On the Issue of an Anomaly Detection Algorithm for Identifying Potentially Generated Entangled Photons

Serhii Aleksieienko, Dmytro Babets, Kamil Wereszczyński and Krzysztof A. Cyran

Abstract—**Detecting entangled photons is critical for quantum technologies but challenging due to weak signals, noise, and rarity. This study proposes a fast, unsupervised anomaly detection algorithm using a multivariate Gaussian distribution to identify potential entangled photon events. Applied to data from spontaneous parametric down-conversion (SPDC) and silicon photomultiplier (SiMP) detectors, it successfully detected rare correlated events, offering a simple, efficient tool for quantum optics.**

I. Introduction

The development of quantum technologies is one of the most promising areas of modern science and technology. Quantum cryptography, quantum computing, and quantum communications open up new horizons for secure data transmission and high-performance computing. One of the cornerstones in the implementation of these technologies is the ability to create and detect entangled photons efficiently. Entangled photons have unique properties due to quantum nonlocality, which makes them indispensable for quantum information transfer tasks [1].

However, detecting entangled photons is associated with a number of technical difficulties. Signals from such photons are often weak and susceptible to noise, and the events of their occurrence are extremely rare. This requires the use of highly sensitive optical systems and intelligent data analysis algorithms. In this context, developing anomaly detection algorithms capable of efficiently identifying events associated with detecting entangled photons is a critical task. The use of advanced data analysis methods contributes to a significant improvement in the characteristics of detection systems and an increase in the reliability of the results. This paper is devoted to the study of methods for solving this problem and assessing the possibility of using the Anomaly Detection Algorithm for Identifying Potentially Generated Entangled Photons.

The structure of the paper is as follows: Section II describes the experimental setup and methodology. Section III presents the anomaly detection approach. Section IV discusses results and analysis. Section V concludes the paper and outlines future research directions.

II. Methodology of Investigation

A. Generation of Entangled Photons

Entangled photons can be generated using nonlinear optical processes such as Spontaneous Parametric Down-Conversion (SPDC) in nonlinear crystals. The generation process involves splitting a high-energy (pump) photon into two lower-energy photons called signal and idle. Such photons have interrelated quantum properties, such as polarization, momentum, and energy, which makes them entangled. The SPDC process requires the use of nonlinear crystals with special optical properties. The most common of these are Beta-Barium Borate (BBO), Lithium Niobate ($LiNbO_3$), Rubidium Titanyl Phosphate (RTP), and Periodically Poled Potassium Titanyl Phosphate (ppKTP). The latter is one of the most efficient devices for generating entangled photon pairs [2]. A high-power laser emits photons that interact with the crystal, randomly splitting them into two photons with double the wavelength. Unlike traditional crystals such as BBO, the ppKTP structure increases the probability of generating entangled states by periodically polarizing the material. As a result of this process, two beams are formed at the crystal output: a strong beam of photons of the original wavelength (e.g., 405 nm, coming from the feed laser) and a weaker beam of photon pairs of 810 nm wavelength, entangled in the polarization channel.

Entangled states arise when orthogonally polarized photons intersect at the same spatiotemporal position. This requires precise calibration of optical paths and the use of high-quality nonlinear crystals. High-gain lasers with short pulses and carefully calibrated optical elements can be used to improve the generation efficiency [3].

B. Detection Principles of Entangled Photons

The basic principle of detecting entangled photons is based on the analysis of their time and polarization correlation. In this case, a polarization beam splitter (PBS) is used to divide the photon flow into two channels depending on the polarization: one channel registers photons with vertical polarization, the other – with horizontal polarization. SiMP (Silicon Photomultiplier) detectors are used to register photons, which have high sensitivity and fast response.

When a photon hits a SiMP detector, an avalanche multiplication of electrons occurs, which generates an electric pulse with an amplitude of about 100 mV and a duration of about 5 ns. These signals are fed to the inputs of an oscilloscope, which records the time characteristics of the signals. Comparison of the time coincidences of pulses in two channels allows us to identify correlated events that potentially indicate the detection of entangled photons.

S. Aleksieienko is with the Department of Mechanical Engineering Technology and Materials Science, Dnipro University of Technology Dnipro, Ukraine ORCID 0000-0003-0320-989X (corresponding author to provide phone: +380504808503; e-mail: aleksieienko.s.v@nmu.one).

Dmytro Babets is with the Department of Applied Mathematics Dnipro University of Technology Dnipro, Ukraine ORCID 0000-0002-5486-9268

K. A. Cyran is with the Department of Computer Graphics, Vision and Digital Systems, Silesian University of Technology Gliwice, Poland ORCID 0000-0003-1789-4939.

K. Wereszczyński is with the Department of Computer Graphics, Vision and Digital Systems, Silesian University of Technology Gliwice, Poland ORCID 0000-0003-1686-472X

C. *Experimental setup*

The studies presented in this paper were conducted on an experimental optical setup created in the photon laboratory of the Silesian University of Technology. The setup diagram is shown in Fig. 1.

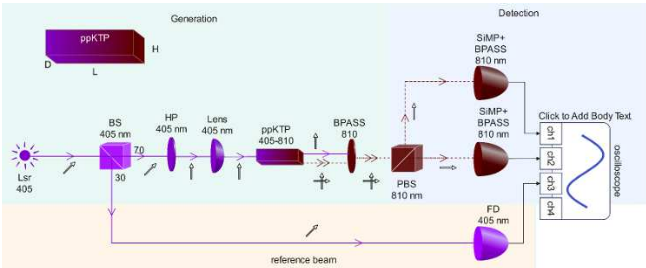


Figure 1. Optical system for detecting entangled photon pairs

The thin arrows show the light beams. The color of the arrows and devices is related to the wavelength of the light – purple means 405 nm, and brown 810 nm, the ppKTP crystal has a gradient fill from purple to brown, as it converts 405 nm photons into pairs of entangled 810 nm photons. The arrows next to the beams show the polarization of the photons: a slanted arrow means that the polarization is unknown, and vertical and horizontal arrows mean their polarization. A single arrow means no polarization pairs and crossed arrows mean an entangled pair. Table 1 provides a list of the symbols used in the diagram.

TABLE I. TABLE TYPE STYLES

Symbols	Device
BS	Beam Splitter
BPASS	Band-Pass Filter, 810±5nm
HP	Half Plate
PBS	Polarizing Beam Splitter
ppKTP	Periodically Pooled Potassium Titanyl Phosphate
SiMP	Silicon Multiplier Amplified Detector

The optical system includes three key components, each of which performs a specific function in the process of photon generation, detection and synchronization (shown in the diagram by three colored rectangles).

The first stage is the generation of entangled photon pairs. This process is initiated by a laser emitting pulses with a duration of 129 ns. The photons generated at this stage have a wavelength of 405 nm and a polarized structure, but their propagation direction remains undefined. A beam splitter (BS) is used to create a reference beam, which allows for the selection of a region of interest (ROI) in subsequent measurements. The polarization of the main beam is adjusted using a half-wave plate (HP) oriented along the H-direction of the ppKTP crystal. The optical system also includes lenses that ensure beam focusing in the ppKTP crystal. As a result of passing through the crystal, two types of photons are present in the beam: (i) photons with a wavelength of 405 nm that do not participate in spontaneous parametric scattering (SPDC) and (ii) pairs of entangled photons with a wavelength of 810 nm. To isolate only entangled pairs, a bandpass filter (BPASS) with a central wavelength of 810 nm and a passband of ± 5 nm is used. At this stage, the generation process is considered complete.

The next stage is detection, which involves measuring the polarization of photons and recording their time correlation. Separation of photon beams by polarization is performed using a polarization beam splitter (PBS), which directs photons with vertical and horizontal polarization to different output channels. SiMP detectors detect photons by generating avalanches of electrons when they hit the sensor, which leads to the formation of a signal with an amplitude of about 100 mV and a duration of 5 ns. The received signals are transmitted to the oscilloscope via the corresponding input channels (ch1 and ch2), where their time correlation is analyzed. It is important to note that along with useful detections, false positives are possible, caused by several factors: (i) external photons hitting the sensor due to imperfections of the system, (ii) dark current, which is spontaneous avalanches occurring in the absence of photons and accounting for about 5% of the total number of responses, and (iii) double avalanches generated by simultaneous excitation of two sensor pixels, which leads to the appearance of a signal with an amplitude of 2 mV, indistinguishable from the true detection of two photons.

The final element of the optical system is the reference signal, which ensures synchronization of the generation of entangled photons with laser pulses. Since entangled photons can only be formed at the moment of irradiation of the system with a laser pulse, the analysis of time correlations is carried out exclusively within the corresponding time windows. For this purpose, a reference beam is used, which serves as a key marker for the time binding of the detected photons.

An example of a signal fragment recorded by SiMP detectors connected to the oscilloscope inputs (channel 1 and channel 2) is shown in Fig. 2. Comparison of the time coincidences of pulses in these two channels can reveal correlated events that may indicate the detection of entangled photons. The main task in detecting pairs of entangled photons is to develop a methodology and software for their accurate recognition based on the recorded experimental signals.

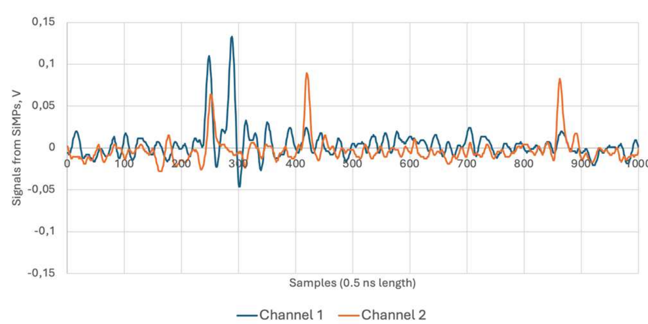


Fig. 2. Example of signals obtained from SiMPs connected to the oscilloscope inputs (channels 1 and 2) with a duration of 0.5 ns

III. Anomaly Detection Techniques

A. Anomaly Detection Types and Methods

The paper attempts to solve this problem using one of the machine learning algorithms, considering the detection of entangled photons in the recorded signal (as a very rare event) as an anomaly.

There are three types of anomalies that can occur [4].

Individual anomalies are considered to deviate from the entire data set. This can be one or more values that deviate significantly from the general trend.

Contextual anomalies are deviations within a specific context or subset of data. Such anomalies may not be deviations at the general level, but within a limited context, they represent anomalous events.

Collective anomalies are where a data set or several related instances are considered anomalous. These are more complex cases where a group of data as a whole violates the expected pattern or model.

Anomaly detection methods are usually divided into three categories [5].

Supervised detection – requires the presence of “normal” and “anomalous” labels in the dataset, which allows the algorithm to learn from examples. Semi-supervised detection – works with partially labeled data when most examples are missing labels, but there are a small number of labels for training. Unsupervised detection – uses completely unlabeled data, where the main idea is to construct a boundary of expectations (or a region of normal data). All instances outside this boundary are considered anomalous.

Since, in this paper, the anomalous event will be a sequence of measurements in channels, it is proposed to consider the anomaly in the problem as a collective anomaly, and since there are no clearly defined cases of entangled photon detection in the analyzed data, it is proposed to use an unsupervised anomaly detection algorithm. Such an algorithm will “search” for rare, unusual events that may be associated with the detection of entangled photons despite the lack of specific labels for such cases.

Based on the above, it is proposed to consider the possibility of using the "Anomaly Detection Algorithm" to solve the problem, the main assumptions of which are an extremely small proportion of anomalous examples in the data set and the assumption that the data are distributed according to the Gaussian (normal) law. The main idea of the proposed algorithm is to consider the probability density function as a bell-shaped curve with a certain arithmetic mean and variance. This curve has two tails, and the rarer the instance, the further it is from the center, which makes it a more likely anomaly. In the context of detecting entangled photons, such rare events falling into the tails of this curve will be identified as anomalies, which may indicate the potential detection of entangled photons.

B. Unsupervised Anomaly Detection Algorithm

The distribution’s probability density function, *p(x)*, is a bell curve centered at the arithmetic mean, μ, and the variance, σ, of the dataset defines the width of the curve.

This curve has two elongated tails on each end. The more rare the instance – the further it is from the center – the more likely it is to be an outlier or an anomaly.

The probability density function, *p(x)*, measures the probability of some outcome x in the dataset

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right). \quad (1)$$

Assuming only one feature that follows a normal distribution, some threshold, ε, can be set that determines whether a case is anomalous or not

$$p(x) < \varepsilon. \quad (2)$$

Threshold ε should be set heuristically, and its value will depend on the use case and the preferred sensitivity to anomalies.

If the features cannot be considered independent, as in our case, a multivariate probability density function must be used.

In the multivariate case, a covariance matrix, Σ, is constructed, and the multivariate probability density function can be shown as [6]:

$$p(x) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} exp\left(\frac{-1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right),$$

$$\mu = \frac{1}{m}\sum_{i=1}^{m} x^{(i)}, \quad (3)$$

$$\Sigma = \frac{1}{m}\sum_{i=1}^{m}(x^{(i)}-\mu\backslash)\left(x^{(i)}-\mu\right)^T.$$

The multivariate Gaussian model was selected for its ability to effectively model correlations between multiple features – crucial for detecting simultaneous peaks in both channels – and its computational efficiency, making it suitable for real-time quantum applications.

C. Additional features

To improve the model, additional features can be engineered to demonstrate values very different from the mean for cases of suspected anomalies.

Since the anomaly that we hope to detect in the existing data set for Channel 1 and Channel 2 (recording the impact of horizontally and vertically polarized photons) is the consistent appearance of peaks in both channels, it is proposed to use

$$x_3 = (x_1 x_2)^2 \quad (4)$$

as a new feature.

The anomaly detection method allows for identifying anomalous measurement values/samples in the dataset. An anomalous event to be detected is the appearance of an entangled photon. Correlated peaks in the data series are interpreted as probable entangled photon detection events.

IV. Results and Analysis

The proposed anomaly detection method was applied to the recorded signals of the experimental setup to identify potential entangled photon detection events from the ENGELBERT dataset, a collection of experimental data, file S001E002T002.csv, which includes data on 10^7 measurements. The data consists of measurements from two channels (Channel 1 and Channel 2) corresponding to horizontally and vertically polarized photons, as well as a reference beam providing synchronization of entangled photon generation with laser pulses (Channel 3). The goal was to determine whether the proposed unsupervised anomaly detection algorithm based on a multivariate normal distribution can distinguish rare events with a probability of detecting entangled photon pairs from background noise and normal fluctuations.

To implement the anomaly detection algorithm, the following procedure was carried out:

- the recorded signals from both channels were normalized to ensure their comparability;

- an additional feature (4) was introduced to increase the sensitivity of the algorithm, improving its ability to distinguish noise from cases of true photon correlations;

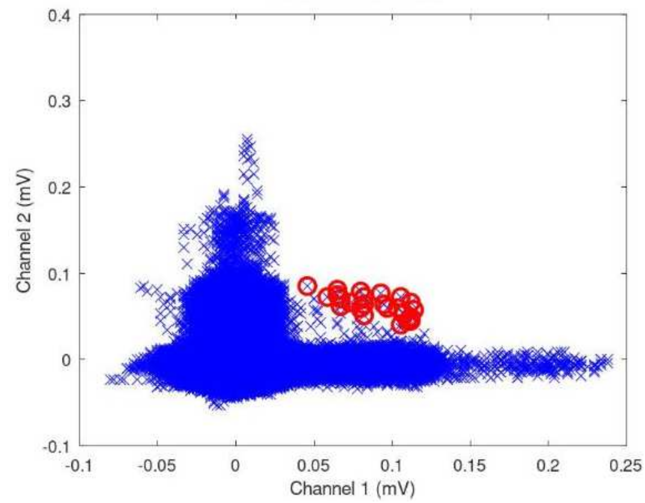
- the probability density function $p(x)$ was estimated using (3), based on pre-calculated values of the mean vector μ and the covariance matrix Σ;

- the threshold ε (2) was set by a heuristic method to separate normal and abnormal events, measurements having probability values below a given ε were classified as anomalies.

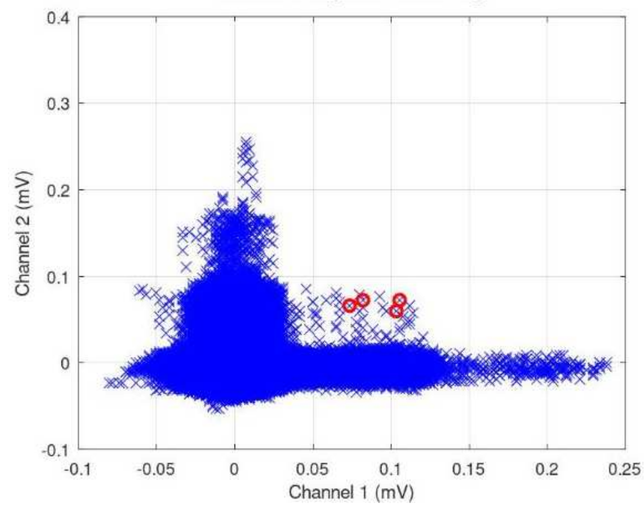
As a result of the calculations, a fairly small part of the combinations of groups of points on two channels (corresponding to electrical impulses recorded by the oscilloscope) from the data set showed significantly lower probability values, which indicates their anomalous nature. Four cases were detected, which are shown in Fig. 3 and 4.

Fig. 3 shows the measurement data contained in the file S001E002T002.csv, displayed with blue crosses and anomalous values highlighted with red circles. Fig. 3a indicates all detected anomalous measurement values, and in Fig. 3b – individual detected anomalous events, defined as peak values from groups of points. For better clarity, the results in Fig. 4 show the probability density function distribution with detected anomalous events in red circles.

Fig. 5 shows the measurement data corresponding to the above-mentioned four dataset elements identified as anomalous, with serial numbers 47670, 5652817, 2414631, and 7123248.



a)



b)

Fig. 3 Visualizing a data set with detected anomalous: a) – measurement values, b) – events

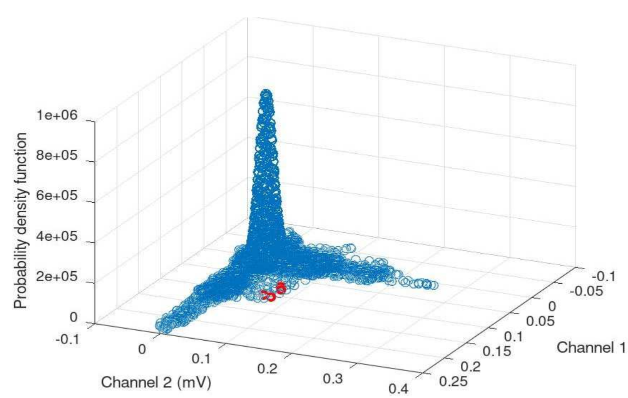


Fig. 4 Probability density function distribution, detected anomalous events

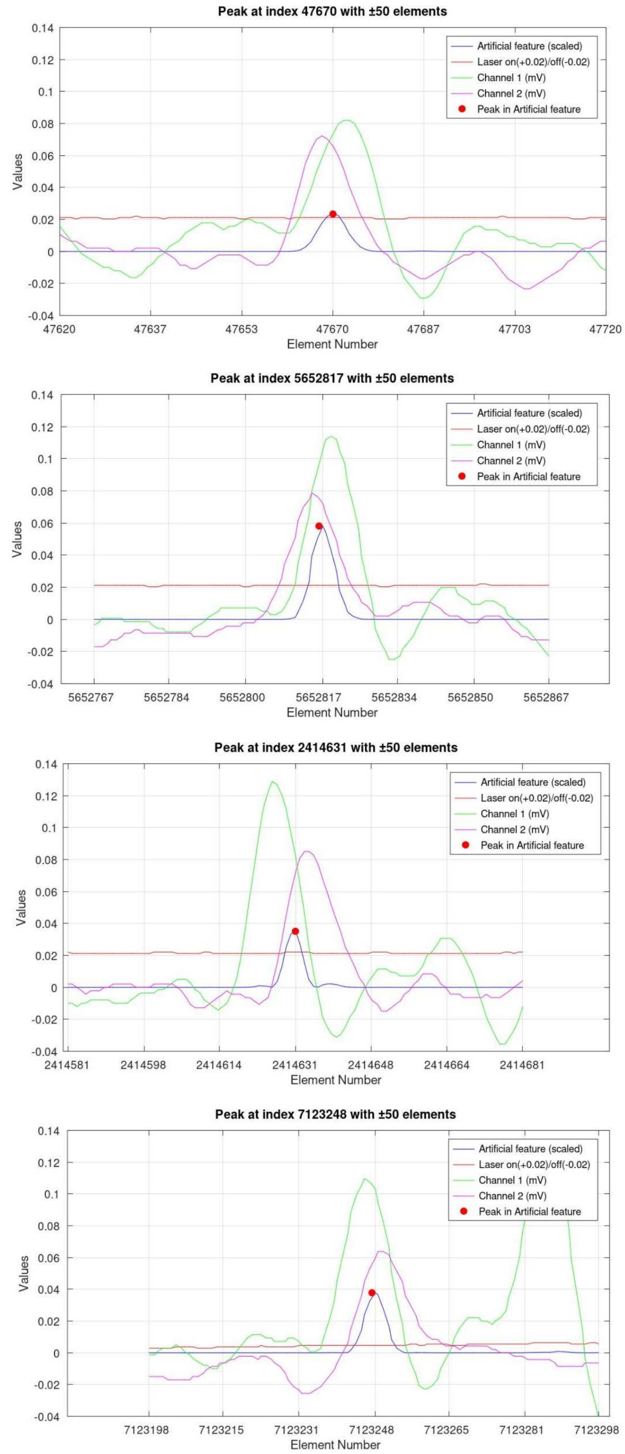


Fig. 5 Measurement data for anomalous dataset elements with serial numbers 47670, 5652817, 2414631, 7123248

The green and magenta lines show the voltage values (mV) in channels 1 and 2, respectively. The red line shows the data in channel 3 – conventionally, the values 0.02 and -0.02 correspond to the presence and absence of the laser beam, respectively, at a given time. The red dot shows the local peak value of an additional artificial feature (4), which can be indirectly considered as a marker for detecting a possible pair of entangled photons. The value of the artificial feature in the figure is shown by a blue line in an absolute scale, chosen for ease of perception on these graphs. Also, to ensure the clarity of the presented results, the graphs show data for 50 points to the left and right of the point with the peak value.

These anomalies corresponded to cases where both channels demonstrated simultaneous peaks, which is consistent with the expected behavior when detecting entangled photons.

The unsupervised nature of the method made it possible to effectively detect anomalies without pre-set marks, which makes it a promising approach for detecting entangled photons in experimental data.

V. Conclusion

This study addresses the challenge of detecting entangled photons by proposing a novel, unsupervised anomaly detection algorithm based on the multivariate Gaussian distribution. Applied to experimental data from two channels, it identified rare correlated events consistent with entangled photon detection. This fast, simple approach enhances quantum optics research, offering potential applications in quantum technologies. This study introduces the first application of unsupervised anomaly detection to entangled photon identification, offering a novel, efficient tool for quantum optics with potential applications in quantum communication and computing. Further optimization, such as threshold tuning and noise resistance, could improve accuracy.

References

[1] Z. Zhang, C. You, O. S. Magaña-Loaiza, R. Fickler, R. de J. León-Montiel, J. P. Torres, T. S. Humble, S. Liu, Y. Xia, and Q. Zhuang, “Entanglement-based quantum information technology: a tutorial,” *Adv. Opt. Photon.*, vol. 16, pp. 60–162, 2024.

[2] Y.-C. Jeong, K.-H. Hong, and Y.-H. Kim, “Bright source of polarization-entangled photons using a PPKTP pumped by a broadband multi-mode diode laser,” *Opt. Express*, vol. 24, pp. 1165–1174, 2016.

[3] H. Kim, O. Kwon, and H. S. Moon, “Pulsed Sagnac source of polarization-entangled photon pairs in telecommunication band,” *Sci. Rep.*, vol. 9, no. 1, p. 5031, Mar. 2019.

[4] R. Foorthuis, “On the nature and types of anomalies: a review of deviations in data,” *Int. J. Data Sci. Anal.*, vol. 12, pp. 297–331, 2021.

[5] M. Bauw, S. Velasco-Forero, J. Angulo, C. Adnet, and O. Airiau, “From unsupervised to semi-supervised anomaly detection methods for HRRP targets,” in *2020 IEEE Radar Conf. (RadarConf20)*, Florence, Italy, 2020, pp. 1–6.

[6] “The Basics of Anomaly Detection with Multivariate Gaussian Distribution,” *Towards Data Science*, 2019. [Online]. Available: https://medium.com/towards-data-science/the-basics-of-anomaly-detection-65aff59949b7.