# Applications of White Noise Stochastic Methods in Machine Learning for Binary Particle Classification

An Outline Submitted to

Sergei Gleyzer, PhD

Department of Physics and Astronomy

The University of Alabama, Tuscaloosa

In Fulfillment of the Requirements for

PH561 (Nuclear and Particle Physics)

Resham Lal Sohal Jr.

December 2023

## 1. Brief Introduction

The Standard Model (SM) is arguably at the core of particle physics, providing a remarkably successful description of any known theoretical or experimental phenomena [1-3]. In efforts to probe and better understand the SM, collaborations over the last two decades have been made to create collider experiments [4,5] that could reach high energy scales, for instance the TeV scale of Large Hadron Collider (LHC) [6,7]. At the reduced Planck scale $M_P^{-1} = (G_N/\hbar c)^{1/2} \sim 10^{19}$ GeV, where quantum gravitational effects become important, a new framework will be certainly required. Existence of new physics occurs because of the vast energy difference ($\sim 10^{16}$ orders of magnitude difference) between the presently explored territory near the electroweak scale $M_W$, and the Planck scale, $M_P$ [8].

The SM potential [9], which gives Higgs mass, is of the form:

$$V = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \tag{1}$$

where $\phi$ is the SM Higgs field, and $\lambda$ is the strength of the Higgs self-interaction with $\lambda > 0$, $\mu^2 > 0$. It is important to note that the Higgs mass $m_H$ and the factor $\mu$ are related by the relation $\mu^2 = m_H^2$.

When there are no loops, i.e., at tree level, there is no problem. At the loop level, however, $m_H$ sustains corrections brought by contributions such as gauge loops, self-interactions, and fermion loops [10]. SM, being renormalizable, entails that for all higher order loop corrections, finite results will be recorded, even if we take our virtual momenta to infinity in the loop integrals.

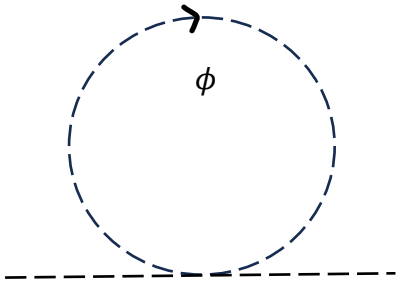


**Figure 1.** One-loop self-energy Feynmann Diagram for a scalar field propagator

In $\phi^4$ — theory, the self-interaction at a one-loop order contributes to the $\phi^\dagger \phi$ term in equation (1). That contribution [10] happens to be proportional to some cutoff scale $\Lambda^2$, where $\Lambda$ is an ultraviolet momentum cutoff value. Thus, the correction to the one-loop corrected 'physical' mass can be approximated as

$$\delta m_H^2 \sim \lambda \Lambda^2 \tag{2}$$

Looking at equation (2), if $\Lambda$ approaches the Planck scale, we see that $m_H^2$ receives enormous quantum corrections (diverges quadratically as dictated by $\Lambda$) from the virtual effects of every particle or other phenomenon that couples, directly or indirectly, to the Higgs field [11]. One solution to this problem requires very fine tuning as this problem affects $m_H$, consequently affecting all masses in the SM. This was one of the major motivations to search for physics beyond SM.

The divergence seems to affect only the scalar sector $\phi$. The motivation then is to find a symmetry that can group scalar particles with either massless fermions or massless vector bosons, so that quadratic divergence brought by $\Lambda$ can be eliminated and that no fine-tuning will be needed [12]. **Supersymmetry** is essentially such a symmetry. It groups scalars together with fermions, and vector bosons with fermions as well. In essence, supersymmetry is a space-time symmetry theory that allows transformations of bosons to fermions and vice versa [13] by the virtue of an anti-commuting spinor operator $Q$ [14]. That is,

$$Q\,|Boson> = |Fermion> \qquad Q\,|Fermion> = |Boson> \tag{3}$$

## 2. A Quick Look at Supersymmetry and Minimal Supersymmetric Standard Model (MSSM)

The simplest supersymmetric theory (SUSY) involves two free fields: **a complex spin-0 scalar field $\phi$ and an L-type spinor field $\psi$** [15]. The SUSY theory contains one massless complex scalar field and one massless Weyl fermion field. Such fields form a SUSY **supermultiplet** called a **chiral supermultiplet** [14,16]. **Vector** (or gauge) **supermultiplet** is also allowed, where a massless spin-1 field, having two on-shell degrees of freedom [10], is partnered with a massless Weyl fermion field. It is important to note that there exist only these two types of supermultiplet in the MSSM.

**MSSM is the supersymmetric extension of the SM** [15]. In this theory, each of the known fundamental particles, either in a gauge or chiral supermultiplet, **must have a superpartner whose spin differs by a half integer unit** [17]. In a theoretical standpoint, **only chiral supermultiplets** can have fermions that undergo distinct transformation within the gauge group. These fermions are said to have left chirality [14]. All SM fermions have this fundamental property, further implying that currently observed fermions are members of chiral supermultiplets. Hence, the bosonic partners of quarks and leptons must have a spin of 0, and not spin-1 like vector bosons [14].

## 3. SUSY Particles

The SM gauge bosons are associated with local $SU(3)_c$ symmetry [18]. Gauginos, on the other hand, are flavor singlets that belong to the eight-dimensional 'adjoint' representation of SU(3). To create a supersymmetric version of QCD, a new SU(3) octet of Weyl fermions is required [19]. These are called '**gluinos** ($\tilde{g}$)', which are the **superpartners of gluons (g)** and fields do belong to the **gauge supermultiplet**. Similarly, one can create SUSY partners of W-bosons by introducing an $SU(2)_L$ triplet of Weyl fermions [14]. The superpartner for W-bosons are called '**winos**' ($\widetilde{W^{\pm}}, \widetilde{W^0}$).

A scalar partner for the neutrino is called **'sneutrino'** and a scalar partner for the electron is called **'selectron'**. Similarly, we also have **'smuons'** and **'staus'**. These are all in chiral supermultiplets, and $SU(2)_L$ doublets, and they all carry the same lepton numbers as their SM partners [14]. Quarks, on the other hand, are a triplet of the $SU(3)_c$ color gauge group that belongs to be the octet representation of SU(3). The supersymmetric partner of quarks is called **'squarks'**, and they are spin-0 particles with the same baryon number as the quarks. In the Higgs sector, the scalar Higgs field need their own SUSY partners '**higgsinos'**.

The chiral and gauge supermultiplets introduced in this section constitute the field content of the MSSM, and the full theory includes supersymmetric interactions and soft SUSY-breaking terms. However, until now, none of the aforementioned '**superpartners'** have been detected experimentally on collider experiments.

| **Names** | | **Spin 0** | **Spin ½** | **$SU(3)_C$, $SU(2)_L$, $U(1)_Y$** |
|---|---|---|---|---|
| Squarks, quarks (x3 families) | Q | $(\widetilde{u_L}, \widetilde{d_L})$ | $(u_L, d_L)$ | $\left(3,2,\frac{1}{6}\right)$ |
| | $\bar{u}$ | $\widetilde{u_R^*}$ | $u_R^\dagger$ | $\left(\bar{3},1,-\frac{2}{3}\right)$ |
| | $\bar{d}$ | $\widetilde{d_R^*}$ | $d_R^\dagger$ | $\left(\bar{3},1,\frac{1}{3}\right)$ |
| Sleptons, Leptons (x3 families) | L | $(\tilde{\nu}, \widetilde{e_L})$ | $(\nu_L, e_L)$ | $\left(1,2,-\frac{1}{2}\right)$ |
| | $\bar{e}$ | $\widetilde{e_R^*}$ | $\widetilde{e_R^\dagger}$ | $(1,1,1)$ |
| Higgs, higgsinos | $H_u$ | $(H_u^+, H_u^0)$ | $\left(\widetilde{H_u^+}, \widetilde{H_u^0}\right)$ | $\left(1,2,+\frac{1}{2}\right)$ |
| | $H_d$ | $(H_d^0, H_d^-)$ | $\left(\widetilde{H_d^0}, \widetilde{H_d^-}\right)$ | $\left(1,2,-\frac{1}{2}\right)$ |

**Table 1.** Summary of Chiral Supermultiplets in MSSM [14]. Each SM field and its superpartner must have the same $SU(3)_C \times SU(2)_L \times U(1)_Y$ quantum numbers.

## 4. SUSY on Experimental Point of View

In the context of LHC experiments, searching for SUSY signatures on decay kinematics is an arduous task to do, as SM signatures are predominantly present almost in every recorded event. There is hope, however, by considering parity conservation. The baryonic number **B**, lepton number **L**, and spin **S** in SUSY formalism are all related to a quantity called R-parity $\boldsymbol{R} = (-\mathbf{1})^{\mathbf{3}(\boldsymbol{B}-\boldsymbol{L})+\mathbf{2}\boldsymbol{S}}$ [20]. Understanding cases where R-parity is conserved or violated is a must on SUSY phenomenology. This allows experimentalists to devise algorithms to better improve existing constraints on the theory and look at possible particle kinematics to study. For instance, discriminating SUSY signal from SM signatures can be done by looking at decay kinematics such as transverse energy $H_T$, missing transverse momentum $E_T^{miss}$, and effective mass $m_{eff} = H_T + E_T^{miss}$ [21]. The effective mass can somehow be reconstructed using techniques such as combined maximum likelihood estimation [22], and the corresponding $m_{eff}$ peak in the distribution for SUSY signal can be correlated to SUSY mass scales. This method works because SM signatures dominate at low $m_{eff}$. However, mass reconstruction for SUSY-related signals imposes difficulties [23] because too much information is carried away by neutrinos that were formed on the decay chain.

As previously stated, even state-of-the-art experiments have not confirmed the existence of SUSY particles yet. Even so, CERN cannot rule out SUSY because as new particle masses are raised, the more the theory agrees with the SM [24]. Thus, while current efforts haven't recorded any SUSY particle, improvements on particle search exclusion limits were made, all of which were at 95% confidence level [25-29]. A concise compilation of search exclusions limits for SUSY particles is maintained by the Particle Data Group (PDG) and can be freely accessed online [30,31].

## 5. Well-known Methodologies in Machine Learning

In this paper, an attempt to improve signal-background discrimination for SUSY phenomenology uses will be carried out using machine learning and available simulated open data [32]. The same goal is aimed to be achieved for an end-to-end – related electron-photon classification task [33]. Modern machine learning models are typically data hungry and for this reason, generating artificial samples from available simulations would be a great idea. This can be carried out through a process called Synthetic Data Augmentation (SDA). For deep learning applications in particle physics, SDA techniques such as small noise injection [34, 35], rotation [36], Generative Adversarial Networks (GANs) [37] and Variational Autoencoders (VAEs) implementations [38] have been proven to increase the performance and robustness of a machine learning model.

Throughout the rest of this text, we will look and limit our methodological approach to using white noise formalism as a noise injection-based SDA technique for the SUSY dataset, and using the same formalism to modify convolutional neural networks.

## 6. A Peek on White Noise Stochastic Methods

White noise analysis methods [39] are intended for datasets that have inherent temporal dependence. In a general sense, data analysis is carried out in the following fashion:

a. Plot the data points.
b. Obtain the best fit line for the plotted raw data.
c. Subtract the data points from the best fit line. This is the fluctuation.
d. Calculate for the empirical MSD.
e. Plot log(MSD) against log(time). The output graph will serve as a reference for the theoretical MSD modeling.
f. Adjust the modulating function $\mathrm{g(T)}$ to fit the theoretical MSD with the empirical MSD.
g. Formulate the corresponding probability density.

Let $x(t)$ be some stochastic variable having no memory of its history and whose behavior obeys white noise. It follows that $x(t)$ can be parameterized starting at some initial point $x_0$ and later on evolves in accordance to Brownian motion $B(t)$

$$x(t) = x_0 + B(t) = x_0 + \int_{t_0}^{t} \omega(s)\ ds \tag{4}$$

where $\frac{dB(t)}{dt} = \omega(t)$ implies the white noise variable. In order to realistically describe physical processes, we need to factor in $x(t)$'s memory of its past by introducing the factors $f(\tau - t)h(t)$ in the integrand to modify the white noise variable $\omega(t)$. Here, $f(\tau - t)$ serves as a memory function and $h(t)$ serves as some general function that is time dependent. This allows us to write equation (4) as

$$x(t) = x_0 + \int_{0}^{\tau} f(\tau - t)\ h(t)\ \omega(t)\ dt \tag{5}$$

where $t$ runs from the first instance $t = 0$ to some later time $t = \tau$. If we take the sum over all possible paths, starting at $x(\tau = 0) = x_0$ and ends at $x(\tau = T) = x_T$, the probability density function (PDF) for the system of interest can be evaluated by integrating the delta function $\delta(x(\tau) - x_T)$ over the Gaussian white noise measure $d\mu(\omega)$. In so doing we are essentially getting the expectation value $E\delta(x(\tau) - x_T)$. That is,

$$P(x_T, T; x_0, 0) = E\delta(x(\tau) - x_T) = \int \delta(x(\tau) - x_T) d\mu(\omega) \tag{6}$$

Here, $\delta(x(\tau) - x_T)$ acts as a constraint that all possible paths must satisfy, and $d\mu(\omega)$ is called the white noise Gaussian measure defined as

$$d\mu(\omega) = N_\omega \, exp\left(-\frac{1}{2}\int \omega(\tau)^2 d\tau\right) d^\infty\omega \tag{7}$$

where $N_\omega$ serves as some normalization constant. Substituting equation (5) to equation (6), one can obtain

$$P(x_T, T; x_0, 0) = \int \delta\left(x_0 - x_T + \int_0^\tau f(\tau - t)\ h(t)\ \omega(t)\ dt\right) d\mu(\omega) \tag{8}$$

The integral can be made more illuminating if we express the delta function in terms of its Fourier representation:

$$P(x_T, T; x_0, 0)\ = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\ e^{ik(x_0 - x_T)} \int_{-\infty}^{\infty} e^{ik\int_0^\tau f(\tau - t)\ h(t)\ \omega(t)\ dt}\ d\mu(\omega) \tag{9}$$

To integrate over $d\mu(\omega)$ we use the definition of the characteristic functional [40]

$$C(\xi) = \int e^{i<\omega(t),\, \xi(t)>} d\mu(\omega) = \int e^{i\int_0^\tau f(\tau - t)\ h(t)\ \omega(t)\ dt} = e^{-\frac{1}{2}\int \xi(t)^2\ d\tau} \tag{10}$$

with $\xi(t) = kf(\tau - t)h(t)$. Hence,

$$P\left(x_T, T; x_{0,} 0\right)\ =\ \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\ e^{ik(x_0 - x_T) - \frac{k^2}{2}\int_0^\tau [f(\tau - t)h(t)]^2\ dt} \tag{11}$$

Equation (11) is a Gaussian integral that can easily be evaluated as

$$P\left(x_T, T; x_{0,} 0\right)\ =\ \frac{1}{\sqrt{2\pi\int_0^T [f(\tau - t)h(t)]^2\ dt}} \times exp\left(\frac{-(x_0 - x_T)^2}{2\int_0^\tau [f(\tau - t)h(t)]^2}\right) \tag{12}$$

with $\tau = T$ being the final time. Moreover, we can obtain another quantity of interest called mean squared deviation (MSD)

$$MSD = (x - < x >)^2 = < x^2 > - < x >^2 \tag{13}$$

To evaluate this, we have to calculate first

$$< x >= \int_{-\infty}^{\infty} xP\left(x_T, T; x_{0,} 0\right) dx \tag{14}$$

and

$$< x^2 >= \int_{-\infty}^{\infty} x^2 P\left(x_T, T; x_{0,} 0\right)\ dx \tag{15}$$

Equations (14) and (15) are known as the first moment and second moment, respectively. The first moment is evaluated to be $< x >= x_0$. On the other hand, evaluating the integral in (12) results to $< x^2 >= x_0^2 + g(T)^2 \int_0^T [f(\tau - t)h(t)]^2\ dt$. Therefore, we can express equation (13) as

$$MSD = \ g(T)^2 \int_0^T [f(\tau - t)h(t)]^2\ dt \tag{16}$$

It follows that the expression for the PDF in equation (12) can be simplified to

$$P\left(x_T, T; x_{0,} 0\right) = \frac{1}{\sqrt{2\pi MSD}}\ e^{-\frac{(x_0 - x_T)^2}{2MSD}} \tag{17}$$

### 6.1. Applying White Noise Techniques using "Pseudo-time" Conceptualization (SUSY Dataset)

Over the years, mathematical models, computer simulations, and machine learning algorithms have been developed in particle physics to cater various needs for data analysis such as, but not limited to, particle searches, regression-related tasks, and detector performance enhancements. White noise techniques are typically not used on any particle searches as available open and simulated datasets are triggered, in a sense that recorded particle kinematics are static in time. For instance, if a machine learning task involves a classic binary classification, any signal or background occurrence is totally independent of past interactions. This means that neighboring events are treated separately as they are generally far away in time.

On a normal LHC operation, about 100 billion of protons are accelerated every 25 nanoseconds, where only about 50 protons collide to produce new particles. While events are not time-inherent, one can treat each recorded particle kinematic as a sequential data, in a general macroscopic outlook, by noting the fact that on an average, the LHC detector needs about 25 nanoseconds to capture an event. Here, I will call the 25 nanoseconds interval between each event as "pseudo-time $(n)$", so that white noise methods can be applied. An attempt to improving the signal-background discrimination for a machine learning-based binary particle identification will be made by using stochastic white noise analysis methods as a feature engineering technique, specifically for data augmentation.

The SUSY dataset [32] has a total of 5M samples with 18 features; the first eight (8) features are low-level (as recorded by detector) particle kinematic properties, and the last ten (10) features are high-level and functions of the first eight features. For consistency, the last 500k samples were used as the test set [32] and out of the remaining samples, 80% was allocated for training while 20% was held for validation.

As part of the data processing step, all samples were normalized after splitting the dataset into training, validation, and test set. Below is a general outline of how SDA will be applied.

a. Separate signal (**S**) and background (**B**) events.
b. Take a particle kinematic quantity or image channel of interest from **S** and **B**.
c. Subtract **B from S**. This is the fluctuation.
d. Calculate for the empirical MSD using numerical methods.
e. Plot log(MSD) against log(n). The output graph will serve as a reference for a theoretical MSD modeling.
f. Adjust the modulating function $g(N)$ to fit the theoretical MSD with the empirical MSD.
g. Formulate the corresponding probability density function (PDF).
**h.** Augment the original data by injecting it with a random value (noise) pulled from the PDF. **Note that this will be done to expand the number of samples on the training set only.**
i. Create a machine learning model, such as Neural Network (NN) or Extreme Gradient Boosting (XGB) model, to cater the needs of the task on hand.
j. Feed the combined original and augmented data to the model and evaluate the performance on unseen data.

### 5.2. White Noise Technique for Electron-Photon Classification

The electron/photon dataset [32] contains about 400k combined samples of 32x32 electron (signal) and photon (background) images. Each pixel in the image represents a detector cell and the intensity indicates energy hit that corresponds to that cell. In addition, there are two channels available; namely energy hit for Channel 1 and energy deposit timing for Channel 2. A white noise inspired formalism, in lieu of a traditional convolution, will be employed to classify images.

## 6. Results and Discussion

Out of the sixteen (16) theoretical memory functions found in [39], a Bessel-type of memory function was used to model the theoretical MSD. Incorporating the notion of "pseudo-time", it was modified it as

$$f(N-n) = \sqrt{J_{\nu}(N-n)} \tag{18}$$

$$h(n) = \sqrt{\frac{J_{\mu(n)}}{n}} \tag{19}$$

The motivation behind choosing a Bessel function as a candidate is grounded on the notion that it appears frequently on physical problems that exhibit cylindrical symmetry, such as in the field of electromagnetism [41], optics [42], thermodynamics [43], and even Quantum Field Theory (QFT) [44]. The CMS detector is no exception as energy deposition exhibits some degree of $\phi$ independence [45] within specific regions defined by pseudorapidity $\eta$ [45].

The MSD was acquired by substituting equations (18) and (19) to equation (16), and evaluated by looking up at a table of integrals [46]

$$MSD = [g(N)]^2 \frac{J_{\mu+\nu}(N)}{\mu} \tag{20}$$

Here, $\mu$ and $\nu$ are some tunable parameters and $N$ is the length of the whole dataset.

## SUSY DATASET

The MSD was calculated from the training residuals and a good visual fit was observed when the modulating function $g(N)$ was chosen to be proportional to an exponential function.

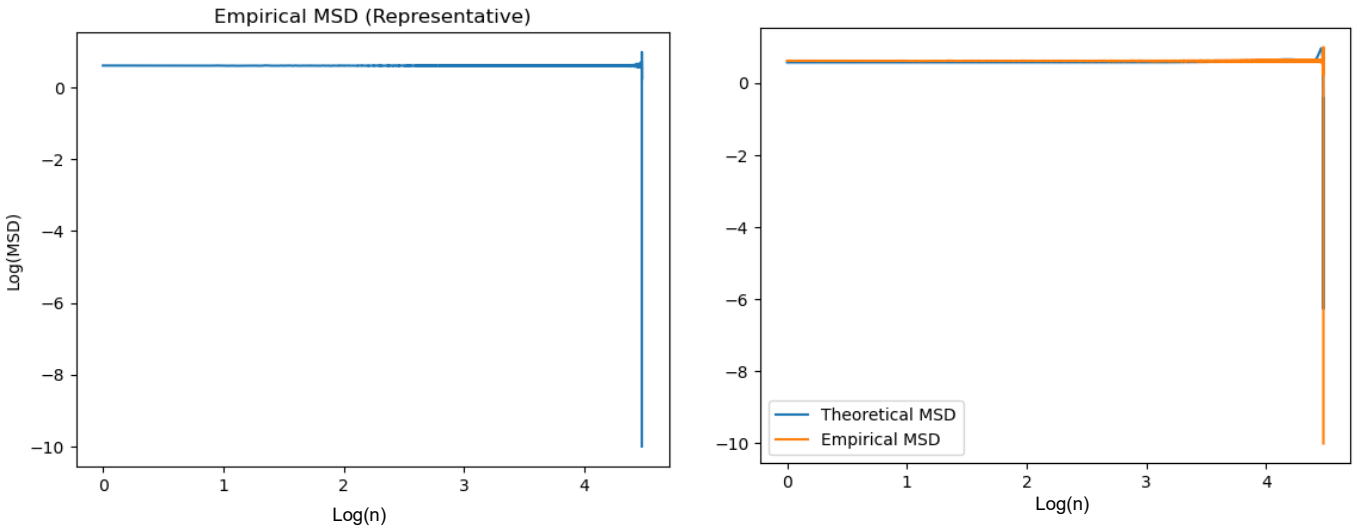


**Figure 2.** Log(MSD) vs Log(n) plot for empirical data (left) and theoretical fitting (right)

An Extreme Gradient Boosting Classifier (XGB) is used to train low-level features. Such a classifier is rooted on building a strong model by averaging learned predictions across a collection of weak models, and the best parameters were chosen for the training.

| **Classifier** | **ROC-AUC** | |
|---|---|---|
| | No Data Augmentation | With 1M Additional Augmented Data |
| XGB | 0.85870 | 0.85873 |

**Table 2.** Summary of results for an XGB Classifier trained on low-level features, where data augmentation was performed on the **lepton1_eta** feature.

As expected, no significant boost on performance was observed because the dataset has no inherent temporal dependence.

## Electron Photon Dataset

As mentioned earlier, a white noise inspired convolution will be done for this dataset. The idea is to use equation (20) to apply transformation both on the learned kernel/filter weights and inputs in the training process. For simplicity, it is convenient to use $g(N) = 1$, $\mu = 1$, and $\nu = 0$. Such values were chosen because $J_1(x = 0)$ is equal to zero, which then preserves the pixel values for sparse inputs and consequently not altering the data too much.

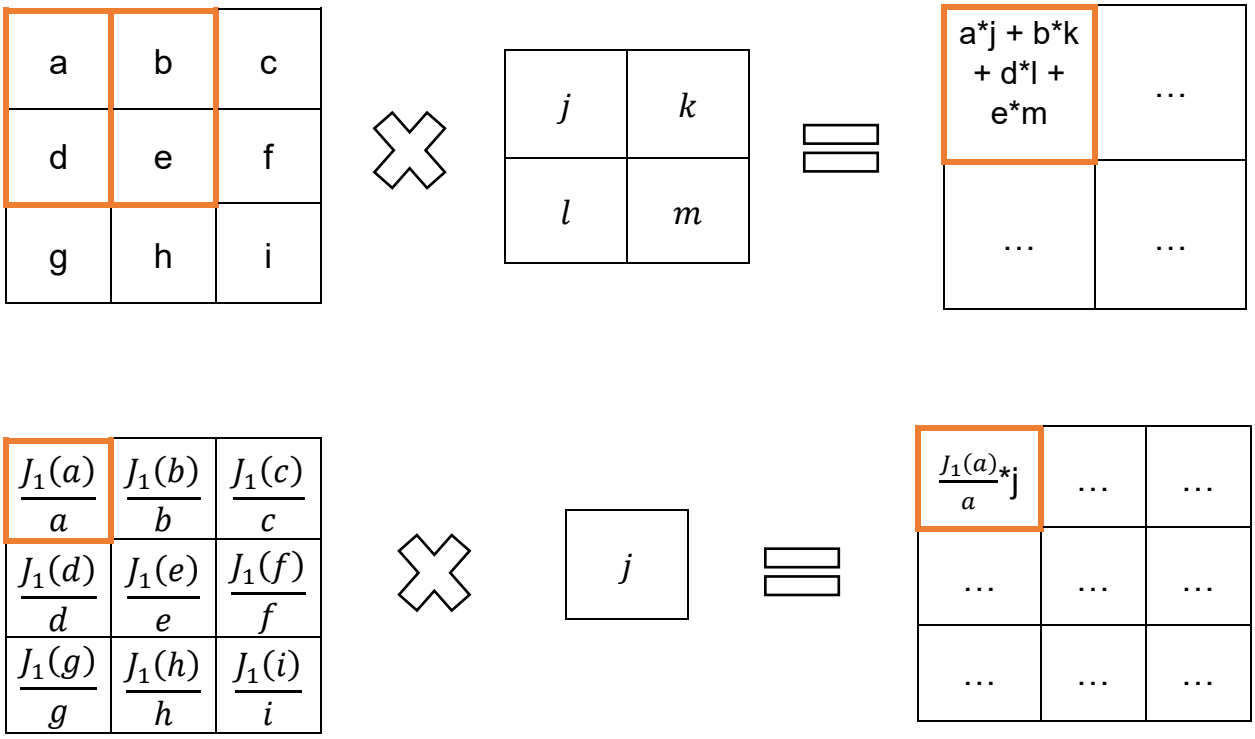


**Figure 3.** An example of how convolutional neural network (CNN) works. From top to bottom: traditional CNN with a 2x2 kernel, and a white noise inspired 1x1 convolution. Here, the input is transformed and the 1x1 kernel $j$ is constrained to have values that are within the range of the Bessel function of zeroth order, which happens to be from -0.4 to 1.0 in this case.

The reason behind why only 1x1 convolution is allowed for a white noise inspired formalism is because it treats the dataset as if pixel values are flattened. In a mathematical standpoint, the memory function $f$ accounts for how neighboring pixels (row-wise) are related to each other and the modulating function $h$ can be thought of as a function that records current values.

| Model | Channel 1 and 2 | Channel 1 Only | Execution time per epoch |
|---|---|---|---|
| VGG (2x2 kernel) | 0.8092 | 0.8124 | 25s/epoch |
| VGG (1x1 Kernel) | 0.7840 | 0.7843 | 25s/epoch |
| White noise inspired CNN (1x1 kernel) | 0.7936 | 0.8031 | 2min/epoch |

**Table 3**. ROC-AUC for the models trained in the Electron-Photon Dataset and execution time taken for each epoch.

For each model, 80% was allotted for training, 10% for validation, and 10% for testing. A simple CNN model was used with two convolutional blocks, each with 32 and 64 filters respectively, and a two-dimensional max pooling operation in between. The output was then fed to a fully connected layer that includes a dropout factor of 0.2 to reduce overfitting. Moreover, the model was trained for 100 epochs with a learning rate that decreases linearly by a factor of 0.2 when the validation accuracy stops improving for six epochs. It was observed that the model performance plateaued after about 40 epochs in the training process, and the results based on averaged three independent runs were tabulated in Table 3.

Based on the tabulated results, no significant boost in performance was observed, and in fact, the performance for a white noise inspired CNN yielded unsatisfactory result as compared to a VGG-type network with a 2x2 filter size. On another note, while keeping the network architecture the same, it is worth noting that the white noise inspired formalism transformation performs better than a VGG-type with a 1x1 kernel counterpart. This implies that the choice of the memory function $f$ and modulating function $h$ does capture neighboring pixel dependencies up to some extent.

## 6. Conclusion and Possible Extension

The underlying methodology on white noise analysis can be used as an input transformation technique for machine learning to capture relevant dependencies among neighboring datapoints that have temporal dependence. This stems from the fact that a measurement is characterized by some intrinsic fluctuations, where particle kinematics of interest can be modelled using a probability density function [47].

On another hand, it is important to note that 1x1 kernel convolution can only be performed with a white noise formalism, which can be thought of as a simple neural network.

However, a possibility of extending it to $nxn$ kernel can be obtained by parametrizing along the $\phi$ and $\eta$ space to capture stochastic dependencies that are similar to Levy's stochastic area. Carrying out such a task is challenging, as the choice of memory function can lead to model performance degradation, which is not an ideal situation.

## References


[1] Kobayashi T. (2021). Experimental verification of the standard model of particle physics. Proceedings of the Japan Academy, Series B, Physical and Biological Sciences **97**(5), pp. 211-235. DOI: 10.2183/pjab.97.013

[2] Gaillard M., Grannis P. Sciulli F. (1999). The standard model of particle physics. Reviews of Modern Physics **71**(2), pp. 96-11. DOI: https://doi.org/10.1103/RevModPhys.71.S96

[3] Tanabashi M. et al. (Particle Data Group). (2018). Physical Review D 98. DOI: https://doi.org/10.1103/PhysRevD.98.030001

[4] Hauptman, J. (2011). Particle Physics Experiments at High Energy Colliders. ISBN: 9783527408252

[5] Bhat P.C. et al. (2022). Future Collider Options for the US. High Energy Physics – Experiment, Accelerator Physics. FERMILAB-CONF-22-144-PPD. [arXiv:2203.08088] DOI: https://doi.org/10.48550/arXiv.2203.08088

[6] LHC: The Guide (2017). CERN Brochure. Retrieved from https://cds.cern.ch/record/2255762/files/CERN-Brochure-2017-002-Eng.pdf. Accessed on October 04, 2023.

[7] Fernandez J.L. et. al. (2012). A Large Hadron Electron Collider at CERN Report on the Physics and Design Concepts for Machine and Detector. Journal of Physics G: Nuclear and Particle Physics **39**(7). DOI: 10.1088/0954-3899/39/7/075001

[8] Weinberg S. (2000). The quantum theory of fields Volume III: Supersymmetry. Cambridge Univ. Press, Cambridge. ISBN: 0-521-66000-9

[9] Salerno, D. (2019) Theoretical Background. In: The Higgs Boson Produced with Top Quarks in Fully Hadronic Signatures. Springer Theses. Springer, Cham. DOI: https://doi.org/10.1007/978-3-030-31257-2_2

[10] Csaki C. and Tanedo P. (2013). Beyond the Standard Model. Proceedings of the 2013 European School of High-Energy Physics. [arXiv:1602.04228v1]

[11] Buchmüller W. and Lüdeling C. (2006). Field Theory and Standard Model. High Energy Physics – Phenomenology, DESY-06-151.[arXiv:hep-ph/0609174] DOI: https://doi.org/10.48550/arXiv.hep-ph/0609174

[12] Baer H. and Tata X. (2022). Weak Scale Supersymmetry: From Superfields to Scattering Events. Cambridge University Press, University of Cambridge. ISBN: 978-1-009-28984-9. DOI: 10.1017/9781009289801

[13] Kane G. and Shifman M. (2000). The Supersymmetric World: The Beginnings of the Theory.

[14] Martin S. (2010). A Supersymmetry Primer. Advanced Series on Directions in High Energy Physics, Perspectives on Supersymmetry, pp. 1-153. DOI: https://doi.org/10.1142/9789812839657_0001

[15] Tong D. (2020). Supersymmetric Field Theory. University of Cambridge Part III Mathematical Tripos, Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, Cambridge. Retrieved from: https://www.damtp.cam.ac.uk/user/tong/susy/susy.pdf

[16] Bilal A. (2001). Introduction to Supersymmetry. High Energy Physics-Theory, NEIP-01-001 pp. 1-77. [arXiv:hep-th/0101055] DOI: https://doi.org/10.48550/arXiv.hep-th/0101055

[17] Fayet P. (1987). Supersymmetric Theories of Particles and Interactions. Physica Scripta **46**(15). DOI: 10.1088/0031-8949/1987/T15/006

[18] Altarelli G. and Wells J. (2017). Gauge Theories and the Standard Model. In: Wells, J. (eds) Collider Physics within the Standard Model. Lecture Notes in Physics, vol 937. Springer, Cham. DOI: https://doi.org/10.1007/978-3-319-51920-3_1

[19] Rodriguez M. (2019). The Minimal Supersymmetric Standard Model (MSSM) and General Singlet Extensions of the MSSM (GSEMSSM): A short review. High Energy Physics–Phenomenology. [arXiv:1911.13043] DOI: 10.48550/arXiv.1911.13043

[20] Farrar G. and Fayet P. (1978). Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry. Physics Letters B **76**(5), pp. 575-579. DOI: https://doi.org/10.1016/0370-2693(78)90858-4

[21] Hinchliffe I., Paige F.E., Shapiro M.D., Soderqvist J., and Yao W. (1996). Precision SUSY Measurements at LHC. **High Energy Physics – Phenomenology, LBNL-39412.** [arXiv:hep-ph/9610544] DOI: https://doi.org/10.48550/arXiv.hep-ph/9610544

[22] Djilibaev R. and Konoplich R. (2008). A new approach for reconstructing SUSY particle masses with a few $fb^{-1}$ at the LHC. Journal of High Energy Physics. DOI: 10.1088/1126-6708/2008/08/036

[23] Vermeulen, A. ATLAS Collaboration (2019). Reconstruction techniques in Supersymmetry searches with the ATLAS detector. Journal of Physics: Conference Series Vol **1390**. DOI: 10.1088/1742-6596/1390/1/012037

[24] Martin S. (2023). Status and Future of Supersymmetry. Retrieved from https://indico.cern.ch/event/1214022/contributions/5461067/attachments/2688809/4665483/SUSY2023_martin.pdf. Accessed on October 5, 2023.

[25] Aad G. et. al. (The ATLAS Collaboration) (2023). Search for direct pair production of sleptons and charginos decaying to two leptons and neutralinos with mass splittings near the W-boson mass in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector. Journal of Energy Physics Vol **31**. DOI: https://doi.org/10.1007/JHEP06(2023)031

[26] Aad G. et. al. (The ATLAS Collaboration) (2023). Search for long-lived, massive particles in events with displaced vertices and multiple jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. High Energy Physics Experiment, CERN-EP-2023-002. [arXiv:2301.13866] DOI: https://doi.org/10.48550/arXiv.2301.13866

[27] Aad G. et. al. (The ATLAS Collaboration) (2023). Search for supersymmetry in final states with missing transverse momentum and three or more b-jets in 139 fb-1 of proton–proton collisions at $\sqrt{\boldsymbol{s}} = \mathbf{13}$ TeV with the ATLAS detector. European Physics Journal C **83**(561). DOI: https://doi.org/10.1140/epjc/s10052-023-11543-6

[28] Tumasyan A. et. al. (The CMS Collaboration) (2023). Search for top squark pair production in a final state with at least one hadronically decaying tau lepton in proton-proton collisions at $\sqrt{\boldsymbol{s}} = \mathbf{13}$ TeV. Journal of High Energy Physics Vol **110**. DOI: https://doi.org/10.1007/JHEP07(2023)110

[29] Aad G. et. al. (The ATLAS Collaboration) (2023). Search in diphoton and dielectron final states for displaced production of Higgs or Z bosons with the ATLAS detector in $\sqrt{\boldsymbol{s}} = \mathbf{13}$ TeV pp collisions. Physical Review D **108**(1). DOI: https://doi.org/10.1103/PhysRevD.108.012012

[30] Tanabashi M. et. al. Particle Data Group (2018). Supersymmetric Particle Searches. Physics Review D Vol **98**. Retrieved from https://pdg.lbl.gov/2018/listings/rpp2018-list-supersymmetric-part-searches.pdf. Accessed on October 06, 2023.

[31] Workman R.L. et. al. Particle Data Group (2022). SEARCHES not in other sections. Progress of Theoretical and Experimental Physics. Retrieved from https://pdg.lbl.gov/2023/tables/rpp2023-sum-searches.pdf. Accessed on October 6, 2023.

[32] Baldi, P., P. Sadowski, and D. Whiteson. (2014). Searching for Exotic Particles in High-energy Physics with Deep Learning. Nature Communications **5**(4038). DOI: https://doi.org/10.1038/ncomms5308

[33] Andrews M., Paulini M., Gleyzer S., and Poczos B. (2018). End-to-End Event Classification of High-Energy Physics Data. Journal of Physics Conference Series **1085**(042022). DOI: doi :10.1088/1742-6596/1085/4/042022

[34] Liu X., Wang H., Zhang Y., Wu F., and Hu S. (2022). Towards Efficient Data-Centric Robust Machine Learning with Noise-based Augmentation. Data-Centric Robust Learning on ML Models, Adversarial Machine Learning and Beyond at AAAI2022. [arXiv:2203.03810] DOI: https://doi.org/10.48550/arXiv.2203.03810

[35] CMS Machine Learning Documentation. (2023). Retrieved from https://cms-ml.github.io/documentation/index.html. Accessed on October 6, 2023.

[36] He K., Zhang X., Ren S., and Sun J. (2016). Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778. DOI: 10.1109/CVPR.2016.90.

[37] Oliviera L., Paganini M., and Nachman B. (2017). Learning Particle Physics by Example: Location-Aware Generative Adversarial Networks for Physics Synthesis. Computing and Software for Big Science **1**(4). DOI: https://doi.org/10.1007/s41781-017-0004-6

[38] Cerri O., Nguyen T., Pierini M., Spiropulu M., and Vlimant J.R. (2019). Variational Autoencoders for New Physics Mining at the Large Hadron Collider. Journal of High Energy Physics Vol **36**. DOI: https://doi.org/10.1007/JHEP05(2019)036

[39] Bernido, C. C., & Carpio-Bernido, M. V. (2012b). White Noise Analysis: Some Applications in Complex Systems, Biophysics and Quantum Mechanics. International Journal of Modern Physics **B** **26**(29), 1230014. DOI: https://doi.org/10.1142/s0217979212300149

[40] Hida, T., Obata, N., & Saito, K. (1992). Infinite dimensional rotations and Laplacians in terms of white noise calculus. Nagoya Mathematical Journal Vol **128**, pp. 65–93. DOI: https://doi.org/10.1017/s0027763000004220.

[41] Hunlap D.H. and Kenkre V.M. (1986). Dynamic localization of a charged particle moving under the influence of an electric field. Physical Review B **34**(6), pp. 3625-3633. DOI: 10.1103/PhysRevB.34.3625

[42] Lentz W. (1976). Generating Bessel functions in Mie scattering calculations using continued fractions. Applied Optics **15**(3), pp. 668-671. DOI: https://doi.org/10.1364/AO.15.000668

[43] Singh S. (2008). Analytical solution to transient heat conduction in polar coordinates with multiple layers in radial direction. International Journal of Thermal Sciences **47**(3), pp. 261-273. DOI: https://doi.org/10.1016/j.ijthermalsci.2007.01.031

[44] Munasinghe R. (2023). Derivation of Bessel function closed-form solutions in zero dimensional $\varphi^4$-field theory. Heliyon **9**(2), e13168. DOI: 10.1016/j.heliyon.2023.e13168

[45] Bayatian G.L. et. al. (The CMS Collaboration). (2006). CMS Physics: Technical Design Report Volume 1: Detector Performance and Software. Report Number: CMS-TDR-8-1.

[46] Gradshteyn I. and Ryzhik I. (2014). Table of Integrals, Series, and Products, Eight Edition. Elsevier Inc., Academic Press. ISBN: 978-0-12-384933-5. DOI: https://doi.org/10.1016/C2010-0-64839-5

[47] Bini C. (2014). Data Analysis in Particle Physics. Department of Physics, Sapienza University of Rome and INFN, Rome.