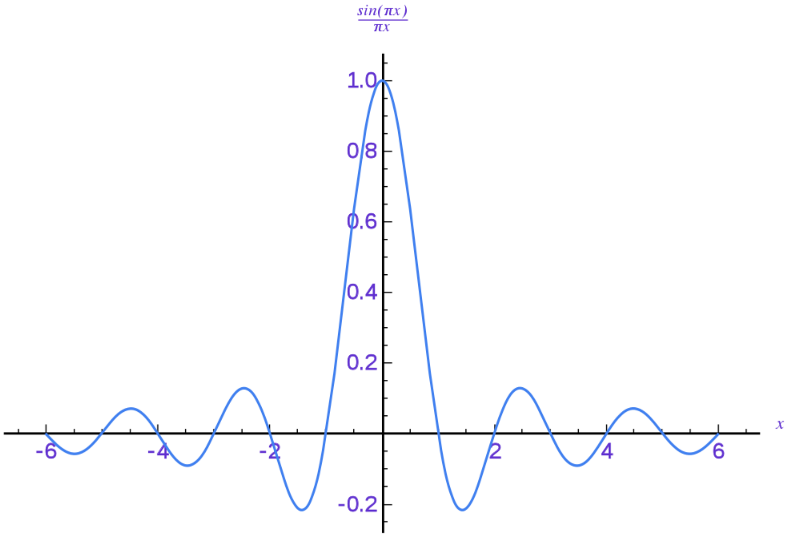

Figure 1.5: Plot of the function sinc(πx)=sin(πx)/πx. Image from [6].

1.2.2 Real Interferometer

In the previous section, we derived the ITF-GW interaction in the simple case of a Michelson ITF. Although the core idea is the same, real world ITF used to detect GWs are much more complex. In this section, a brief overview of how a real Gravitational Waves Interferometers works will be presented.

Let's recall the expression for the GW dephasing in a ITF (1.31). The magnitude of the dephasing effect dependence on the length of the arms L is two folds: a proportionality modulation factor, and a term modulated by the sinc() function:

$$\Delta\phi_{gw} \propto L \cdot \mathrm{sinc}\left(\frac{\omega_{gw}}{c} L\right).$$

The first term suggests that the longer the arms are, the greater would be the effect. However, for very high L values (for a fixed ω_{gw}), the sinc() function goes to zero, suppressing the phase effect. Ideally, we would like to build our ITF with the optimal length value which maximizes the $\Delta\phi_{gw}$. By maximizing the above expression in L, the optimal length is found to be

$$L_{opt} = 750\mathrm{km}\left(\frac{100\mathrm{km}}{f_{gw}}\right),$$

which is a very hard configuration to realize on Earth. Indeed, the ground based detectors LIGO and Virgo that were built and are now detecting gravitational waves have a much shorter arm length: 4km and 3km respectively. How was it possible for them to detect GWs with such small arms?

Fabry-Perot Cavities

The solution lies in the use of Fabry-Perrot cavities, from which follows their name of *Fabry-Perrot Interferometers.* A Fabry-Perrot cavity is an optical cavity made of two mirrors in which light is made to bounce back and forth many times before being sent back to the BS. In this way, compared to a simple MI, photons travels a much longer effective distance in the detectors frame. Equivalently, in the TT gauge frame this can be seen as the light travelling for a much longer time, hence increasing the GW interaction time. The scheme of the Virgo ITF, is pictured in figure 1.6. Comparing this scheme with the classical MI (fig.1.4), it can be seen that rather than a simple path we now have the presence of two cavities in the two arms.

The magnification factor of the effective length of the ITF depends on the finesse, $\mathfrak{F}$ of the cavity, which is defined as

$$\mathfrak{F} = \frac{\pi\sqrt{r_1 r_2}}{1 - r_1 r_2},$$

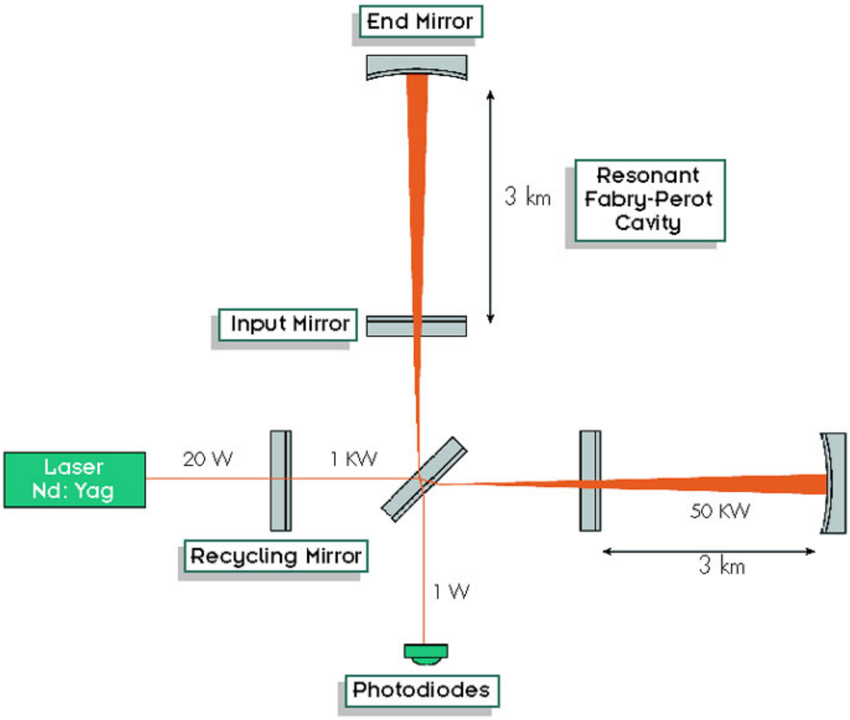

Figure 1.6: Scheme of the Virgo ITF. In the x and y arms, the Fabry-Perot cavities can be seen. Image from [7].

where $r_{1,2}$ are the reflectivities of the two cavity mirrors. One can see that the average storage time of the photons inside the cavity linearly depends on the finesse:

$$\tau_s = \frac{2L}{c}\frac{\mathcal{F}}{\pi}.$$

As a consequence, we can express this relation in terms of the effective distance that a photon travels inside the arm (the effective length of the interferometer) as:

$$L_{eq} = \tau_s c = \frac{2L}{\pi}\mathcal{F}.$$

This simple relations mean that by increasing the finesse of the cavity, we also increase the effective length of the arm and the interaction time between the photon and the GW. This is the reason why high finesse cavities are used in the LIGO and Virgo ITF: as an example, Advanced Virgo uses a $\mathcal{F} = 450$ finesse FP cavity in its arms, reaching an effective arm length equal to $L_{eq} = 860$km [1].

Equivalently, this increased sensibility can be seen as in the same fashion as a consequence of using a FP cavity: the phase shift induced by the passage of a GW increases too, and the dependence is linear in the finesse:

$$\Delta\phi_{gw}^{FP} \approx \frac{2\mathcal{F}}{\pi}\Delta\phi_{gw}^{MI} \approx \frac{4\mathcal{F}}{\pi}h_0 k_l L,$$

where k_l is the wavelength of the laser.

Power Recycling

One further improvement with respect to the Fabry-Perot ITF is the *power recycling*. This idea starts from the fact that we have chosen to work in the dark fringe configuration: when no GW is passing, the output power sent from the BS to the PD is zero. This means that all the light circulating in the ITF arms is sent back to the laser input port, and is eventually lost. In addition, as will be presented in the noise source discussion (Sec.1.2.3), there are a series of advantages to work with an higher laser power. However, we can not buy a more powerful laser right away so easily, since the power for a continuous and stable laser is limited to O(10-100W) and technical difficulties limits the maximum power obtainable. [1].

A brilliant solution to this problem came by matching the two issues presented above. In fact, by placing a highly reflective mirror (called *Power Recycling Mirror*, PR) on the laser input port of the BS, the light that would have been lost is reflected back into the ITF increasing the effective laser power circulating inside the ITF. The ITF can then be modelled as an equivalent mirror and, by adding the PR, an additional FP cavity (Power Recycling Cavity, PRC) is formed between the PC and the equivalent ITF mirror. The PR mirror can be seen in Fig.1.6. The power circulating inside the cavity can be increased with a gain of O(100): in fact, as can be seen from the figure, the PR mirror allows us to increase the effective laser power from 20W provided by the laser to the 1kW power arriving at the BS.

In addition, other cavities are used to enhance even more the performances of the ITF. On the PD port of the BS, a Signal Extraction Cavity is used, which enhance GW sidebands and can be tuned to change the response of the ITF. Moreover, Input and Output Mode Cleaner cavities are used in the input and output signal respectively, and allow to remove undesired laser modes from the beam [8].

1.2.3 ITF Noise Budget

Sensitivity Curve

The performance of a GW ITF are limited by a variety of different noise contributions, and the overall sensitivity of the detector is described by one characteristic figure: the sensitivity curve. The sensitivity curve is defined as the output strain of the detector h obtain by summing all the noise sources in input without any real GW; and it is reported over a certain frequency range.

In general, the effect of a noise source can be described in terms of its Noise Spectral Density $S_n(f)$ as a function of the frequency (Power Spectral Density PSD). The PSD for a given stationary process[6] $x(t)$, represented by $S_x(\omega)$ is defined as the Fourier transform of the auto-correlation function $R(\tau)$

$$S_x(\tau) = \int_{-\infty}^{+\infty} d\tau R(\tau) e^{i\omega\tau},$$

where the auto-correlation function is $R(t,t') = \langle x(t)x(t')\rangle = R(\tau \equiv t - t')$. The last equality comes from the stationary of $x(t)$, and the $\langle\rangle$ notation represents the mean value of the quantity in brackets:

$$\langle\alpha\rangle = \int_{-\infty}^{\infty} d\alpha f(\alpha)\alpha,$$

where $f(\alpha)$ is the probability density. From the PSD, the Amplitude Spectral Density (ASD) can be computed as

$$S_x^{1/2}(\omega) = \sqrt{S_x(\omega)}.$$

In our case, the process x will be the noise, and the output will be express in terms of the strain h caused by this noise only (assuming no GW is passing). The sensitivity curve is therefore expressed as an ASD, $\tilde{h}(f)$, in units of $1/\sqrt{\mathrm{Hz}}$. The design sensitivity curve for Advanced Virgo (AdV) is reported in figure 1.7.

From figure 1.7 one can see that the frequency bandwidth of the ITF can be divided into three different regions, which are dominated by different noise sources.

The Low Frequency (LF) region of the detector is dominated by gravity gradients, seismic movements of the ground and thermal agitation of the mirror suspensions. Gravity gradients noise, also known as Newtonian Noise (NN), are produced by fluctuations in matter density and position around the detector, for example seismic vibrations (Lowe waves and compression waves) and variation in atmospheric pressure (infrasound). Such fluctuations will generate fluctuations in the gravitational field at the detector, which will eventually sum up as noise in the measurements [10]. Contrary to the other noise sources, NN is actually a fluctuation of the metric, which is exactly what we want to measure

---

[6] A stationary process is a stochastic process for which statistical properties (e.g. the mean or the variance) are time independent.

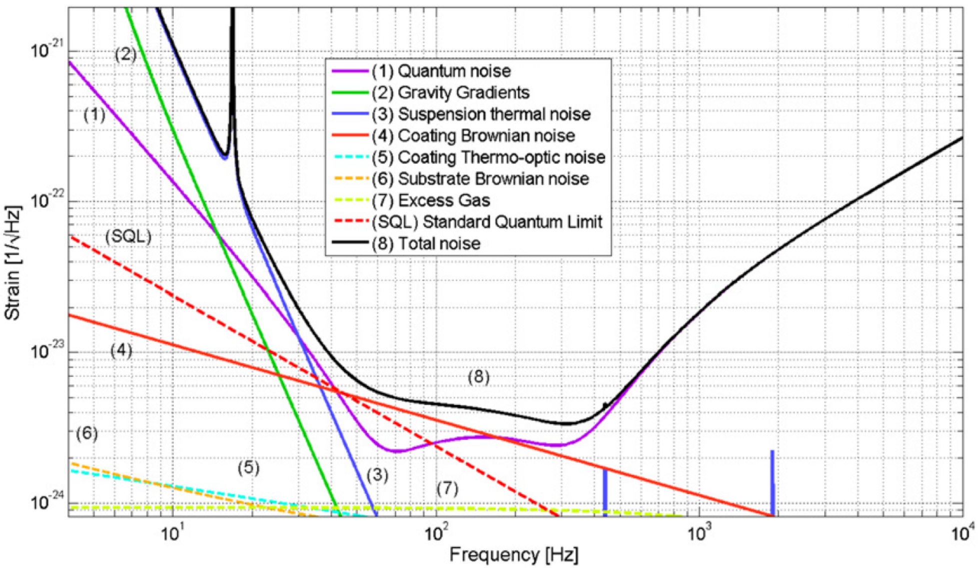

Figure 1.7: Design sensitivity curve for the AdV detector [9]. The different colored lines refer to different noise sources.

when detecting GW: only the source is different, but the physical quantity that fluctuates is the same. Since NN can not be shielded, the only viable option is an accurate modelling with the deployment of a network of sensors to measure ground displacements and atmospheric pressure variations, and modelling the effect of the measured disturbances on the mirrors in order to subtract it from the GW signal. Seismic noise is constituted by Earth's ground motion (in the order of $O(10^{-6}m)$), caused by many different sources, in particular by human activity, which affects the 1-10 Hz band. In order to suppress the seismic noise, a complex suspension system named Superattenuator (SA) has been design and realized. It is composed of a five-stages pendulum where the last stage of the suspension (the payload) includes the End Mirror [11]. Eventually, mechanical thermal noise of the mirror suspension wires is shown to depend on the fraction of energy lost for each oscillation [12], for this reason materials with low mechanical losses must be chosen in building mirrors and suspensions.

The Medium Frequency (MF) region is dominated by thermal noise of the high-reflectivity coating layers deposited on the End Mirrors surface. Reduction of this noise is possible by using new high refractive index material that exhibit the same optical characteristic of the currently employed ones but featuring lower mechanical losses [13].

The High Frequency region (HF) is dominated by quantum noise, in particular by shot noise: statistical fluctuations in the detected photon arrival rate by the readout apparatus, which translates in fluctuations of the detected output power. The other manifestation of Quantum Noise is the radiation pressure noise, the uncorrelated photon impacts on mirrors.

Quantum Noise

From figure 1.7 it can be seen that the most stringent bound at the highest frequency is the Quantum Noise, and even in the LF region it represents one of the leading noise sources. Quantum noise is the noise imposed by quantum mechanical effects arising from the quantized nature of light interacting with the mirrors and readout systems. It dominates most of the design sensitivity and limits the current ground based detectors at frequencies above 150Hz. It is composed of two main effects: *shot noise* and *radiation pressure noise*.

Shot noise It is caused by statistical fluctuations in the detected photons arrival rate by the readout system. Let N_γ be the number of photons that arrives on the photodetector (PD) in a time interval

T. The average power measured by the PD is

$$P_0 = \frac{1}{T} N_\gamma \hbar \omega_l,$$

where ω_l is the laser frequency. However, when we measure the average output power we are measuring discrete independent events, which follows a Poissonian distribution

$$p(N;\bar{N}) = \frac{1}{N!} \bar{N}^N e^{-\bar{N}},$$

where $\bar{N}$ is the average value of the variable N. For large N values the Poissonian distribution becomes a Gaussian distribution with standard deviation $\sigma = \sqrt{N}$, therefore the fluctuation in the number of photons is given by

$$\Delta N_\gamma = \sqrt{N_\gamma}.$$

The fluctuations in the readout power becomes

$$\Delta P^2 = \frac{\Delta E^2}{T^2} = \frac{(\Delta N_\gamma \hbar \omega_l)^2}{T^2} = \frac{N_\gamma \hbar^2 \omega_l^2}{T^2} = \frac{P_0 \hbar \omega_l}{T}.$$

The same standard deviation of the output power as above, ΔP, can also be found from its definition as the integral of the Spectral Density of power fluctuations over the whole frequency domain delimited by the time interval T

$$\Delta P^2 = \frac{1}{2} \int_0^{1/T} S_P(\omega) d\omega = \frac{S_P(\omega)}{2T},$$

where we used the fact that $S_P(\omega)$ is constant because the autocorrelation function is $\delta(t)$ (since photon arrival events are uncorrelated). Moreover, counting photons for a time T means sampling at frequency $f_s = 1/T$, and the single sided spectrum extends from 0 to $f_s/2 = 1/(2T)$. By putting together the two different expressions for ΔP, we can write

$$S_P(\omega) = 2 P_0 \hbar \omega_l,$$

where the factor 2 comes from the fact that we are using a single sided PSD. The output of the detector is proportional to P_0, so the relative noise is

$$S^{1/2}_{\Delta\Phi,\mathrm{shot}} = \frac{C}{P_0}\sqrt{2P_0\hbar\omega_l} = C\sqrt{\frac{2\hbar\omega_l}{P_0}},$$

where C accounts for the working point and PD efficiency, and it is of order unity. Referring the noise to the input we obtain [1]:

$$S^{1/2}_{h,\mathrm{shot}} = \frac{C}{8\mathcal{F}L}\sqrt{\frac{4\pi\hbar c\lambda}{P_0}}\sqrt{1+\left(\frac{f_{gw}}{f_p}\right)^2}. \tag{1.32}$$

$\mathcal{F}$ is the Finesse of the FP cavity, L is the length of the ITF arm, $f_p = 1/4\pi\tau_S$ is the Cavity Pole Frequency and $\tau_s = \frac{2L}{c}\frac{1}{1-r_1^2}$ is the Cavity Storage Time. From this expression, one can see that the shot noise intensity is decreasing for increasing laser power and it is also increasing with the frequency of the GW, f_{gw}.

Radiation Pressure Noise is originated from the uncorrelated photon impacts on the mirrors. Each photon that arrives to the mirror exchanges a momentum $2\hbar\omega_l/c$. Reasoning as above, one can obtain the noise referred to the input [1]

$$S^{1/2}_{h,rp} = \frac{16\sqrt{2}\mathcal{F}}{ML(2\pi f_{gw})^2}\sqrt{\frac{P_0\hbar}{2\pi c^2\lambda_l}}\frac{1}{\sqrt{1+\left(\frac{f_{gw}}{f_p}\right)^2}}, \tag{1.33}$$

where M is the mirror mass. It can be seen that, contrary to shot noise, the radiation pressure noise decreases with increasing GW frequency, and can be reduced by reducing the laser power.

Standard Quantum Limit As was shown from equations 1.32 and 1.33, shot- and radiation pressure-noise have opposite dependence on laser power:

$$S_{h,\text{shot}}^{1/2} \sim \frac{1}{\sqrt{P_0}}, \qquad S_{h,\text{rp}}^{1/2} \sim \sqrt{P_0}.$$

It seems that we can only lower one at the expenses of the other. This is a direct manifestation of the Heisenberg uncertainty principle. We are measuring the position of an object (the mirror), but photons impart a non-deterministic recoil to it. The less photons we have (low P_0), the smaller the imparted recoils will be but, on the other hand, measurement accuracy will be lower. In opposite way, by having an higher number of photons (high P_0) we will have smaller uncertainty on the position measurement, but higher recoils of the mirrors. The total quantum noise is obtained by summing together the shot- and quantum radiation pressure-noise:

$$S_{h,qn}^{1/2} = \frac{1}{L\pi f_0}\sqrt{\frac{\hbar}{M}\left[\left(1+\frac{f^2}{f_p^2}\right)\frac{f_0^4}{f^4}\frac{1}{1+f^2/f_p^2}\right]}, \tag{1.34}$$

where the frequency f_0 is defined as

$$f_0 = \frac{4\mathcal{F}}{\pi}\sqrt{\frac{P_0}{\pi\lambda_l c M}}.$$

For each value of the frequency f, the total quantum noise in equation 1.34 can be minimized by changing P_0, so that the two contributions are equal. The envelope of the minima is named *Standard Quantum Limit*:

$$S_{h,SQL}^{1/2}(f) = \frac{1}{2\pi f L}\frac{8\hbar}{M}. \tag{1.35}$$

The sum of the shot- and radiation pressure-noise over the GW frequency, and the Standard Quantum Limit are reported in figure 1.8. Different values of the total quantum noise are reported for different laser powers too.

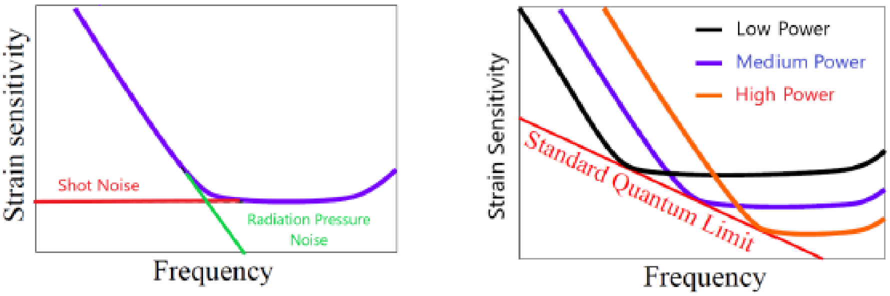

Figure 1.8: Shot- and radiation pressure-noise (red and green curves respectively) are reported over frequency, and the sum of the two (total quantum noise) is reported as the purple curve (left panel). On the right, the total quantum noise is reported for three different values of the laser power. The Standard Quantum Limit is pictured too (red line). Picture from [14].

Quantum Squeezing To lower quantum noise in the ITF, *quantum squeezing* is used. From quantum mechanics, the classical wave equation is equivalent to the evolution of an harmonic oscillator with angular velocity $\omega = kc$. From this equivalence we can substitute the position p and momentum q observables in the harmonic oscillator with two quadrature operators:

$$\hat{X}_1 = \frac{1}{2}(\hat{a} + \hat{a}^\dagger)$$

$$\hat{X}_2 = \frac{i}{2}(\hat{a} - \hat{a}^\dagger),$$

where $\hat{a}, \hat{a}^\dagger$ are the photon creation and annihilation operators respectively. One can now represent any quantum state of light in a quadrature representation, namely in a Cartesian plane where the axis coordinates are the two quadratures $\hat{X}_{1,2}$. The states that we are dealing with in an ITF are the coherent state (laser light) and the vacuum state. These two states are graphically displayed in the quadrature representation in figure 1.9. Both states are represented as a cloud-shaped probability region, rather than a point-like source: this is due to quantum mechanical vacuum state fluctuations that introduce uncertainties in our states. The difference between the two states is in the magnitude of the displacement from the origin vector: being displacement magnitude $\langle n \rangle$, the vacuum state is centred in the origin ($\langle n \rangle = 0$). The bra-ket notation on n refers to the average number of photons present in a particular state.

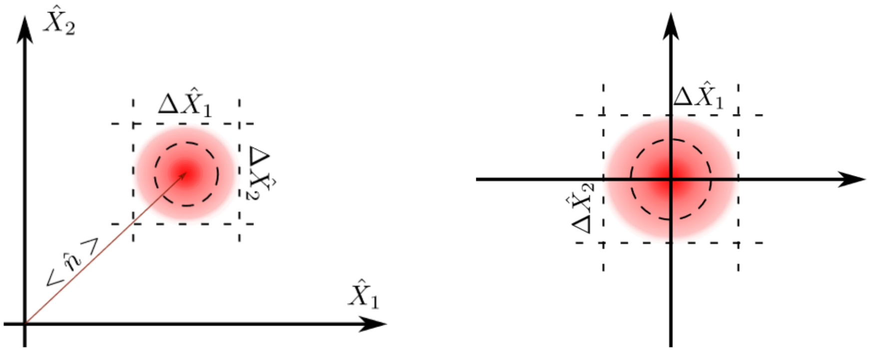

Figure 1.9: Quadrature representation of the coherent state (left panel) and vacuum state (right panel) [15].

However, in principle one could be interested in measuring only one of the two quadratures, neglecting the other one. This is the analogous case of a quantum mechanical particle: one may be interested in measuring only the position p with arbitrary precision and not caring about the momentum q: in this way one can have an (almost) arbitrary small uncertainty on the position, at the expense of increasing the uncertainty on the momentum. In terms of quantum light states: we can deform the cloud-shaped uncertainty region in order to minimize the error on the quadrature we are interested in, and increase the error on the other one. The state that we obtain after this first squeezing will the be an ellipse, with the shorter arm corresponding to the lower uncertainty quadrature. The uncertainties on the two quadratures can be rewritten as [15]

$$\Delta\hat{X}_1'^2 = \Delta\hat{X}_1^2 e^{-2r}$$

$$\Delta\hat{X}_2'^2 = \Delta\hat{X}_2^2 e^{+2r},$$

where r is named the squeezing factor.

In addition, we can also add a rotation to our state after the squeezing, in order to turn the elliptic state and obtain different uncertainties levels on the quadratures. In particular, this last operation is what allows us to beat the Standard Quantum Limit in practice. By changing the squeezing angle as a function of the frequency, we can obtain different squeezed states over the whole frequency range of the detector, each one of them being optimized for that specific frequency.

The difference between the coherent and a rotated squeezed vacuum in the quadratures representation can be seen in figure 1.10.

In practice, implementing a quantum squeezing technique in an ITF is far from easy, and the technical complexity of the required setup is very high. However, despite the many difficulties, those quantum squeezing systems are build and running in the current ITF detectors, allowing to reduce the impact of quantum noise and obtain an higher sensitivity [15].

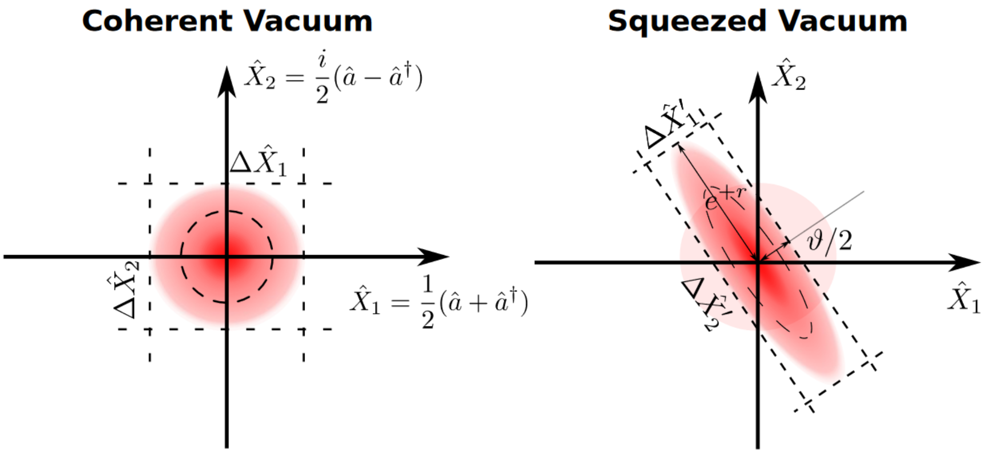

Figure 1.10: A coherent light vacuum state (left panel) and a squeezed light vacuum state (right panel) [15].

Other Noise Sources

In addition to the noise sources modelled in the design sensitivity curve (fig.1.7), there are other noise source present which cause the measured noise to be higher that the modelled one. A strong noise candidate for this excess noise is the stray light. Stray light is defined as all the light that follow a different path rather than the intended one and after recoupling with the main beam it causes random phase fluctuations which eventually are read as noise. In addition it may also spoil control loops, additionally decreasing the sensitivity of our measurements.

The stray light is a highly complex phenomenon and it is hard to experimentally model. The study of the stray light in Virgo is the main focus of this thesis, and a will be analysed in depth in Chapter 2.

1.3 GW Detections

After having seen how faint gravitational waves amplitude is from the theory, and (some of) the many difficulties in building and operating a GW Interferometer, doubts on the feasibility of a GW detection may start to arise. Those doubts are completely legit: Albert Einstein himself was one the first wondering whether or not a gravitational wave signal could be ever detectable by human kind, not only on a experimental and technical standpoint, but if they were actually *real* physical signals or just a mathematical artifact. Not an hopeful start. Later on, it was demonstrated that, in principle, GWs can be measured, although the physical effect is extremely small, and the experimental apparatus was full of difficulties and adversities. Nevertheless, thanks to the efforts of all the scientific community aiming toward the goal of detecting such signals, we eventually ended up detecting not one, but 90 of them.

1.3.1 GW Transient Catalog: GWTC-3

The third Gravitational-wave Transient Catalog (GWTC-3) describes signals detected with the Advanced LIGO (aLIGO) and Advanced Virgo (aVirgo) up to the their third observing run (O3), which ended on the 27th of March 2020, 17:00 UTC [16]. Counting all the events that were ever detected, we have a total of 90 signals from coalescing Black Holes Binaries (BBH), binary Neutron Stars (BNS) and Neutron Star - Black Hole binary (NSBH).

Among all the detections, a series of ”first time” milestones events are worth mentioning:

- **GW150914**: the first ever GW signal detected [2]. The event was detected by the two LIGO detectors on September 14, 2015. At the time, the Virgo detector was not operating on observation mode, as it will join the interferometers network during the second observing run O2. The source was identified to be a coalescing black holes binary, with components of $36M_\odot$ and $29M_\odot$.

Thanks to this historical detection, Rainer Weiss (1/2 price share), Barry C. Barish (1/4 price share) and Kip S. Thorne (1/4 price share) were awarded the 2017 Nobel Prize in Physics "for decisive contributions to the LIGO detector and the observation of gravitational waves" [17]. In figure 1.3 are reported the filtered waveforms of GW150914 in the Hanford and Livingston LIGO interferometers and the comparison with the numerical relativity predictions. It is also reported the power spectral density (PSD) of the signal: in this representation the wave can be clearly seen as an excess power at increasing frequency for times reaching the merger time.

- **GW170814**: the first event detected by both aLIGO detectors and the aVirgo detector together, detected on August 14, 2017. This joint three-detector detection allowed for an improved sky localization of the source, reducing the area of the 90% credible region from 1160deg^2 using only the two LIGO detectors to 60deg^2 using all the three detectors. In addition, for the first time it was possible to test the polarization of GWs from the antenna response of the LIGO-Virgo network [3]. The source was a BBH coalesce with masses $30^{+5.7}_{-3.0}M_\odot$ and $25.3^{+2.8}_{-4.2}M_\odot$.
- **GW170817**: not only being the first detected event of coalescing neutron stars binary (BNS), but for the first time the same event was detected using two different messengers: gravitational waves and gamma rays. In fact, a short duration gamma-ray burst, named GRB 170817A, was independently detected by the FERMI and INTEGRAL spacecraft beginning 1.7s after the GW merger signal. This event marked the birth of the *Multimessenger Astronomy Era*. The components mass were estimated to be in the range 0.86 to 2.26 $M_\odot$ [18].
- **GW190521**: this detection is consistent with a BBH merger source, but with unusually high component masses $85^{+21}_{-14}M_\odot$ and $66^{+17}_{-18}M_\odot$ compared to previous events [5]. The final mass of the merger, $142^{+28}_{-16}M_\odot$, classifies it as an intermediate-mass BH. Those high masses make GW190521 the heaviest GW source ever detected.
- **GW200105** and **GW200115**: separated by a 10 days span, they are the first observations of GWs from a neutron star - black hole merger (NSBH). The first event (GW200105) was detected only by LIGO Livingston and Virgo, while the second event (GW200115) was detected by all three LIGO-Virgo detectors [19]. For GW200105 the component masses were estimated to be $8.9^{+1.2}_{-1.5}M_\odot$ and $1.9^{+0.3}_{-0.2}M_\odot$, whereas the source of GW200115 has component masses $5.7^{+1.8}_{-2.1}M_\odot$ and $1.5^{+0.7}_{-0.3}M_\odot$.

Chapter 2

Stray Light

With *Stray Light* we refer to all the light that follows a different path rather than the intended one. After experiencing scattering or reflection, stray light may recouple with the main interferometer beam and introduce additional phase noise on local control systems, spoiling the performances of control loops. Furthermore, optical and mechanical elements experience uncontrollable motion. This causes stray light effects to show an high degree of variability depending on environmental conditions. Overall, stray light causes the phase to change in a non-linear way.

Noise in aLIGO and aVirgo manifests most notably as transient noise features (i.e. glitches) affecting the sensitivity at low frequencies. In fact, scattered light affecting the PSD during the observing runs was recorded as noise arches in spectrograms: an example of this is reported in the GWTC-3 catalogue [16], and reported here in figure 2.1. In the two spectrograms, slow (top panel) and fast (bottom panel) scattering glitches are reported. Slow scattering appears as long arches (2-2.5s long) in the time-frequency plane, while fast scattering appear as short-duration (0.2-0.3s long) rapidly-repeating arches. It is reported that, during O3, slow scattering (fig. 2.1, top panel) was the most frequent and second most frequent source of transient noise at LIGO Handford and LIGO Livingston respectively. As can be seen in figure 2.1, the slow and fast scattering glitches affects the ITF sensitivity in the $\approx 20-100$Hz frequency region. In addition to this well recognizable glitches due to oscillating scattering elements, stray light might also contribute to the more broadband, almost stationary excess noise registered in the mid-low frequency band, but this effect would be harder to isolate and characterize.

The Gravitational-Waves community is putting a large effort towards increasing the sensitivity of the instruments at the lowest frequencies, and it is therefore crucial to characterize and possibly mitigate the stray light noise contribution. The advantages of an increased sensibility in the low-frequency region is two-fold. First of all, we would be able to observe at an earlier time the inspiral phase of merging binaries, as can be seen from figure 1.3, and by inverting equation 1.20:

$$\tau \equiv t - t_{coal} = \left(\frac{1}{f_{gw}}\right)^{8/3}.$$

Being able to observe a signal for a larger time window of its inspiral phase will allow for a better reconstruction of source parameters. The second important advantage of increasing the sensibility at the lowest frequencies is the possibility to observe the signal from the merging of binaries with higher masses, since the maximum GW frequency reached at coalescence decreases with the chirp mass of the binary.

Being able to mitigate the stray light sources in the Virgo and LIGO interferometers will allow us to increase their sensitivity, allowing the detection of signals at lower frequency. In addition, the next-generation ground based interferometer, Einstein Telescope (ET) is designed to have a much higher sensitivity in the low-frequency band. For this reason, in ET stray light noise could have a greater impact compared to nowadays detectors, and its mitigation is therefore crucial [20].

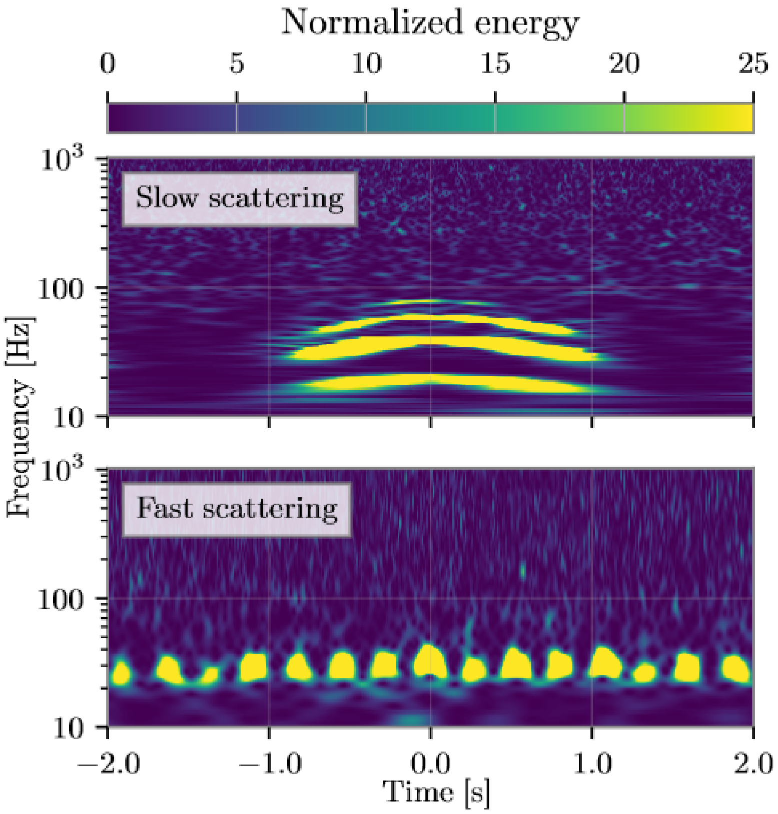

Figure 2.1: Top panel: spectrogram of glitches caused by slow scattering. Multiple arches are due to multiple reflections between the test mass optics and the scattering surface. Bottom panel: fast scattering. Image from [16].

2.1 Stray Light from Dust

In Virgo, the two main causes of scattered light are the quality (e.g. surface roughness) of optical elements and contamination from dust particulates. Due to optics having an extremely low roughness, dust in the 0.1μm $- 50\mu$m diameter range can be the leading contributor to stray light even in the clean environments of GW detectors [21]. Most of the working environments in the ITF are inside Clean Rooms, and usually this is the range of dust particles that are found inside those rooms [22]. It is possible to encounter particles as big as 100μm in diameter that may have been caused by anthropic activity. The purpose of the work presented in this thesis is to characterize the stray light noise contributed by dust particles in the Quantum Noise Reduction (QNR) environments.

2.1.1 Basic Radiometry

The final goal of our work is to quantify how much scattered light is generated by dust contamination in our optical systems. This is done by measuring the *Bidirectional Scattering Distribution Function* (BSDF). The BSDF can be directly computed from measurable quantities, so it constitutes a convenient tool to compute scattered light. From the measured dust distribution we can model the BSDF, and from this quantity we compare different scattering in our environments.

Before digging into the analysis of the light scattered by dust particulates, it is necessary to introduce some basic radiometric quantities and concepts that will be used in the future. To begin with, we define the *radiant flux*, defined as the energy per unit time

$$\Phi = \epsilon/t,$$

where ϵ is the energy and t is the time interval. From the radiant flux, and following Fest's notation [23], the *radiance* is defined as the radiant flux emitted by a surface per unit solid angle and projected area:

$$L(\theta) = \frac{d^2\Phi}{dA\cos(\theta)d\Omega},$$

where $d\Phi$ is the differential power emitted by the differential projected area of the source $dA\cos(\theta)$ into the differential solid angle $d\Omega$. θ is the angle between the normal to the emitting surface with respect to the source-observer direction. The radiance quantifies the amount of "brightness" of a surface: the more flux is emitted by a surface per unit area or per projected solid angle, the greater its radiance is. In the same fashion, one can also define the *irradiance* as the radiant flux received by a surface per unit area (also know as flux density):

$$E = \frac{d\Phi}{dA},$$

where $d\Phi$ is the differential flux incident on the surface, and dA is the differential area of the surface. Eventually, the *Bidirectional Scattering Distribution Function* (BSDF) is defined as the radiance of a scattering surface normalized by the irradiance incident of the surface [23]

$$\mathrm{BSDF}(\theta_i, \phi_i, \theta_s, \phi_s) = \frac{dL(\theta_i, \phi_i, \theta_s, \phi_s)}{dE(\theta_i, \phi_i, \theta_s, \phi_s)},$$

where θ_i and ϕ_i are the elevation and azimuth angles of the incident ray, while θ_s and ϕ_s are the elevation and azimuth angles of the scattered ray. The angles and geometry upon which all the previous definitions are based is illustrated in figure 2.2. The differential quantities dL and dE are the differential radiance of the scattering surface and differential incident irradiance, respectively [24].

The BSDF describes completely the scattering properties of a surface and it is a very useful quantity that allows to characterize the scattering behaviour of a system, since stray light performance of an optical system is usually a strong function of the BSDF of its components.

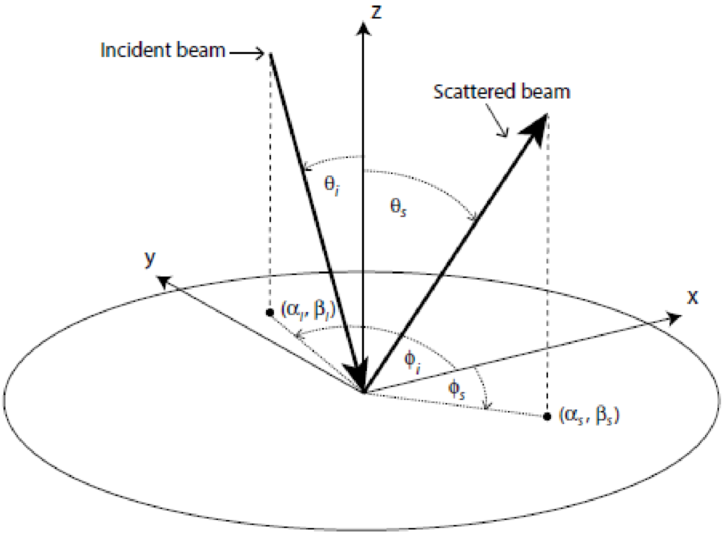

Figure 2.2: Angles used in the definition of the BSDF. Image from [23].

Our goal is to estimate the BSDF for a given dust distribution deposited on the optics: an analytical solution will be provided in the next section using *Mie Theory.*

2.1.2 Scattering from Dust Particulates

The theory for light scattering from small particulates was developed by Gustav Mie in 1908 by solving Maxwell's equations with the appropriate boundary conditions [25]. Mie Scattering assumes that particles are spheres with a complex index of refraction m, and have a diameter D comparable to the wavelength of the light. Since we are interested in the total scattered light from the whole particles distribution, we assume that particles are spatially separated enough (in the order of few diameters length) so that different contributions do not sum up coherently. In this way we can incoherently sum up the contributions coming from different particles and obtain the total scattered light from the whole distribution.

Starting from an arbitrary shaped particle irradiated by an arbitrary polarized, monochromatic, plane wave, the field at a distant point can be computed from the scattering matrix of the system:

$$\begin{bmatrix} E_s^p \\ E_s^s \end{bmatrix} = \frac{\exp(ikr - ikz)}{-ikr} \begin{bmatrix} S_2 & S_3 \\ S_4 & S_1 \end{bmatrix} \begin{bmatrix} E_0^p \\ E_0^s \end{bmatrix}, \tag{2.1}$$

where the $\hat{z}$ axis is the direction of beam propagation, $\vec{k}$ is the beam wavevector and r is the distance from the point in question and the center of the particle (for clarity sake, the radius of the particle will be labelled as R, and its diameter D accordingly). The superscript p and s refer to the polarizations parallel and perpendicular to the plane of incidence, and the subscript s and 0 refer to the scattered and incident field respectively. The terms $S_{1,2,3,4}$ are the amplitude scattering functions and they depend both on the geometry of the system and physical properties of the dust particle (a deeper, and practical, analysis will be performed in sec.2.3.1). In the special case of scattering from an isotropic sphere, S_3 and S_4 are null. In this case, if the incident radiation is polarized parallel to the scattering plane (only p polarized terms) then the scattered light will be polarized parallel to the scattering plane. The scattered irradiance I_s^p can be written in terms of the incident irradiance I_0 [26]

$$I_s^p = \frac{1}{k^2 r^2} |S_2|^2 I_0.$$

Similarly, for an incident light polarized perpendicular to the scattering plane, the scattered light is still polarized perpendicular to the scattering plane and the scattered intensity will be

$$I_s^s = \frac{1}{k^2 r^2} |S_1|^2 I_0.$$

The scattering amplitude terms $|S_{1,2}|^2$ depend on several Legendre polynomials, spherical Bessel functions and their derivative. Their analytical expression can be found in Eq.4.74 of Boheren [25], but from a practical standpoint, there are several computational tools (FRED [27], PyMieScatt [28], MiePython [29] etc...) which allow to easily calculate them. Rather than their analytical expression, what we are really interested in is the physical parameters upon which they depend. The scattering matrix elements $S_{1,2}$ are depending on the scattering angle θ_s (defined as the polar angle between the scattered radiation and the normal to surface, 2.3), the complex refractive coefficient of the particles and the size parameter x, which is defined as

$$x = \frac{2\pi R}{\lambda}, \tag{2.2}$$

where R is the particle's radius and λ is the wavelength of the laser. The λ dependence is particularly important since in general the same dust particle distribution interacting with lasers with different wavelength may not generate the same amount of stray light. A detailed study on the BSDF dependence on λ will be performed in section 5.5.

Let us now consider the case of a spherical and isotropic particle deposited on a surface, where R_f[1] is the surface reflectance. The geometry of the system is illustrated in figure 2.3. Two basic assumptions are made to compute the total BSDF [26]. Firstly, particles scatter radiation according to Mie theory and do so independently. Second, the radiation scattered in the forward direction reflects from the mirror and does not interact again with the particle. This is a key feature of the BSDF that is important to underline. The final BSDF comprehends two terms: the *Bidirectional Transmission Distribution Function* (BTDF) and *Bidirectional Reflectance Distribution Function* (BRDF).

---

[1] Often times the reflectance is indicated with R in literature, but in this case R is also the radius of the particle. For this reason, the reflectance will be indicated with R_f instead.

The BRDF corresponds to the light that is reflected back either by the particle (backward-scattered beam, fig.2.3) or by the surface (forward reflected-scattered beam, fig.2.3). On the other hand, the BTDF corresponds to the light that is transmitted by the surface. Depending on the system under study, one contribution may be neglected. However, in our study we will only focus on systems were we can neglect the BTDF contribution, and the BSDF will be constituted only by the BRDF.

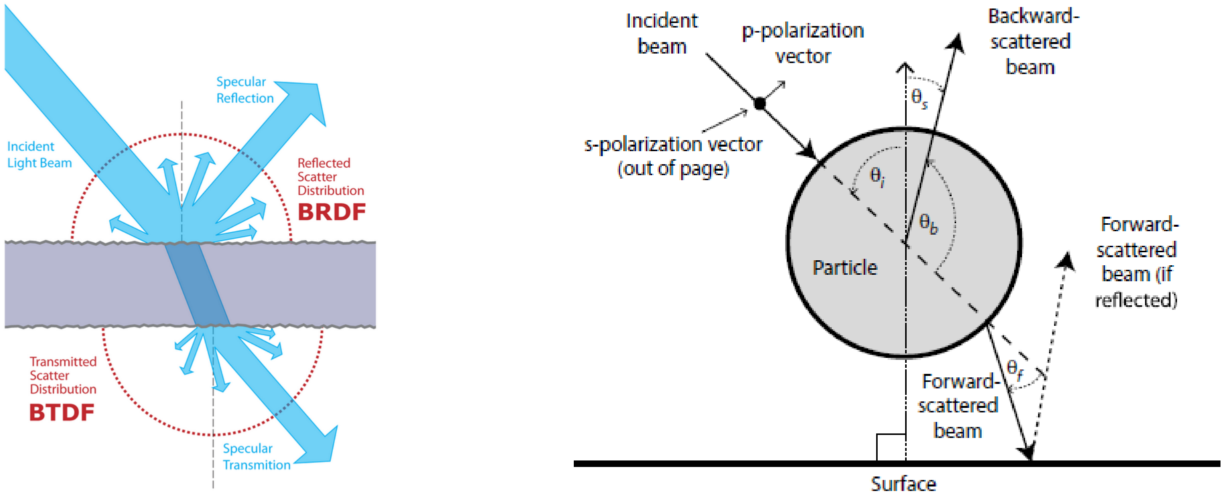

Figure 2.3: Left: pictorial representation of the BRDF and BTDF contributions separately; the BSDF is the sum of the two (image from [30]). Right: geometry of the particle on a surface scattering system (image from [25]). In this case, both the backward- and forward-scattered beams constitute the BRDF.

From those assumptions, the BSDF for a given dust particles distribution is given as [26] [23]

$$\begin{aligned}\mathrm{BSDF}(\lambda, m, \theta_s, \theta_i) = &\frac{1}{k^2\cos(\theta_s)}\sum_D f(D)\\ &\times\left[\frac{R_f|S_1(x,m,\theta_f)|^2 + R_f|S_2(x,m,\theta_f)|^2 + |S_1(x,m,\theta_b)|^2 + |S_2(x,m,\theta_b)|^2}{2}\right],\end{aligned} \tag{2.3}$$

The angle θ_i is the angle between the incident radiation and the surface normal, θ_f and θ_b are the angles between the forward and backward scattered beams with respect to the direction of the incoming radiation (fig.2.3). If one assumes the incident angle θ_i to be positive and the positive angle direction is defined as the angle from the surface normal to the specular beam, the following trigonometric relations allow to calculate: $\theta_f = \theta_s - \theta_i$ and $\theta_b = 180 - \theta_s - \theta_i$. The summation is performed over every diameter D in the dust population. The quantity $f(D)$ is the *particles density function* and represents the number of particles with diameter D per unit area. In the limit case of a continuous distribution, the summation over the dimensions will tend to a integral.

It can be noted as the $|S_{1,2}|^2$ terms inside Eq.2.3 are both appearing two times: once computed over the forward scattering angle and once computed over the back scattering angle. The only difference (a part for the argument angle) is the modulation by the transmittance of the surface for the forward scattered contributions. For example, if the dust particulate is suspended in air there would be no mirror underlying ($\eta = 0$), hence we will not have any forward-scattered beam reflected by the surface. In this case the BRDF will be given only by the backward-scattered beam.

2.2 Dust Particles Distribution

Due to the strong dependence of the BSDF from $f(D)$ in equation 2.3, differences in how one approximates the real particle distribution $f(D)$ can lead to significant differences in the estimated BSDF. It is therefore important to model $f(D)$ as accurately as possible. Two main way of expressing $f(D)$ are used. First, an approximate model where $f(D)$ is represented using a simple equation by following the

IEST standard for estimating the dust contamination on a surface [31]. As a second, more accurate method, $f(D)$ is computed as a table of the exact number of particles being present on the surface, obtained by a detailed inspection of the surface. In this section the IEST Standard and how $f(D)$ can be modelled accordingly will be presented. The second way to model $f(D)$ will be discussed in section 5.5, after having discussed our experimentally measured dust distributions.

2.2.1 IEST CC1246D Standard

The most widely used method for estimating $f(D)$, when a direct measurement is not possible, is provided by the Institute of Environmental Sciences and Technology (IEST) CC1246D standard [31]. In this standard, the number of particles on the surface N_p of a given diameter D is estimated starting from a single quantity, called the *cleanliness level* (CL) of the surface. The number of particles N_p over an area of 0.1m^2 whose diameters are greater than or equal to D, is defined as

$$N_p(S, CL, D) = 10^{|S|\left[\log_{10}^2(CL) - \log_{10}^2(D)\right]}, \tag{2.4}$$

where S is the particle distribution slope, CL is the surface cleanliness level and D is the diameter of the particles in microns. The particle distribution slope S is a function of the environment in which the particles are and on the cleanliness of the surface (the cleaner the surface, the lower S). The CC1246D standard provides a values of the slope to be $S = -0.926$, which is representative of particle distributions on cleaned surfaces. This value is used for describing freshly cleaned surfaces, and other studies conducted in Clean Rooms (CRs) indicated a more frequently observed slope of $S = -0.383$ [32]. The different estimated on S present in literature would results in drastically different distributions (and BSDF, eventually), and this may be a source of discrepancies in the results. However, in many works S=-0.926 is assumed to be the reference value for pristine surfaces.

The physical meaning of the CL value is the diameter, expressed in μm, of the particle such that there is only 1 of them with diameter bigger than CL per 0.1m^2. As an example, CL=200 does mean that there is only 1 particle over 0.1m^2 with a diameter equal to or larger than 200μm. In general, surface's cleanliness can be broadly classified as:

- CL<200: pristine surfaces
- CL=600: visibly clean surfaces
- CL>1000: visibly dirty surfaces

In literature, it is common practice to represent equation 2.4 for a given distribution in a log-log^2 diagram. The distributions corresponding to three different CL values (CL=200, CL=400 and CL=600) and assuming S=-0.926 are reported in figure 2.4.

Eventually, from the definition of N_p and the IEST standard, the distribution of particles is computed as

$$f(D) = -\frac{d}{dN} N_p(S, CL, D). \tag{2.5}$$

Once N_p is known, the particle distribution density $f(D)$ can be computed, leading to an estimation of the BSDF according to equation 2.3.

In an ideal Clean Room without human activity, the dust deposit rate onto a surface is assumed to be constant in time, and the resulting number of particles on the surface increases linearly with time. If the particles count on the surface increases but the distribution is not modified, it is straightforward to see that in a $\log - \log^2$ representation this corresponds to a simple upward translation of the line representing the particle distribution. As a consequence, CL will increase too. Assuming now that the dimension distribution of the particles does not change (the slope will remain constant), the x-axis intercept, which is the CL value, will increase too. An analytical expression for the CL dependence in time inside a clean room in stationary conditions was found in [33]:

$$\log_{10}(\mathrm{CL}) = \sqrt{\frac{1}{S}\left[\log_{10}(h) + \log_{10}(\rho) + \log_{10}(t) + 0.773\log_{10}(X_c) - 1.24\right]}, \tag{2.6}$$

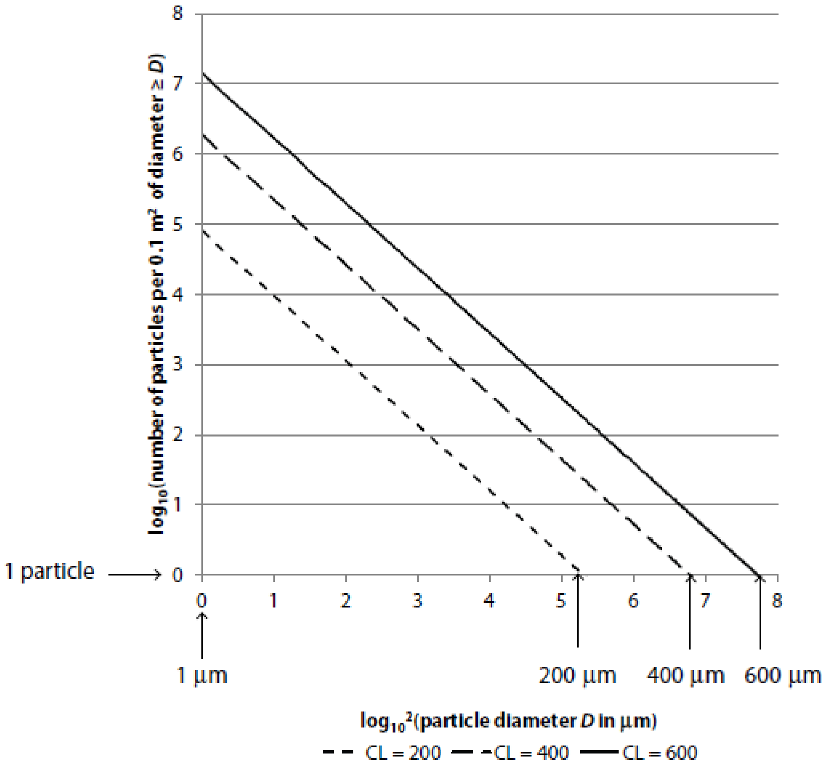

Figure 2.4: $\log - \log^2$ graphical representation of three different N_p distribution, corresponding to three different CL values: CL=200, CL=400 and CL=600 according to IEST-STD-CC1246D. The slope is S=-0.926. The x intercept of each line occurs at the particle whose diameter in microns is equal to CL. Figure from [23].

ISO 14644-1	FED 209
ISO 3	1
ISO 4	10
ISO 5	100
ISO 6	1'000
ISO 7	10'000
ISO 8	100'000

Table 2.1: Conversion table between selected ISO classes and FED 209 standard for Clean Rooms classification.

where h is the ratio between the horizontal and vertical dust deposit rate. Here I just report the literature values: $h = 1$ for an horizontal surface and $h = 0.1$ for vertical surface, from [34], and an in-depth analysis will be performed in 5.2. ρ is the number of air change per hour, and, for a typical non-laminar flow Clean Room, we can assume $\rho = 2851$ [34]. t is the time in days and S is the slope of the distribution.

X_c is the Federal Standard 209 Classification (FED 209) of clean rooms levels [35]. Nowadays, the most diffuse classification standard is the ISO: it replaced the FED 209 standard in 1999 and it was revised in 2015. However, the FED standard is sill in use in many companies, and can also be found in old research papers (before the ISO classification become the standard). The ISO and FED numbers defines the quantity and size of particles permitted per volume of air in the room, with ISO 1 being the cleanest and ISO 9 being the dirties classes. In Virgo, the cleanest environments is inside the Central Building Clean Room "Perugia". The comparison between ISO Standards 14644-1 [36] ($N_{ISO} = 3, 4, 5, 6, 7, 8$) and the FED 209 levels ($X_c = 1, ..., 100000$) is reported in table 2.1.

It should be noted that, while the CL increases with time, the ISO and FED classification of a room are assumed constant in time (if no technical issues or problems occur). An analysis of the Virgo environments will be performed in Section 5.4. In particular, by measuring the CL value from our samples, for a known exposure time t, equation 2.6 can also be used to estimate the ISO level of the room. From the comparison with the declared ISO level, we can see whether or not we are still matching its cleanliness requirements or cleaning or maintenance is required. However, our clean rooms are not stationary environments, since human activity is present. This could lead affect their the matching between the predictions from eq.2.6 and our data.

2.2.2 Compute $f(D)$ in practice

As presented in the previous section, following the IEST standard, the $f(D)$ can be sampled from the CL and S values, according to equation 2.7. Those parameters are obtained from the linear fit in the $\log - \log^2$ representation. However, such parameters are computed from a representation which is based on the number of particles with a diameter $\geq D$, $N_p(D)$ over an area of 0.1m^2. In fact, $N_p(D)$ is the quantity in the left-hand-side of equation 2.4 and upon which the IEST standard is defined. The function $f(D)$ is instead defined as the *exact* number of particles with diameter D (and not greater) over the same area of 0.1m^2.

To compute $f(D)$ from CL and S, we must compute the difference between two consecutive points in $N_p(D)$. The exact number of particles having the i-th diameter value, $N(D_i)$, reads

$$N(D_i) = N_p(D_i) - N_p(D_{i+1}) = 10^{|S|[\log^2_{10}(CL) - \log^2_{10}(D_i)]} - 10^{|S|[\log^2_{10}(CL) - \log^2_{10}(D_{i+1})]}. \tag{2.7}$$

The diameters are quantized with a uniform spacing between two subsequent values. The function $f(D)$ is then computed by normalizing over the area, and eventually converting it in μm^2 to directly compare it with the usual scale of the laser's wavelength:

$$f(D) = \frac{N(D)}{0.1\text{m}^2} = \frac{N(D) \cdot 10^{-11}}{\mu\text{m}^2}.$$

As a second option, the distribution $f(D)$ can be exactly computed from the measured dust distributions. This second method, and a comparison with the IEST standard (described above) will be discussed in 5.5.

2.3 BSDF Numerical Estimates

As reported in Eq.2.3, the BSDF expression for a given dust distribution $f(D)$ is analytically known, but the complexity of the $S_{1,2}$ terms requires numerical algorithms to be used. In this section I will present an overview of the software packages that were used in this work to compute the BSDF of the measured dust distributions.

2.3.1 Scattering Matrix Elements

A Python script was developed based on equation 2.3. Some of the quantities appearing in the BSDF expression are easily computed: $k = 2\pi/\lambda$ is the wavevector of the beam (where λ is the beam wavelength) and the angular quantities $\theta_{f,b}$ can be computed from $\theta_{s,i}$ (ref. to section 2.1.2 for the details). The particles density function $f(D)$ is computed from the particles distribution data, and will be studied later on. The key elements that are left are the quantities $|S_1|^2$ and $|S_2|^2$. Among the many packages available online for Mie Scattering calculations, MiePython was chosen [29], because in our simulation it was the fastest one[2]. This package allows us to compute the $S_{1,2}$ from the quantities upon which they depend:

- the complex refraction index of the dust: m. We assume that all the dust particles in our samples have the same optical properties, hence the same m.

---

[2] To obtain the BSDF in the same conditions, MiePython was faster than PyMieScatt [28]. Concerning the accuracy, their estimates were compatible with a 10^{-6} relative error.

- the size parameter normalized over the wavelength: $x = 2\pi r/\lambda$, where r is the radius of the particle,
- the cosine of the forward or backward scattering angle, $\mu \equiv \cos(\theta)$; with $\theta = \theta_{f,b}$, according to 2.3. The scattering angles were defined in Sec.2.1.2.

Note that, since they depend on the size of the particle x, the elements $|S_{1,2}|$ must be computed for every available size value, corresponding to each step of the summation Σ_D. This follows from the assumption that the BSDF is computed by summing all the contributions from each different particle and assuming that there is no interference between them.

The complex index of refraction m is a physical property of the particles population and depends on their material. Moreover, for environmental dust particulate, m is also found to depend on the wavelength of the laser. A reasonable list of refractive indexes for environmental dust, with respect to different light wavelength, can be found in literature and it is provided by Jennings et al. [37] and is reported in table 2.2. As a convention, MiePython requires the imaginary part of the refractive index sign to be negative, hence reading

$$m = m_{\text{Re}} - i \cdot m_{\text{Im}},$$

where m_{Re} and m_{Im} refer to the real and imaginary part of the refractive index respectively. For example, accounting also for the sign convention assumed in tab.2.2, for a laser wavelength of $\lambda = 1.15\mu$m the complex index of refraction reads $m = 1.5 - i \cdot 0.001$.

Wavelength [μm]	Complex Refractive Index m_{Re}	 m_{Im}
0.2	1.53	0.05
0.3	1.53	0.01
0.4	1.53	0.005
0.6328	1.53	0.0005
1.15	1.5	0.001
3.39	1.5	0.02
10.6	1.7	0.2
20	1.9	1

Table 2.2: Particulate complex index of refraction values at different wavelengths. m_{Re} and m_{Im} are the real and imaginary part of the complex refraction index. Table from [37].

2.3.2 Wavelength λ dependence

Particularly interesting is the dependence of $S_{1,2}$, and eventually the BSDF, from the wavelength of the laser. In fact, it is not trivial how the same dust distribution behaves in terms of scattered light under different laser wavelength, as anticipated in Sec.2.1.2. This is the case of the Virgo and ET interferometers. Although the dust monitoring campaign is currently undergoing in the Virgo QNR environments, as I reported in Chapter 1, the next generation interferometer Einstein Telescope is designed to have a higher sensitivity at the lowest frequencies, hence the impact of stray light on the performances of the detector may be much more important compared to Virgo. However, Virgo and ET uses two different laser wavelength [14] [38]:

$$\lambda_{\text{Virgo}} = 1064\text{nm}$$

$$\lambda_{\text{ET}} = 1550\text{nm}.$$

Because of this, it is not trivial to directly translate BSDFs computed in the Virgo environments to ET environments. In equation 2.3 not only the wavelength directly enters from the numerical factor as $1/k^2$, but it is also hidden inside the $S_{1,2}$ values and in the dust complex refractive index m (tab.2.2).

To estimate the effect of the λ dependence on the BSDF in the Virgo ($\lambda = 1064$nm) and ET ($\lambda = 1550$nm), the BSDF curves for different wavelength are reported in figure 2.5. The dust particles distribution is the same in all of the cases and it sampled from CL=400 and S=-0.926. As can be seen, in the three cases the BSDFs are almost the same, while greater discrepancies are found at the smallest angles. Focusing on the BSDF at 1550nm (ET) and 1064nm (Virgo), the ratio between the two is computed and reported in figure 2.6. For lower angles the ratio significantly increases, suggesting that (as shown in fig.2.5 too) the BSDF at 1550nm, starting from the same dust particles distribution, seems to be higher than the BSDF at 1064nm. It can also be noted that near 0° the tip of the BSDF flatten. This is due to the discrete angular resolution used in the simulations (we have simulated angles in [-90°,+90°], with a $\Delta\theta = (1/6)^{\circ}$ interval).

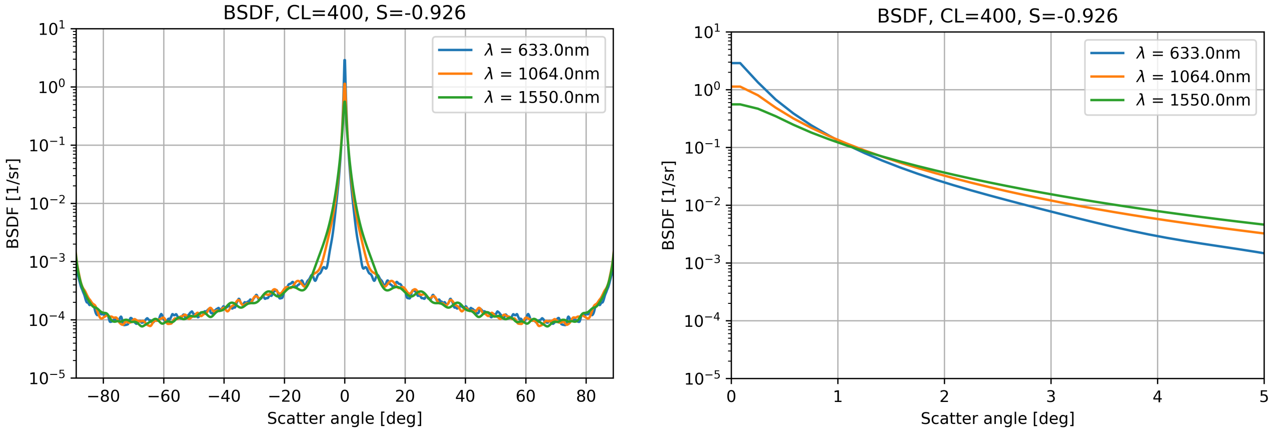

Figure 2.5: BSDF corresponding to three different wavelengths computed over the same dust particles distribution. In the right panel, a zoomed version of the curves is reported (only the positive angles are reported since the BSDF is symmetric with respect to the y-axis).

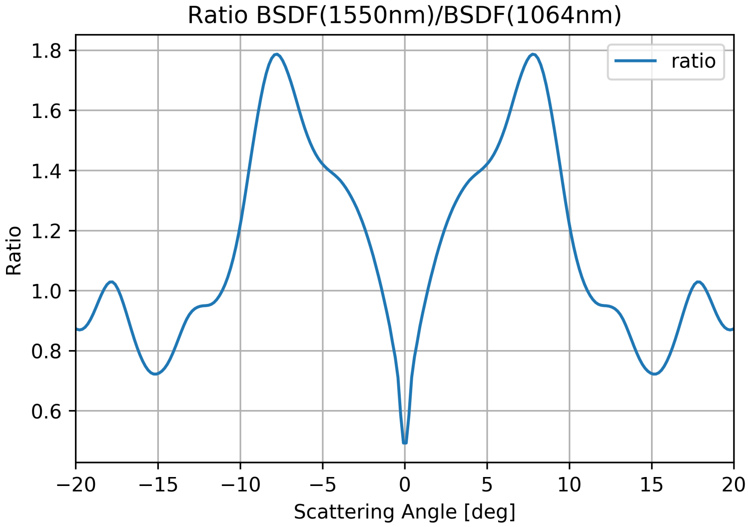

Figure 2.6: Ratio between the BSDF at 1550nm and 1064nm in figure 2.5.

Those simulations seem to suggest that in ET we may experience an higher stray light contributions at equal dust contamination level. However, the difference is very little (a factor 0(1)) and may be overcome by experimental uncertainties in the estimation of the contaminations. In addition, we have assumed the same dust particles distribution in both cases. This may not be true, as further study will be performed on assessing the expected dust particles contamination in the ET environment, which may differ with what we have measured in Virgo.

Chapter 3

Dust Monitoring Campaign in Virgo

In order to estimate the stray light due to dust particulates on optics and the associated noise in the measurements, a monitoring campaign of the dust contamination in the Virgo Quantum Noise Reduction (QNR) and Clean Room environments has been started as a collaboration of the Virgo groups in Genova and Padova [21]. The goal of this campaign is to keep monitoring in time the contamination to understand how, where and when the dust accumulates the most, and how fast. From the dust deposited on the samples we compute the overall cleanliness level of the optics and eventually the stray light contribution in terms of the BSDF. Moreover, this allows us to also measure whether or not the cleanliness standards required are matched, and plan mitigation interventions when needed, such as dedicated cleaning.

For the monitoring campaign, clean 3" Si wafers are used as witness samples that are left exposed in the tested environments for a given period of time. They are then analysed in order to measure and count the dust particles deposited, from which their distribution and expected BSDF are estimated. The analysis is performed by photographing the wafers, and the photos are analysed with a Matlab algorithm. The details of the Matlab algorithm will be presented and discussed in Chapter 4.

From the analysis we will measure differences in the dust distributions for different exposure condition: estimate an average deposit rate in different environments and how human activity may affect the dust deposit. Moreover, we also expose the wafers in two different orientations: horizontally and vertically. Although the optics are in general vertically oriented, a vertical surface (the optics) collects less dust compared to a horizontal one. Because of this, horizontally oriented wafers are used to collect an higher dust statistics and, in some cases, because it is not possible to place it vertically due to limited space in the bench. From the dust distributions in horizontal and vertical wafers exposed under the same conditions (time period, location) we extrapolate a general rule to translate the results from horizontal wafers to vertical ones where the latter cannot be used. The results of the samples analysis will be presented in Chapter 5.

In addition, we study how the stray light caused by dust contamination compare with the stray light caused by the optics surface roughness. A previous estimate [21] showed that given the high quality of the optical components used in Advanced Virgo, even in the clean environments, the dust contribution to stray light can be higher than the contribution from roughness: an example of this is shown in figure 3.1 which compares the BSDF estimated for a surface with CL=200 (normally representative of pristine surfaces, see 2.2.1) and from the roughness of the viewport used in the FIS (Frequency Independent Squeezing) lab. Although numerical simulation suggests that dust can be the most important contribution to stray light at least in the in-air benches, in Virgo the knowledge of the cleanliness of the different environments is limited and sparse. For this reason, the Virgo group of Padova has decided to start a systematic campaign, now limited mainly to the QNR environment but with the view of extending this further.

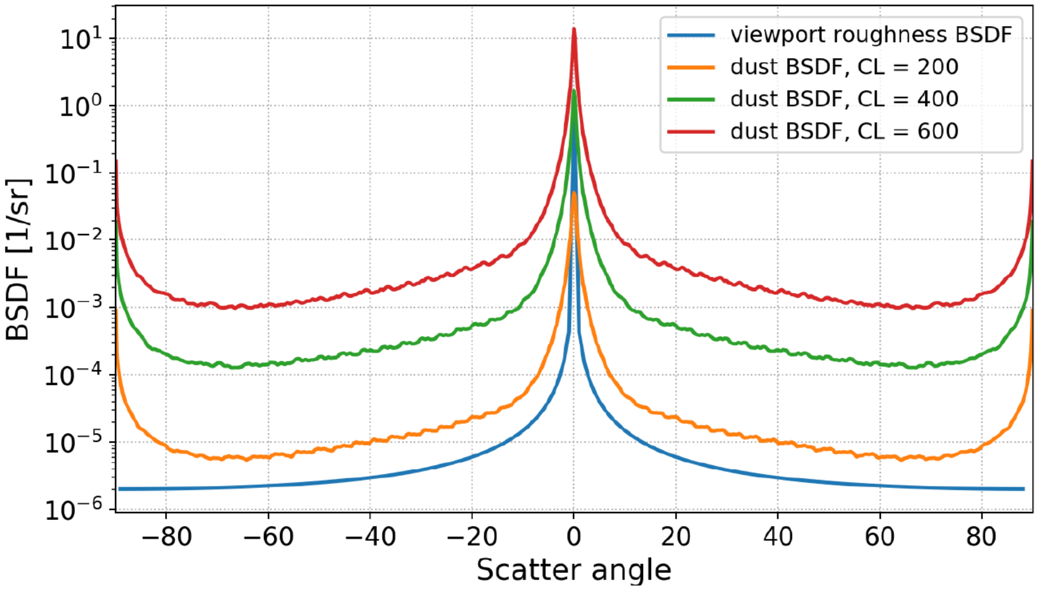

Figure 3.1: Comparison between BSDF from roughness (generated by the optical viewport of the FIS lab, blue curve, assumed to be perfectly clean) and BSDF from dust calculate for cleanliness distributions levels: CL=200, 400 and 600. The BSDF are numerical simulations of stray light performed with FRED. The viewport connects the ITF to the FIS, and has a roughness $\sigma < 0.3$nm declared by the vendor (in the simulations, $\sigma = 0.3$nm is used as an upper limit). Simulation and plot from [21].

3.1 Wafer Handling

In this section, we will present the handling procedure that were followed when the wafers were used in the monitoring campaign. As already said, the monitoring campaign was performed by exposing to QNR and Clean Room environments 3" silica wafers that served as witness samples. Brand new wafers were used each time: they came in sealed plastic boxes that were opened only just before installing the wafer on the desired spot, in order to minimize the in-air movements of the wafer and spoil the dust deposit measurement.

After opening the lid of the box, to pick up and handle the wafer a set of small tweezers was used. They were used in order to minimize the contact surface (that may leave dust particles or marks on the surface). In some cases, there was room enough to place the wafer on site with its whole box (the bottom part). In this way there was no need to pick up and transport the wafer with the tweezers. An example of this is reported in Fig.3.2: since the wafer was not directly placed on the optical bench and there were no major impedance, it was kept on the bottom of its box. Moreover, it can be noted as the box cover lid (placed on the side of the wafer) is placed with its inner side facing down to avoid dust particulates to deposit on its inside.

When the wafer had to be removed, the above procedure was performed in the same fashion (but reverse time order, of course). If the wafer was not placed inside its box it was picked up and stored back in it with the tweezers, while otherwise (e.g. in figure 3.2) it was only necessary to close the box lid and it was ready to be transported to the camera system. The only thing left to do now is to remove again the wafer from its box and place it under the camera system. In this procedure however it was only possible to use the tweezers. The height gap between the wafer plane and the illuminating LED ring was approximately 5cm. Moreover, the LED ring surface extended way more than the wafer. Because of this geometry it was difficult to operate the wafer by hand, and the tweezers were the only viable option.

Up to now, only brand new wafers were used: the pristine cleanliness of the witness samples before being exposed if of course crucial to avoid unwanted spoiling contamination. However, in this way the wafers can be used only once. This requires a large amount to be used, and the periodic need to buy

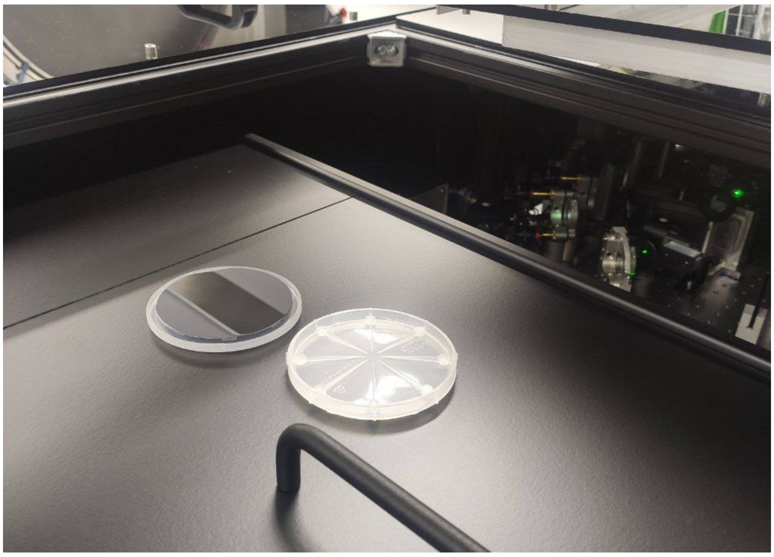

Figure 3.2: A wafer that were left exposed on top of the bottom part of its box. The cover lid is placed on the right, with the inner side facing down, ready to be closed for the shipping to the PG lab for the imaging. The wafer in this photo was exposed on the squeezing cover box in the EQB1 bench.

sets of new wafers. As a future improvement, we are planning to re-use photographed wafers after being cleaned and, to do so, a study of the optimal cleaning procedure to restore the wafer cleanliness as close as possible to the brand-new level is undergoing.

3.2 Photographic setup

In this section, the photographic setup used for imaging of the samples will be presented. In addition, the samples used for the calibration and the standard imaging procedure will be discussed.

3.2.1 Camera system setup

The photos are taken with the procedure proposed in [21] [39] [40]. The camera system used to image the samples is located in the Virgo Central Building Clean Room "Perugia" (in the following it will be referred to as "Perugia" Clean Room or PG Lab). The camera system setup is shown in figure 3.3. The system is composed by

- camera: PixeLINK camera, model number: PL-D729MU. Resolution of 9.6 MegaPixels, image size (H x V) 3840x2484 pixels and pixel size (H x V) of 2.4μmx2.4μm. [41].
- lens: aperture (f/#): f/2-f/6, fixed focal length f = 35mm [42] and 10 MegaPixel resolution.
- illumination: for illuminating the sample a white light LED illumination ring (by Edmund Optics) is used. The LED ring illuminated the samples with the dark field illumination technique [43].

Following the tuning of the system done in [21], the camera system parameters are set to

- aperture = 5.6
- exposure = 40ms
- distance lens-sample = 5cm.

From the calibration of our camera setup [21] we have a pixel-to-length ratio in the photos of

$$\eta = 1.41 \cdot 10^5 \text{px/m}, \tag{3.1}$$

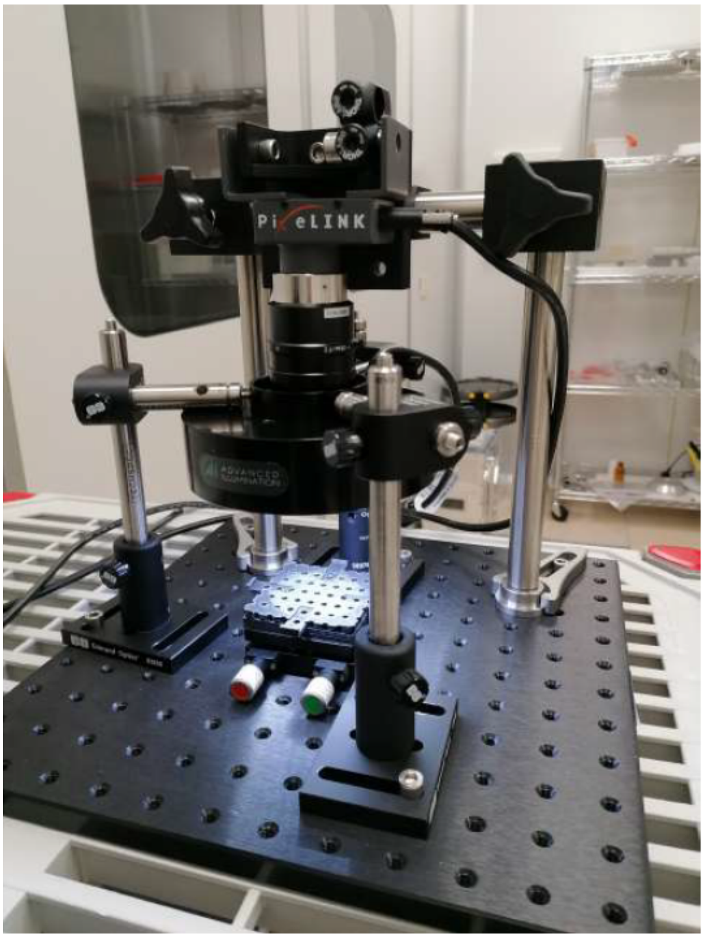

Figure 3.3: Camera system setup in the Virgo Central Building Clean Room "Perugia". The sample is placed in the holder under the camera and lens. The LED ring is visible too. The height and tilt of the sample can be adjusted with the red and green nobs on the base. Photo taken from [21].

which means that in the final photo, one pixel corresponds to

$$1\text{px} = 7\mu\text{m},$$

and the photo are 2.7cm × 1.7cm in size, for a photographed area $A = 4.8 \cdot 10^{-4}\text{m}^2$.

In general, it may be expected that the handling and imaging procedures will add a dust contamination contribution to the exposed samples. In order to quantify this contribution, a series of tests were performed. Different wafers were exposed to only the handling procedure (opening the box, removal and replacing the wafer in the box) and imaging (placing under the camera and taking the photos) and eventually analyzed. We have seen that those procedures do not affect significantly the samples. The detailed analysis and discussion will be given in Chapter 5.

In addition, in order to increase the photographed area of each sample, every wafer is photographed in five different regions by manually shifting the wafer holder under the camera as described in [21].

3.2.2 Samples

The samples used for the photos are 3" silica wafer. They were exposed to the environments that we wanted to monitor, or near a particular optical element. After exposing the wafer for a certain period of time, in which dust deposits, the wafer is placed under the camera system and photographed. The handling procedures were discussed in Sec.3.1.

For the purpose of calibrating and tuning the algorithm, three different samples of metallic dust powder used in 3D printing (AlSi [44], Ti [45] and Mo [46]) are employed. In this case, a small quantity of dust was deposited on the wafer. For each one of them the 10th, 50th and 90th percentiles on the diameters (D10, D50 and D90) are provided by the vendor. The values of the known percentiles are reported in table 3.1.

	D10 [μm]	D50 [μm]	D90 [μm]
Mo	10.9	16.2	24.1
Ti	27	38	55
AlSi	27.22	44.72	68.83

Table 3.1: Known percentiles provided by the vendors of the three metallic dust samples used in the calibration of the apparatus and algorithm.

It should be noted that, in general, the metallic dust powders may not be representative of the environmental dusts that are contaminating our environments and we want to monitor. They are however the only dust powders of known dimension available at the moment, and hence we use them to calibrate our setup. As a future improvements, As a possible future development, we might use other dust powders, of known size but with different materials, keeping in mind that here the important material parameters are the index of refraction and the isotropy

3.2.3 Stack of photos

In order to reduce noise from background fluctuations fluctuations, instead of processing a single image with the algorithm, a series of 20 photos are taken in rapid succession, and the stacked, final image, is constructed by assigning to each pixel the average value of the luminosity that the same pixel has across the whole stack. This procedure reduces the chance that random fluctuations are mistakenly read as real dust particles.

In the figures 3.4 and 3.5 the difference between a single image and a stacked image obtained by averaging 20 photos is shown. As can be seen, in both cases the noise fluctuations are drastically reduced. As a numerical estimate of this reduction, the standard deviation of the pixels luminosity levels for the images in 3.4 is computed, obtaining

$$\sigma_{single} = 0.28 \qquad\qquad \sigma_{stack} = 0.05.$$

Where σ_{single} refers to the single image (left picture) and σ_{stack} refers to the stack of 20 images (right). The ratio of the variances $\sigma_{single}/\sigma_{stack} \sim 5.6$ is the same order of magnitude reduction that was expected as $\sqrt{N_{stack}} = \sqrt{20} \sim 4.5$. For background fluctuations, the luminosity level of the pixels in general is $I_{bkg} \lesssim 4-5$, in a [0,255] grayscale.

Where the particle is present (fig.3.5), the separation between the particle and the background is sharper compared to the single image. This difference is crucial especially for the smaller, darker particles in which the luminosity is closer to the background level.

The optimization of the stacking process will be analysed in depth in section 4.3.1.

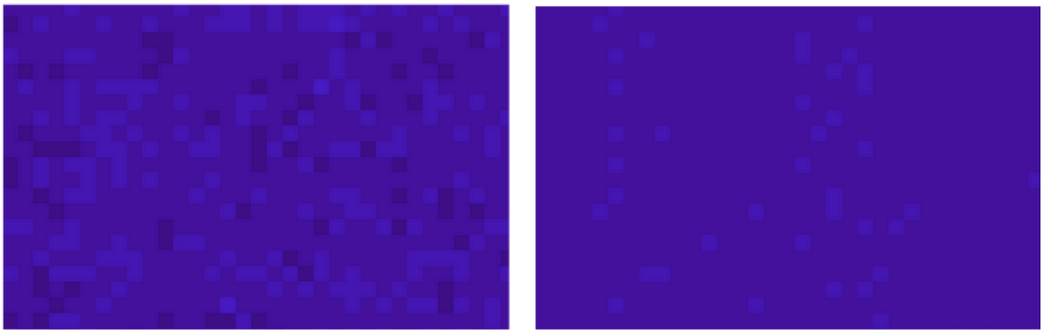

Figure 3.4: Comparison between a single photo (left) and the averaged stack of 20 photos (right) on the background only. As can be seen, the fluctuations are significantly reduced in the stack of 20 photos.

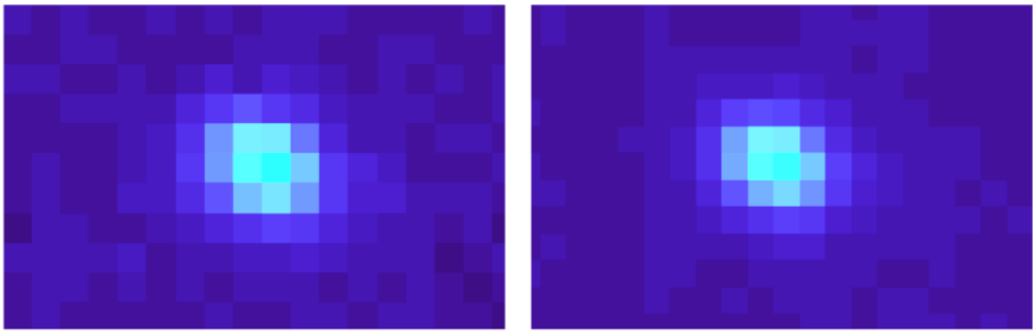

Figure 3.5: Comparison between a single photo (left) and the averaged stack of 20 photos (right) on a particle. As the can be seen, the fewer background fluctuations result in a sharper separation between the particle and the background.

The final image is eventually read by the algorithm as an 8-bit grayscale image and the image is processed by Matlab as a 3840 x 2484 (H x V) matrix of uint8 data. More details on the Matlab algorithm will follow in Chapter 4.

3.3 Monitored Environments in Virgo

The environments that were tested in the monitoring campaign in Virgo are the Quantum Noise Reduction (QNR) and Clean Room environments. In this section a brief description of the different environments and benches is given. The maps and photos of the tested environments, reporting the positions of the samples are reported in Appendix A.

- Detection (DET) Lab: the DET Lab is a ISO 7 Clean Room located in the Virgo Central Building. Inside the DET Lab the Frequency Dependent Squeezing (FIS) was located, which featured the ESQB in-air bench during O3. At the end of O3, a massive update was performed to upgrade FIS to a Frequency Dependent Squeezing (FDS) system. For this reason the ESQB bench was replace by another in-air bench: EQB1. The EQB1 bench is closed on all its sides by plexiglass cover. In addition, inside the DET Lab it is located the in-vacuum suspended bench SQB1. Since inside the DET there are very different conditions (e.g. air exposure, human activities, airflow,...) between different benches, we have decided to test each on of the separately [21].
- LAPP: located in France, the LAPP laboratory hosted the building of the SQB1 bench, which was then shipped to EGO for installation. Wafers referring to LAPP were installed on SQB1 in four different positions: L1,L2,T1,T2 (see Appendix A). Two of them (L1,L2) were removed just before the shipping, while T1 and T2 wafers were left mounted during shipping and handling of the bench [47].
- 1500W: located in the west arm of the ITF, it hosts the clean rooms dedicated to the squeezing R&D laboratory, the payload assembly and to the fused silica fiber facility [48].
- Perugia Laboratory: clean room located in the Virgo Central Building, it hosts the camera setup described in sec.3.2.1.

Chapter 4

Algorithm for Dust Characterization

In Chapter 3, it was presented how, using silica wafer as witness samples, the dust monitoring campaign is performed in the Clean Room an QNR environments in Virgo. In addition, the camera system used to photograph the sample was introduced. In this Chapter, we will present a detailed analysis of the Matlab algorithm that was used to detected and measure dust particles in the photos.

To detect and size dust particles on a photographed sample an existing MATLAB algorithm uses an user-input luminosity threshold to binarize the image and compute the number and size of the particles. The threshold value converts the image from grayscale to binary, by setting each pixel either to 0, if their luminosity is below threshold, or 1, if their luminosity is above the threshold [49].

The threshold choice results to be critical for the analysis [21]. If the threshold is too high, the algorithm will mistakenly classify certain particles, or parts of particles, as background. This causes an underestimation of the particle size, or the loss of the particle identification, as it is mistaken as background. On the other hand, if the threshold is too low, the algorithm will read bright parts of the background as real particles. This causes fake particles to be detected. As shown in [21], experimentally we are not able to find a threshold providing an high detection efficiency and an accurate dimension reconstruction at the same time.

An improved version of the existing algorithm was developed in this thesis: the particles sizing is now done with a local binarization of each particle found, with the local threshold that depends on the luminosity of the cluster of pixels. In particular, the particle-specific threshold is taken to be a fraction of the maximal luminosity of the particle. This allows us to separate the steps of particle detection and of particle sizing: the detection is done with a first global threshold (which can now be taken relatively low), while the size reconstruction is done locally for each particle.

We will describe the algorithm that we have developed, reporting each step of its workflow and the optimization procedure that was performed. The algorithm calibration and the estimation of the performance of the whole setup are reported at the end of this chapter. The following one (Chapter 5) reports the results related to the monitoring campaign in Virgo obtained with this setup and algorithm. The Matlab code used in our analysis can be found in [50].

4.1 Original LIGO Algorithm

As reported in the introduction, the original code [49] works with a user-input luminosity threshold to binarize the entire image. However, this simple approach is unable to perform equally good on both the detection and sizing of the particles. The scheme of the algorithm is the following:

1. background subtraction: the background illumination due to the illumination ring in the image is removed. The detailed analysis of this procedure is reported in Sec.4.2.1.
2. Binarization of the image with a given single user-input threshold, T: the image is converted from grayscale to binary, by setting each pixel either to 0, if their luminosity is below T, or 1, if

their luminosity is above T instead.

3. Identification and measurement of particles' area: those steps are performed using standards Matlab algorithms [51] [52].

From the measured area, the effective diameter, d, of the particles is computed assuming they have a circular shape, namely:

$$d = 2\sqrt{\frac{A}{\pi}}.$$

The optimization of this algorithm consisted in finding the threshold value that leads to best estimation of the (known) particle dimensions, and it is performed using the known dust samples (AlSi, Mo and Ti) described in section 3.2.2. In fact, in addition to excluding the background fluctuations, the choice of the threshold has an important impact on the size of the reconstructed particles. As an example, in figure 4.1 the reconstruction of the same particle is shown for different luminosity threshold values. Both the binarized and the non-binarized image (along with the reconstructed perimeter, red line) are displayed. It can be clearly seen how different threshold choices affect the particle's dimension. In particular, for higher threshold values it can be seen that the particle tails tend to be considered as background. The value $T = 80$ found by the previous calibration [21], seems to perform well in the reconstruction of the dimensions.

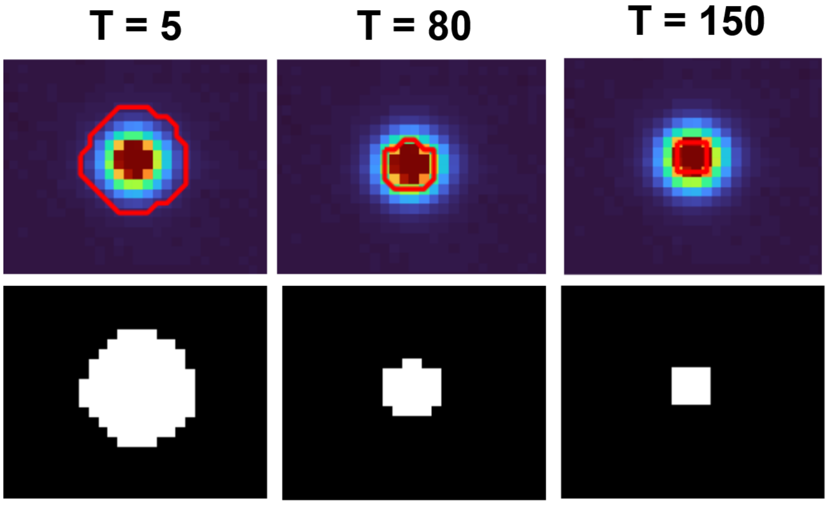

Figure 4.1: In the panels the reconstruction of the same particle dimensions is shown for different values of the threshold, T=5 (left), T=80 (center) and T=150 (right). In the upper row the original image (non-binarized) is reported, with the perimeter reconstructed by the algorithm. In the bottom row the image binarized with the respective threshold is shown. By default setting in Matlab, the boundary line (in red) is drawn by connecting the centers of the boundary pixels.

Given the percentiles reconstructed by the algorithm ($D10_{exp}$, $D50_{exp}$, $D90_{exp}$), the residual from the percentiles provided by the vendors (D10, D50, D90) are computed as

$$res_{D10} = \frac{D10_{exp} - D10}{D10}, \tag{4.1}$$

in the case of the D10 percentile. The D50 and D90 are computed similarly. From the three residuals, the Sum of the Squared Residuals (SSR) is computed:

$$\mathrm{SSR} = \sqrt{res^2_{D10} + res^2_{D50} + res^2_{D90}}. \tag{4.2}$$

The SSR provides an estimate on how close the reconstructed distribution percentiles are to the known percentiles. The optimal threshold corresponds to the value giving the lowest error on the percentiles, hence minimizing the SSR. However, this simple approach causes a series of limitations, due to the criticality of the threshold choice, causing shortcomings in the algorithm performances with the aforementioned optimal threshold:

- lower thresholds overestimate the number of particles (background pixels) and the particle's dimensions,
- higher threshold underestimate the number of particles and their diameter.

In previous analysis [21], the threshold value

$$T = 80.$$

was found. This relatively high threshold has the drawback that many small particles are discarded: being less luminous, such small particles do not survive the binarization, and are not recognized by the algorithm. Indeed, a visual comparison between the results of the particles search with $T = 80$ and $T = 5$ is shown in figure 4.2. The particles highlighted in red are the ones found by the algorithm, while the gray ones are discarded. Some of the discarded ones are highlighted with arrows. It is clear by the visual comparison that a significant portion of particles is discarded by the algorithm for the high threshold, which on the other hand minimizes the SSR.

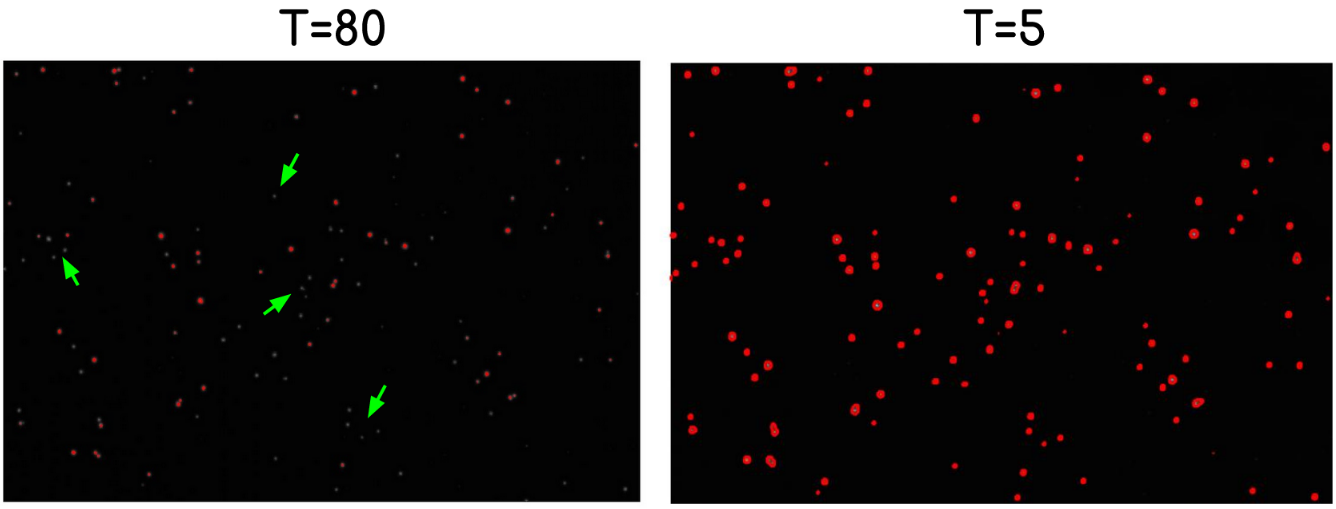

Figure 4.2: Visual comparison between the particles found with $T = 80$ (left) and $T = 5$ (right). The picture is a zoomed portion of a photo containing Ti dust.

We find that, the limitations of the existing algorithm are due to the single-threshold binarization of the image, therefore an improved version of the algorithm was developed as part of this work. In this new version of the software, the tasks of detection and sizing of the particles are separated: the detection is done after binarizing the whole image with a single threshold T_1, while sizing of the detected particles is done with particle-specific threshold T_2 which depend on the particle luminosity. In addition, the estimation of the optimal first threshold T_1 is now done automatically by the algorithm. In this way there is no need for human supervision and the threshold choice can adapt to each sample individually.

4.2 The Improved Algorithm

In this section the improved version of the code that was developed is presented. Firstly an introduction of the general workflow is given, while in the next sections each step is analysed in depth.

The core upgrades of the code are the separation of the detection and sizing of the particles in a 2-step binarization and the introduction of an automatic threshold choice. Firstly, a global search with a luminosity threshold T_1 is performed. The task of this step is the detection of all the particles. Then a second search, iterated over each particle (i.e. bright cluster of pixels) found, with a threshold T_2.

This second threshold is adaptive to the luminosity of each cluster and allows for a more accurate reconstruction of the dimension.

Schematically, the algorithm works as such:

1. removal of the background luminosity.
2. Determination of the optimal initial threshold, T_1
3. Binarization of the entire image with the threshold T_1.
4. Identification of bright pixel clusters over the whole image
5. For each cluster found:
    (a) Select the portion of the original (i.e. not binarized) image which contains the cluster
    (b) compute a second luminosity threshold as the fraction α of the maximal intensity of the cluster, I_{max}:
    $$T_2 = \alpha \cdot I_{max}$$
    (c) binarize the selected image portion using T_2
    (d) computes particle properties (position, area, perimeter, diameter, peak luminosity)

The flowchart of the improved algorithm, compared with the original LIGO code, is reported in figure 4.3.

In the next sections, each step is described in depth. At the end of each section, the input and outputs are explicitly reported too.

4.2.1 Background light removal

The algorithm starts from the image obtained by the averaged stacking of 20 images, as said in section 3.2.3. The first step of the algorithm removes the background light noise.

The unwanted background light can be divided into two categories: a halo of light (which is evident from visual inspection of the photo), and a background diffused light (which causes an overall luminosity offset in the image). Both those contributions are removed with a morphological opening procedure [53]:

1. create a background image by removal of small details from the image (i.e. the particles). The algorithm is set to remove all the features that can not contain a disk with diameter of 15 pixels. From the pixel-to-meter ratio obtained from our calibration $\eta = 1.41 \cdot 10^5$px/m (eq.3.1) we can see that this means that all the disks that have a radius smaller than $r = 105\mu$m are removed. In our samples, no particle with such dimension was ever found. Moreover, a particle this big can be seen even with the naked eye both on the wafer and on the photo. If this is removed from the photo by the erosion, it can be seen from a direct comparison of the photos. As a safety check, both the photo before and after the morphological opening are shown by the algorithm for a visual comparison. The morphological opening is performed with the Matlab function `imopen()`.
2. Subtract the obtained image (the background) from the original one: now we are left with an image where the background is subtracted and we are only left with particles.

As an example, in figure 4.4, we report the same image (containing Mo dust) before (left) and after (right) the background light removal: the background light contribution can be seen to be removed by this process.

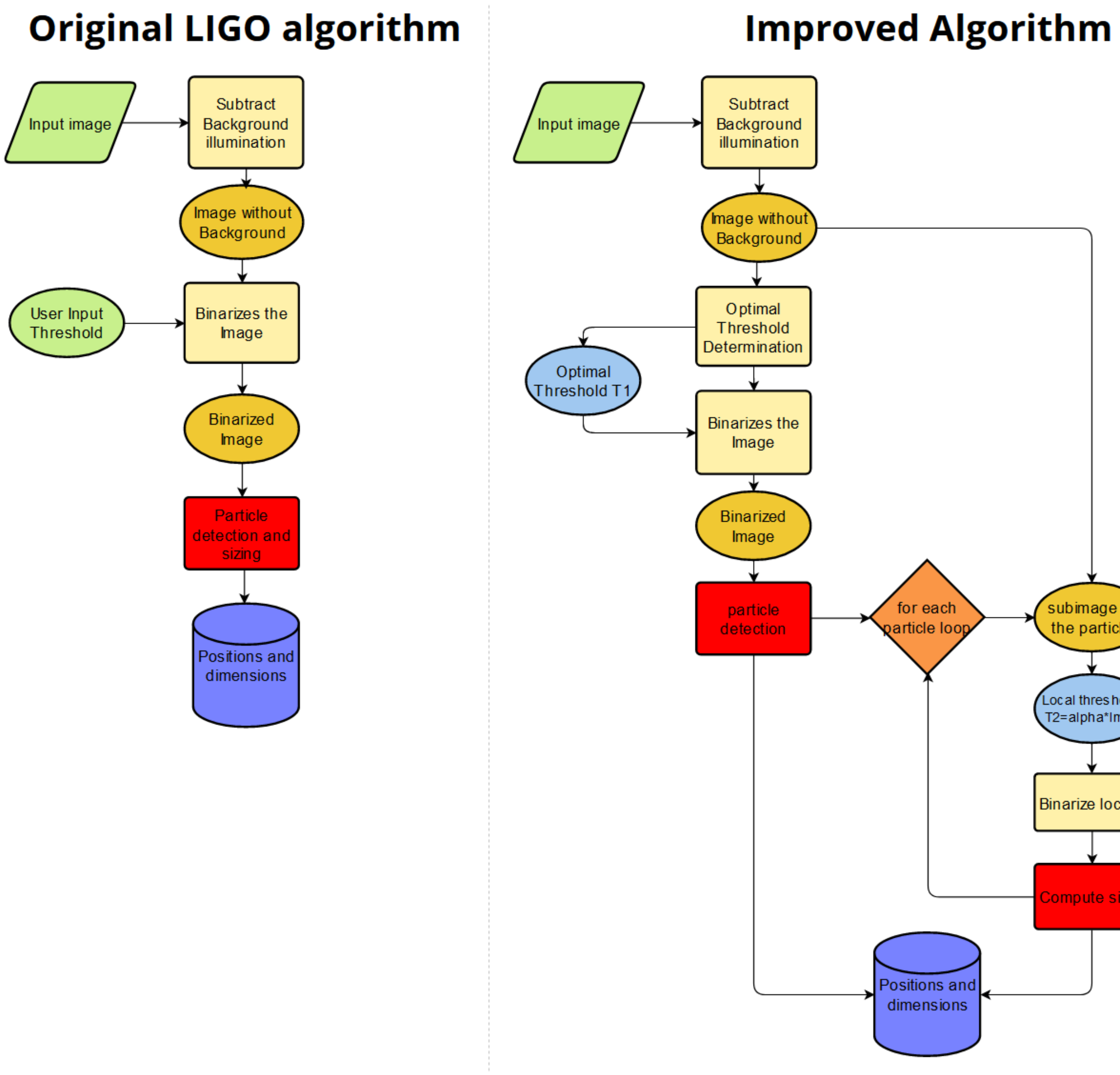

Figure 4.3: Flowcharts of the original LIGO algorithm (left) and the improved algorithm here presented (right). The elements in light green are the inputs which are provided by the user. The dark yellow blocks are the different processed version of the input image. In red the task of detection and sizing of the particles are shown: the differentiation of the tasks in the improved algorithm can be seen. The two different thresholds (T_1 and T_2) are explicitly highlighted (light blu circles). Eventually, a data table of the position and dimensions of all the particles are returned.

Input & output

- Input: the stacked image (e.g. fig.4.4, left)
- Output: the image without background luminosity (e.g. fig.4.4, right)

4.2.2 Automatic initial threshold determination

After the background light is removed from the image, the algorithm needs to find the optimal luminosity threshold T_1 in order to perform the first binarization. This choice is based on the Number of Particles (NoP) found in the image: the algorithm binarizes the image for different threshold values (e.g. T=1,2,...,20) and for each threshold computes the NoP. The NoP as a function of the threshold, for different samples, is reported in figure 4.5. While decreasing the threshold, there is a certain

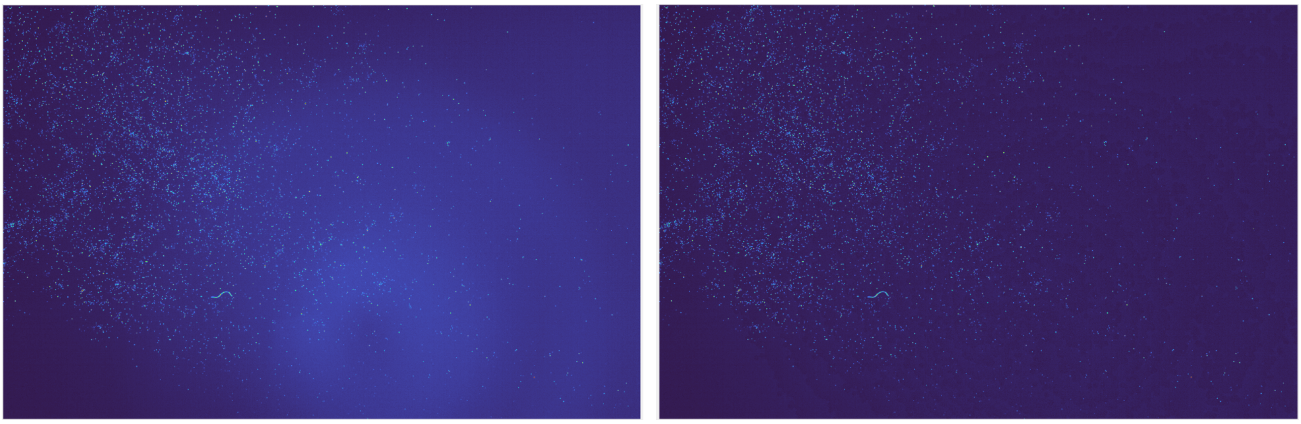

Figure 4.4: The same image before (left) and after (right) the background light removal. The image on the left is the result of the average stacking of 20 nominally identical photos, as in section 3.2.3. The reported image is a photo containing Mo dust, with size 3840x2484 pixels. A small fiber can be seen in the photos, in the bottom-left part of the central region, which survives the background subtraction. In fact, the fiber is 130 pixel long, but being very narrow (5 pixel wide) it can not fully contain a disk of 15 pixel. For this reason it is removed when creating the background image (imopen() function of Matlab) and therefore, after performing the subtraction of the background image from the original photo, it is not removed from the latter.

value at which the NoP drastically increases, due to random, slightly more luminous than average background pixels being counted as particles. This can be seen as the peak in the NoP (fig. 4.5) at $T \lesssim 3.$. The previous reasoning is however valid because we assume to work in a regime where few particles are expected; and, in addition, such an abrupt increase cannot be explained by more small (i.e. less luminous) particles being detected as we expect the particle distribution to vary smoothly with the diameter.

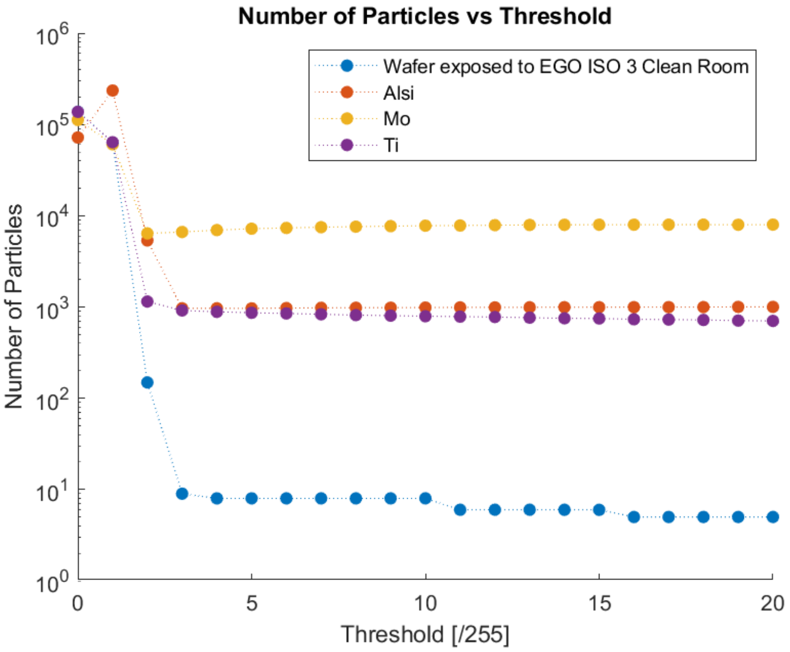

Figure 4.5: NoP as a function of the threshold for different samples: AlSi, Mo and Ti are the three known dust samples, while the light-blu one is a clean wafer that was left exposed in the Perugia Clean Room (labelled "EGO-Virgo ISO 3 Clean Room") for 15 hours.

The optimal threshold should be the lowest possible value which apparently does not include a substantial fraction of background pixels. Being as low as possible, we limit the discarding of the smallest particles, which are less luminous. Indeed, this was one of the limitations found in the original algorithm: the reconstruction of the particles dimensions required a relatively high threshold, which however discarded the smallest particles from the analysis.

To do so, once the values of NoP as a function of the threshold are obtained, the algorithm starts from the highest value of threshold, $t_{\max}$ (e.g. $t_{\max} = 20$ for the above plots) and moves backwards to lowest values. At each i-th step, the algorithm computes the average NoP, $\overline{NoP}(t_i)$, for all the values of the threshold $t_i \leq t \leq t_{\max}$. Namely:

$$\overline{NoP}(t_i) = \frac{1}{t_{max} - t_i} \sum_{t=t_i}^{t_{max}} NoP(t), \tag{4.3}$$

The standard deviation at the i-th step, $\sigma(t_i)$, over the same threshold range ($t \in [t_i; t_{\max}]$) is computed as well

$$\sigma(t_i) = \sqrt{\frac{1}{t_{max} - t_i - 1} \sum_{t=t_i}^{t_{max}} \left| NoP(t) - \overline{NoP}(t) \right|}. \tag{4.4}$$

In addition, a lower bound of $\sigma(t_i) \geq 1$particle is set: if from the above equation $\sigma(t_i) < 1$ results, the standard deviation is set to $\sigma(t_i) = 1$ instead. The importance of this bound will be explained below.

To determine if the number of particles found increases anomalously, the algorithm checks if the NoP of the next steps, (i+1)-th, is contained or not in a $5\sigma(t_i)$ interval from the average value computed on the previous step $\overline{NoP(t_i)}$. Namely, it checks if the following relation holds

$$NoP(t_{i+1}) < \overline{NoP}(t_i) + 5 \cdot \sigma(t_i). \tag{4.5}$$

If this condition is satisfied it iterates again the procedure with a smaller threshold. However, if the condition is not satisfied, it marks the $NoP(t_{i+1})$ as an outlier: noise is being binarized in the image. It then stops the search, and assumes t_i as the optimal threshold to be used in the first binarization

$$T_1 = t_i. \tag{4.6}$$

An issue that may be noted right away, by looking at fig.4.5, is that if the NoP values corresponding to the highest values of the threshold are all equal, the first point which differ would be marked as an outlier since the previous standard deviation is null. The algorithm would then stop at the wrong threshold value. This is avoided by limiting the standard deviation to be $\sigma(t_i) \geq 1$ at every iteration. In this way, if the NoP changes slightly ($\sigma < 1$) the algorithm is not triggered at the first outlier: the procedure continues by updating the mean NoP and σ, but keeping the check on the σ lower bound. Only if, at a given step, $\sigma(t_i) > 1$, and the NoP exceeds the mean value by 5σ, the algorithm triggers and the threshold is chosen. In this way, the chance of detecting an outlier which exceeds the 5σ range by random fluctuations is reduced.

The values of the optimal threshold found by the algorithm for the samples reported in fig.4.5 are reported in table 4.1.

Sample	T_1
PG Lab	3
Alsi dust	3
Mo dust	2
Ti dust	3

Table 4.1: Values of the optimal threshold automatically found by the algorithm for different samples. The PG Lab sample is a wafer that was exposed to the Perugia Lab for 15 hours.

Comparing the values reported in tab.4.1 with the NoP values reported in fig.4.5, it can be seen that the algorithm correctly stops when the noise peak is reached. It can also be noted that the

Molibdenum sample accepts a lower threshold compared to the other samples. This can be seen both from the table and from the image. This shows another advantage of this procedure: the algorithm is able to adapt the threshold to each photo individually.

Input & output

- Input: the image after the background removal, see section 4.2.1
- Output: the optimal luminosity threshold, T_1.

4.2.3 Global Binarization with Threshold T_1

This step identifies the particles and separates them from the background.

Given the threshold T_1, the binarization maps the value of each pixel to either 0 (black) or 1 (white) if, respectively, the luminosity level is below or above T_1. The binarization of the image is performed using the `im2bw()` function [54].

In figure 4.6 an example is given: the original picture is shown in the original grayscale (left) along with the result of the first binarization (right). The particles are identified as the white cluster which survives.

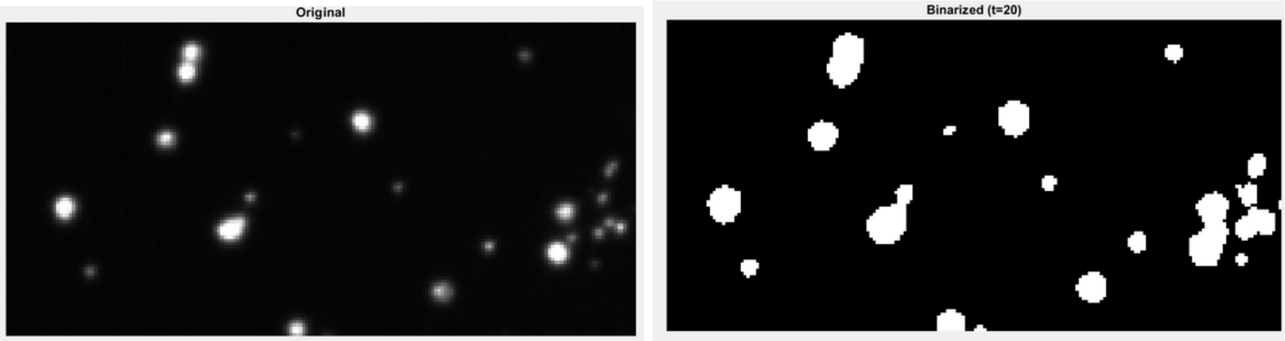

Figure 4.6: The initial image before (left) and after the binarization (right). As an example, a threshold of T=20 is used.

Input & output

- Input:
    1. image from section 4.2.1
    2. threshold for the binarization, given by 4.2.2
- Output: binarized image.

4.2.4 Bright pixel cluster search

Once the image is binarized with the given threshold T_1, a first bright cluster search is performed. This is done with the function `bwboundaries()` [51]. This function looks for white clusters of pixels (the particles) and labels them. The labelling is crucial, since it allows us to iterate over each different cluster found and perform a local luminosity-adapted binarization, which is the most important improvement to the algorithm. Moreover, for each one of the clusters it also returns and stores the coordinates of its perimeter pixels. In this way, we are also able to locate the clusters in the original figure.

Input & output

- Input: binarized image from sec.4.2.3

- Output: list of all the clusters found and their perimeter coordinates

4.2.5 Local binarizarion

Once all the bright clusters are detected and labelled, the algorithm iterates over each one of them. The detailed steps which are performed at each cluster are reported in the following paragraphs.

Sub-image cut

For each cluster found, a new local binarization is performed. For this step, the original image is used (after the background light removal, i.e. the output of section 4.2.1), and not the binarized one. This is done by retrieving on the original (i.e. non-binarized) image the area corresponding to the boundary found on the binarized image from sec.4.2.3. This is done by cutting the rectangle circumscribed to the perimeter of the cluster.

This aroused an issue: the higher the threshold, the more tails of the particles are erased in the binarization, resulting in a smaller particles (as reported in fig.4.1) since increasing the threshold means that pixels more and more luminous are removed from the photo.

To avoid this, the perimeter of the circumscribed box is enlarged in order to restore back the erased tails and compensate for the cut due to the threshold. This enlarging is done automatically by the algorithm. The perimeter is expanded iteratively by 1 pixel in each direction (up/down, left/right and diagonals for the corners), until the luminosity of the successive pixel increases. If, in at least one pixel of the new perimeter, the luminosity is higher than the corresponding previous pixel, the expansion stops. This can happen either because of a nearby particle being encountered, or due to background fluctuations. In both cases, there is no need to expand further, and the enlarging stops. In this way we make sure that in this box there is only the initial cluster found with its tails.

As an example, consider the top-right corner of the boundary box. Let x and y be its coordinates, and $I(x, y)$ is the luminosity of the pixel. Being the corner, the perimeter is expanded in three directions by 1 unit: top, right and top-right diagonal. The "expanded" pixels luminosities are then $I(x, y+1)$, $I(x+1, y)$ and $I(x+1, y+1)$. The check on the decreasing luminosity reads:

$$\begin{cases} I(x,y) > I(x+1,y); \\ I(x,y) > I(x,y+1); \\ I(x,y) > I(x+1,y+1); \end{cases} \tag{4.7}$$

If all the three conditions are satisfied, it means that the the luminosity is decreasing in all the directions from the starting pixel.

The above procedure is performed for each one of the boundary pixels in the same way. The only difference is in the directions taken for the expansion: the expansion is only made in the direction exiting from the boundary box (e.g. for a pixel in the right edge we can only move right, not up-down or to the left). If in all the directions going outside from all the pixels the luminosity is decreasing the expansion is safe: the boundary box is enlarged by 1 pixel and the process is iterated again. When in at least one direction going outside from one pixel the luminosity is found to increase, the process stops and the boundary box is taken as the last one that was computed. Moreover, to save computation time, the enlarging is limited to 5 pixels in each direction.

In figure 4.7, we report pictorially the enlarging of the boundary box procedure.

Second threshold computation

Starting from the selected region given by the previous step, the second local binarization with a new threshold is performed.

In order to allow for a more accurate dimension reconstruction, the second threshold, T_2, is computed by the algorithm according to the maximum intensity of the cluster. Namely, it is taken to be the

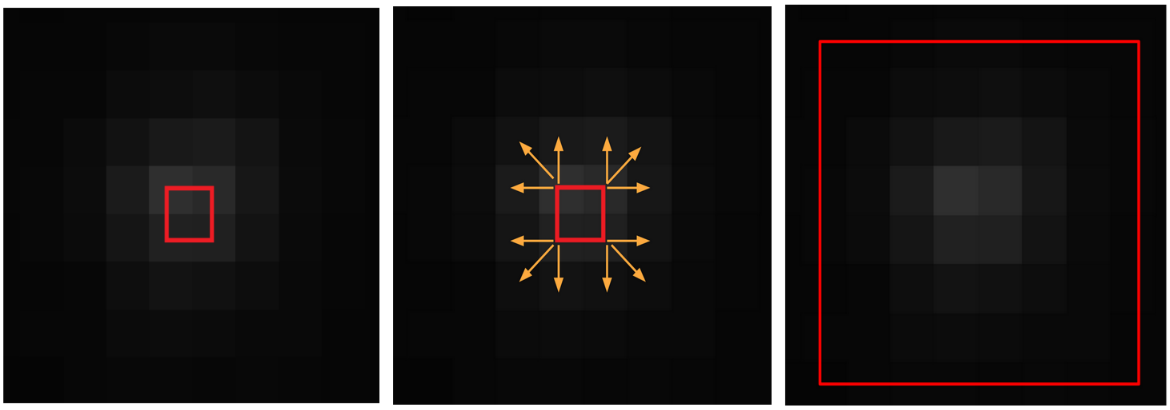

Figure 4.7: The left image is how a particle is initially seen: the red square is the result of what survives the first binarization. As can be seen, due to the fact that the threshold was high relative to the luminosity of the particle, most of the tails are erased. This box is then enlarged (center), until all the tails are restored back (right). The region highlighted in the right figure is the one used in the second analysis.

fraction α of the maximum intensity, I_{max}:

$$T_2 = \alpha * I_{max} \tag{4.8}$$

The optimal α parameter is computed by tuning the algorithm on the known dust samples. The performances of the algorithm are based on how closely the D10, D50 and D90 percentiles are reproduced: i.e. by minimizing the SSR (defined in section 4.1). The tuning of the algorithm is done by running the code on the known dust samples with different hyperparameters (threshold T_1 and α) and an heatmap of the SSR is built. Additionally, the heatmap for the number of particles (NoP) found in the picture is constructed too. This latest one allows us to monitor the eventual noise entering in the picture, which would be read as an anomalous increase in the NoP.

A deeper discussion on the tuning of the algorithm is done in 4.3.2. Meanwhile, I report the final α parameter chosen from the calibration:

$$\alpha = 0.15. \tag{4.9}$$

This value means that the second threshold T_2 is computed as the 15% of the maximal intensity of the particle.

Input & output

- input: the expanded box from section 4.2.5
- output: second threshold, T_2.

Second particle search

Given the luminosity-adapted threshold computed the last step, the local binarization and cluster search steps as in 4.2.3 are performed, but locally on the subfigure given by section 4.2.5.

In figure 4.8 the step of detection and sizing of the particles are visually reported. The background removal and optimal T_1 determination are not reported in the figure.

As can be seen, by comparing the results of the search in the steps (2) and (5) of fig.4.8, in the first search (step (2)) almost all the tails of the particles are classified as being part of the particle, while in the second reconstruction (step (5)) the boundary are closer to the core of the cluster, removing the scattering tails.

Another advantage of such local binarization, is that the algorithm is now able to separate two different particles that would have been merged together from the first binarization. One example of this effect

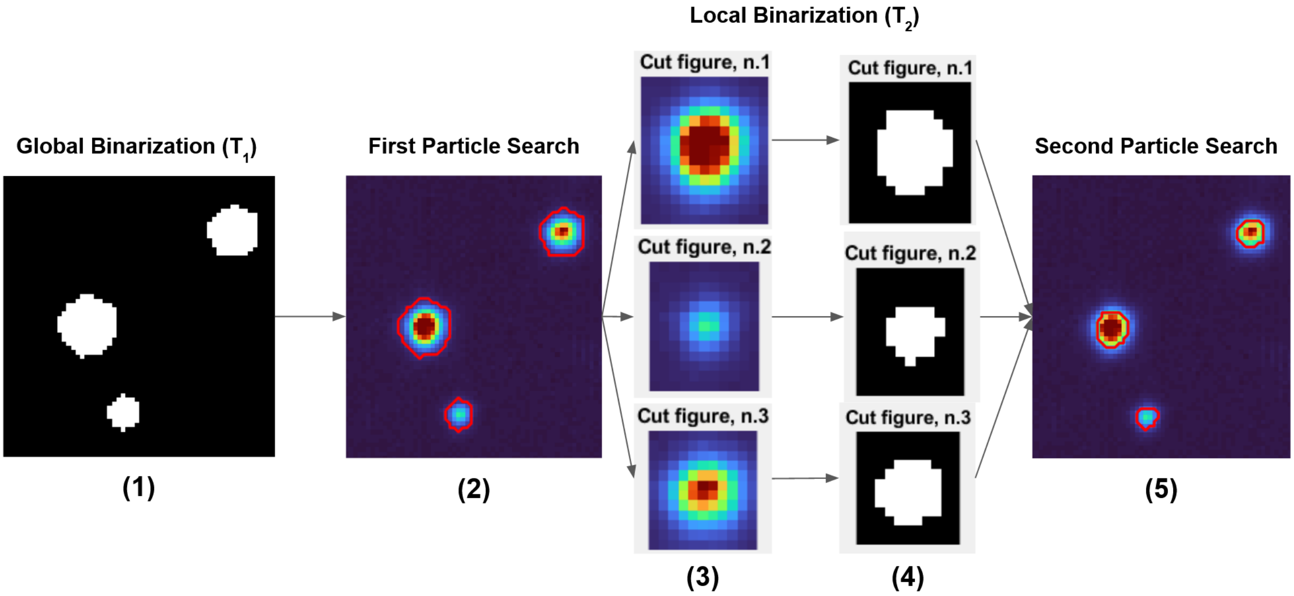

Figure 4.8: Visual representation of the detection and sizing of the particles. In the region interested in the pictures, three particles are present. Step (1): global binarization with the optimal threshold T_1. Step (2): particles detection (on the black-white, BW, image) and their boundaries reported on the original image. Step (3): selection of each cluster found singularly. Step (4): local binarization with the luminosity-adapted T_2. The properties of each particle are then computed. Step (5): the boundary of the single clusters are reported on the original image.

is reported in figure 4.9. The two images report the same two particles with the result (red perimeter) of the global search (left) and the local one (right). As can be seen, after the global binarization and the first particle search, the two particles are fused together as a unique big cluster. This is due to the fact that the pixels between the two particles are all above threshold and the binarization is not able to separate the two particles. Indeed, it can be noted that such connection pixels appear brighter that the rest of the background. The first threshold $T_1 = 3$ is too low in this case, and the separation of the two particles would then require an higher threshold: this is in fact done with the luminosity-adapted local binarization (right image). Thanks to the high luminosity of the particles, $I_{max} = 255$, the second threshold is higher than the first one:

$$T_2 = \alpha \cdot I = 0.15 \cdot 255 = 38.$$

Because of this higher threshold, the pixels in between are now below T_2 and the second binarization correctly separates the cluster in two. The two separated particles are individually reconstructed by the algorithm, as can be seen from the right image.

Regarding the example discussed above, it should be noted that the value I=255 is the maximum luminosity value available, and indicates that the particle saturated in luminosity. In our photos, there were cases of saturated particles, and they can also be seen in figure 4.8 and 4.9. However, in our study we have seen that the saturation of the particles do not cause any issue. In fact, a set of equal photos was taken for different exposure values, obtaining different saturations conditions. The algorithm was not affected by those differences and the results were comparable. For this reason we have kept 40ms of exposure time, as from the previous calibration.

This feature allows for a more accurate estimation of the number and size of the particles, since otherwise the two smaller particles are read as a unique big particle.

Input & output

- input:

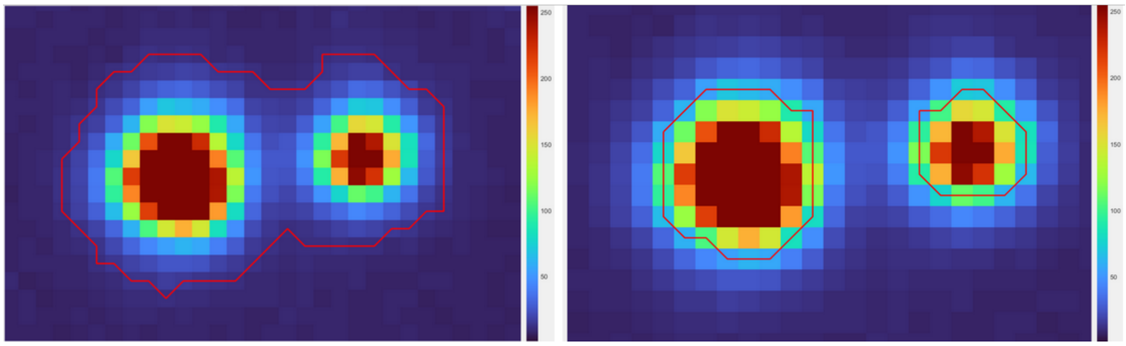

Figure 4.9: Comparison between the same two particles seen from the first global binarization and particle search and the second local luminosity-adapted binarization and search. The first binarization is performed with the optimally found T_1 threshold, while for the second search an $\alpha = 0.15$ parameter is used.

1. expanded box and figure from section 4.2.5
2. second threshold T_2, computed as 4.8

- output: the final clusters.

4.2.6 Particles properties

Once the local binarization is performed for one particles, the dimensions of the particle (area and perimeter) are computed with the `regionprops()` function [52]. Both area ($A/1\text{px}^2$) and perimeter ($P/1\text{px}$) are initially computed in pixels. To convert the area and perimeter in microns, the pixel-to-meter ratio η (eq.3.1) is used

$$\left(\frac{A}{1\mu\text{m}^2}\right) = \left(\frac{A}{1\text{px}^2}\right)\frac{10^{12}}{\eta^2}, \tag{4.10}$$

$$\left(\frac{P}{1\mu\text{m}}\right) = \left(\frac{P}{1\text{px}}\right)\frac{10^{6}}{\eta}. \tag{4.11}$$

The effective diameter, d, is computed by assuming that in the image the particles are circles, namely:

$$\left(\frac{d}{1\mu\text{m}}\right) = 2\sqrt{\frac{(A/1\mu\text{m}^2)}{\pi}}. \tag{4.12}$$

Finally, the positions and maximal luminosity of the particle are stored too. For the location of the particles the coordinates of the geometrical center of the particle are returned: h_c and w_c. The geometrical center is computed as the mean between the h and w extrema of the boundary. h refers to the height (y-direction) and w to the width (x-direction) of the pixel in the image, with the origin of the coordinate system in the upper-left corner of the image [55].

All the observables are stored in a table and then used for the analysis.

Input & output

- input: particles binarized clusters after the local binarization (fig.4.8, point (4))
- output: table with the observables of all the particles (label, dimension, position, maximal luminosity)

4.3 Algorithm tuning

4.3.1 Stacks

The stacking of the photos, described in sec.3.2.3, had to be optimized both on the number of photos used, N, and the method used for assigning a luminosity to each pixel (i.e. the average luminosity or the minimum). In order to have a better background rejection, we have tested how increasing N or changing the method would affect the performances.

For the purpose, a set of photos of calibrated dusts and a set of photos of the wafer exposed to the EGO "Perugia" Clean Room were taken. For each sample, 150 nominally identical images were taken, so that we could test the stacking with a maximum of N=150 photos. To assess differences due to the choice of N, we have tested how N affected the first automatic threshold (T_1) choice and eventually the NoP found in the picture. The first threshold T_1 chosen by the algorithm for the average and minimum stacks are reported in table 4.2. The NoP as a function of N for the different samples is reported in 4.10.

Average				
N	CR Wafer	AlSi	Mo	Ti
2	4	5	4	4
5	4	4	3	3
10	3	3	2	4
20	3	3	2	3
30	3	3	2	3
40	3	3	2	3
50	3	3	2	3
100	3	3	2	3
150	3	3	2	3

Minimum				
N	CR Wafer	AlSi	Mo	Ti
2	5	5	4	5
5	4	5	3	4
10	4	4	3	4
20	4	4	3	4
30	4	4	3	4
40	3	4	3	11
50	3	4	3	11
100	3	4	3	4
150	3	4	3	4

Table 4.2: The first threshold T_1 automatically found by the algorithm for two different stacking methods: average (left table) and minimum (left table). The second to last columns refers each to a different sample.

As can be seen, the threshold T_1 stabilizes for all the samples after $N_{avg} = 20$ and $N_{min} = 100$ for the average and minimum respectively. Regarding the NoP, except for very low number of images, no significant fluctuations are found when using the average, and minor ones are found when using the minimum.

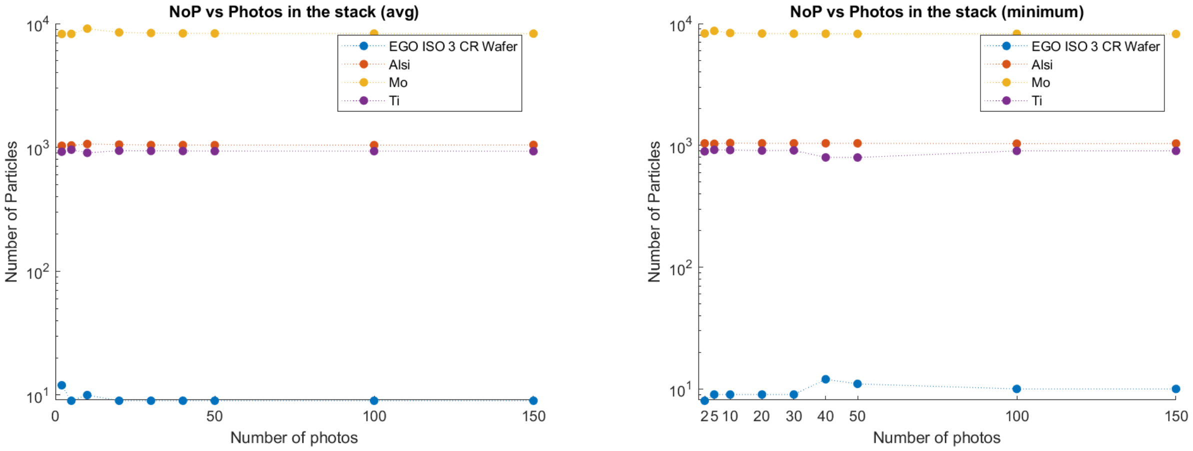

Figure 4.10: NoP as a function of N with the average (left) and minimum (right) stacking method.

In principle, the idea behind taking the minimum is that only the pixels corresponding to a particles would scatter light directly and with an approximately constant intensity, while the background randomly fluctuates. By taking the minimum we then assign to each pixel fluctuation their minimum value, while the particles pixel would get the approximate value of the incoming intensity over the different photos. However, when working with a large number of photos, the minimum could be negatively affected by external factors, such as a dead pixel in the camera, drop or changes in the illumination due to human activity etc... For this reason, the average is chosen as stacking method:

this provides equally good performances and a more robust statistical basis.

Regarding the number of photos in the stacks, ideally the maximal possible number of photos (which can also be greater than 150) for the stack would be to best choice. However, since from tab.4.2 and fig.4.10 no significant increment in performances is seen above N=20, we have chosen this value as the one giving us the best performances with the minimum amount of computational resources and time.

For the above reasons, all the working procedures are performed by using 20 stacked photos, where at each pixel the average value across the stack is associated.

4.3.2 Dimension reconstruction

To optimize the dimension reconstruction a tuning of the α parameter is required. This is done through the heatmaps build by running the algorithm for different T_1 and α, as mentioned in section 4.2.5. Although the final code chooses the threshold T_1 automatically, in the tuning of α it was set to cover a range of values around the optimally found T_1. This is done to test if the dimension reconstruction is indeed independent from the first threshold (which is only responsible for the detection of the particles). The threshold intervals are set in order to have the optimal T_1 as second value of the threshold (i.e. the second row of the heatmaps corresponds to the optimally found T_1).

Since the only samples for which a reference on the dimensions is given are the dust samples, for the tuning of α the three samples AlSi, Ti and Mo are used. For each sample the SSR computed as in eq.4.2 and the NoP are reported in the heatmaps, which can be seen in figure 4.11 and 4.12. In this way we can both track how well the dimensions are reconstructed with the SSR and how the NoP behaves to see whether noise is entering in the image. The heatmaps cover the range $\alpha = [0.01,\ 0.90]$, with a $\Delta\alpha = 0.01$ increment.

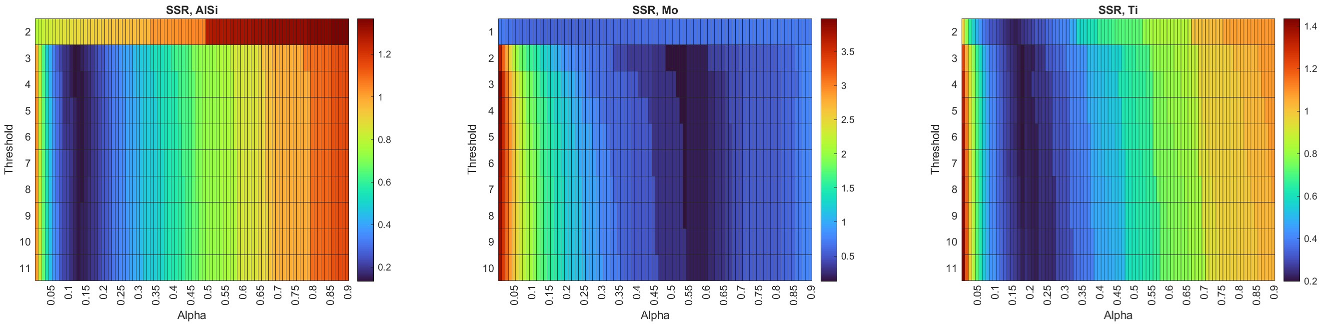

Figure 4.11: Heatmaps for the three different dust samples: AlSi (left), Mo (center) and Ti (right). α runs over the horizontal axis, T_1 in the vertical axis. The SSR is shown as colorcoded, the color legend is reported at the right of the map. An higher SSR value means an higher total relative error on the percentiles. The SSR values in the vertical scale are decimal values. For example, a vertical value of 0.2 in the heatmap corresponds to SSR = 0.2 = 20%. Similarly, a value greater than 1, such as 2 corresponds to SSR = 2 = 200%.

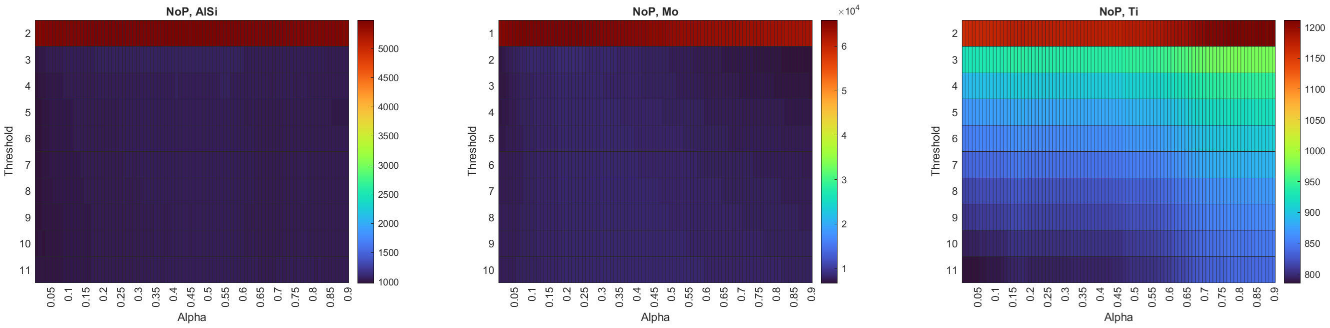

Figure 4.12: Heatmaps for the three different dust samples: AlSi (left), Mo (center) and Ti (right). α runs over the horizontal axis, T_1 in the vertical axis. The colorcode represents the NoP found in the picture.

From the reported heatmaps, a series of considerations can be made. Firstly, a significantly higher NoP can be seen in the first row of the three samples. Since the optimal T_1 value (which is chosen by the algorithm as the lowest threshold value that removes noise) is set to be always the second row, it means that the first row of the heatmaps corresponds to a threshold value in which noise is being binarized in the picture. Indeed, the high NoP is a consequence of random background fluctuations being counted as fake particles. In terms of dimensions, noise binarization can be seen as an increased SSR, especially in the AlSi sample. In the Mo and Ti samples the noise do not highly affect the dimension reconstruction, but nevertheless the minimal SSR configuration can always be found from the second row and below.

Secondly, it can be noted that the total error on the sizing of the particles is now largely independent from T_1, as expected. This is due to the recovering of the tails performed with the expansions of the particles boundary box described in section 4.2.5. Likewise, the NoP is mostly independent from α.

The main effect that α has in the count of the particles is the separation of nearby particles that would have been clustered together by the first binarization, as shown in 4.9. An higher α value would better separate nearby particles, leading to an increasing NoP. Indeed, an increasing NoP for increasing α values can be seen in the right map in figure 4.12 (Titanium sample). The Ti sample is more affected by this effect compared to AlSi and Mo. In the Ti sample, from $\alpha = 0.1$ to $\alpha = 0.9$ there is a relative increment in the NoP of

$$\frac{\Delta NoP}{NoP} = \frac{NoP(\alpha = 0.9) - NoP(\alpha = 0.1)}{NoP(\alpha = 0.1)} = 5\%.$$

However, although particles are better separated, the particles sizing is heavily affect: the SSR at $\alpha = 0.9$ is SSR=100%. This is due to the fact that with an higher α value the particle size is underestimated. A side effect of increasing α is that the increasing of T_2 leads the possibility that, in the case of two nearby particles with a wide luminosity gap, one of the two may not survive the second binarization. This is present for the highest values of α, and results in a decreasing of the NoP. For a clearer visual representation, the NoP as a function of α is plotted for the three dust samples and their optimally found T_1. The plots are reported in figure 4.13.

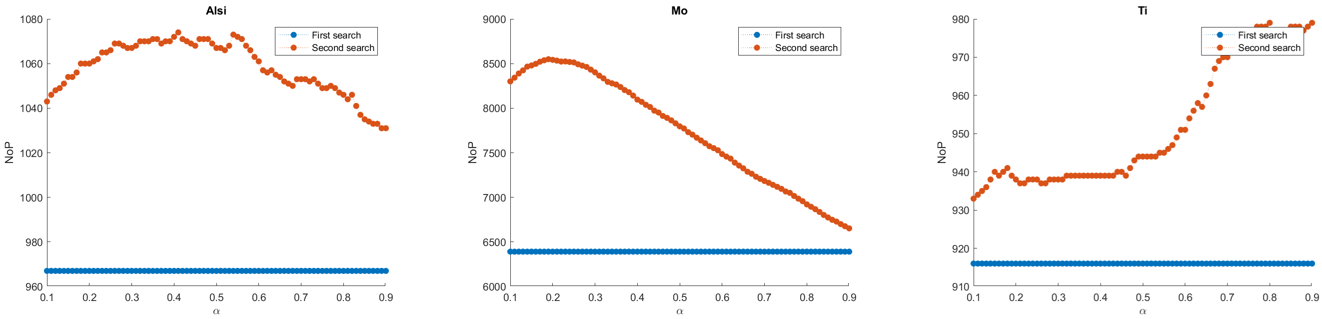

Figure 4.13: NoP as a function of α for the three dust samples: AlSi (left), Mo (center) and Ti (right). The threshold T_1 used are respectively $T_1^{AlSi} = 3$, $T_1^{Mo} = 2$ and $T_1^{Ti} = 3$. In the plots, are reported the NoP found from the first (global) search (blu curve) and from the second (local) search. Since the first binarization and search is independent from α, the corresponding NoP is constant, while the NoP from the second search is the only quantity which changes.

In the AlSi and Mo cases, both contributions can be seen: in the left part of the plots the NoP increases (due to a better particles separation), while in the right part, for higher α values, the NoP starts to decrease (due to the too high threshold value, which deletes the fainter particles nearby more luminous ones). In the Ti case, the first effect is predominant on the latter, and the NoP is almost always increasing. Also to be noted that the first search finds considerably less particles than the second one. This seems to indicate that the "fusion" of nearby particles is quite relevant in those samples, while we don't expect this to be true for the much cleaner samples we expect as witness of the Virgo clean environments.

Eventually, since a trade-off between the NoP detected and the dimensions reconstructed has to be found, for the optimization of α the SSR is chosen. In fact, the NoP dependence on α is significantly weaker that SSR dependence, as shown by the heatmaps (fig.4.11 and fig.4.12). In addition, the α effect on the separation of the particles originates from the large number of particles in the AlSi, Ti and Mo samples ($NoP \approx 10^3$). Having a large number of particles, it is more likely that some of them are close enough to be clustered together and, as reported above, we expect to have lower particles densities in clean room samples.

α estimation

From the heatmaps and using the T_1 optimally found by the algorithm, the α value corresponding to the minima SSRs is computed. The three different values are

$$\alpha_{AlSi} = 0.12 \qquad \alpha_{Mo} = 0.52 \qquad \alpha_{Ti} = 0.18, \tag{4.13}$$

which corresponds to an SSR value of

$$\text{SSR}_{AlSi} = 13\% \qquad \text{SSR}_{Mo} = 13\% \qquad \text{SSR}_{Ti} = 20\%. \tag{4.14}$$

Computing an averaged value between all the three samples, $\alpha = 0.27$ would result. However, this value results in a worsening of the SSR by $\Delta SSR_{alsi} = +27\%$, $\Delta SSR_{Mo} = +53\%$ and $\Delta SSR_{Ti} = +14\%$. Since the Mo seems to behave differently from the other two (as will be discussed in depth in the next section), this sample was discarded from the average. Finally, as a final estimation of the optimal α parameter, the value which optimizes the dimension reconstruction in AlSi and Ti is:

$$\overline{\alpha} = 0.15, \tag{4.15}$$

which corresponds to an SSR of

$$\text{SSR}_{AlSi} = 17\% \qquad \text{SSR}_{Ti} = 25\%. \tag{4.16}$$

The SSR for the Mo sample with $\alpha = 0.15$ is $\text{SSR}_{Mo} = 125\%$.

Mo discrepancy

From the heatmaps in figure 4.11, it can be seen that while AlSi and Ti samples behaves similarly in terms of the SSR dependence from α, the Mo sample requires a significantly higher α value in order to minimize the error on the diameters percentiles. By looking at the diameter distributions reconstructed by the algorithm (figure 4.14) with $\alpha = 0.15$, Mo particles dimensions would appear to be bigger than the provided ones, since all the percentiles are on average 70% larger than what we expected from the data provided by the vendor [46].

A series of factors could explain this behaviour:

1. particles merging together: the photographed Mo sample is the one with the highest number of particles ($NoP \approx 8000$) compared to AlSi ($NoP \approx 1000$) and Ti ($NoP \approx 900$). Due to the higher particles density, particles are on average closer to each other and separating different particles might require a larger value of α (and eventually an higher T_2);
2. different scattering properties: different production processes or intrinsic material properties of Mo particles, compared to AlSi and Ti, may cause different optical properties, in particular in light scattering. A different intensity of scattered light leads to particles to appear differently in the image. E.g. if Mo particles are more reflective and scatter more light they would appear brighter and bigger in the image (due to the bigger light tails), requiring an higher α value to determine small dimensions.

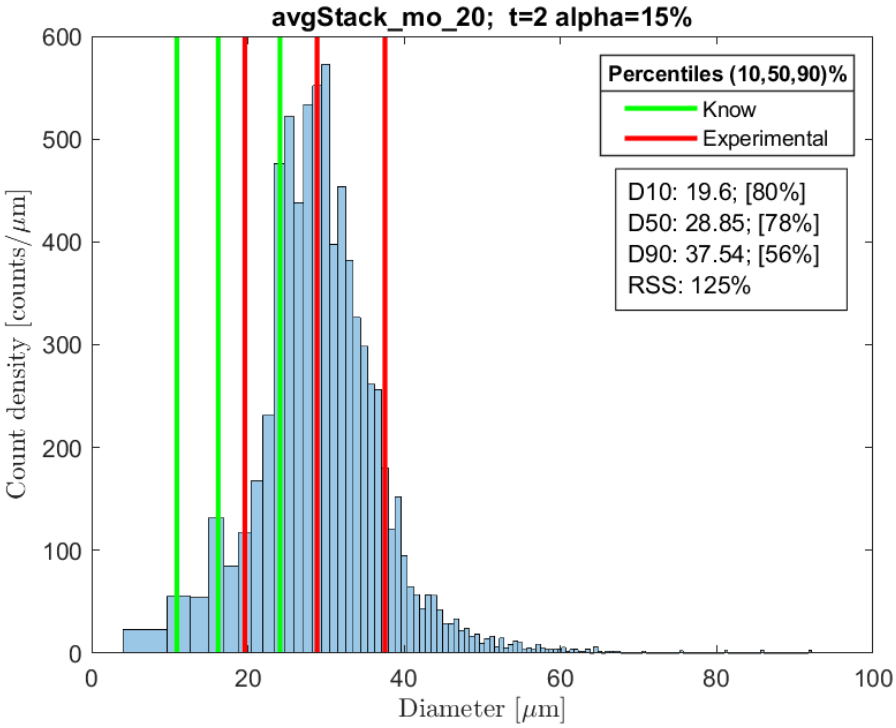

Figure 4.14: Distribution of the particles diameters for the Mo sample. The non-uniform binning is due to quantization of the area in units of $A = 1\text{px}^2$ and the fact that the diameter is computed following the circular-particle approximation as in eq.4.12. Indeed, following the quantization of the area, it can be seen that the possible diameter values are not uniformly spaced. For this reason, the histogram reports the count density: the normalization of the counts by the width of each bin. The vertical lines are the known (green) and reconstructed (red) 10th, 50th and 90th percentiles. The threshold value is the optimal threshold found by the algorithm ($T_1 = 2$). α is given by the average of the α values minimizing the SSR of AlSi and Ti (fig.4.11). As can be seen, the particles are reconstructed bigger than they should be and an higher value of α is needed for a better particle sizing.

To test whether the merging of particles in the pictures could be the cause of the discrepancy, the result of the particles search is displayed and as an example, a zoom of the Mo sample image is reported in figure 4.15. The particle search is performed with $T_1 = 2$ and $\alpha = 0.15$, which is the average between the α minimizing the SSR in AlSi ($\alpha_{AlSi} = 0.12$) and Ti ($\alpha_{Ti} = 0.18$). In particular, a region in which the particles density is particularly high (i.e. particles are closer together) is reported. The red boundaries are the perimeters of the particles reconstructed by the algorithm (after both the global and local search). As highlighted by the arrows, some cases of particles being merged together are present. Others, which are not highlighted, can be seen too. However, the limited amount of merging cases does not seem to explain the large residual discrepancy.

To further investigate if the overdensity can be the cause of the discrepancy, a new set of photos in which fewer particles were present was taken. In the new picture, less particles were found ($N \approx 1200$) and the heatmaps run over this new picture led to the same result: α compatible with was previously found and higher than α_{AlSi} and α_{Ti}. The heatmaps on the less populated Mo sample are reported in figure 4.16.

Since the same SSR trend is seen in both the high and low populated samples, we conclude that the merging of particles together due to the higher particles density is not the cause of the higher α value. Moreover, by looking at the reconstructed perimeters in figure 4.15, the particles do not seem to be found $60\% - 80\%$ larger than the real size, as would be suggested by the errors on the percentiles in the histogram 4.14. This would suggest that the discrepancy may be due to different reflecting properties of the Mo particles, that may change how they appear in the images. If, for examples, Mo

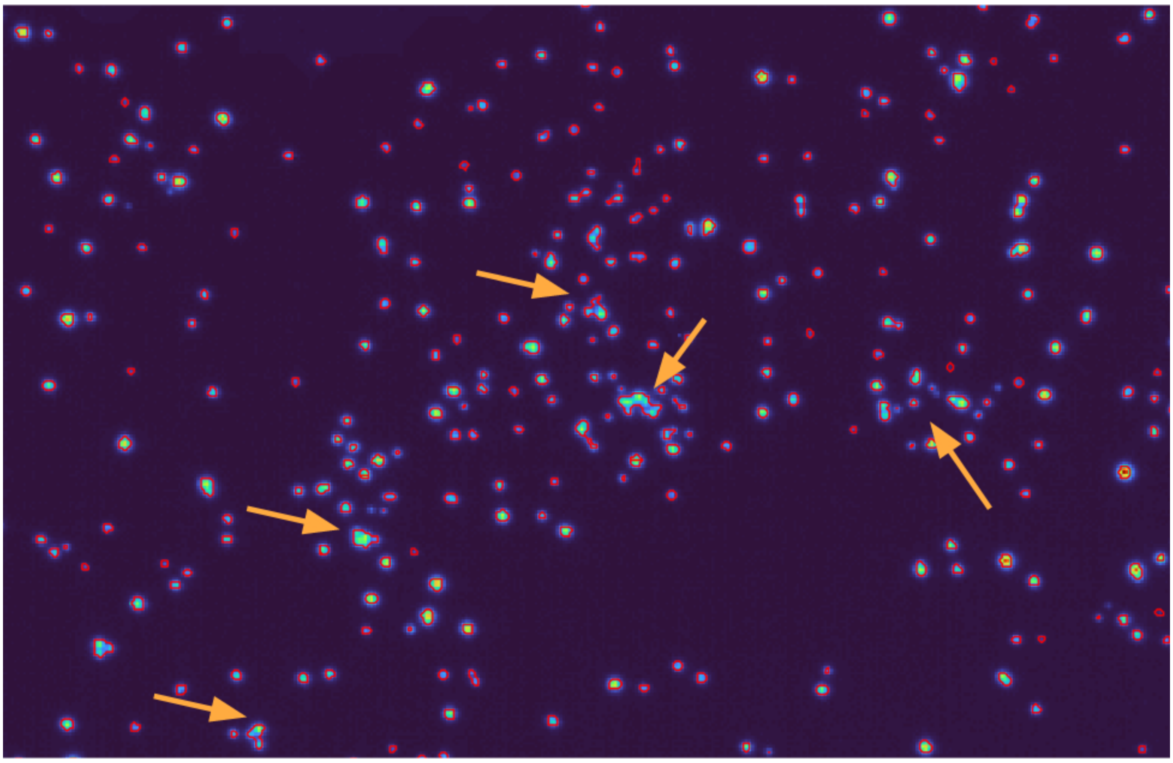

Figure 4.15: Zoomed-in view of the Mo sample. In red the reconstructed perimeters of the particles are reported. As can be seen, in some cases nearby particles are identified as a single particle. This lead to an excess of bigger particle, while removing smaller one.

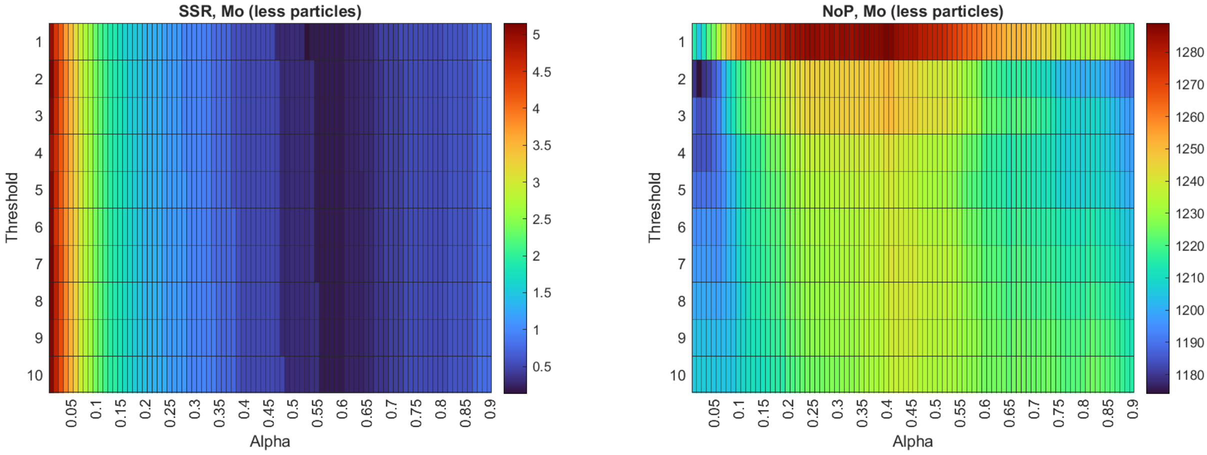

Figure 4.16: Heatmaps for the SSR (left) and NoP (right) for the new Mo sample, containing $N \approx 1200$ particles. The same trend of the previous sample is found. This shows that the α discrepancy is not due to particles merging.

particles appear bigger in the images, this would result in an higher optimal α value for this sample.

However, with out apparatus we were not able to test different scattering properties and we could not further test if the different metallic dusts scatter differently. Another possible explanation could be that the particles deposited on the wafer may not have the distribution declared by the vendor, maybe because of the way they are shipped, handled or deposited. To test this further, in the future those samples will be analysed with a microscope with an higher resolution compared to our actual camera system.

For those reasons, and the deeper analysis presented in the following section, we have discarded the Mo sample in the estimation of α.

4.4 Performances

From the optimization of the original version of the code [21], a threshold of $T = 80$ was found to be the optimal trade-off between particle detection and sizing. In this improved version of the algorithm, we have separated the tasks of detection and sizing of the particles. The detection is done globally with a relatively low threshold: this allows us to detect the majority, if not all, of the particles. The sizing is instead performed locally, with a threshold T_2 which depends on the luminosity of each cluster.

Regarding the accuracy in particle detection, in table 4.3 the comparison between the old version of the code and the new one is reported. For the three calibrated dust compounds I report the threshold used (T_1) and the NoP found in the picture. In the case of the old algorithm, the threshold was set to $T_1 = 80$ as the optimal value found by the previous tuning. Instead, in the new algorithm the threshold T_1 is the one which is optimally found by the algorithm automatically.

Sample	Old		New		Ratio
	T_1	NoP	T_1	NoP	
Alsi	80	970	3	1054	1.09
Mo	80	4484	2	8479	1.89
Ti	80	573	3	940	1.64

Table 4.3: Comparison between the detected NoP with the original version of the code with the improved one. As can be seen, in all the three samples more particles are found. The ratio between the NoPs, $NoP(new)/NoP(old)$ is also computed and reported in the last column.

Thanks to its adaptive threshold which is able to adapt to the single sample, the improved algorithm performs better in the detection of the particle. In particular, in the Mo and Ti samples, almost two times more particles than before can be found. A visual comparison between the original version of the algorithm and the improved one is displayed in figure 4.17. The image is a zoomed portion of a Ti dust sample. The two algorithms use a different global threshold: $T = 80$ the orginal algorithm, while $T_1 = 3$ for the improved algorithm (see Sec.4.2.2). The particles highlighted in red are the particles which are detected by the algorithm. The figure shows clearly that the new, improved algorithm allows to detect a larger number or real particles, if not all of them.

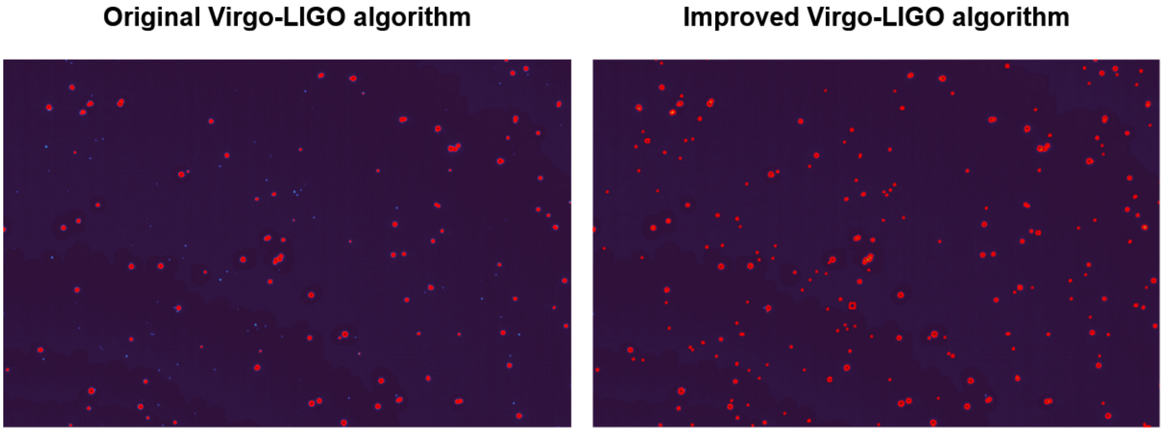

Figure 4.17: Visual comparison between the detection efficiency of the original algorithm (left) and the improved algorithm (right). Thanks to the lower threshold of the new algorithm ($T_1 = 3$, instead of $T_1 = 80$), more small particles can be correctly detected. The improved algorithm search is performed using $\alpha = 0.15$.

For what concerns the dimensions, the count density histograms of AlSi, Ti and Mo are reported in figure 4.18. These distributions are reconstructed using T_1 automatically found by the algorithm and $\alpha = 0.15$ from the tuning in the previous section. In general, a good agreement between the known and reconstructed percentiles is found. The only sample which shows an higher error, compared to the other two samples, is the Mo, for the reasons discussed in section 4.3.2.

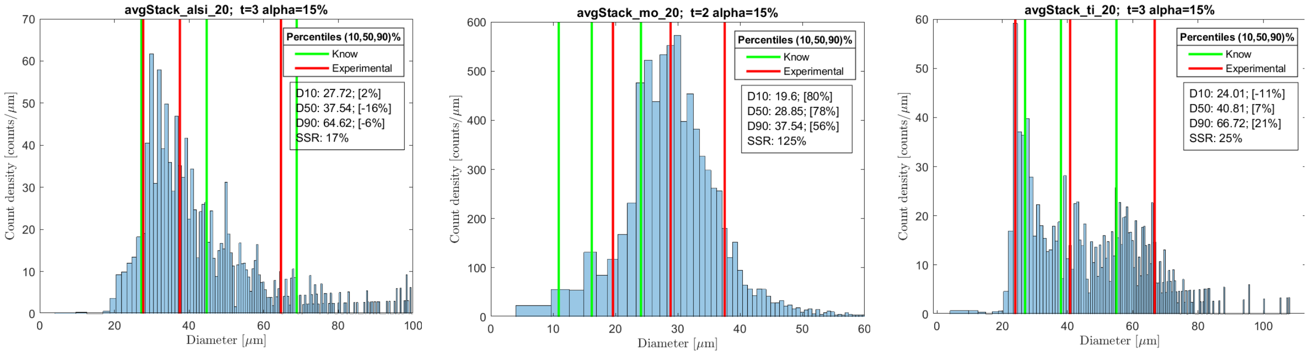

Figure 4.18: Count density histograms of the AlSi (left), Mo (center) and Ti (right) dust compounds. The reconstructed (red) and known percentiles (green) are reported too. In the box, the numerical value of the experimental percentiles are reported. In square brackets the percentiles residuals are also reported.

Chapter 5

Analysis of the Samples

After having described how the monitoring campaign is performed in Chapter 3 and the image processing algorithm that we have used in Chapter 4, in this chapter the results of the analysis of the exposed samples will be presented.

5.1 Anomalous Dust Particles Detected

When photographing the samples, a standard imaging procedure was followed, which is described in Sec.3.1. However, it may happen that anomalous structures were found in the images. With the term *anomalous*, we refer to objects which are clearly not caused by dust particles belonging to the same distribution and may be caused by accidental factors spoiling the sample. From an eye inspection of all the photos in the database, those anomalies come at least in three cases:

- Finger print: this anomaly is caused by a finger mark on the sample. An example of a finger print anomaly is reported in Fig.5.1. The finger mark can be seen as the clear finger-shaped cluster of particles in the bottom-right part of the image. In this particular sample, the finger mark in this position was also detected and recorded by an eye inspection performed on the sample. This anomaly is found in 5% of the photos (21% of wafers).
- Fiber: this anomaly in the picture can be seen as the long trail of supposed detected particles. An example is reported in Fig.5.2: the trail of particles is present in the upper-center part of the photo. Since one may expect a random deposit of particles over a give area, it is unlikely that particles line in as in this photo by pure change. Instead, we suppose this kind of events may be due to fibers depositing on the wafer and being read as a trail of particles or because of particles falling of during activities carried out over the wafer. This anomaly is found in 10% of the photos (25% of the wafers).
- Cluster: this anomaly in the picture can be seen as those big and brighter clusters that are clearly not reconstructed as a whole by the algorithm, an example is reported in fig.5.3. Instead, only the brighter parts are detected as non-background, resulting in a much higher particles count. We suppose this type of anomaly is caused by a bigger particles that, due to luminosity changes, is not reconstructed as a whole by the algorithm. This anomaly is found in 4% of the photos (13% of the wafers).

In some cases (e.g. in the case of the fingermark shown in figure 5.1), those anomalies were seen with the naked eye on the wafer, while in general this was not possible and they were only seen in the subsequent image processing (as in the case of the cluster reported in figure 5.3, whose size is $O(10\mu\text{m})$). In general, however, it is unlikely that they were caused by natural falling of dust particulate, which is expected to be uniformly distributed over the photographed area. For this reason, it was decided to discard such photos from the analysis. To do so, before every analysis, each photo (after the stacking procedure) was inspected by eye in search of such anomalies. However, since we take photos of different (disjoint) regions of the wafers (see sec.3.2.1), if one anomaly was found in a particular region of a

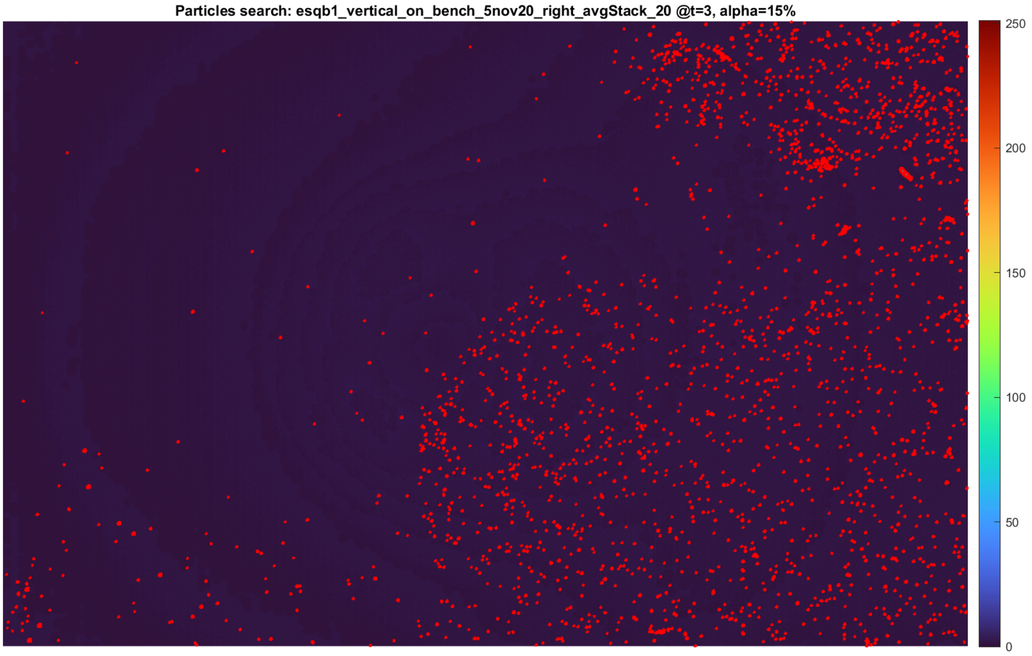

Figure 5.1: Anomaly detected in the samples caused by a finger mark on the sample. The image dimensions are 2.7cm × 1.7cm.

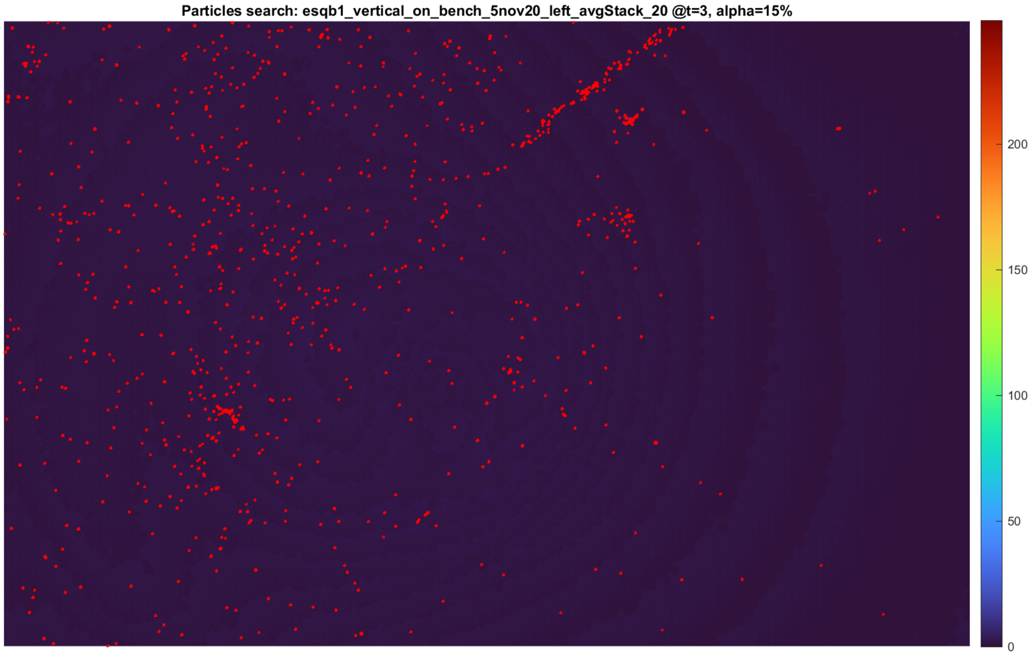

Figure 5.2: Fiber type of anomaly detected in the samples. The image dimensions are 2.7cm × 1.7cm.

sample, that region is discarded from the analysis and only the remaining regions were used. In the following, the NoP is expressed as an average over the different regions available (with no anomalies), and the error associated as standard deviation.

For the reasons discussed in this section, they were discarded from our quantitative analysis, but it should be noted, however, that if we have detected the aforementioned anomalies in our sample, it

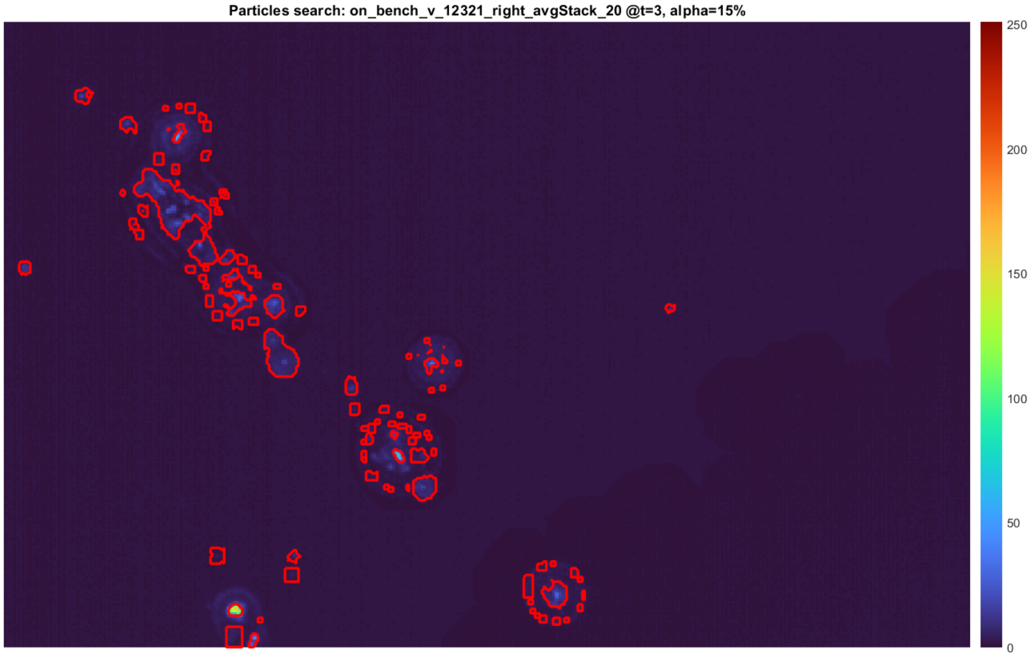

Figure 5.3: Cluster type of anomaly detected in the samples. The image is 3.0mm × 1.9mm.

may be expected that they are also present on the optics. In future improvements, the presence and impact of this features on the performance of our system may be studied further.

5.2 Wafer Orientation

Comparing the dust distributions between the Horizontally and Vertically oriented wavers (hereafter they will be referred only with H or V) we will test any possible discrepancies in the dust accumulation between horizontal and vertical surfaces, with the latter being more representative of the actual lenses and optical elements mounted on the bench. In fact, since our optics are in general V oriented we are more interested in modelling the dust distribution deposited on V surfaces. However, it is not always possible to place a V oriented wafer on an optical bench: in those cases we have to use H oriented wafers, which in addition also collect more statistics. From the comparison of V and H wafers, we can then translate the results from H to V oriented surfaces.

From the whole database of wafer exposed, there were four couples of wafer which had the same location and exposure time, but differ only by orientation: one wafer was exposed horizontally, while the other was exposed vertically. In photo 5.4 one of the V/H couple of wafers is shown. In literature, an estimation of the ratio between the number of particles (NoP) deposited on a V surface compared to a H surface is found to be [34]

$$h \equiv \frac{\mathrm{NoP}_V}{\mathrm{NoP}_H} = 10\%, \tag{5.1}$$

where NoP_H and NoP_V are the number of particles deposited on a V and H surface respectively.

The four couples of V and H wafers are all exposed in the EQB1 bench, but in different locations and different exposure times:

- 1 minute, both wafers were removed from their boxes and have been left exposed (H and V) from 1' in order to monitor the contamination due to handling [56]. They were left exposed in the middle of the bench.
- 34 days, exposed near the Squeezing (SQZ) box on EQB1 [56]

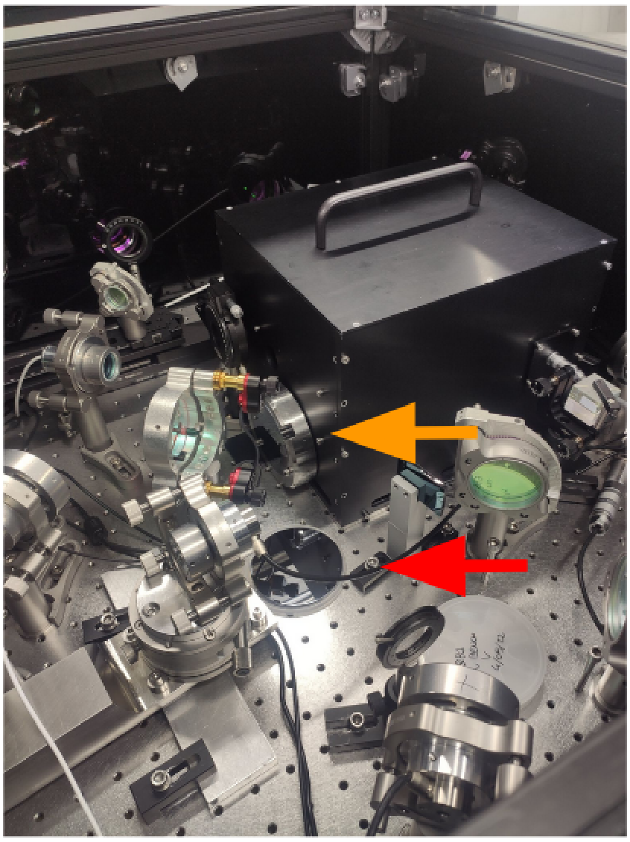

Figure 5.4: A couple of wafers differing only by orientation. The wafers are highlighted by two arrows: the red one referring to the H oriented wafer and the orange arrow to the V wafer. They are positioned in the EQB1 bench (DET lab), near the homodyne.

- 89 days, exposed near the SQZ box on EQB1 [57]
- 418 days, exposed on the EQB1 bench, near the box of the electronics of the homodyne detection. [58]

The number of particles (NoP) detected on the samples are reported in table 5.1, while the reconstructed dust particles distributions are reported in figure 5.5.

Among the four different pairs of wafers, only two of them are truly representative of a real dust deposit. One of the defective pairs is the 1' exposed pair, while the other pair is the pair left exposed for 34 days. The 1' exposed pair is anomalous due to the very short time of exposure. In fact, those wafers were used to monitor the dust deposit due to handling of the sample, hence they may not reflect the real dust fallout on optics. The second couple, the 34 days exposed wafers, shows anomalous particles count. With reference to the data reported in table 5.1, in 34 days the H wafer collected

$$\mathrm{NoP^H}_{34d} \approx 700\mathrm{prt/A_{photo}},$$

over a photographed area $A_{photo} = 4.8 \cdot 10^{-4}\mathrm{m}^2$. Without any external factor altering the dust deposit on the wafer, and if the deposit is only due to environmental dust fall out, one should expect the amount of dust deposit to be at least increasing with time. Instead, we observe a smaller NoP on the 89 days exposed H wafer

$$\mathrm{NoP^H}_{89d} \approx 270\mathrm{prt/A_{photo}},$$

and the approximately the same order of magnitude NoP for the 418 days exposed wafer

$$\mathrm{NoP^H}_{418d} \approx 1000\mathrm{prt/A_{photo}},$$

although they are left exposed for 3 and 10 times longer time, approximately. In fact, the ratio between

Room	Position	Exposure [days]	Orientation	NoP/A_{photo}	h
EQB1	center	0 (1')	H	6 ± 3	$(100 \pm 70)\%$
EQB1	center	0 (1')	V	6 ± 3	
EQB1	near SQZ	34	H	700 ± 50	$(3 \pm 3)\%$
EQB1	near SQZ	34	V	20 ± 20	
EQB1	near SQZ	89	H	270 ± 20	$(30 \pm 20)\%$
EQB1	near SQZ	89	V	90 ± 70	
EQB1	near homodyne	418	H	1000 ± 170	$(36 \pm 7)\%$
EQB1	near homodyne	418	V	370 ± 20	

Table 5.1: NoP of all the wafers constituting the four different H/V couples. In the last columns, we also reported the h ratio of the NoP between the V and H oriented wafers, as defined in 5.1. The NoP reported is the average NoP among the regions of the wafer used in the analysis, and is normalized over an area $A = 4.8 \cdot 10^{-4}\text{m}^2$, which is the area captured in a photo (as reported in Sec.5.1). The associated error is calculated as the standard deviation of the NoPs in the aforementioned regions.

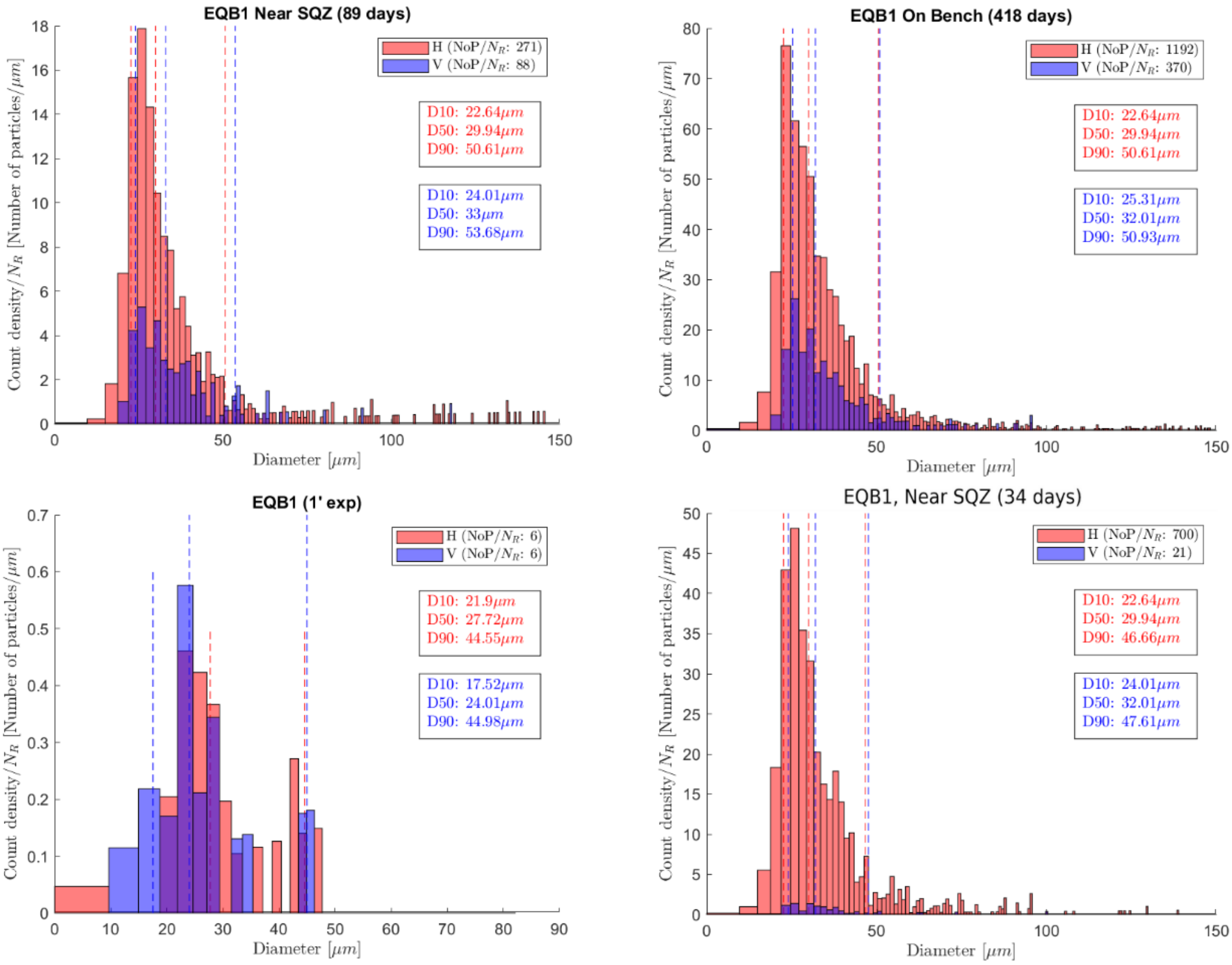

Figure 5.5: Diameter distributions of the dust particles detected on the four couples of V/H wafers, each couple is reported in a single panel. In the titles it is reported the place and time exposure of each couple. The red and blue histograms refer to the H and V wafers respectively. In each panel, the 10th, 50th and 90th percentiles of the two different distributions are reported (labelled D10, D50, D90, respectively). The percentiles are also shown as dashed vertical lines. The histograms are compute by averaging together the counts of particles in the different regions used in the analysis (N_R). The counts are normalized over the photographed area A_{photo}.

the NoP and the exposure time (neglecting the surface area for clarity) are:

$$\left(\frac{\mathrm{NoP}^{\mathrm{H}}_{34\mathrm{d}}}{34\text{ days}}\right) \sim 21\text{ prt/day},\qquad \left(\frac{\mathrm{NoP}^{\mathrm{H}}_{89\mathrm{d}}}{89\text{ days}}\right) \sim 3\text{ prt/day},\qquad \left(\frac{\mathrm{NoP}^{\mathrm{H}}_{418\mathrm{d}}}{418\text{ days}}\right) \sim 2.4\text{ prt/day}.$$

The last two values are compatible (89 and 418 days), while the ratio of the H wafer exposed for 34 days is an order of magnitude larger. However, considering the V wafer of the same couples, one can see that the NoP scales in the same fashion as the exposure time in all the three samples (neglecting the surface area for clarity):

$$\left(\frac{\mathrm{NoP}^{\mathrm{V}}_{34\mathrm{d}}}{34\text{ days}}\right) \sim 0.6\text{ prt/day},\qquad \left(\frac{\mathrm{NoP}^{\mathrm{V}}_{89\mathrm{d}}}{89\text{ days}}\right) \sim 1.0\text{ prt/day},\qquad \left(\frac{\mathrm{NoP}^{\mathrm{V}}_{418\mathrm{d}}}{418\text{ days}}\right) \sim 0.9\text{ prt/day}.$$

In this case, we observe an approximately the same trend between the NoP and the exposure time (as will be discussed in more details in the next section) in all the three samples. For this reason, we suspect the H wafer exposed for 34 days was negatively affected by human activity or other external factors and that couple of wafers is discarded by the analysis.

For what concern the handling contamination (1' exposed wafers), from table 5.1 and the distribution in figure 5.5 (bottom-left plot) no notable difference in size is present, and the same NoP is counted in both cases ($h = 100\%$). This suggests the contamination contribution due to handling to be the same for H and V wafers, as may be expected.

For a quantitative analysis on the differences between H and V wafers, we are left with only the couple of wafers exposed for 89 and 418 days. The V and H distributions report no notable difference in size, and the only difference is the a smaller particles count which appears to be constant over the whole diameter range. Eventually, by averaging the h values for the non anomalous couples of wafers, the experimental ratio between V and H wafers is

$$h_{exp} = (30 \pm 10)\%.$$

This value is larger to the expected value found in literature (eq. 5.1), but could be negatively affected by the limited statistics of samples in the analysis. Other couples of V/H wafers are currently exposed in the same conditions, and more of them will be exposed in the future. This will increase the statistics of our samples and allow for a more accurate and precise estimation of h_{exp}.

5.3 Dust Deposit in Time

Since we have collected samples being left exposed for different amount of time, we can also monitor the increase of dust deposit in time and provide an estimate on the expected dust deposit rate on the surfaces. This is particularly important because with an estimation of the number of particles deposited in a given time we can estimated, through BSDF simulations, when the dust level will become critical and consequently evaluate the need for cleaning procedures or improvement in the environmental cleanliness.

The NoP for the whole database of wafers collected as a function of the exposure time are reported in figure 5.6 [59]. They are grouped according to the four environments where the wafers were exposed (DET lab, FIS lab, and the EQB1 and SQB1 benches). Hereafter, the NoP reported is computed as the average NoP among the different (non-anomalous, ref.5.1) regions of the sample. It is reported as the absolute number of particles in a region of area A_{photo}.

As a general trend, it can be seen that the NoP deposited on the wafer is approximately linearly increasing with time, as expected.

For each sample, the dust deposit rate $\eta_{H,V}$, expressed in prt/day/m^2 (where prt is the abbreviation of particles) is computed separately for all the H and the V oriented wafers. This is because, as

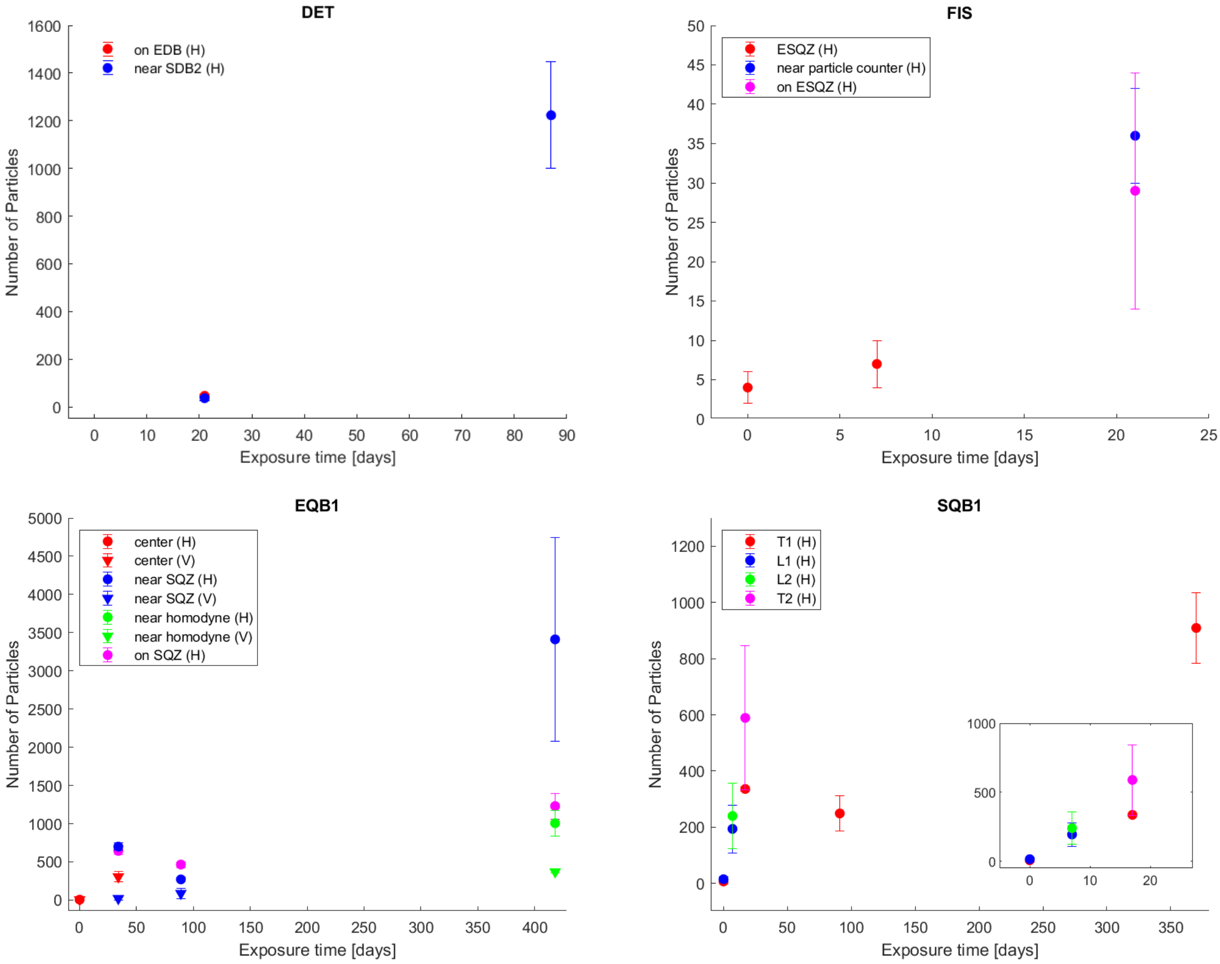

Figure 5.6: NoP as a function of the exposure time (in days) for the collected samples. The four different panels refer each one to a different environment. The label of each datapoint refer to its location and the orientation (in parenthesis), and they are also color coded based on the position (wafer exposed in the same place has the same color). The positions labelled as T1,T2,L1,L2 in the SQB1 bench are reported in figure A.4 in the Appendix. In addition, H wafer are represented as dots, while V wafers (present only in EQB1) are represented with downwards triangles. The NoP is normalized over the photographed area, $A_{photo} = (4.8 \cdot 10^{-4})\text{m}^2$.

shown in section 5.2, differently oriented surfaces experience a different amount of particulate deposit. Regarding the H wafers, the dust deposit rate η_H is found to be varying in the range

$$\eta_H \approx (0.1 \; - \; 7) \cdot 10^4 \text{ prt/m}^2\text{/day}.$$

This range is very broad, and includes all kind of wafers, in very different environments with or without human activity or external factors that may have contaminated the sample. To study how this dust deposit rate is distributed over the samples, the histogram of the number of wafers showing different dust deposit rates is reported in figure 5.7. As can be seen from the histograms, although the large range of values, almost half of the H oriented wafers (13 out of 27) recorded a dust deposit rate

$$\eta_H \lesssim 1 \cdot 10^4 \text{ prt/m}^2\text{/day}.$$

Moreover, by plotting the histogram in the first bin with an higher resolution (fig.5.7, right), it can be seen that the dust deposit rates $\eta_H \lesssim 1 \cdot 10^4$ prt/m^2/day are uniformly distributed.

Since the wafers are exposed in environments where routinely human activities (for example, commissioning operations) are performed, it is expected that those factors can affect the sample by depositing additional dust on the surface than what is expected from environmental dust fall. The distribution of our results may be a sign that the difference between the samples with an higher or lower dust deposit rate may be due to human activity. In particular, for horizontally oriented

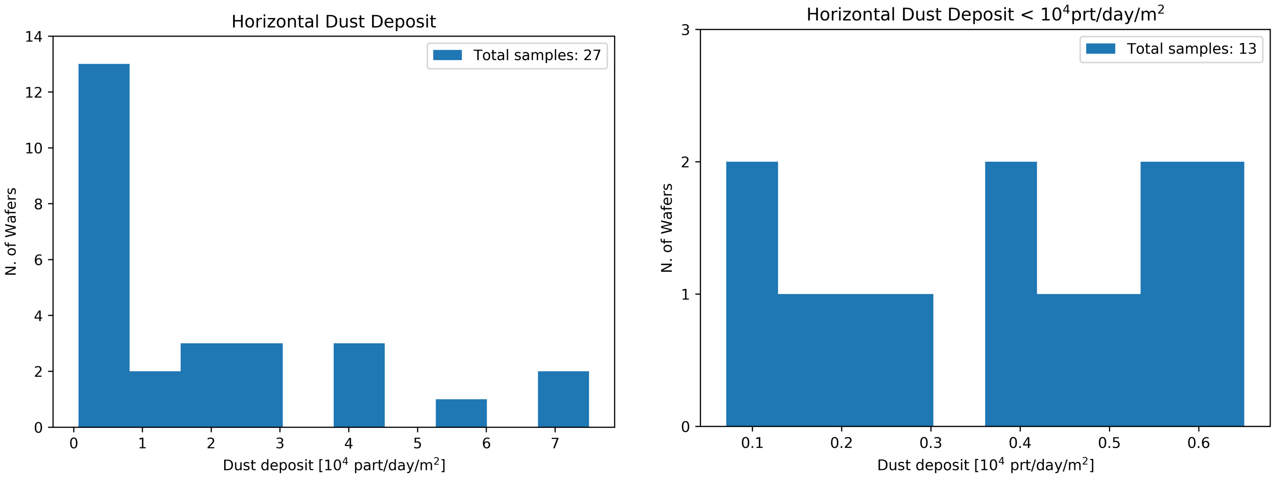

Figure 5.7: Histogram reporting the number of wafers with respect to different values of the measured dust deposit rate. In the left plot, all the H oriented wafers are reported (for a total of 27 samples). In the right panel is shown a zoomed histogram corresponding to the first bin in the left plot to display with higher resolution the distribution for the lowest dust deposit rates (dust deposit rates $< 1 \cdot 10^4$ prt/day/m^2).

wafers, the difference seems to be made by dividing the wafers with a dust deposit rate greater or smaller than $\eta_H = 1 \cdot 10^4$ prt/m^2/day. However, this range should be used to estimate the deposit rate due to environmental dust fall only, rather than separate between wafers experiencing human activity or not in general. In fact, a wafer exposed to human activity may still return a deposit rate $\eta_H < 1 \cdot 10^4$ prt/m^2/day if exposed long enough: a key factor is also represented by how frequently each wafer is exposed to human activity.

For this purpose, we can consider two test samples that were placed in the PG Lab, where human activity can be tracked in a shorter period of time (O(days)), since fewer commissioning operations are performed inside it compared to the optical benches. The first wafer was exposed for 7 days (from 11/12/2020 to 18/12/2020) [60], while the second wafer was exposed for a 3 times longer time (21 days) but in between of the Christmas break (from 18/12/2020 to 08/01/2021) [61]. They were both exposed with H orientation. As reported in [60] [61], they differ in the sense that when the first wafer was left exposed, external manipulation from other people outside our group was recorded in the CR: the wafer has been moved and its box lid put upside down with its inner side facing up. In this way the lid itself become a dust collector, possibly contaminating the wafer when closing the box. On the other hand, the second wafer was left exposed during the Christmas break, hence in a period when the usage of the room is very limited. The dust deposit rates normalized over the exposure time of the two wafer are

$$\eta_H^{\text{activity}} = 2.7 \cdot 10^4 \text{ prt/m}^2\text{/day}$$

$$\eta_H^{\text{xmas}} = 0.2 \cdot 10^4 \text{ prt/m}^2\text{/day},$$

where the superscript "activity" refers to the first wafer, which was manipulated, and "xmas" refers to the wafer left exposed during the Christmas break. As can be seen, the second wafer recorded a lower particle deposit rate than the first one, which was affected by human manipulation. More in detail, the η_H values seem to be compatible with the cut off found from the histograms over the whole database (fig 5.7), $\eta_H \approx 1 \cdot 10^4$ prt/m^2/day. The analysis over all the database and the samples exposed in the PG Lab would suggest that the value $\eta_H \approx 1 \cdot 10^4$ prt/m^2/day determined by the above analysis roughly divides between environmental dust deposit and human intervention. Unfortunately, reporting of activities in the logbook, down to the details needed to infer the passage and work of people in the rooms, is not available and may not be realistic. In addition, due to the limited statistics, wafers exposed to different environments (with different ISO and cleanliness levels) are compared, and the aforementioned value is a rough average among them. In the future, with the monitoring campaign undergoing, we will collect more sample, allowing for an higher statistics in the single environments and a more accurate estimate of the dust deposit rates.

For what concerns instead the V oriented wafers, the analysis is performed in the same fashion as above. All the samples show a dust deposit rate approximately

$$\eta_V \approx (0.1 \; - \; 0.2) \cdot 10^4 \text{ prt/m}^2\text{/day},$$

with only one wafer reporting a dust deposit rate almost time times as large:

$$\eta_V^{max} \approx 1.9 \cdot 10^4 \text{ prt/m}^2\text{/day}.$$

The histograms of the number of V wafers as a function of the dust deposit rate values are reported in figure 5.8. In the same fashion as for the H wafers, although the smaller statistics, there are more samples at the smallest dust deposit rate ($\eta_V \lesssim 0.2 \cdot 10^4$ prt/m^2/day).

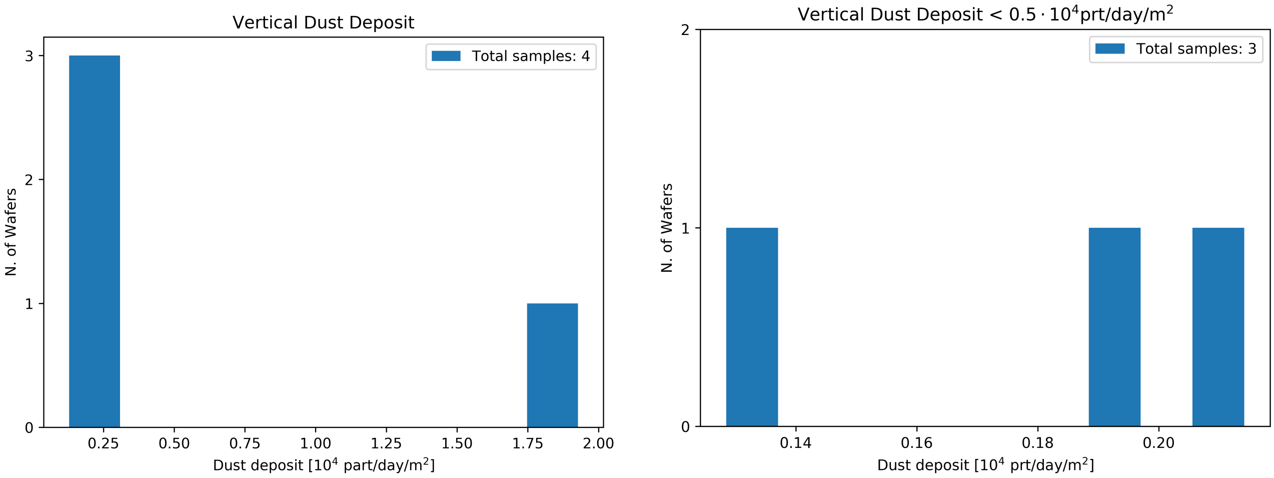

Figure 5.8: Histogram reporting the number of wafers with respect to different values of the measured dust deposit rate. In the left plot, all the V oriented wafers are reported (for a total of 4 samples). In the right panel is shown a zoomed histogram corresponding to the first bin in the left plot to display with higher resolution the distribution for the lowest dust deposit rates (dust deposit rates $< 0.5 \cdot 10^4$ prt/day/m^2).

By comparing the dust deposit rates of the V and H oriented wafers, it can be seen as there is an approximately 25-30% ratio between η_V and η_H values distributions. Those values are computed using all the H wafers available (31 samples, discarding the samples left exposed for 1 minute to monitor handling), even the ones with no V counterpart. Nevertheless, this is compatible with the value $h_{exp} = 30\%$ that was found in section 5.2, which instead was computed only on the H wafers with a V counterpart (3 samples, discarding the sample left exposed for 1 minute to monitor handling). This seems to confirm that in our vertical surfaces it was collected ≈ 30% of dust particulate compared to horizontal surfaces.

5.3.1 Perugia Clean Room Measurements

In the monitoring campaign particular attention was put in the characterization of the Virgo Central Building Clean Room "Perugia", where the camera system is placed. The main reason is that, being the room where the imaging setup is hosted, it is crucial to have an accurate characterization of this environments to monitor additional contaminations added in correspondence of taking the pictures

The first samples to be analysed are the two wafers exposed between December 2020 and January 2021, and discussed in the previous section [60] [61]. As reported, they represent an useful reference being a couple of wafers for which one of them was exposed to human activity nearby the sample, while the second one was not.

A second monitoring was performed on February 2022 [62]. During the shift it was tested how the photographing procedure and general handling of the sample affects the dust distribution. For this purpose, one brand new wafer was opened and a series of photos were taken after each step of a series

Entry	Air Exposure Time	Actions	N. of Particles $[/(4.8 \cdot 10^{-4} \text{m}^2)]$
1	1min30s	box opening, placing the wafer under camera	1
2	2min30s	(left under camera, no activity)	1
3	5min00s	(left under camera, no activity)	1
4	9min00s	(left under camera, no activity)	1
5	12min20s	removed, close and reopen box, placed under camera	1
6	17min	previous point repeated x3 times	1
7	37min	20min of worktime near wafer	5
8	14h47min	left exposed during night (no activity)	5
9	17h47min	left exposed during the morning (with no activity)	7

Table 5.2: NoP detected on the wafer after each one of the listed actions. For the photos taken after the night and after the morning the NoP is calculated as an average among four different regions of the wafers. The entry columns serves as a label to refer to the single photo.

of operations. The NoP detected on the wafer and the list of performed operations are reported in table 5.2.

During the exposure times in which the wafer was not left under the camera (the 20min worktime nearby it and during the night and morning, the wafer was left exposed on the table, between the camera system and the computer. The wafer placement is shown in the photo in figure 5.9.

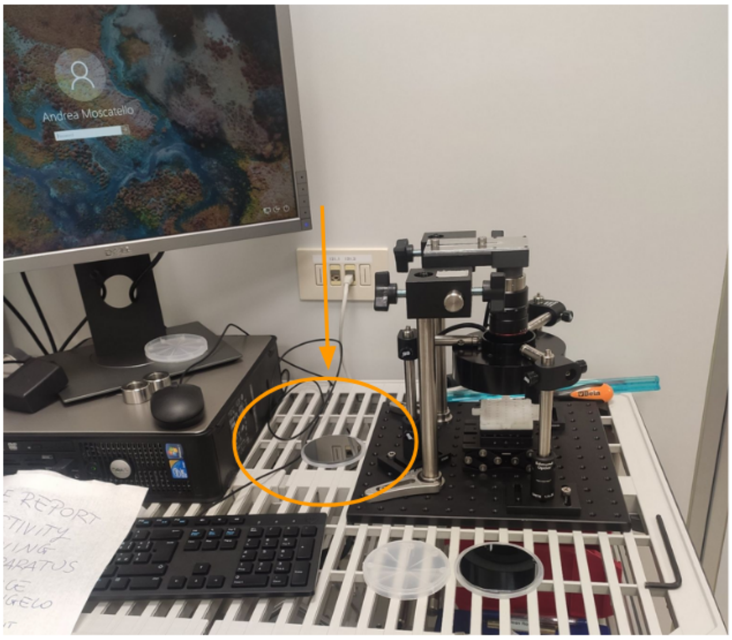

Figure 5.9: Placement of the wafer (highlighted by the orange arrow) used in the characterization of the Perugia Clean Room and analysed in table 5.2. In this photo, on the right of the wafer, the camera system for photographing the samples can be seen.

From the data reported in table 5.2 it emerges that the mere photographing procedure does not affect the sample in an significant way. In addition, it can be seen that the NoP increases when human activity is performed (entry 7) and after having left the wafer exposed during the night and morning (entry 8-9). From the photos taken in this characterization, data suggest that external dust contamination can be related to human activities nearby the wafer rather that handling and photographing the sample.

From the data reported in table 5.2, the dust deposit rates (normalized over an area A_{photo}) corresponding to the following conditions can be computed:

- imaging procedure (and handling): $\eta_{1-6} = 0.06$ prt/min;
- people working nearby: $\eta_7 = 0.2$ prt/min;
- overnight: $\eta_8 = 0$ prt/min;
- morning imaging: $\eta_9 = 0.01$ prt/min;

where the subscript refers to the corresponding labelled entry in tab.5.2. In can be seen as the event with the highest deposit rate is the 20 minutes of human activity ($\eta_7 = 0.2$ prt/min). However, this single event becomes less relevant over a long time period. In fact, the total dust deposit rate, computed over the cumulative air exposure time is

$$\eta_{tot} = 0.006 \text{ prt/min/A}_{\text{photo}} \approx 1.8 \cdot 10^4 \text{ prt/day/m}^2,$$

which is compatible to the periods with no activity of the same wafer. Moreover, this value is compatible with the results found in 5.3, suggesting that a dust deposit rate for an horizontally oriented wafer $\eta_H \gtrsim 10^4$prt/day/m^2 may be an indication of exposure to human activity. However, as shown in this analysis and stated before, the total exposure time of the sample and the amount (and distribution in time) of human activity both impacts the estimation of η, and as a consequence this division has to be taken as an indication rather than a rigid cut-off.

5.4 Cleanliness of the Environments

For the analysis of the cleanliness levels in the monitored environments the IEST Standard described in section 2.2 is used. For each wafer, the diameter distribution of the deposited dust particles is computed by merging together the distributions found in each one of the non-anomalous photographed regions (according to the procedure described in sec.5.1).

The distributions for all the samples used in the analysis are reported in Appendix B. We can then perform a $\log - \log^2$ fit following equation 2.4. In particular, the linear fit is performed by setting $x = \log^2_{10}(D)$ and $y = \log_{10}(N_p)$, resulting in

$$y = -|S|x + |S| \log^2_{10}(CL).$$

The slope S is directly derived as the slope of the fit, while the CL level is trivially computed from the above equation as

$$CL = 10^{\sqrt{q/|S|}},$$

where q is the constant term in the linear fit. In particular, it should be noted how the CL is depending on the value of the slope S. This means that CL values should be correctly compared only between distributions having a similar S. If two distributions show a slope which is too different the comparison between the cleanliness levels will no longer be representative of the cleanlinesses.

The $\log - \log^2$ representations of the samples distributions are reported in figures 5.10 and 5.11. In general, it can be seen that in the central part of the distribution ($2 \lesssim x \lesssim 4$), corresponding to $25\mu\text{m} \lesssim D \lesssim 100\mu\text{m}$) there is an agreement with a linear trend expected from theory. For smaller ($x \lesssim 2$) and larger diameters ($x \gtrsim 4$) the distribution deviates significantly from the expected behaviour. In particular, a plateau can be observed at the smallest diameters, while for larger particles the tail of the distribution can be see to deviate from the linear trend.

For smaller diameters ($D < 25\mu$m) a plateau can be seen in almost all of the samples: this may be due to approaching the lower limit sensitivity of our apparatus for smallest particles. In fact, the limit $D = 25\mu$m corresponds approximately to particles with an area in the photo of 6px^2, which approaches our system sensitivity (the pixel range). It should be noted, however, how this range is limited compared to the ideal range upon which we would like to perform our analysis (approximately $1 - 100\mu$m). This may negatively affect our results, as will be discussed in the future sections, since our reconstructed distributions may not be truly representative of the deposited distributions.

On the other side of the dimensions range ($D > 75\mu$m), most wafers display a discrepancy from the linear trend as an overcount of big particles. From our sample, no correlation between this feature and a particular exposure condition (place, time of exposure, orientation) is found. One possible explanation could be that this discrepancy is caused by human activity affecting the sample. For example, consider the couple of wafers exposed in the PG Lab between December 2020 and January 2021, that were introduced in section 5.3. In particular, one of them was exposed 7 days but human manipulation and activity were recorded, while the second one was left exposed for 21 days during the Christmas Break. Their $\log - \log^2$ distributions are reported in figure 5.12 (as well as C.6). As can be seen from the fit, in this couple of wafer, the one exposed to human activity is the only one showing a discrepancy at the larger diameters. This may suggest that indeed this discrepancy is caused by an excess of bigger particles deposited after human activity (refer to Sec.3.2.1). In fact, with no human activity, the dust particulate in a clean room is filtered by the filters of the camera and, the higher the ISO level, the smaller are the particles that passes through the filters [40]: for a given class, the bigger the particle is and the less likely it is to be found in the room. On the other hand, there is no real limit to the size of the particles caused by human activity which leads to particles release from local materials (gloves, coveralls, skin etc...), which then deposit on the samples. However, in the whole database of samples, we are not able to establish or not the presence of a correlation between this discrepancy and the presence of human activity because, as previously reported in Sec.4.2.2, the history of recorded human activities or manipulations of the wafers is not always present and reliable. This requires for the future a more precise and detailed tracking of human activity.

Due to the features appearing in the distribution for the smaller and bigger particles, the real distributions deviate from the linear trend in the aforementioned ranges. In order to extrapolate the $|S|$ and CL values from the distributions, a linear fit is performed in the range $x \in [2;\ 3.5]$. This range is kept fixed for all the sample, as it removes the plateau found at the smallest diameters and the tail discrepancies at the larger diameters. The fit performed over all the samples are reported in the Appendix C. In terms of dimensions, the linear behaviour is found in the range

$$25\mu\text{m} < D < 75\mu\text{m}.$$

However, the history of recorded human activities or manipulations of the wafers is not always present and reliable, as discussed in Sec.5.2.

5.4.1 CL and Slope values

By applying the linear fit from equation 5.4 to the distributions (fit plot reported in Appendix C), the CL and Slope parameters characterizing the distributions are computed. However, as discussed in the previous section, the linear fit range is limited and we are neglecting the contributions of particles with diameter $D < 25\mu$m and $D > 75\mu$m. Their values are reported, divided in environments, in figure 5.13.

Following grouping as in figure 5.13, the average values of S in the different environments are reported in table 5.3. The compatibility, λ, with the expected value S=-0.926 [31] is computed as $\lambda = |\overline{S} - S|/\sigma_{\overline{S}}$. As can be seen from the table, most $\overline{S}$ values are compatible with the expected value S=-0.926 [31] within 1σ, with the worst compatibility being $\lambda = 1.2$. In general, it can also be noted how the variability of S (expressed in the table as the standard deviation $\sigma_{\overline{S}}$) among different samples can be seen to change over different environments.

Regarding the CL values, it can be seen that overall the samples are in the range

$$\text{CL} \approx 200\ -\ 600,$$

and no anomalously dirty CL levels (computed over particles with $25\mu\text{m} < D < 75\mu\text{m}$) are detected. However, as previously stated, CL values should be compared only if the slopes are not varying much between them. From the data reported in figure 5.13, the FIS samples have a CL with a much higher associated error, and are also the samples showing an higher variability in the slopes.

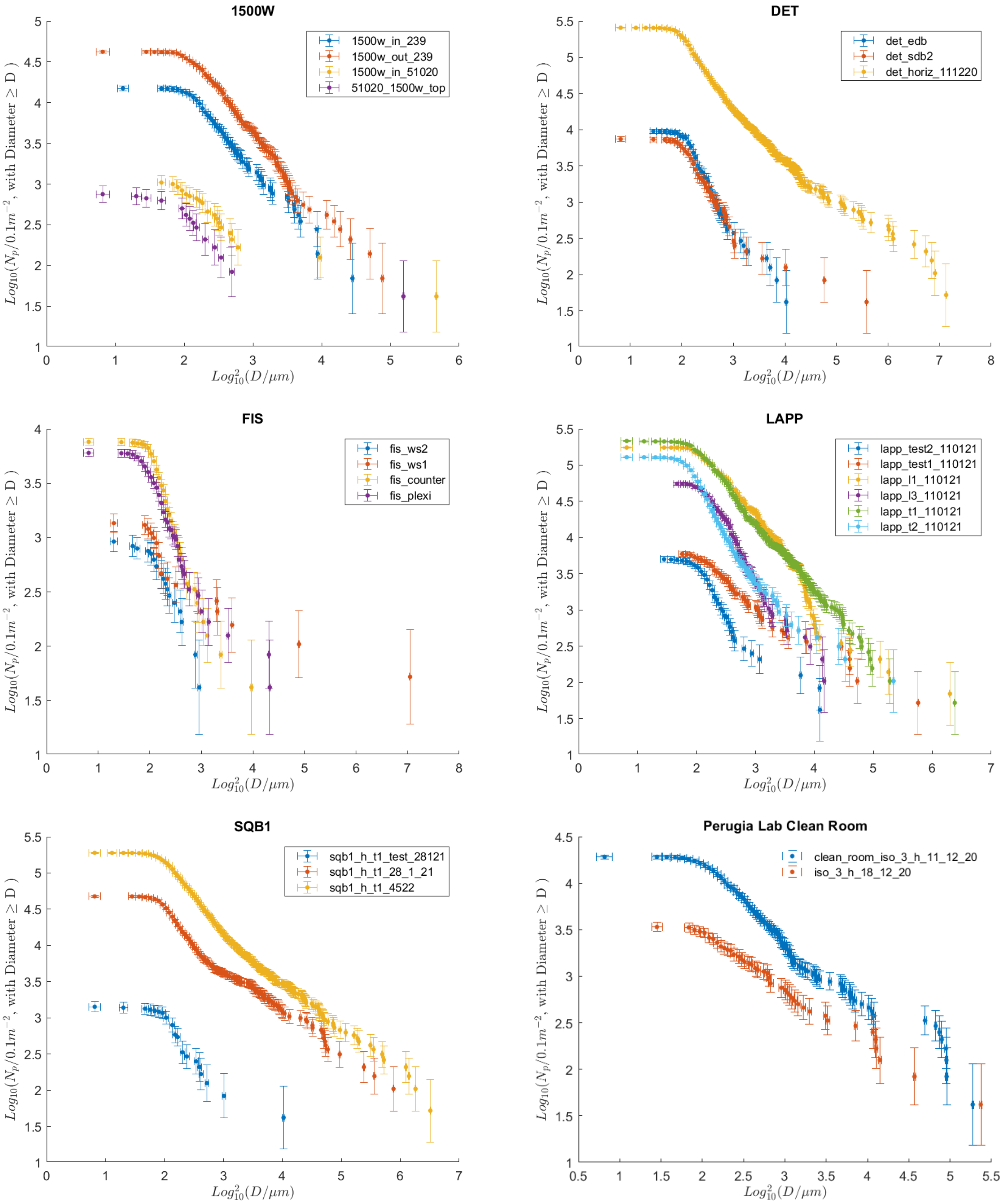

Figure 5.10: $\log - \log^2$ representation of the wafers dust distributions according to the IEST Standard. Each graph refers to a single testes environment, reported in the title. In each plot, the labels of the datapoints report the name of the sample. Note that the distribution are not normalized for the exposure time.

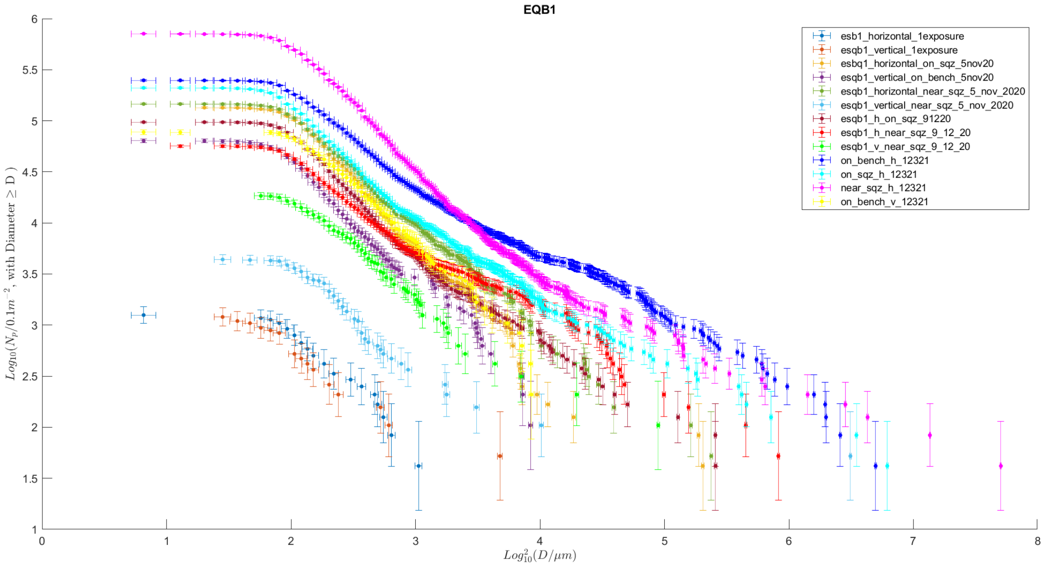

Figure 5.11: log $-$ log^2 representation of the wafers exposed on the EQB1 optical bench.

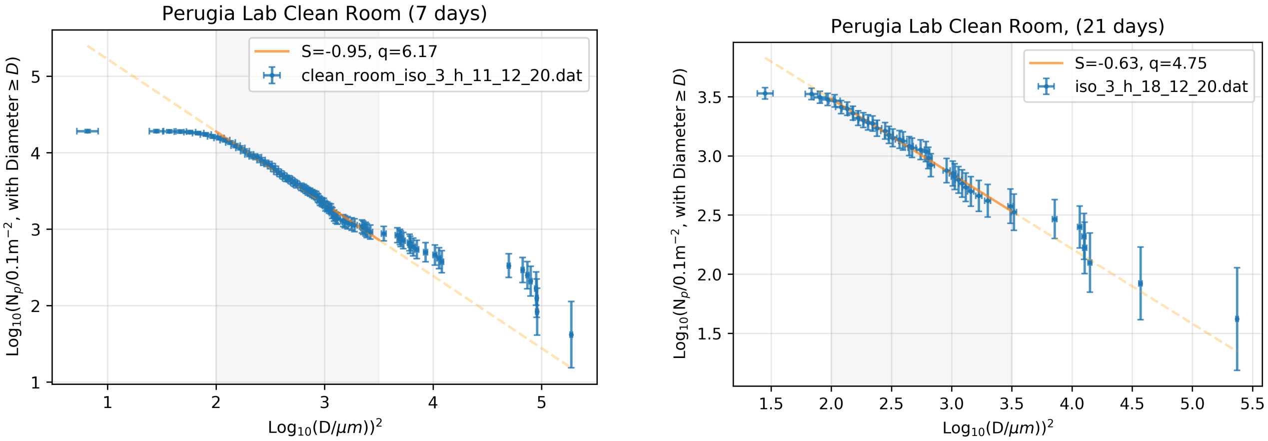

Figure 5.12: From the dust distributions on the samples exposed in the PG Lab between December 2020 and January 2021 (Fig.B.6) the linear fit on the log $-$ log^2 representation is reported. The range of the linear fit is $2 < x < 3.5$. The data reported in the left panel refer to the wafer exposed for 7 days and for which human activity was recorded. The data reported in the right panel refer instead to the wafer left exposed for 21 day during the Christmas Break. The significant tail discrepancy can be seen only in the wafer (at left) which experienced human activity additional to the simple exposure and photographing procedures.

Environment	$\overline{S}$	$\sigma_{\overline{S}}$	λ
1500W	0.98	0.09	0.6
PG Lab	0.80	0.10	0.3
DET Lab	1.20	0.22	1.2
EQB1	1.05	0.13	1.0
FIS	1.1	0.4	0.5
LAPP	1.17	0.27	0.9
SQB1	1.0	0.20	0.3

Table 5.3: Average value of S, $\overline{S}$ and error, $\sigma_{\overline{S}}$, divided in different environments. The error is computed as standard deviation on the different S populations. In the last columns, the compatibility with the expected value $S = -0.926$ [31] is reported.

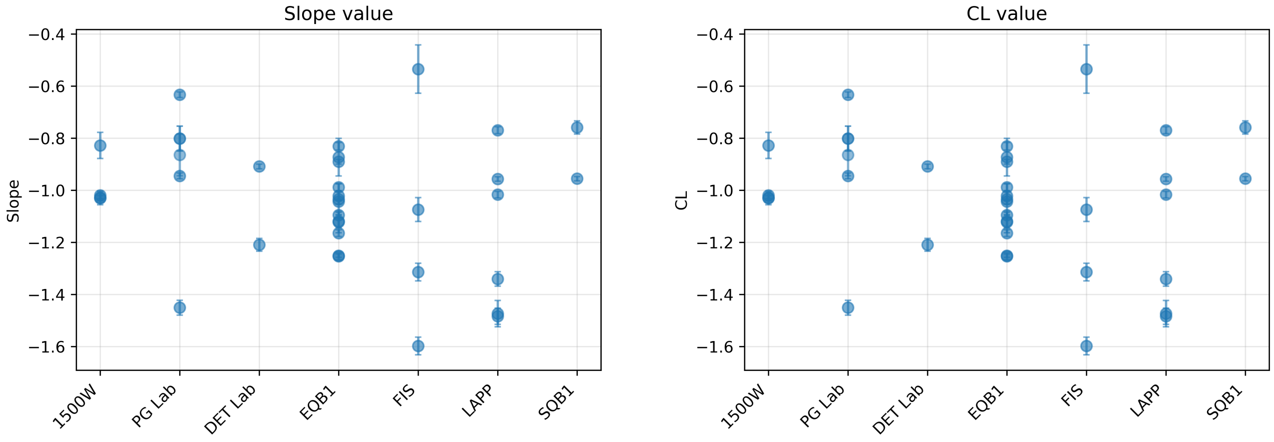

Figure 5.13: CL and Slope (S) values obtained from the linear fit on the dust particles distributions. Each datapoint refer to one sample, and on the x-axis are divided in different environments. The errorbars on S are directly obtained from the fitting algorithm as an error on the parameter, while the errorbars on CL are computed from error propagation following eq.5.4. Be reminded that CL levels corresponding to substantially different slopes cannot be directly compared.

Apart from a direct comparison between the cleanliness of different environments, the CL and S are particularly useful because they allow us to compare the BSDF computed over a particular sample with the BSDF provided in literature. In fact, since we know the exact dust particles distribution on a particular sample (the histograms in Appendix B), we can use this exact distribution as input to compute the BSDF. However, in literature, the BSDF are often times given in terms of CL and S: we can compute the BSDF both by using as input the distribution and the CL-S values, but with the latter we can compare it with the literature. However, as we have reported in the analysis, such parameters may not be truly representative of the real dust particles population because of the sensitivity limitation and discrepancies that may be linked to human activity affecting the sample. This will discussed in Section 5.5.

5.4.2 CL in time and ISO levels

As mentioned in Section 2.2, the CL value increases with time: as dust deposits on the surface, the x intercept in the fit moves to higher values, leading to an higher value of CL. This was shown in equation 2.6. In addition, CL is also found to be dependent on the cleanliness class of the clean room X_c (or equivalently, the ISO number). This allows us to perform a dynamic monitoring of our CL values, and not only a static analysis as performed in Section 5.4.1. We can verify if the CL values that we have obtained vary in time as in equation 2.6 and if the time dependence is compatible with the nominal ISO standards of the tested Virgo environments.

We investigate the dependence of CL with time in two environments: the EQB1 bench, inside the Detection Lab and the PG lab inside the Central Building Clean Room, which contains the camera system. The DET Lab (and EQB1 bench as a consequence) is a ISO 7 clean room [21]. The ISO level of the PG Lab is instead not known: the only ISO class information available is the ISO 3 classification of a plexiglass closed environment inside the laboratory [21]. In any case, for the PG lab many samples taken at different exposure times are available: this allows us to verify the CL time dependence, and derive from our data an estimate of the effective ISO class inside.

The CL values for the EQB1 exposed samples (both H and V), as a function of time are reported in figure 5.14. Along the data, the prediction CL(t) following equation 2.6 is shown too (dotted line). The shaded region is the 1σ uncertainty region computed from error propagation of the variables h and S, while $\rho = 2851$ [34] and $X_c = 10000$ (ISO=7) (from Tab 2.1) are assumed known without uncertainty. For the horizontal samples $h = 1$, while for vertical samples the experimental value $h = 0.3 \pm 0.1$ was used (which was derived in section 5.2). For what concerned the slope, the value $S = 1.05 \pm 0.13$ is used, which is obtained by averaging the S values for all the EQB1 samples (table 5.3). As can be seen from the figure 5.14, the time dependence of both our H and V oriented samples matches the prediction from 2.6 within the uncertainty region computed from error propagation.

However, the 1σ compatibility between the data and the prediction is manly due to the high uncertainties on S, which corresponds to a relative error in the order of 10% on the value of S. In fact, the CL values from the samples show high fluctuations for a given exposure time. For example, this can be seen by the CL values of the H wafers at ≈ 100 and ≈ 100 days (fig.5.14, left), for which we have $\Delta CL \approx 200$, but still every one of them stands in the 1σ compatibility. This high uncertainty, may be caused by the limited fit range used in the analysis ($25\mu\text{m} < D < 75\mu\text{m}$), as discussed in the previous section. In addition, it should also be noted how those data refer to different places and at different times in the EQB1 bench. This causes an high degree of variability: although being exposed (nominally) in the same environment (EQB1), the cleanliness in it may change significantly. For example, wafers in one particular position may experience more commissioning activities compared to other wafers, eventually collecting more dust. As reported in fig.5.14 (left) the H oriented wafers located on the squeezing box (green datapoints) seems to be below the expected CL level, while the H wafers located near the squeezing (purple datapoints) seems to be higher, expect the last point. On the other hand, for V oriented wafers (fig.5.14, right), the samples located near the squeezing are below the expected CL level. However, this analysis is still lacking a robust statistics, which will increase in the future. Moreover, EQB1 itself may not be representative of a ISO 7 cleanliness class. As reported in Sec.3.3, EQB1 is closed inside plexiglass boxes, which limit the air-flow and also affect the dust deposit on this bench. Linked to this previous point, it should also be noted how to model provided by eq.2.6 is limited, since the different assumptions made and the different values adopted are only representative of an average clean room and may not reflect our environments.

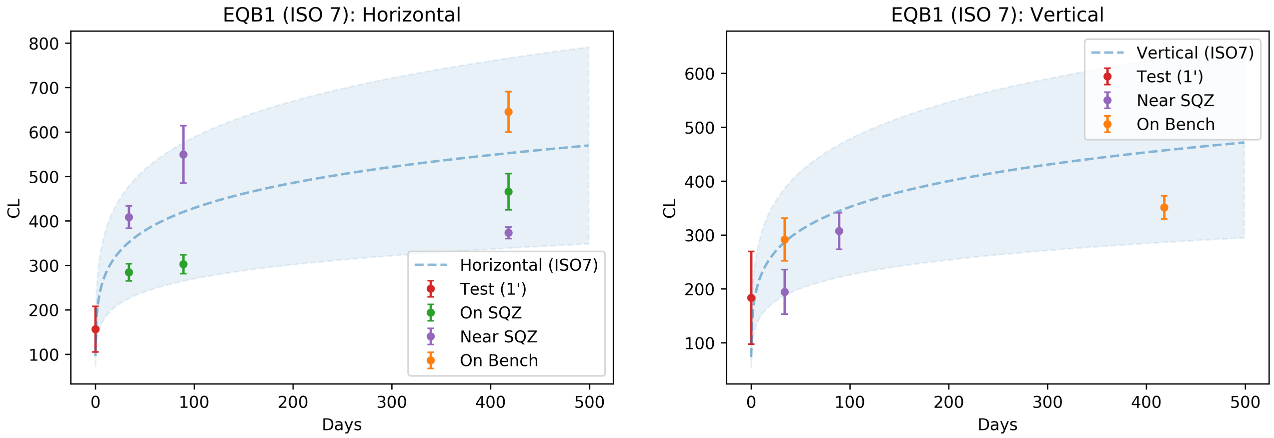

Figure 5.14: Experimental data (points) of the CL values of the samples exposed in EQB1 as a function of their exposure time t. The different samples are colorcode with respect to their location on the bench (reported in the legend). The dotted lines are the theoretical predictions computed from equation 2.6, while the shaded area is the 1σ uncertainty region computed from error propagation on h and S. The left plot refers to the H oriented wafers ($h = 1$), while the right one refers to the V oriented wafers ($h = 0.3 \pm 0.1$).

The data of the samples exposed in the PG Lab (containing the camera system) are reported in figure 5.15. The value of h was kept fixed to $h = 1$ with no uncertainty since we only have H oriented wafers in this case. The slope value was derived by averaging over the Ss of all the samples exposed in the PG Lab, obtaining $S = 0.80 \pm 0.10$ (table 5.3). In this case, since the no precise information on the ISO class of the room was available, we used our data to estimate the ISO class. Among different ISO levels (ISO=4,5,6,7,8), the best fitting one is ISO 6, as can be seen by the residuals plot reported in figure 5.15 (right). The residuals are computed as $CL(t)_{samples} - CL(t)_{prediction}$, where the subscript indicates the CL measured on our sample and the CL predicted from Eq.2.6, respectively. It can be seen how the ISO 6 residuals are the one closer to 0, hence the ISO 6 class is the best matching class to our data. In addition, as can be indeed seen form the figure 5.15 left, all of the data are compatible within 1σ with the ISO 6 class. Moreover, every datapoint (expect one) is compatible with the ISO 6 class within its uncertainty.

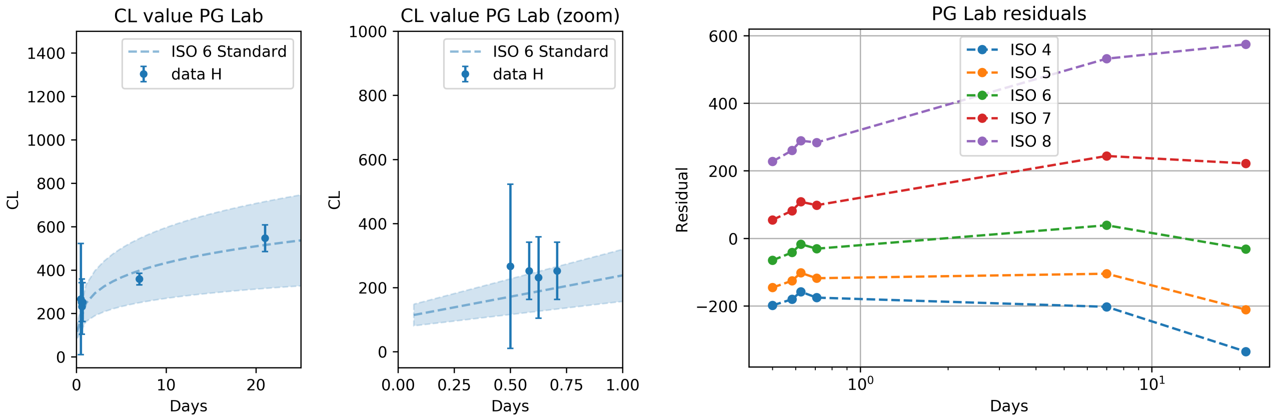

Figure 5.15: Experimental data (points) of the CL values of the samples exposed in the PG lab as a function of their exposure time t (left panel). The dotted lines are the theoretical predictions computed from equation 2.6, while the shaded area is the 1σ uncertainty region computed from error propagation on S ($h = 1$ is constant since we only have H oriented wafers). In the right panel, the residuals from the data and the CL values predicted by Eq.2.6 are reported for different ISO levels (reported in the legend).

From our analysis we can see that the nominal ISO level of the EQB1 bench is compatible with the time evolution of the CL of samples exposed to this environment. As mentioned above, it should be however noted that the compatibility is mainly due to an high uncertainty in the estimate of S. In the case of the PG Lab, the CL dependence in time is matched, and from our data we were able to extrapolate an effective ISO 6 cleanliness level for this environment.

5.5 BSDF Estimation

Once the analysis algorithm computes the dust particles distributions deposited on the samples, we can use such population to compute an estimate on the expected BSDF. The BSDF is calculated from its expression in equation 2.3. As already mentioned, the two main steps are the modelling of the dust particles distributions (the particles density function, $f(D)$), with its optical properties, and the computation of the scattering matrix elements $S_{1,2}$. The scattering matrix elements are computed using the MiePython package, as already discussed in Sec.2.3.

We are only left to provide an estimate on the particles distribution function $f(D)$. This function is the core quantity inside the BSDF since this is what models the dust particles population deposited on the samples. In practice, two different ways of computing $f(D)$ are present. The first was discussed in Sec.2.3: following the IEST standard, $f(D)$ is computed from the CL and S values obtained from the fit. In this case, we assume the distribution to be defined in the range $D = 1 - 170\mu$m (with intervals $\Delta D = 0.1\mu$m), even though the fit were only done in a smaller range, for the reasons discussed in Sec.5.4.

As a second option, we can estimate $f(D)$ directly from the histograms of the dust distribution returned by the algorithm, as in figure 4.18. By doing so we would not need to sample $f(D)$ from an extrapolated distribution (through CL and S) but we can directly use the particle's counts as input. Starting from the histograms, we can determine exactly how many particles for each diameter value we have detected, and by normalizing over the area we obtain the particles density. Since the histograms return the cumulative NoP of all the area in which no anomalies were detected (ref. section 5.1), the total area used for the normalization is the photographed Area ($A_{photo} = 4.8 \cdot 10^{-4}\text{m}^2$) times how many regions are used for the analysis, following the procedure in Sec.5.1. Eventually, $f(D)$ has to be scaled to μm^{-2}.

In practice (as it is often the case) differences between the two methods arises. Conceptually, the biggest difference is that by computing $f(D)$ directly from the histograms allows to follow the measured particles distribution of the sample. On the contrary, by sampling $f(D)$ following the IEST method, we have to perform a linear fit on data that are not perfectly linearly distributed, and will result in an approximation. As shown in section 5.4, two main issues are present: an undercount of particles at the smallest diameters and an overcount of the biggest particles. By sampling $f(D)$ from CL and S we do not account for such discrepancies, and the resulting distributions do not correspond to the reconstructed ones. Because of this, the $f(D)$ computed with the two methods are not the same, leading to differences in the BSDF. In addition, it should also be noted how the histograms of the reconstructed distributions, reported in Appendix B, do not have a uniform diameter binning. In fact, as discussed in Appendix B, the measured area, A, in quantized in units of 1px^2 but the effective diameter, d, depends on A as $d \propto \sqrt{A}$. For this reason the quantization of d is not uniform, contrary to the quantization of the diameters that can be arbitrary chosen in the IEST method (Sec.2.2.2).

As an example, the two BSDFs computed from the two different $f(D)$ are shown in figure 5.17. Starting from a particular wafer[1] which was H oriented, exposed in EQB1 for 34 days the linear fit is performed and from CL and S the $f(D)$ according to the IEST standard is computed. The particles distribution count density histogram and the $\log - \log^2$ fit are reported in figure 5.16 (The $\log - \log^2$ representation is also reported in figure C.8, central row, left). Then, $f(D)$ is also computed directly from the particles distribution histogram. This particular wafer was chosen because the only discrepancy with respect to the linear trend appearing is the plateau at the smallest particles, as can be seen in its $\log - \log^2$ representation, figure C.8.

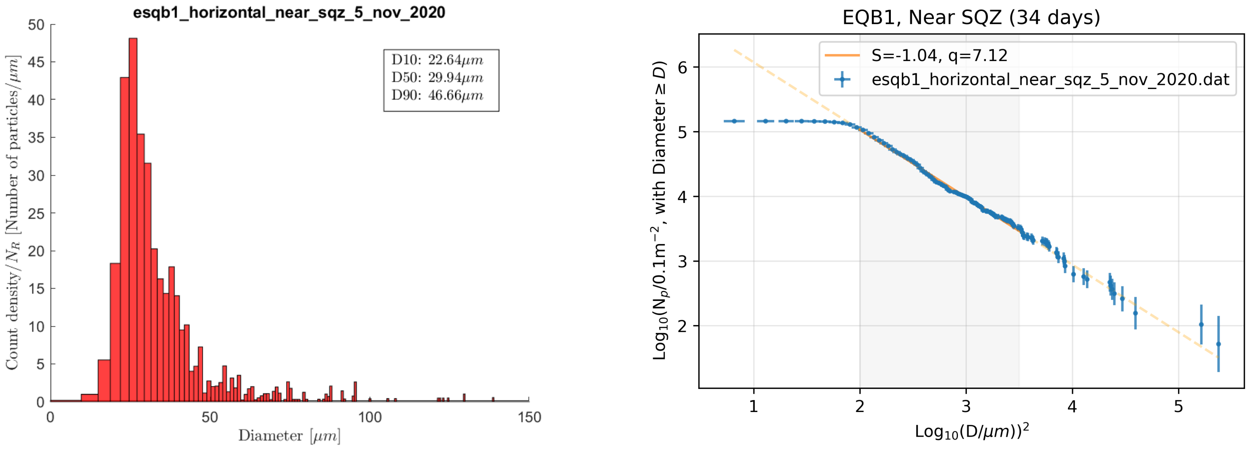

Figure 5.16: Histogram of the dust particles distribution on the sample[1] (left) and linear fit on the $\log - \log^2$ representation (right).

From the two different particles density functions the respective BSDFs are computed. In figure 5.17 the difference between $f(D)$ (left) and the BSDF (right) is observable. From the distributions comparison (fig. 5.17, left) it can be seen that the two distributions matches for diameters $D > 25\mu$m, which was the lower boundary of the linear fit range. Below this value the difference due to the plateau is significant: the histogram data drastically undercount the particles compared to the estimation provided by the IEST method, up to a factor 10^4. As a consequence of this discrepancy, the BSDF computed from the histogram is smaller than the BSDF computed following the IEST method (figure 5.17, right). Eventually, although the IEST method is an approximation of the reconstructed distribution, the latter is probably affected by the sensitivity limit and underestimates the scattering due to the smallest particles limitation in sensitivity by a factor in the order of the unity. For this reason we make a conservative choice, and the following BSDF results are computed by sampling $f(D)$ from CL and S as in the IEST standard method.

By referring to the CL values found in the analysis of the samples and reported in figure 5.13, the CL are found to be overall in the range CL=200-600. To see how this difference reflects in terms of stray light, the BSDFs corresponding to different CL (CL=200, 400, 600) are computed with the Python script described in Sec.2.3, and reported in figure 5.18. CL=200 is approximately the lowest value

---

[1] labelled: esqb1_horizontal_near_sqz_5_nov_2020.dat.

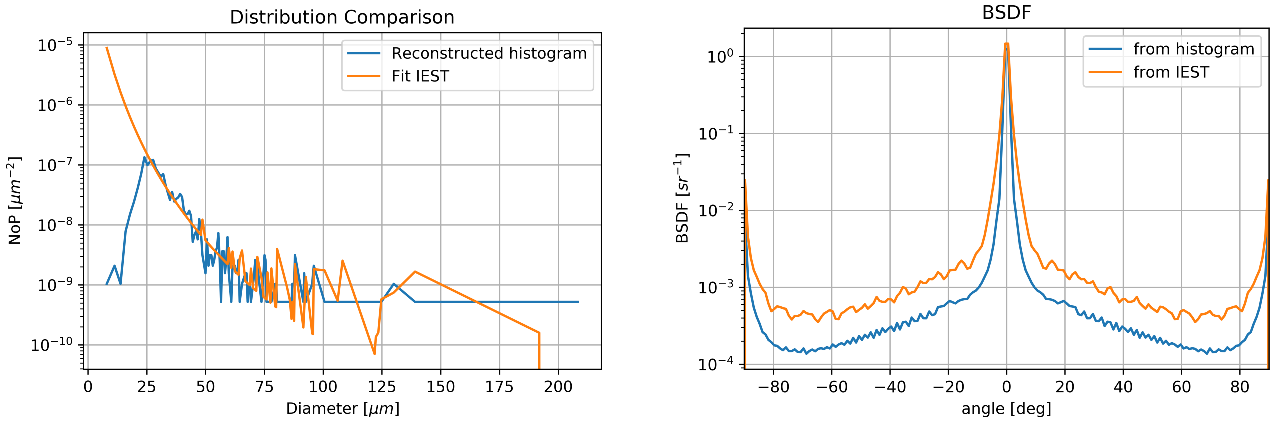

Figure 5.17: Comparison between the particles density distributions and BSDF obtained by following the IEST standard method (orange curves) and computed from the exact number of particles detected (blue curve). The NoP in the histogram data (orange) is directly taken as the y-axis values in the histogram. In the IEST case instead, the NoP is computed from the sampled $f(D)$ according to equation 2.7. The distribution's cleanliness parameters used to sample $f(D)$ are CL=414 and S=-1.04.

found, while CL=600 is the largest. By comparing the lowest and highest curves, it can be seen that there is a factor 10^2 difference between the lowest and highest values of CL obtained. In general, the values reported in Fig.5.18 are compatible with the values returned by FRED and reported in figure 3.1.

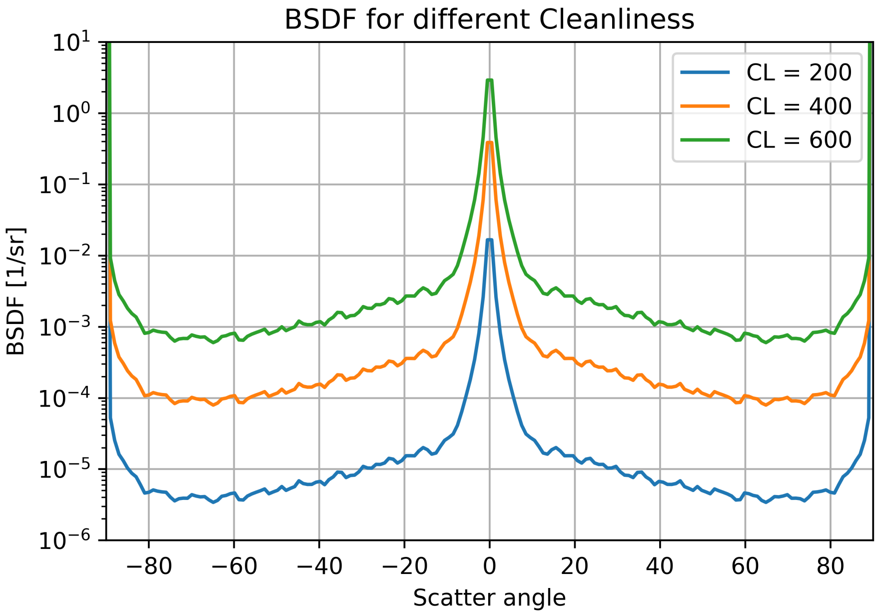

Figure 5.18: BSDF curves corresponding to different CL values: CL=200 (blue), CL=400 (orange) and CL=600 (green). The simulations are made by following the IEST standards, and assuming a slope S=-1.0, corresponding to the average value found in our analysis. The curves are obtained from the Python script that we have written (Sec.2.3).

Finally, comparing the BSDFs found in fig.5.18 to the BSDFs reported in fig.3.1, even in the environments showing the the best cleanliness (i.e. CL $\approx$ 200) the BSDF contribution for dust scattering is greater than the contribution provided by the roughness of the viewport [21]. The viewport under study is the viewport connecting the ITF to the Frequency Independent Squeezing (FIS) in Virgo during the O3 observation run (ref. to [21], Sec.4.1). This results seems to suggest that, for the dust contamination measured in the Virgo environments, the BSDF contribution of dust is comparable (if not higher) to the estimated contribution of the roughness of the viewports. However, as a future improvement there will be a deeper study on the roughness of other optical elements, and eventually the estimation of their BSDF. In addition, in the future the camera system will be replace with a microscope, which will have an higher resolution. With an increased imaging resolution, in addition with the increasing statistics of dust samples, the estimation of the BSDF caused by dust contamination will be more realistic, and overall the comparison between roughness and dust BSDF will be more accurate.

Chapter 6

Conclusion

Stray light manifests as an excess noise in the low-frequency region of the present LIGO-Virgo ITF and it causes a loss in sensitivity both in terms of excess noise in the measurements and also by spoiling control loops. In addition, the next generation ITF, Einstein Telescope (ET), is designed to have an higher sensitivity in the low-frequency region, hence stray light noise may be even more important in the future. For those reasons, it is crucial to be able to monitor and possibly mitigate stray light sources. The main contributions to stray light are surface roughness of the optics surface and scattered light from dust particulates deposited on the optics. However, it was shown that, thanks to the very low surface roughness of the optics, the dust contribution to stray light can overcome the effect caused by surface roughness [21] even in clean environments. In this thesis, we focused on the study of the stray light caused by dust deposit on optics.

We have upgraded and improved the original Matlab algorithm which was routinely used to analyse the images in the LIGO-Virgo collaboration [49]. In particular, a two-step particle search is implemented. In the first step, a global binarization of the image with a threshold T_1 is performed in order to detect the bright clusters. We have designed our algorithm to automatically estimate the optimal first threshold value (Sec.4.2.2). Then the algorithm iterates over each cluster found and performs a second, local binarization at each cluster found. The second threshold, T_2 is computed as a fraction $\alpha = 0.15$ of the cluster maximal luminosity, I_{max}:

$$T_2 = 0.15 \cdot I_{max}, \tag{6.1}$$

where the value $\alpha = 0.15$ follows from calibration on dust powders with known diameter percentiles (Sec.4.3.2). The separation between the phase of detecting the particles from that of estimating their size allows us to obtain both a very high detection efficiency (thanks to the relatively low threshold T_1) and a more accurate estimation of their sizes (by accounting for each particle's luminosity). We have solved a relevant issue in the existing algorithm, that causes a considerable fraction of the smallest particles (up to about 1/3 for certain samples) not to be detected when the parameters are adjusted for the most accurate size reconstruction. With our algorithm, we obtain similar performances in particle reconstruction accuracy and, at the same time, detecting all the particles down to the minimum detectable size of 8μm, corresponding to a 1×1px^2. (Sec.4.2.6). In terms of sizing, the average total error on the diameters of the new algorithm is SSR = 21% (Sec.4.4).

As part of a long term dust monitoring campaign, we first performed a few tests to understand the limits of applicability of our monitoring methods: a series of wafers were exposed to measure the impact of the handling and imaging procedures (Sec.5.2 and Sec.5.3.1). Those measurements showed no signification impact on the samples. From the results of horizontally (H) and vertically (V) oriented wafers we computed the ratio between the number of particles deposited on V wafers with respect to H wafers (Sec.5.2)

$$h_{exp} = 0.3 \pm 0.1.$$

This value is not compatible with the value found in literature ($h = 0.1$, [34]) and further samples are exposed now and will be exposed in the future to increase the statistics on this measurement.

We then performed a preliminary analysis in the still limited amount of samples available so far. In terms of dust deposit rates, we found that most of the exposed samples show a deposit rate

$$\eta_H < 10^4 \text{ prt/m}^2\text{/day},$$

for a H oriented wafer. Higher values were found, with a maximum of $\eta_H \sim 7 \cdot 10^4$ prt/m^2/day (Sec.5.3).

By approximating our data with the model provided by IEST CC1246D standard, we have calculated the cleanliness levels (CL) of our environments, and we found CL to vary in the range CL = $[200; 600]$ (Sec.5.4.1), which indicates "pristine" to "visibly clean" surfaces. The time evolution of CL on the EQB1 bench matches with the expected prediction for the nominal cleanliness class of the environment (ISO 7) within 1σ (Sec.5.4.1).

Finally, we estimate the BSDF for the cleanliness values found in our environments (Sec.5.5). The BSDF we have obtained are compatible with what was previously found in [21]. In particular, it was found that the BSDF level for dust deposited in the cleanest of our samples (CL $\sim$ 200) can be higher than the contribution from optics surface roughness.

As an estimate of the impact on the future generation detectors, we also compared how a typical dust distribution measured in Virgo reflects on ET, which uses an higher wavelength laser compared to Virgo. We show that the BSDF in ET is higher at smaller scattering angles ($\theta_s < 1°$), the discrepancy from our simulations is only O(1), and may be overcome by experimental uncertainties in the estimation of the dust distribution parameters (Sec.2.3.2).

6.1 Future Perspectives

In the future, several improvements and analysis are planned.

Regarding the imaging, different improvements are under study. The first update concerns the dust powder samples used in the calibration of the algorithm: we will use other dust powders of known size but with different materials, keeping in mind that here the important material parameters are the index of refraction and the isotropy. The advantages are two-fold: firstly, we will increase the statistics upon which the algorithm is tested, allowing for a more accurate and precise calibration; secondly, we may use samples that better reflect the optical properties of environmental dust particles, in order to tune the algorithm specifically for the dust particulate that deposits on the wafers (and on the optics). For what concerns the images, we are also planning to upgrade the camera system, allowing us to reach an higher resolution and increase the detection efficiency the small particles ($d < 10\mu$m).

Regarding the dust monitoring campaign, new wafers will be periodically exposed in the different environments to continue the monitoring and increase the statistics both in terms of repeated measurements and in terms of time evolution, and in a more accurate monitoring of specific human activities (e.g. commissioning operations) will be performed to measure their impact on the dust deposit. In addition, the BSDF of the samples will be experimentally measured with a scattering measurements facility to see if the experimental measurement matches the BSDF estimates computed from the measured dust distribution.

Finally, studies will also be performed focusing on Einstein Telescope: there is the need to estimate the amount of stray light from dust contamination that is expected in ET; and from those estimates compute the amount of stray light. Periodic or specific cleaning could be eventually scheduled in order to keep the stray light level under a critical level.

Appendix A

Wafers Locations

In this Appendix section maps with the location of the exposed wafers are provided.

A.1 DET Lab

The Detection Lab (DET Lab) is located inside an ISO 7 clean room in the Virgo Central Building. It hosts the Quantum Noise Reduction (QNR) system that is monitored in the campaign. Inside DET Lab, the FIS, EQB1 and SQB1 environments are located. Locations are labelled as (graphically reported in figure A.1):

- Counter: the corner of the room, near the particle counter
- SDB2: near the SDB1 and SDB2 towers
- EDB: on the cover of EDB

Figure A.1: Photo with highlighted the exposed wafers inside DET Lab. Photos from [21].

A.2 FIS

One of the tested environments was the Frequency Independent Squeezing (FIS) system, which was eventually replaced by the Frequency Dependent Squeezing (FDS) system after O3. The FIS bench was located inside the DET Lab. The locations of wafers are labelled as (graphically reported in figure A.2):

- WS1 (and WS2): on the bench
- SQZH (and SQZV) on the bench, horizontally, H (and vertically V).

- Plexi: on the bench, but the wafer is on its plexiglass cover

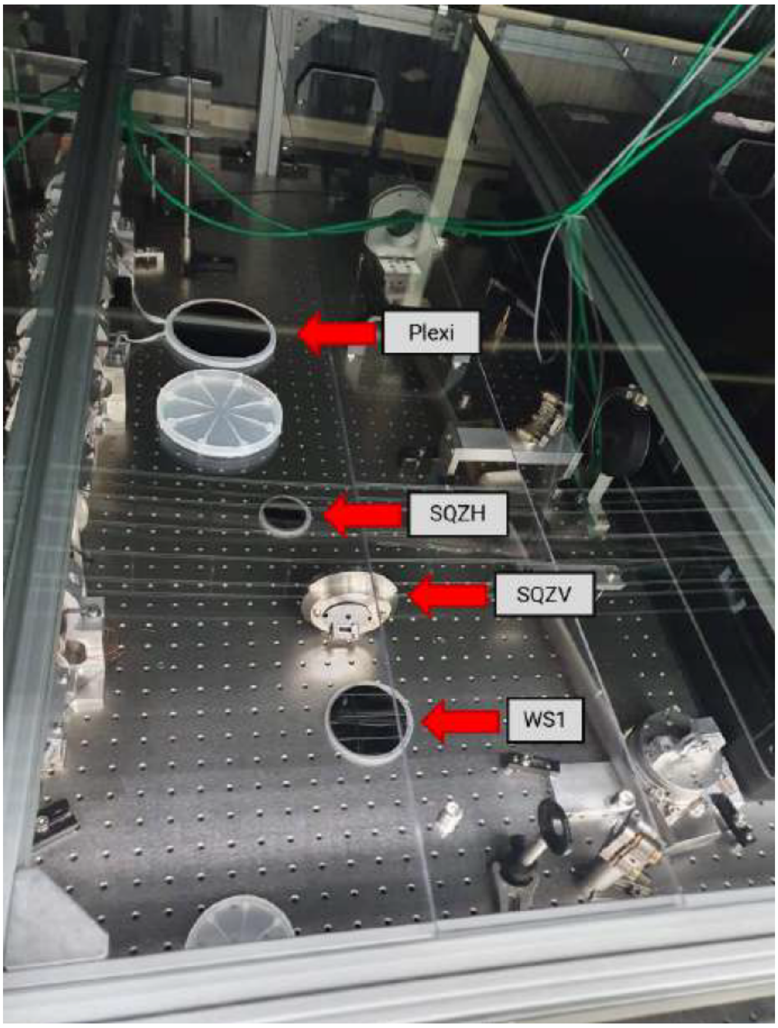

Figure A.2: Photos with highlighted the exposed wafers on the FIS bench. Photo from [21].

A.3 EQB1

The EQB1 bench is an optical bench located inside the DET Lab, in Virgo. It is an in-air bench which is part of the Quantum Squeezing System. In figure A.3 the position of the wafers exposed on the EQB1 bench are reported. In particular, they are placed (graphically reported in figure A.3)

- on top of the squeezing box (On SQZ), Horizontally (H)
- near the squeezing box (Near SQZ), both H and Vertically (V)
- approximately in the middle of the bench (rectangular part of EQB1), both H and V

A.4 SQB1

SQB1 is a suspended and in-vacuum optical bench in Virgo, installed inside the DET Lab. The wafers exposed in this bench were positioned in specific locations, labelled T1,T2,L1,L2, as was reported, for example, in figure 5.6. In figure A.4, a scheme of the SQB1 bench with indicated the wafer positions is reported.

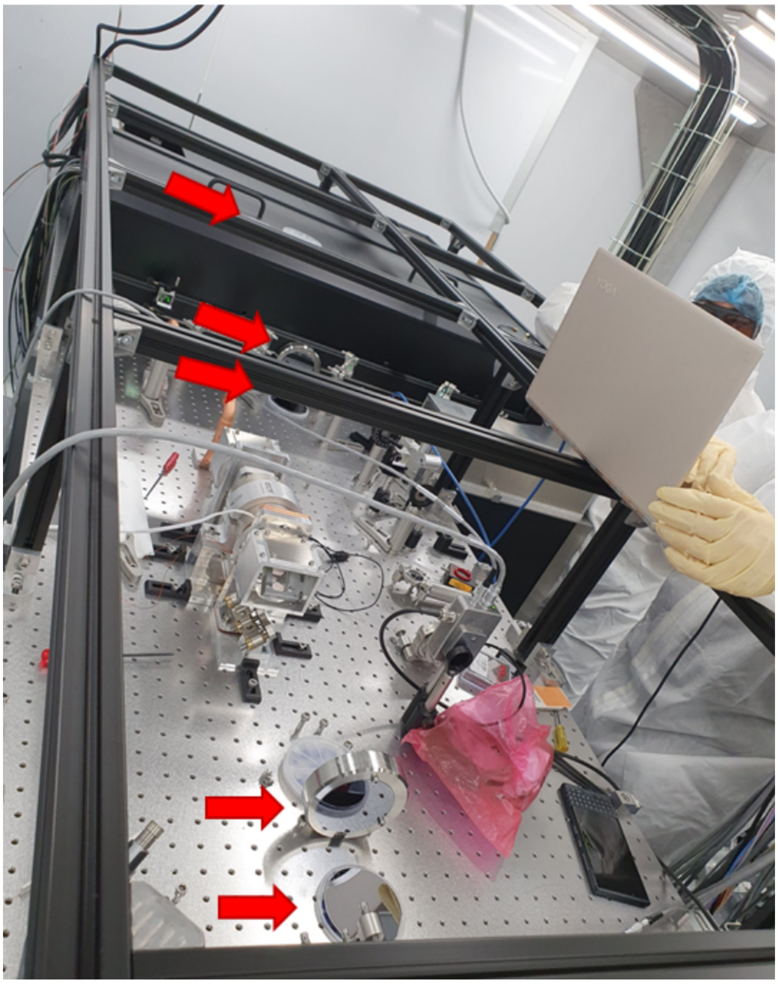

Figure A.3: Photo with highlighted the exposed wafers on the EQB1 bench. Photo from [57].

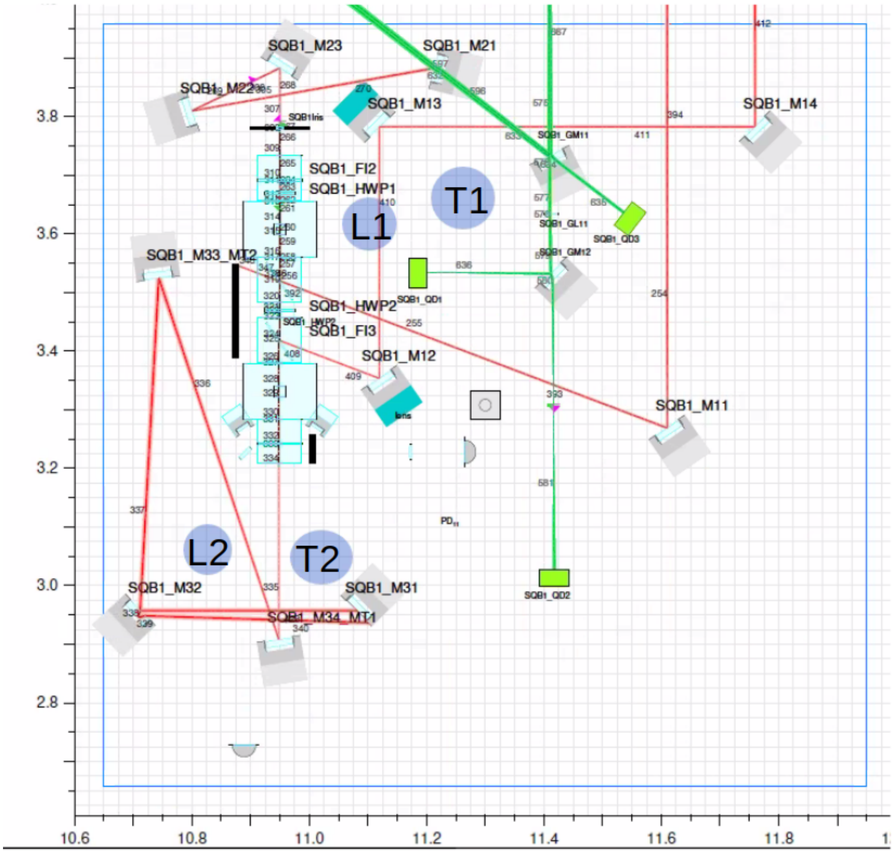

Figure A.4: Scheme of the SQB1 bench where are indicated the positions of the exposed wafers (T1,T2,L1,L2). Image from [47].

Appendix B

Dust Distributions

In this Appendix chapter, the dust distributions reconstructed from the exposed samples are reported. The distributions are reported as count density histograms because of the non homogeneous bin sizes. In fact, from the effective diameter equation 4.12, it can be seen that, while the area is quantized in terms of $\Delta A = \pm 1\text{px}^2$, the diameter quantization is non linear, since $d \propto \sqrt{A}$. This causes diameters values to be non-equally spaced, and this non-uniform quantization is applied to the distribution representations too.

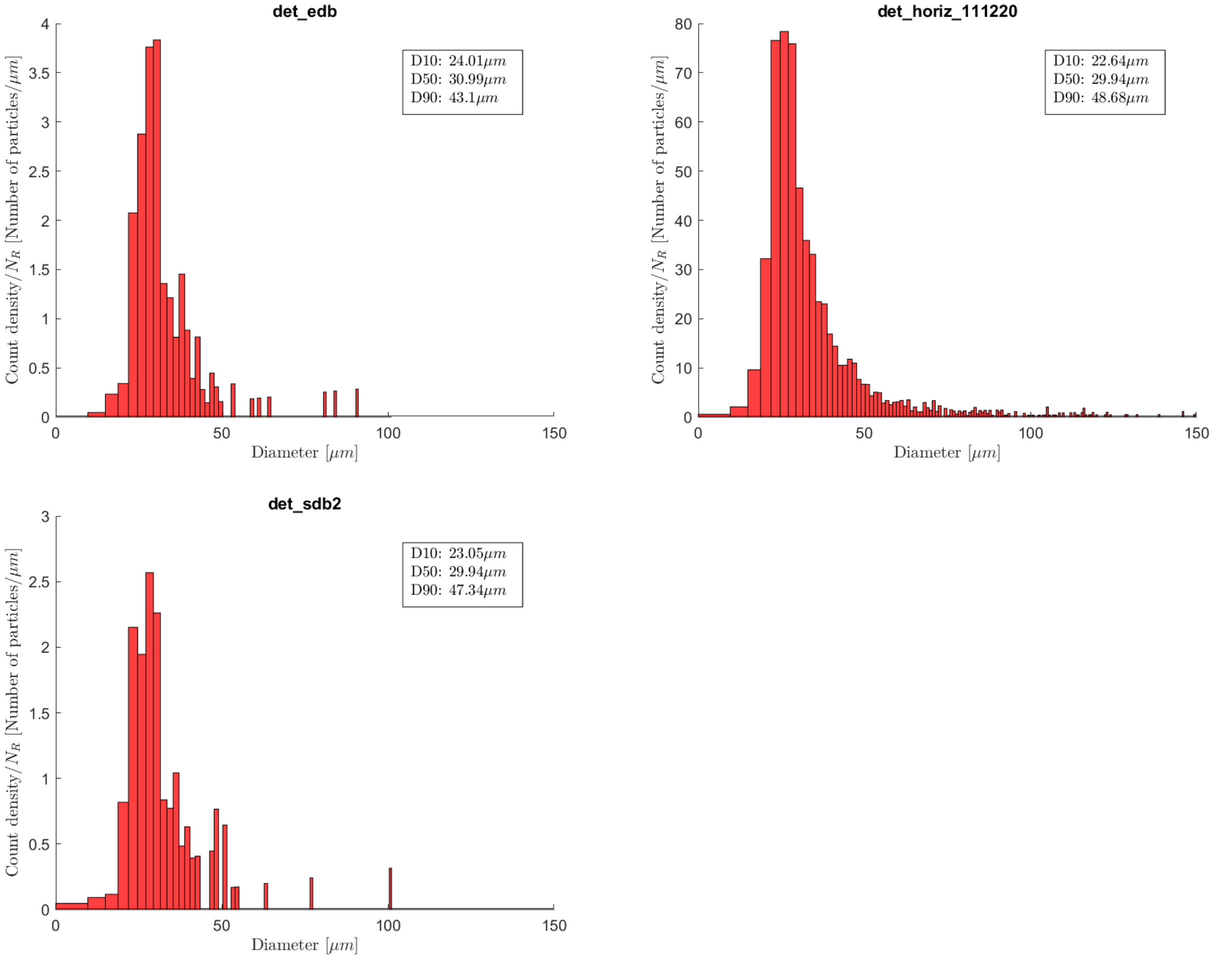

Figure B.1: Reconstructed dust distributions on the sample exposed in the Det Lab. The count densities are normalized over the number of regions used in the analysis, N_R (according to Sec.5.1). The 10th,50th and 90th percentiles are reported in the top-right panel.

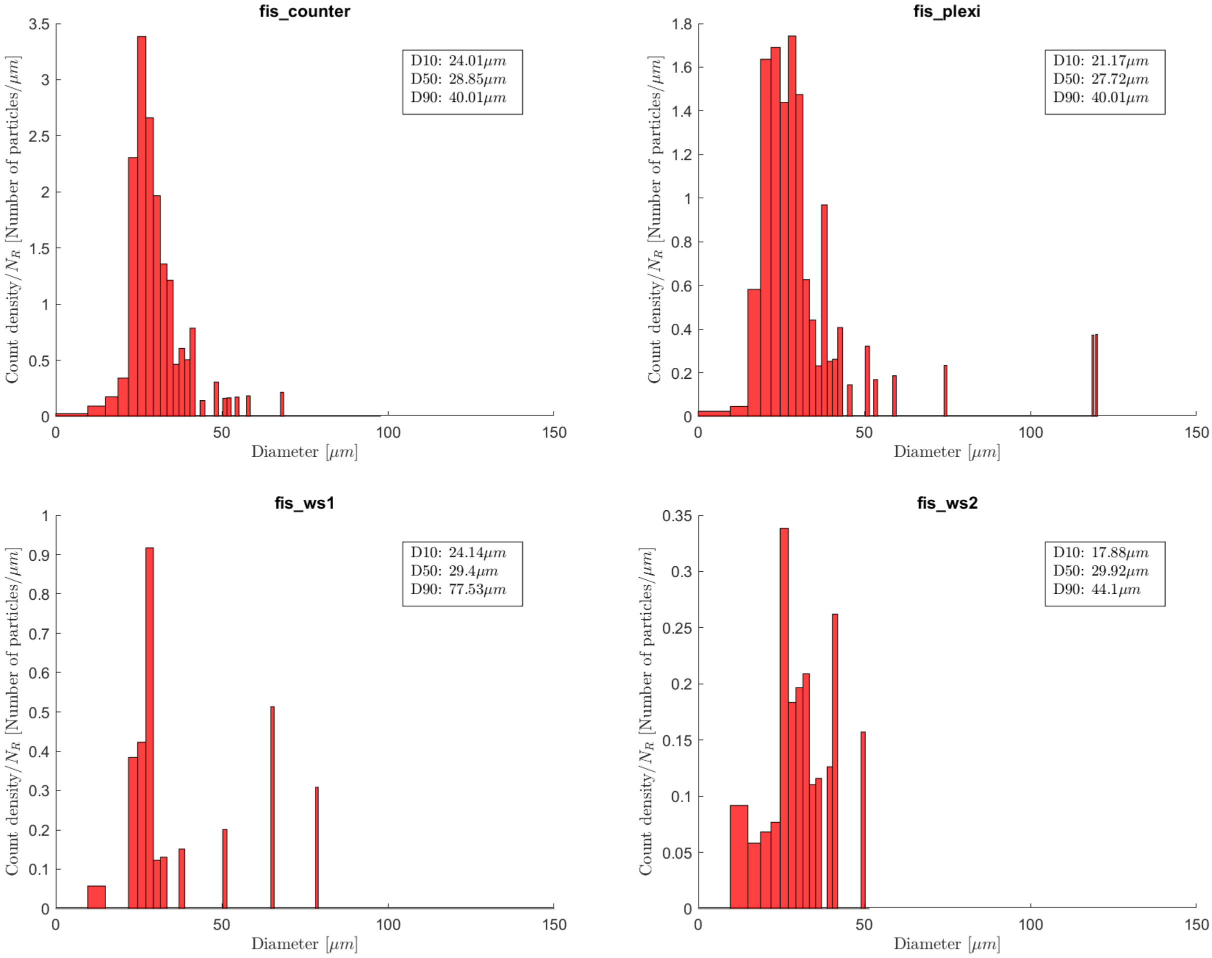

Figure B.2: Reconstructed dust distributions on the sample exposed in the FIS Lab. The count densities are normalized over the number of regions used in the analysis, N_R (according to Sec.5.1). The 10th,50th and 90th percentiles are reported in the top-right panel.

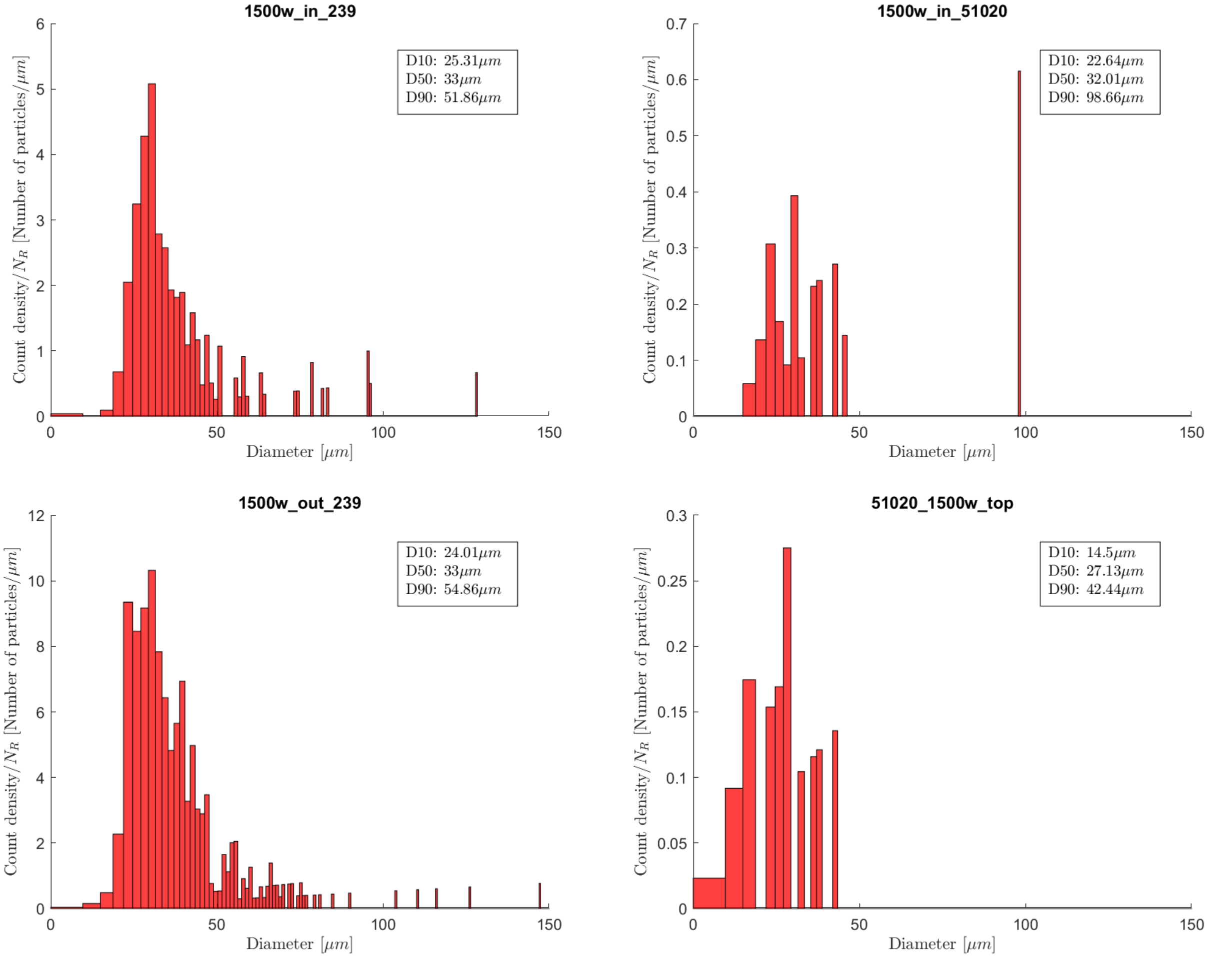

Figure B.3: Reconstructed dust distributions on the sample exposed in the 1500W Lab. The count densities are normalized over the number of regions used in the analysis, N_R (according to Sec.5.1). The 10th,50th and 90th percentiles are reported in the top-right panel.

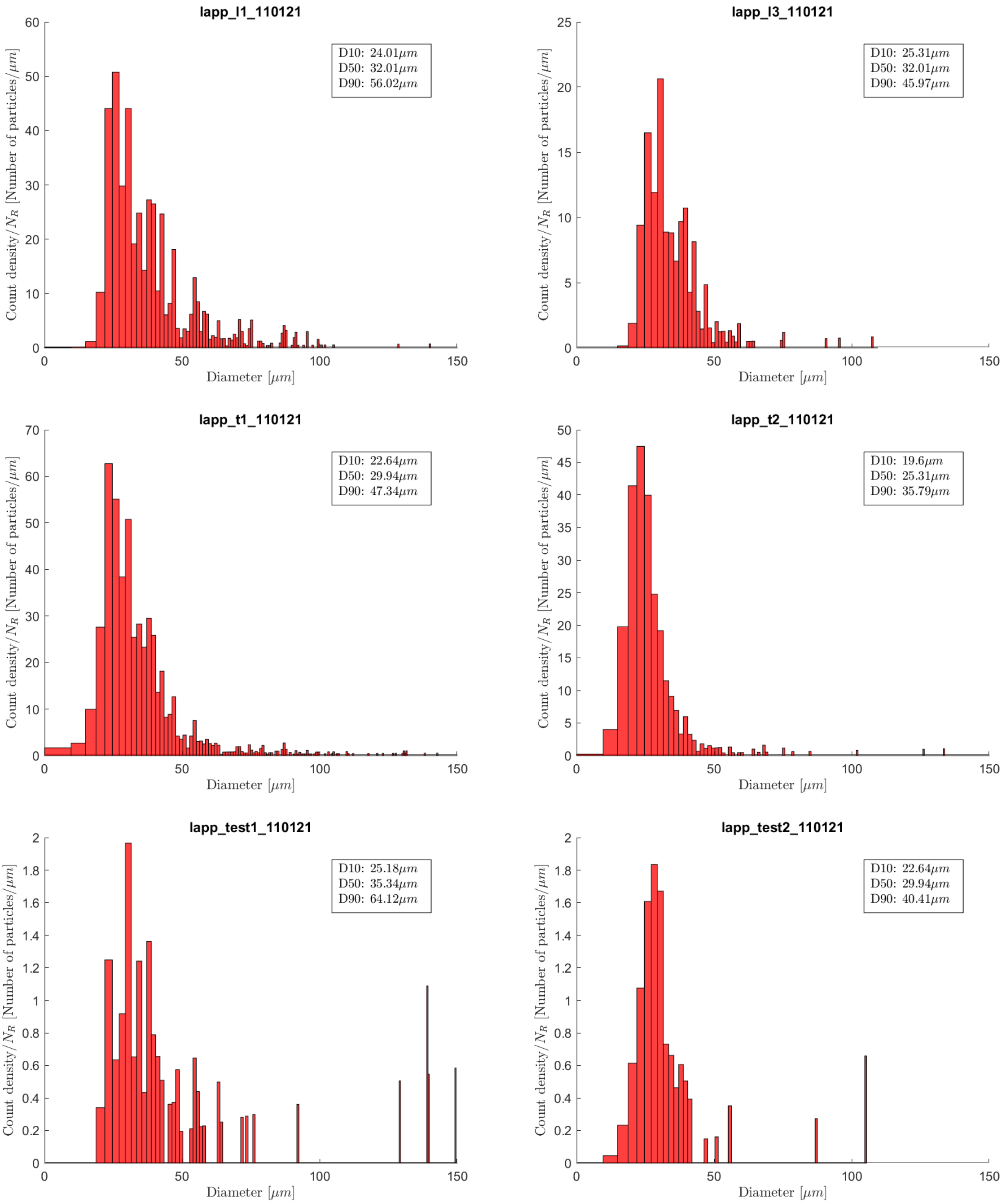

Figure B.4: Reconstructed dust distributions on the sample exposed in the LAPP Lab. The count densities are normalized over the number of regions used in the analysis, N_R (according to Sec.5.1). The 10th,50th and 90th percentiles are reported in the top-right panel.

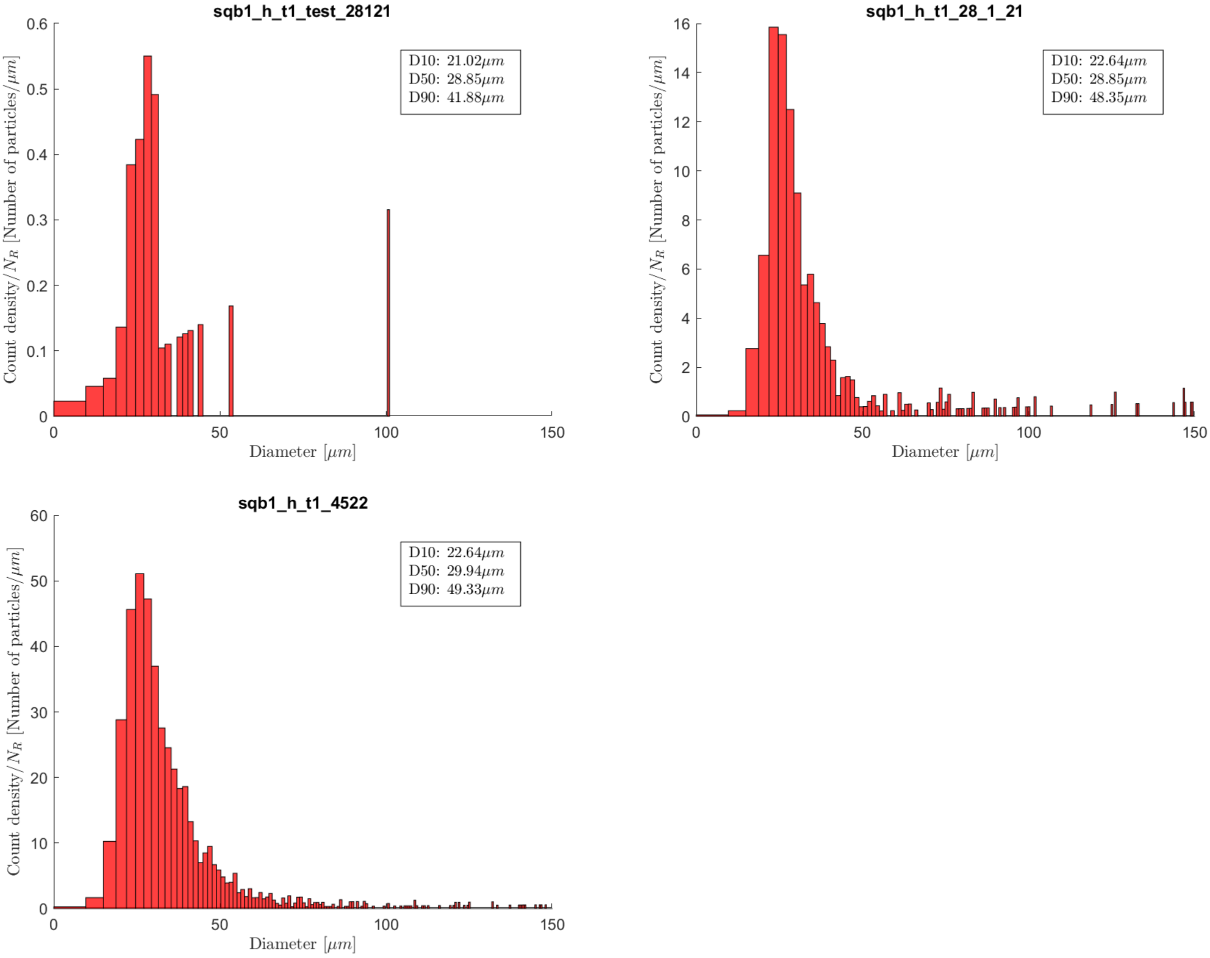

Figure B.5: Reconstructed dust distributions on the sample exposed in the SQB1 bench. The count densities are normalized over the number of regions used in the analysis, N_R (according to Sec.5.1). The 10th,50th and 90th percentiles are reported in the top-right panel.

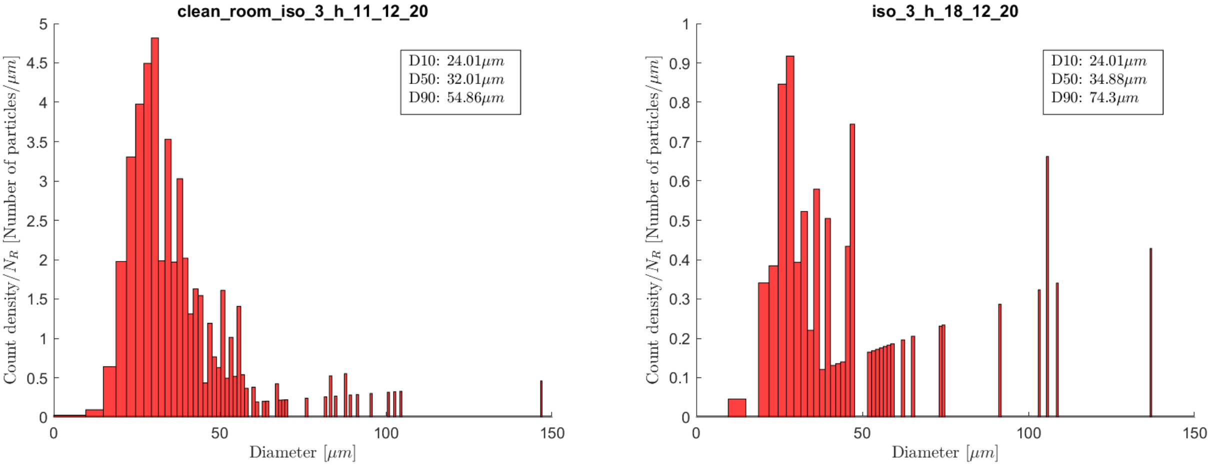

Figure B.6: Reconstructed dust distributions on the sample exposed in the PG Lab between December 2020 and January 2021. The count densities are normalized over the number of regions used in the analysis, N_R (according to Sec.5.1). The 10th,50th and 90th percentiles are reported in the top-right panel.

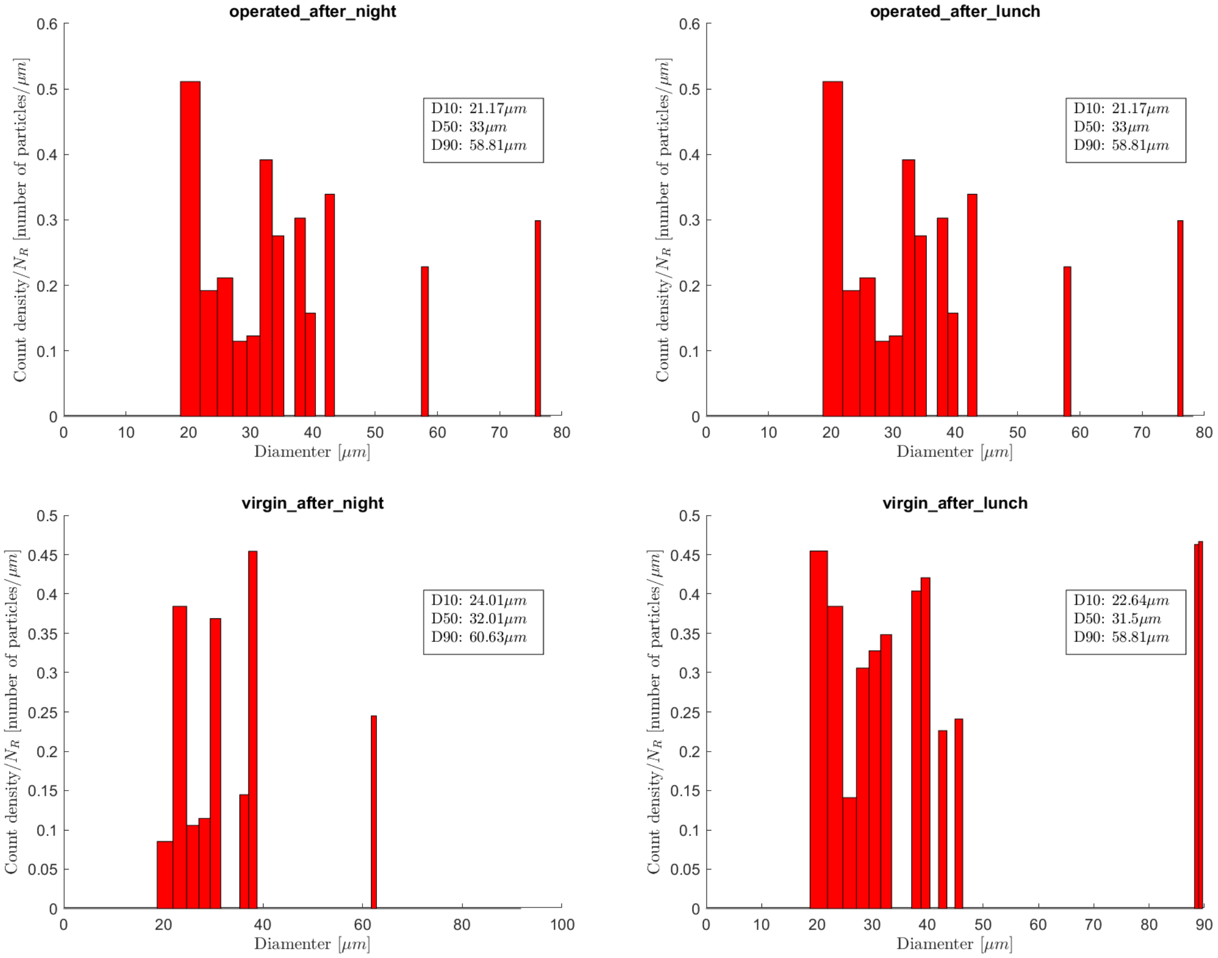

Figure B.7: Reconstructed dust distributions on the sample exposed in the PG Lab in February 2022. The count densities are normalized over the number of regions used in the analysis, N_R (according to Sec.5.1). The 10th,50th and 90th percentiles are reported in the top-right panel.

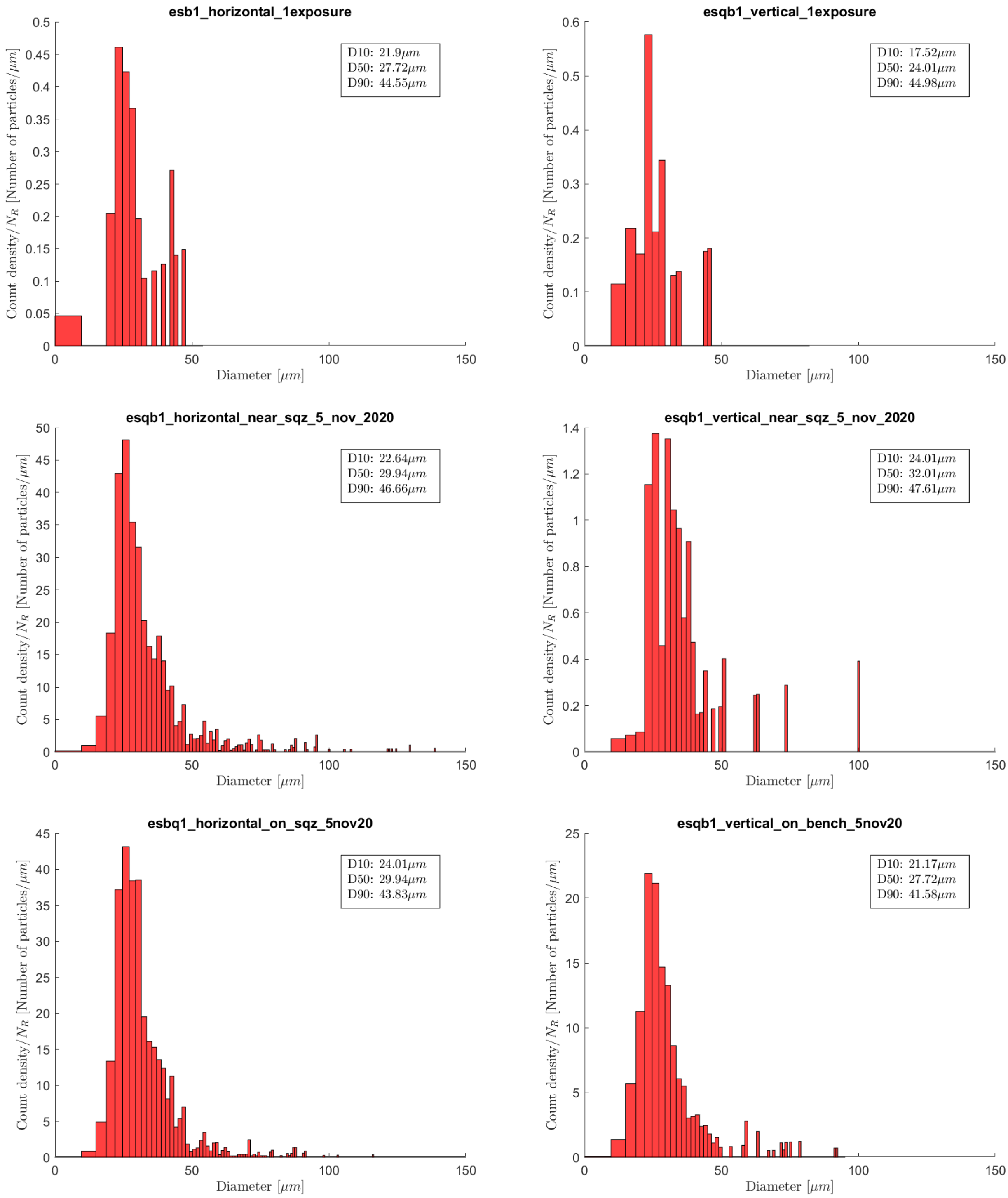

Figure B.8: Reconstructed dust distributions on the sample exposed in the EQB1 bench. The count densities are normalized over the number of regions used in the analysis, N_R (according to Sec.5.1). The 10th,50th and 90th percentiles are reported in the top-right panel.

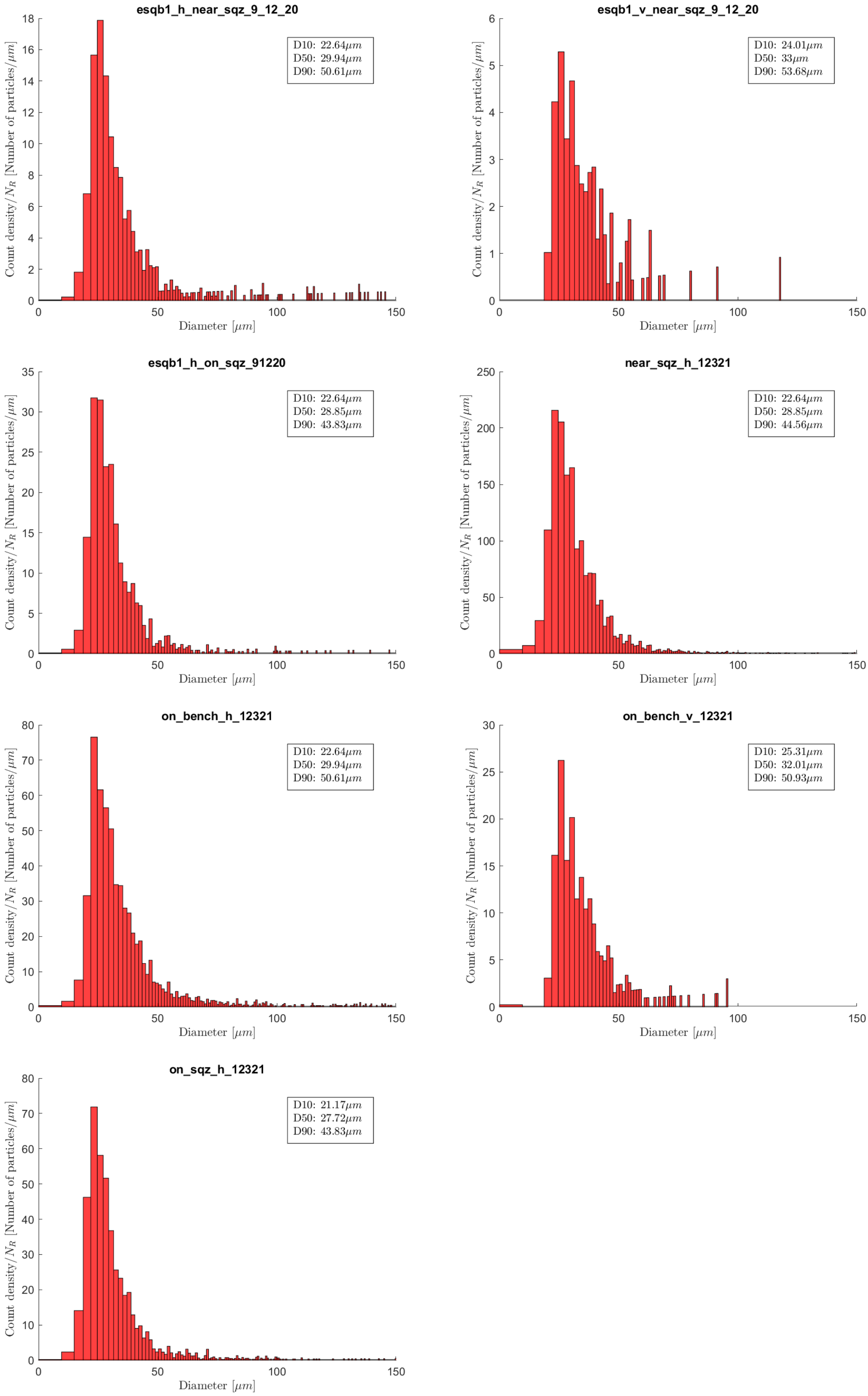

Figure B.9: Reconstructed dust distributions on the sample exposed in the EQB1 bench. The count densities are normalized over the number of regions used in the analysis, N_R (according to Sec.5.1). The 10th,50th and 90th percentiles are reported in the top-right panel.

Appendix C

Linear Fit over the Distributions

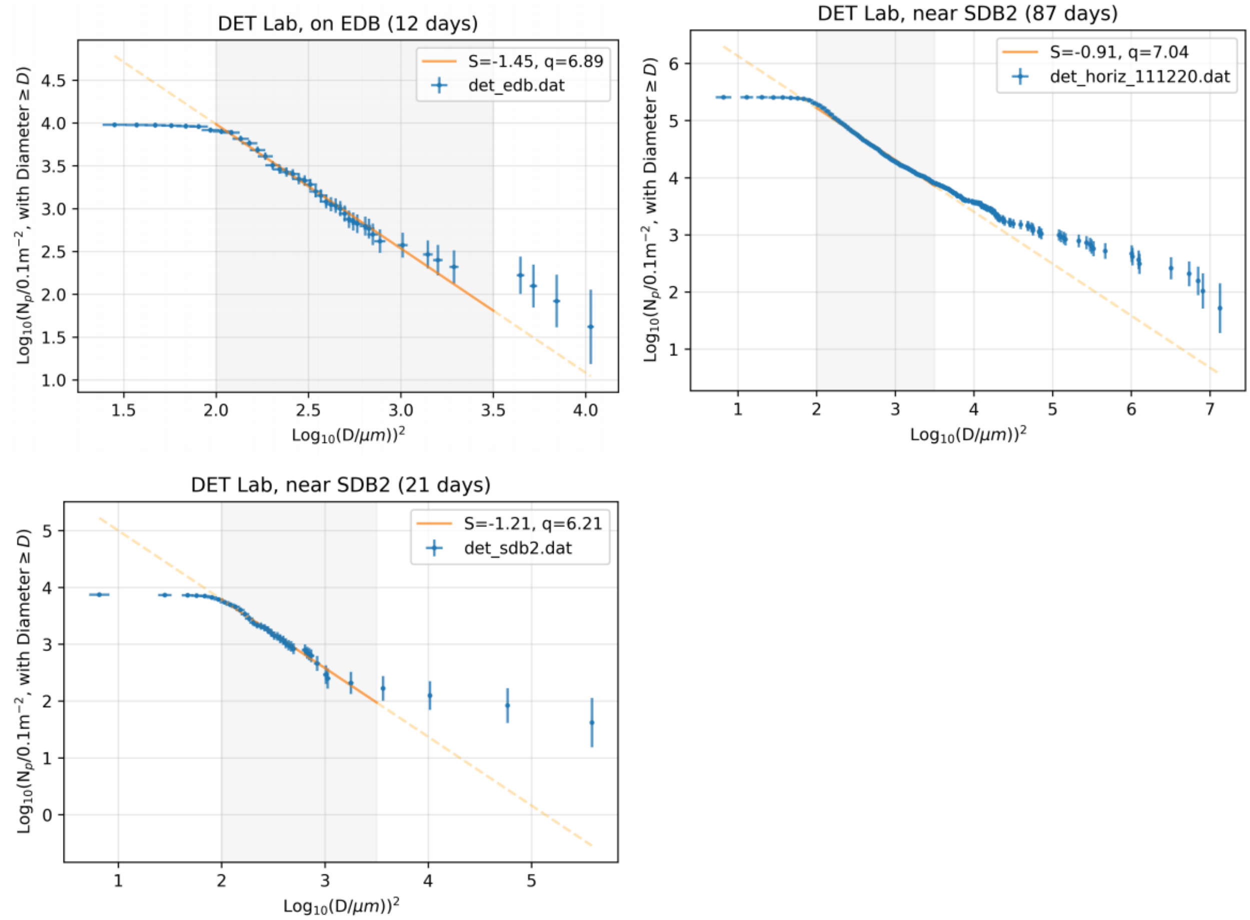

Figure C.1: Linear fit on the $\log - \log^2$ representation of the wafers dust distributions in the DET lab.

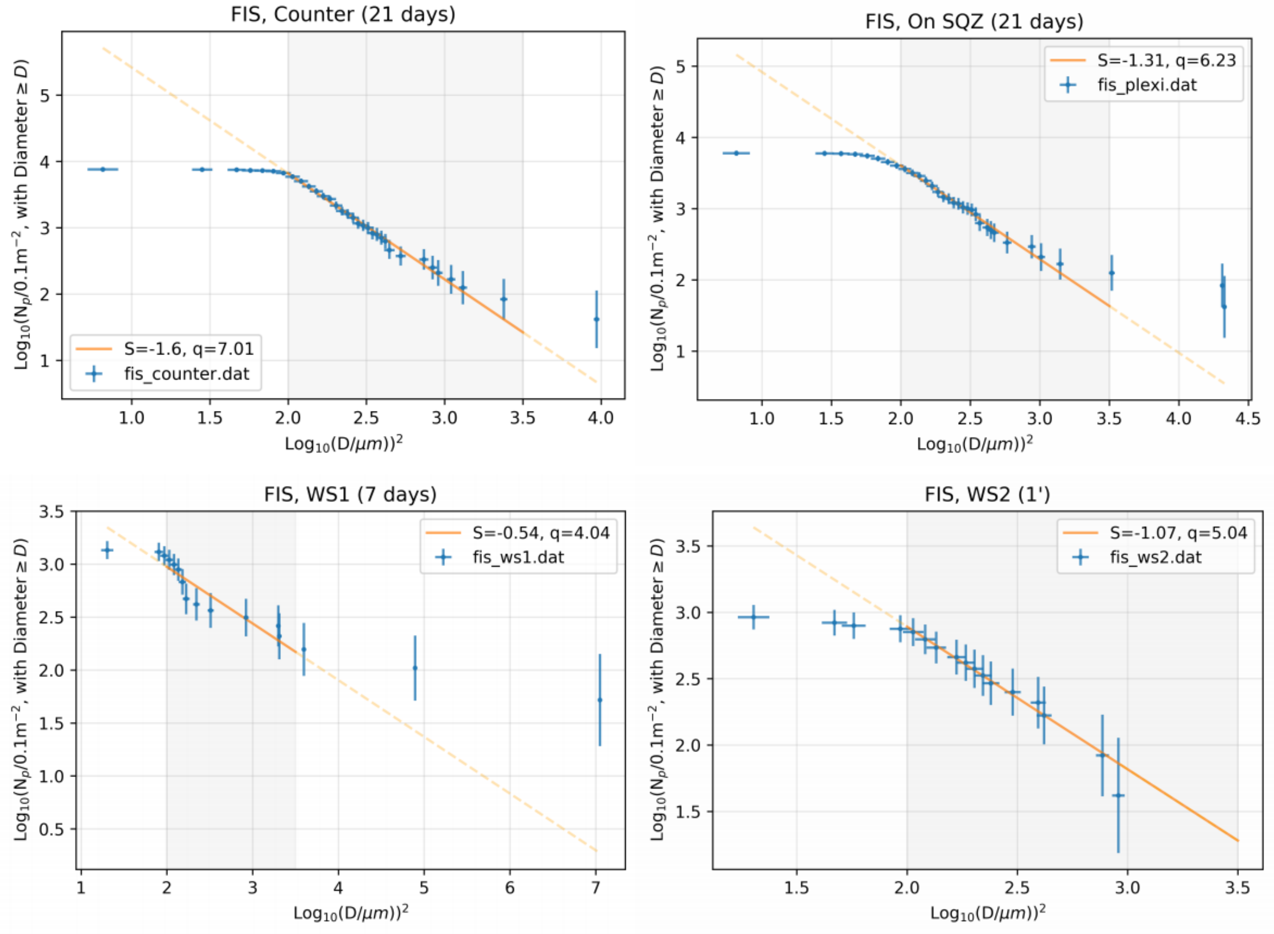

Figure C.2: Linear fit on the $\log - \log^2$ representation of the wafers dust distributions in the FIS lab.

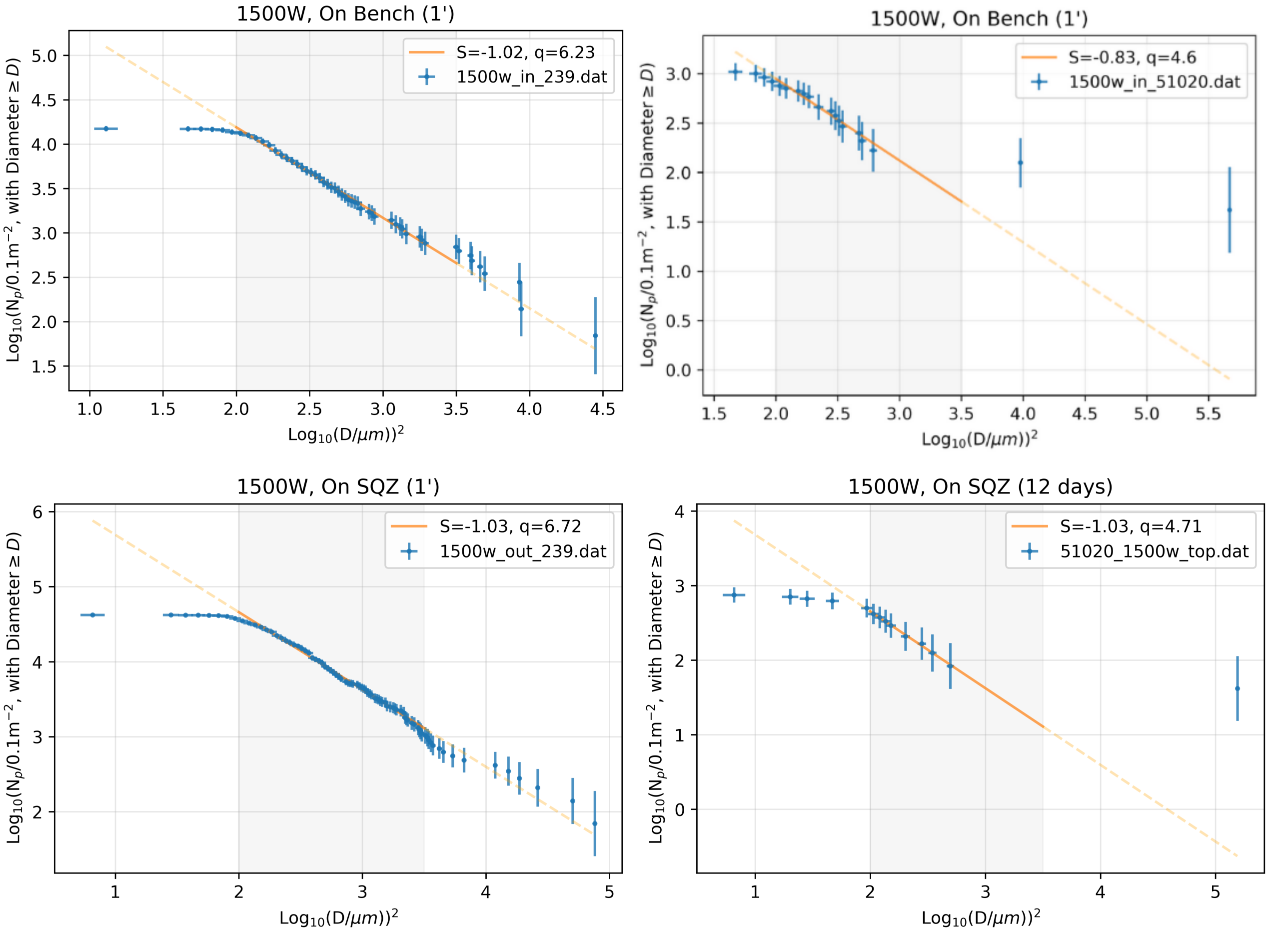

Figure C.3: Linear fit on the $\log - \log^2$ representation of the wafers dust distributions in the 1500W lab.

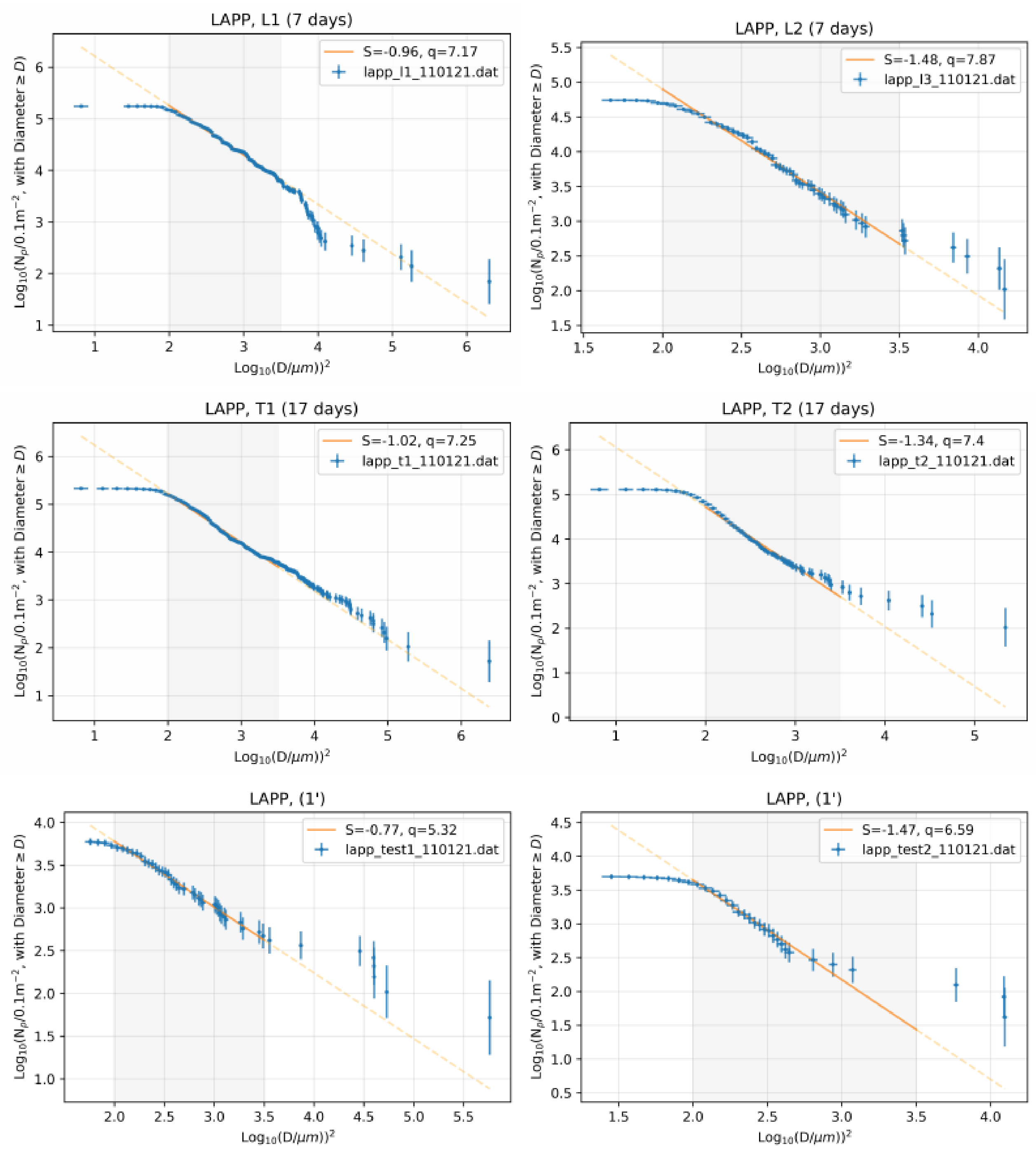

Figure C.4: Linear fit on the $\log - \log^2$ representation of the wafers dust distributions in the LAPP lab.

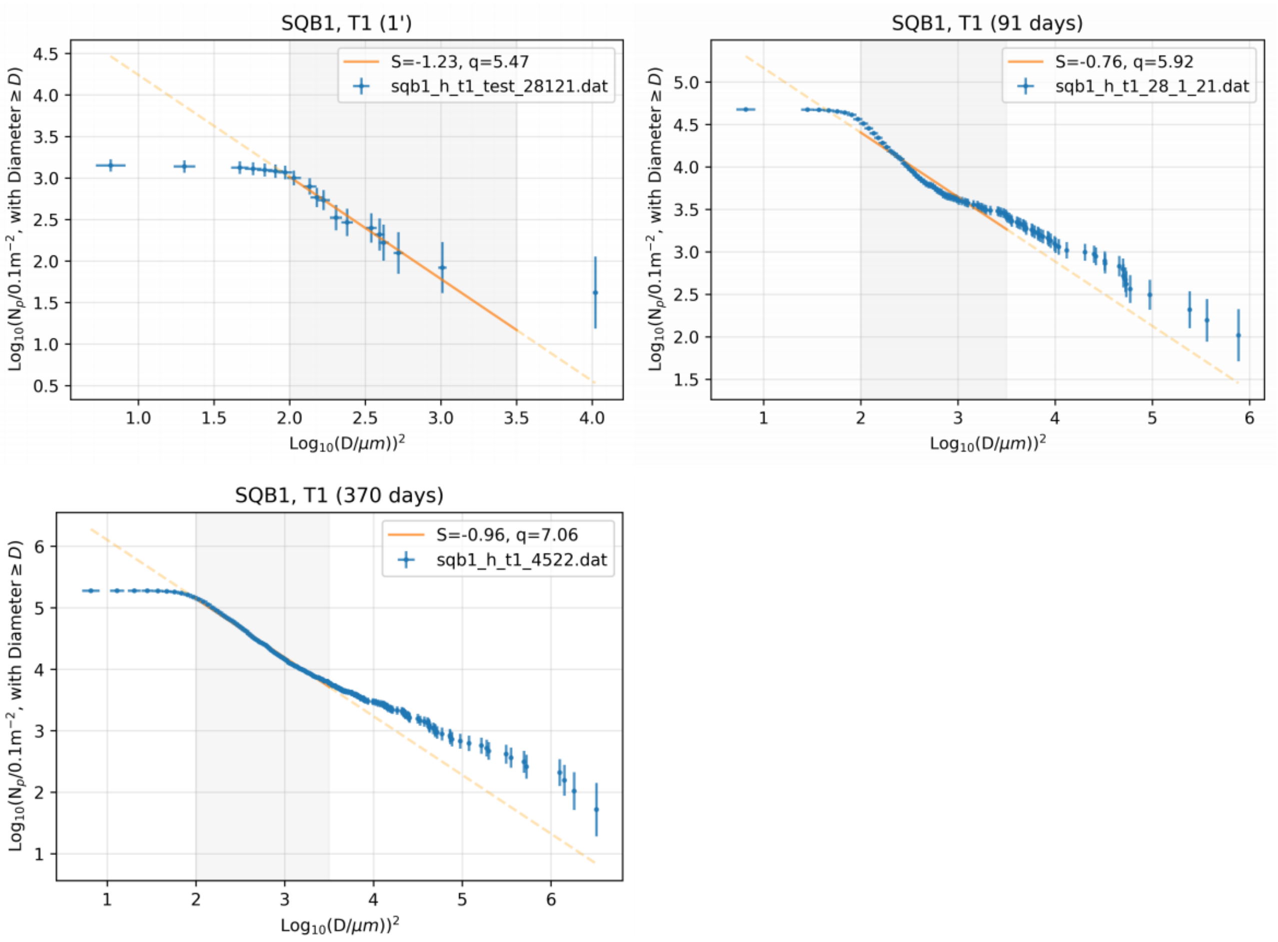

Figure C.5: Linear fit on the $\log - \log^2$ representation of the wafers dust distributions in the SQB1 bench.

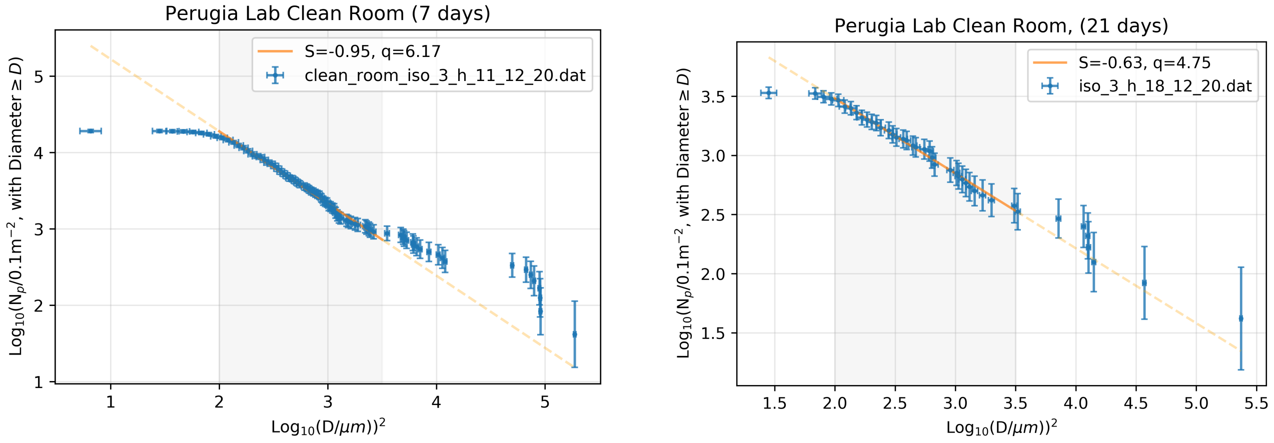

Figure C.6: Linear fit on the $\log - \log^2$ representation of the wafers dust distributions in the PG Lab between December 2020 and January 2021.

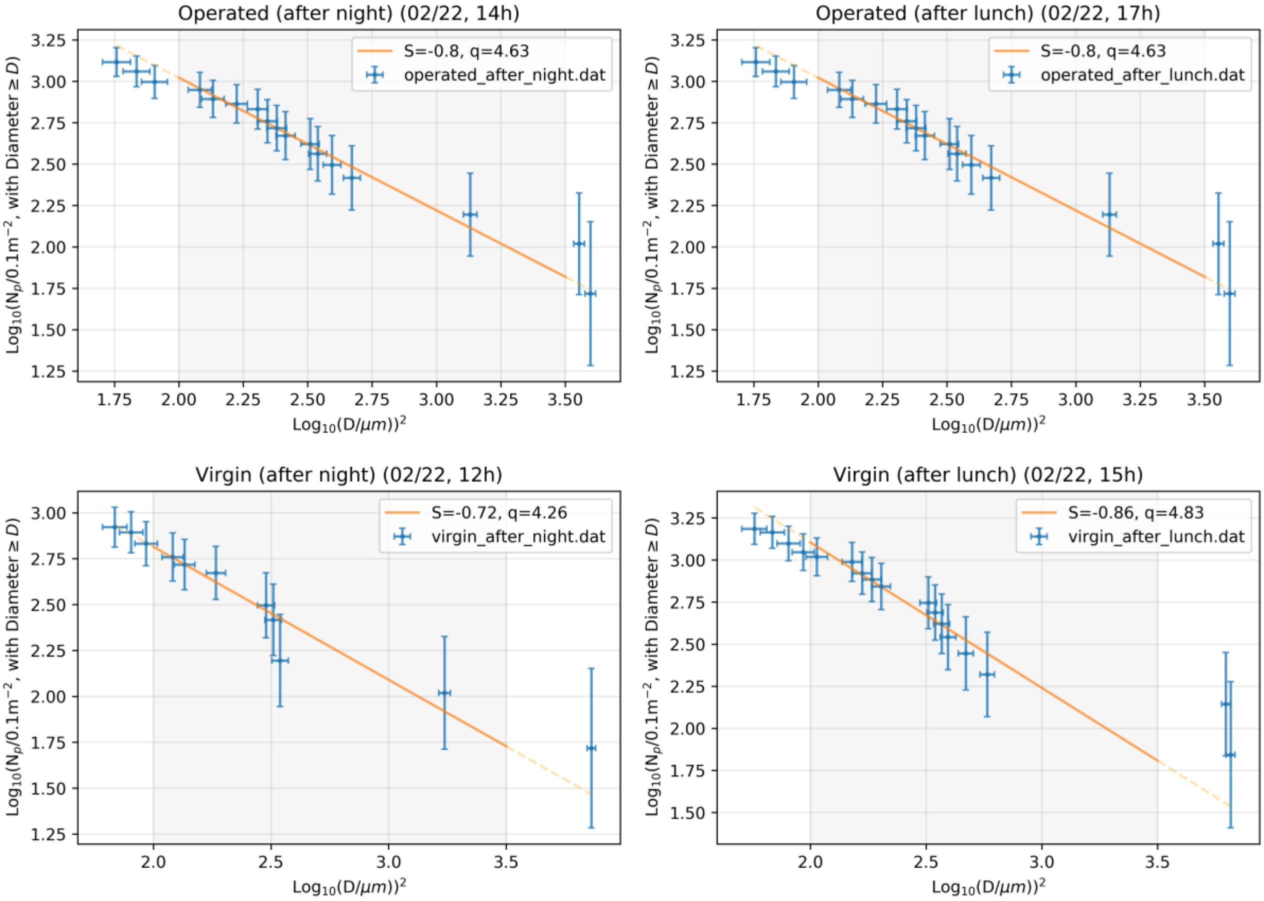

Figure C.7: Linear fit on the $\log - \log^2$ representation of the wafers dust distributions in the PG Lab in February 2022.

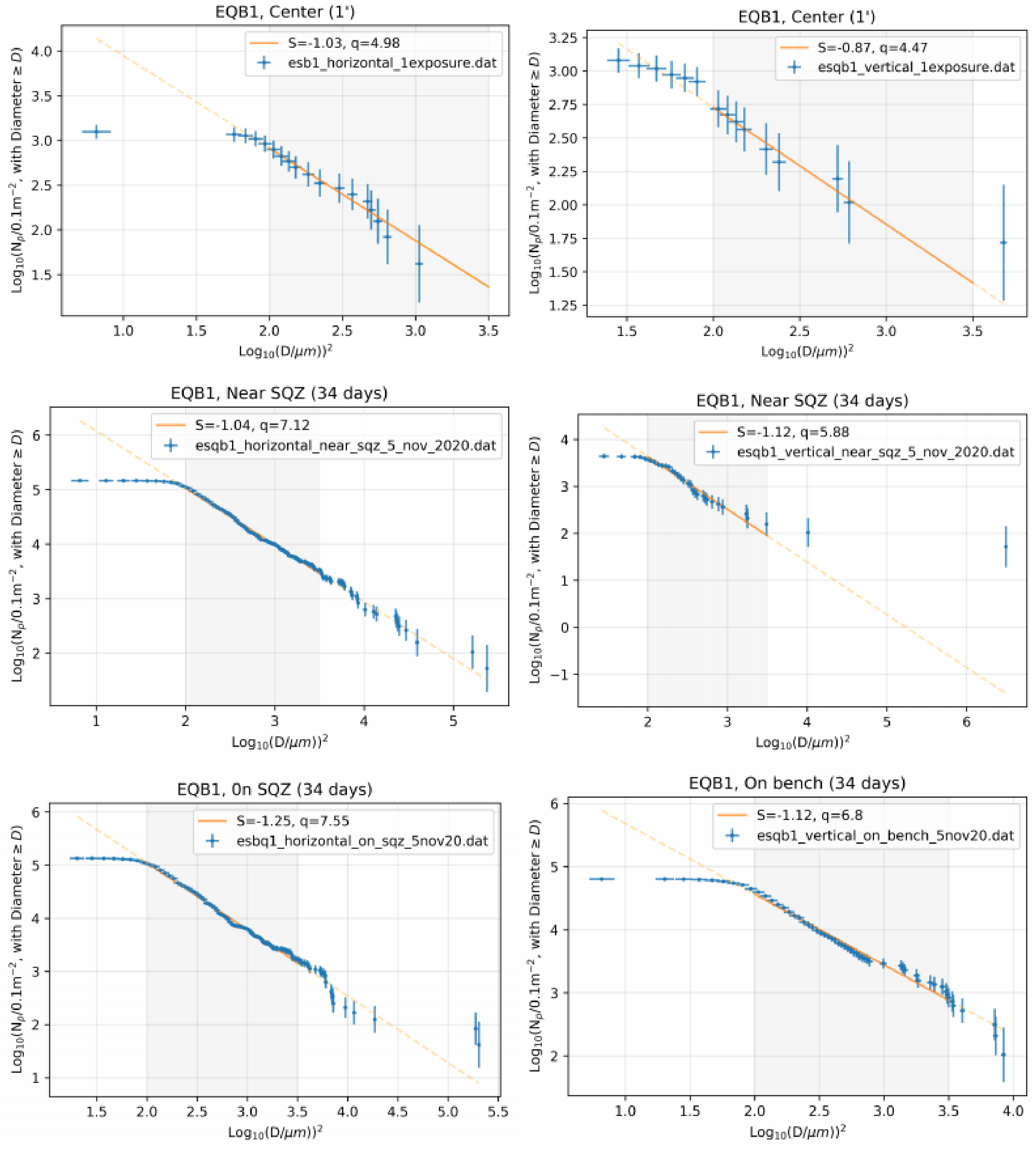

Figure C.8: Linear fit on the $\log - \log^2$ representation of the wafers dust distributions in the EQB1 bench.

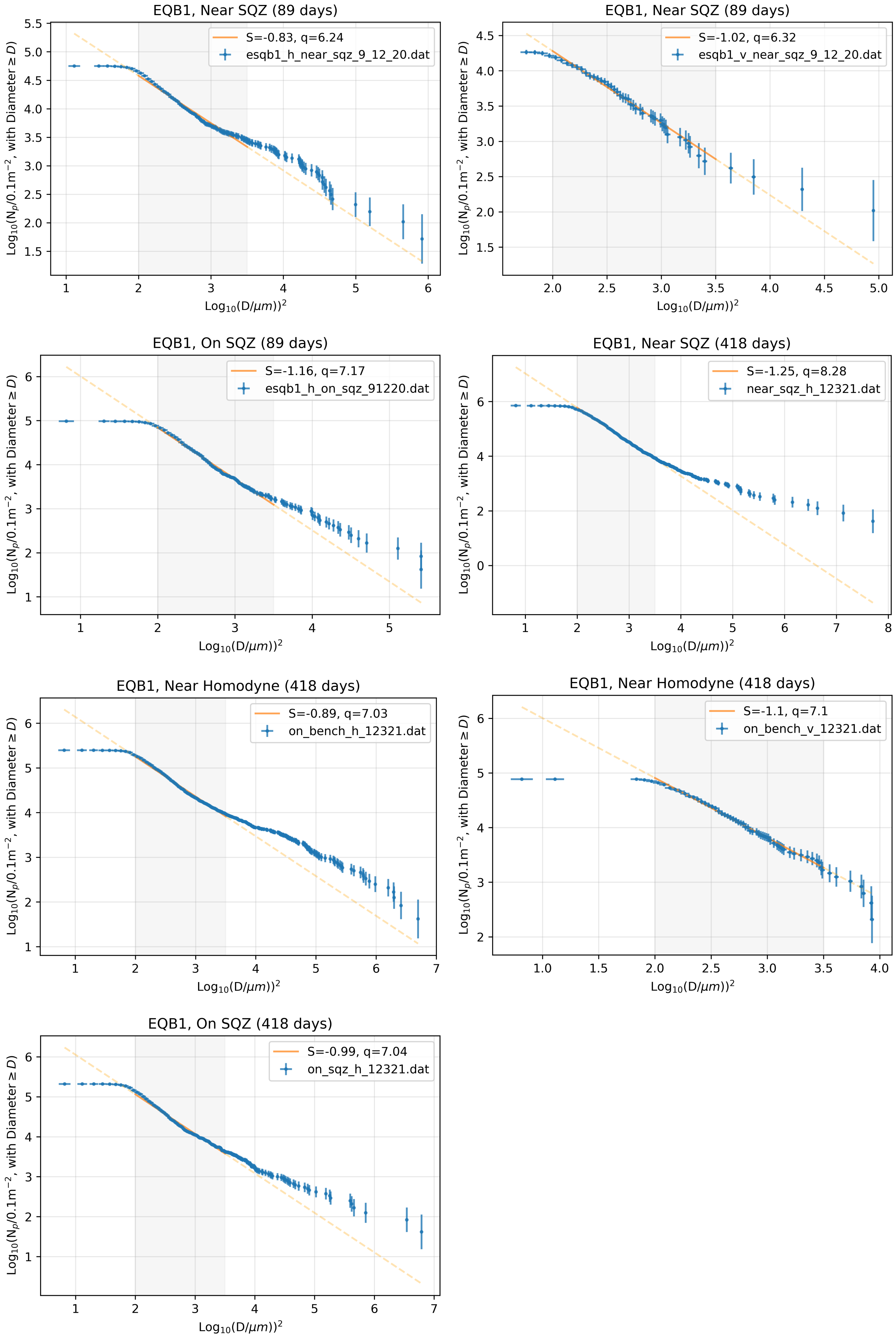

Figure C.9: Linear fit on the $\log - \log^2$ representation of the wafers dust distributions in the EQB1 bench

Bibliography

[1] M. Maggiore, *Gravitational Waves: Volume 1: Theory and Experiments.* Oxford University Press, USA, 2007.

[2] Abbott, B. P. et al. (LIGO Scientific Collaboration), "Observation of gravitational waves from a binary black hole merger," *Phys. Rev. Lett.*, vol. 116, p. 061102, Feb 2016.

[3] Abbott, B. P. et al. (LIGO Scientific Collaboration and Virgo Collaboration), "Gw170814: A three-detector observation of gravitational waves from a binary black hole coalescence," *Phys. Rev. Lett.*, vol. 119, p. 141101, Oct 2017.

[4] M. P. Hobson, G. P. Efstathiou, A. N. Lasenby, *General relativity: an introduction for physicists.* Cambridge University Press, 2005.

[5] R. Abbott et al. (LIGO and Virgo collaborations), "Properties and Astrophysical Implications of the 150 solar masses Binary Black Hole Merger GW190521," *The Astrophysical Journal*, vol. 900, p. L13, sep 2020.

[6] "Sinc Function (normalized)." `https://commons.wikimedia.org/wiki/File:Sinc_function_(normalized).svg`.

[7] "Virgo Optical Scheme." `https://commons.wikimedia.org/wiki/File:VirgoOpticalScheme.jpg`.

[8] G.Caini, "Gravitational Physics Course," 2020.

[9] M. Punturo et al., "The Third Generation of Gravitational Wave observatories and their science reach.," vol. 27, p. 084007, apr 2010.

[10] F. Acernese et al. , "Status of Advanced Virgo (EPJ Web of Conferences),"

[11] G. Ballardin et al., "Measurement of the virgo superattenuator performance for seismic noise suppression," *Review of Scientific Instruments*, vol. 72(9), pp. 3643–3652, 2001.

[12] G. I. González and P. R. Saulson, "Brownian motion of a mass suspended by an anelastic wire," *The Journal of the Acoustical Society of America*, vol. 96, no. 1, pp. 207–212, 1994.

[13] F. Piergiovanni, "The research on amorphous coatings for future gw detectors.," *Journal of Physics: Conference Series*, vol. 1468(1):012216, 2020.

[14] S. Gozzo, "Measurement of ultra-low optical absorption in mirror substrates for the next-generation gravitational wave detectors (Master Degree Thesis)." `http://hdl.handle.net/20.500.12608/21218`.

[15] M. Vardaro, "The quantum enhanced virgo interferometer." VIR-0632A-19, 2019.

[16] LIGO-Virgo-Kagra Collaborations, "GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run," 11 2021.

[17] "The nobel prize in physics 2017." `https://www.nobelprize.org/prizes/physics/2017/summary/`.

[18] B. P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration), “Multi-messenger observations of a binary neutron star merger,” *The Astrophysical Journal*, vol. 848, p. L12, oct 2017.

[19] R. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration), “Observation of Gravitational Waves from Two Neutron Star-Black Hole Coalescences,” *The Astrophysical Journal Letters*, vol. 915, p. L5, jun 2021.

[20] S. H. et al., “Sensitivity studies for third-generation gravitational wave observatories,” *Classical and Quantum Gravity*, vol. 28, p. 094013, apr 2011.

[21] B.D’Angelo, “Numerical Simulations of Stray Light in Virgo.” PhD Thesis, 2022.

[22] M. Ciardelli, V. Dattilo, N. Menzione, “Information to users regarding correct use and safety in Clean Rooms.” Virgo TDS: VIR-1128A-21.

[23] E. Fest, *Stray Light Analysis and Control (SPIE Press PM229).* Press Monograph, SPIE Press, 2013.

[24] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis, *Geometrical considerations and nomenclature for reflectance.* 1977.

[25] D. R. H. Craig F. Bohren, *Absorption and scattering of light by small particles.* Wiley science paperback series, Wiley-VCH, 1998.

[26] P. R. Spyak and W. L. Wolfe, “Scatter from particulate-contaminated mirrors. part 1: theory and experiment for polystyrene spheres and wavelength = 0.6328 microns,” *Optical Engineering*, vol. 31, no. 8, pp. 1746 – 1756, 1992.

[27] CBS INC., “FRED webpage.” `https://technixbycbs.com/page/fred-overview/`.

[28] Benjamin Sumlin, “PyMieScatt documentation.” `https://pymiescatt.readthedocs.io/en/latest/index.html#install-pymiescatt`.

[29] “MiePython documentation.” `https://miepython.readthedocs.io/en/latest/`.

[30] Wikipedia, “Bidirectional Scattering Function.” `https://en.wikipedia.org/wiki/Bidirectional_scattering_distribution_function`.

[31] Institute of Environmental Sciences and Technology (IEST), *IEST-STD-CC1246: Product Cleanliness Levels - Applications, Requirements, and Determination.* 2013.

[32] O. Hamberg and E. Shon, “Particle size distribution on surfaces in clean rooms,” 1984.

[33] P. R. Spyak and W. L. Wolfe, “Scatter from particulate-contaminated mirrors. part 4: properties of scatter from dust for visible to far-infrared wavelengths,” *Optical Engineering*, vol. 31, pp. 1775–1784, 1992.

[34] J. D. Buch and M. K. Barsh, “Analysis of particulate contamination buildup on surfaces,” 1987.

[35] C. Lipeles, “How Classification Impacts the Design of Cleanroom.” `https://www.rdworldonline.com/how-classification-impacts-the-design-of-a-cleanroom/.`, 2018.

[36] Institute of Environmental Sciences and Technology (IEST), “ISO 14644-1:2015, Part 1: Classification of air cleanliness by particle concentration.” `https://www.iso.org/standard/53394.html`, Dec 2015.

[37] S. G. Jennings, R. G. Pinnick, and H. J. Auvermann, “Effects of particulate complex refractive index and particle size distribution variations on atmospheric extinction and absorption for visible through middle ir wavelengths,” *Appl. Opt.*, vol. 17, pp. 3922–3929, Dec 1978.

[38] G.Ciani and L.Conti; private communications.

[39] F.Sorrentino, "Optical system for dust contamination monitoring via Si wafer inspection (CRQ2017/008)." Virgo internal document VIR-0463A-17, 2017.

[40] F.Sorrentino, "Surface dust measurements with a camera system." Virgo internal documento VIR-0944A-17, Dec 2017.

[41] PixeLINK. `https://pixelink.com/products/industrial-cameras/usb-30/23-sensors/pl-d729/`.

[42] E. O. Lens. `https://www.edmundoptics.com/p/35mm-fl-ultra-high-resolution-fixed-focal-length-23305/`.

[43] Advanced Illumination. `https://www.advancedillumination.com/lighting-education/bright-field-dark-field-lighting/`.

[44] Electro Optical Systems Finland Oy, "EOS Aluminum AlSi10Mg." Mill Test Certificate. Inspection certificate, EN10204, type 3.1., 2017.

[45] Electro Optical Systems Finland Oy, "EOS Aluminum Ti64." Mill Test Certificate. Inspection certificate, EN10204, type 3.1, 2017.

[46] Tekna Advanced Materials INC., "Tekmat Mo-25." Material Test Certificate, 2016.

[47] B. D'Angelo, E. Polini, "Virgo logbook entry 50608." `https://logbook.virgo-gw.eu/virgo/?r=50608`.

[48] L. Naticchioni, F. Paoletti, I. Fiori, V. Mangano, "Seismic noise at the clean room floor - 1500WA building (VIR-0575A-22)." `https://tds.virgo-gw.eu/ql/?c=18221`.

[49] S. Waller, "Process for aLIGO Particle Imaging Using a Nikon D7100 and Matlab." LIGO-T1400515-v9, July 2015.

[50] A.Moscatello, G.Ciani, L.Conti, "Improving the algorithm for finding and sizing dust particles (vir-0475a-22)." `https://tds.virgo-gw.eu/ql/?c=18121`.

[51] Matlab Help, "Bwboundaries." `https://it.mathworks.com/help/images/ref/bwboundaries.html`.

[52] Matlab Help, "Regionprops." `https://it.mathworks.com/help/images/ref/regionprops.html`.

[53] Matlab Help, "Imopen." `https://it.mathworks.com/help/images/ref/imopen.html`.

[54] Matlab Help, "Imopen." `https://it.mathworks.com/help/images/ref/im2bw.html`.

[55] Matlab Help, "Image Coordinate Systems." `https://it.mathworks.com/help/images/image-coordinate-systems.html`.

[56] B. D'Angelo, F.Sorrentino, G.Ciani, L.Conti, "Virgo logbook entry 49870." `https://logbook.virgo-gw.eu/virgo/?r=49870`.

[57] B. D'Angelo, G.Ciani, "Virgo logbook entry 50211." `https://logbook.virgo-gw.eu/virgo/?r=50211`.

[58] B. D'Angelo, F.Sorrentino, "Virgo logbook entry 51094." `https://logbook.virgo-gw.eu/virgo/?r=51094`.

[59] A.Moscatello, L.D.Bonavena, B. D'Angelo, G.Ciani, L.Conti, "Stray Light from Dust Contamination: Image Reconstruction and Monitoring Campaign (Virgo SQZ Meeting presentazion)." Virgo TDS: VIR-0785A-22.

[60] B. D'Angelo, G. Ciani, "Virgo logbook entry 50212." `https://logbook.virgo-gw.eu/virgo/?r=50212`.

[61] B. D'Angelo, "Virgo logbook entry 50270." `https://logbook.virgo-gw.eu/virgo/?r=50270`.

[62] B. D'Angelo, G. Ciani, "Pd-tn logbook entry 951." `http://virgolab.physics.unitn.it/logbook/?r=951`.