New statistical method for enhancing the real-time recognition of astrophysical neutrino burst

PhD Candidate
Marco Mattiazzi

Università di Siena, and INFN Padova

December 15, 2021

Overview

1. Introduction
2. Clustering algorithms
3. Standard fixed-threshold method
4. Beta filter method
   - Dynamic clustering
   - Shifted clustering
5. Performance Evaluation
6. Conclusions

Main objective of this seminar

Aim

Enhance the **real-time** recognition of astrophysical neutrino burst in a **statistical** fashion

Requirements

- **short-latency** $\longrightarrow$ prompt alert to ν, EM and GW telescopes
- **low** false alarm rate $\longrightarrow$ crucial in the multi-messenger era

How-to

Exploit the different **time structure** of the **signal** with respect to the (*poissonian*) **background**

Main objective of this seminar

Aim

Enhance the **real-time** recognition of astrophysical neutrino burst in a **statistical** fashion

Requirements

- **short-latency** $\longrightarrow$ prompt alert to ν, EM and GW telescopes
- **low** false alarm rate $\longrightarrow$ crucial in the multi-messenger era

How-to

Exploit the different **time structure** of the **signal** with respect to the (*poissonian*) **background**

Main objective of this seminar

Aim

Enhance the **real-time** recognition of astrophysical neutrino burst in a **statistical** fashion

Requirements

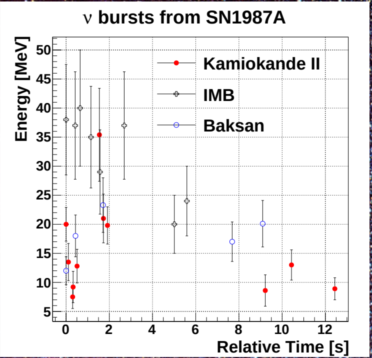
- **short-latency** $\longrightarrow$ prompt alert to ν, EM and GW telescopes
- **low** false alarm rate $\longrightarrow$ crucial in the multi-messenger era

How-to

Exploit the different **time structure** of the **signal** with respect to the (*poissonian*) **background**

CCSNe

- **last stage** lifecycle of **massive stars**
- released **energy** $\sim 10^{53}$ erg $\longrightarrow$ 99% emitted as **neutrinos**
- models based on highly complex simulations but **exact mechanism is still unknown**
- so far, **only** ν-bursts from **SN1987A**



Potential **scientific outcomes** from the next CCSN neutrino detection

- neutrino **mass** limit + **mass hierarchy** + constraints on BSM physics
- unique **probe** for black-hole forming SNe or other *exotic* explosions

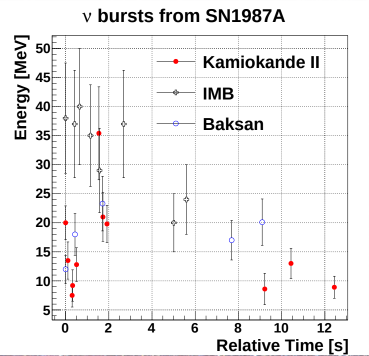
Real-time detection

- guide the optical telescopes to study **EM signal** from its **onset**

Application : ν-driven Core-Collapse Supernovae

CCSNe

- **last stage** lifecycle of **massive stars**
- released **energy** $\sim 10^{53}$ erg $\longrightarrow$ 99% emitted as **neutrinos**
- models based on highly complex simulations but **exact mechanism is still unknown**
- so far, **only** ν-bursts from **SN1987A**



Potential **scientific outcomes** from the next CCSN neutrino detection

- neutrino **mass** limit + **mass hierarchy** + constraints on BSM physics
- unique **probe** for black-hole forming SNe or other *exotic* explosions

Real-time detection

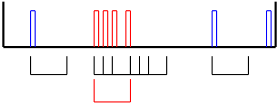
- guide the optical telescopes to study **EM signal** from its **onset**

Clustering algorithms

Dynamic fixed-size sliding time windows start from each event

Background

Signal



Static fixed-size sliding time windows without overlapping

Shifted same as static + *second translated scan* to reduce search bias

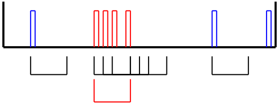
1^{st} Scan

2^{nd} Scan

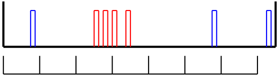
Clustering algorithms

Background
Signal

Dynamic fixed-size sliding time windows start from each event



Static fixed-size sliding time windows without overlapping



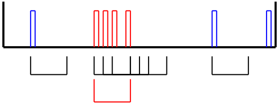
Shifted same as static + *second translated scan* to reduce search bias

1st Scan
2nd Scan

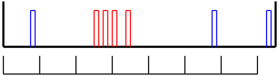
Clustering algorithms

Dynamic fixed-size sliding time windows start from each event

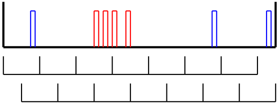
Background
Signal



Static fixed-size sliding time windows without overlapping



Shifted same as static + *second translated scan* to reduce search bias



1$^{\text{st}}$ Scan
2$^{\text{nd}}$ Scan

Standard method based on Imitation Frequency

The *standard method* for the real-time **identification** of **candidate SN cluster** is based on the so-called **imitation frequency**

$$\mathrm{F}_{im}(m \mid r, w) = N_{\mathrm{windows}}(t_{FAT}) \times P_{\mathrm{survival}}(m \mid rw)$$
$$= N_{\mathrm{windows}}(t_{FAT}) \times \sum_{k=m}^{\infty} \frac{(rw)^k e^{-rw}}{k!}$$

and the **fixed-threshold** $\hat{M}$ is defined such that

$$\mathrm{F}_{im}(\hat{M} \mid r, w) < 1$$

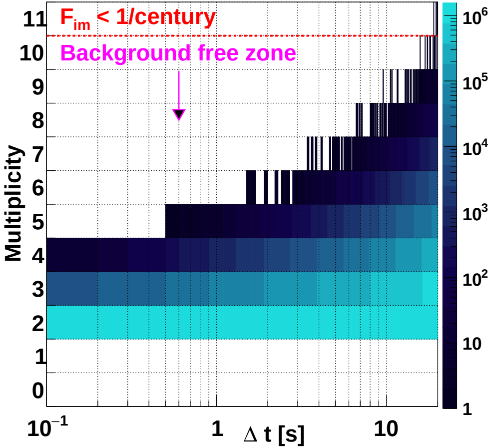
where:

- $w \longrightarrow$ time window size (*e.g.* 20 s)
- $t_{FAT} \longrightarrow$ required *false alert time* (*e.g.* 1 century)
- $r \longrightarrow$ Poisson background rate
- $m \longrightarrow$ multiplicity, *i.e.* number of events within the analysed time window

Standard method based on Imitation Frequency

Main Issue

Cut **only** on *multiplicity* $\longrightarrow$ no sensitivity on the **time structure** of the SN burst



Simulation parameters

- $w = 20$ s
- Rate $r = 0.03$ Hz
- $t_{FAT} = 1$ century
- $N_{\text{simulations}} = 10 \times t_{FAT}$

Hypothesis Only Poisson-distributed background

Δt *duration* of the cluster, *i.e.* time difference between the first and the last event

The statistical basis behind the newly proposed β method

Hypothesis: **stationary Poisson process** with constant rate r

$$P_{\text{survival}}(m \mid rw) = 1 - \sum_{k=0}^{m-1} \frac{(rt)^k \, e^{-rt}}{k!} = \sum_{k=m}^{\infty} \text{Pois}\,(k, rt)$$

probability of observing at least m events in a given time interval $[0, t]$

Key Idea: Order Statistics

In a fixed time window $W = [0, w]$, the m events are uniformly distributed with

$$0 < t_1 < t_2 < \cdots < t_m < w$$

Normalising to [0,1] via

$$x_k := t_k / w$$

then the normalised time of the k-th event follows $f(x_k \mid m, w) = \text{Beta}(x_k; k, m+1-k)$
i.e. the **k-th order statistic** of **uniform distribution**

Transforming back in time domain

$$f(t_k \mid m, w) = \frac{1}{w} \times \text{Beta}\left(\frac{t_k}{w}; k, m+1-k\right)$$

The statistical basis behind the newly proposed β method

Hypothesis: **stationary Poisson process** with constant rate r

$$P_{\text{survival}}(m \mid rw) = 1 - \sum_{k=0}^{m-1} \frac{(rt)^k \, e^{-rt}}{k!} = \sum_{k=m}^{\infty} \mathrm{Pois}\,(k, rt)$$

probability of observing at least m events in a given time interval $[0, t]$

Key Idea: Order Statistics

In a fixed time window $W = [0, w]$, the m events are uniformly distributed with

$$0 < t_1 < t_2 < \cdots < t_m < w$$

Normalising to [0,1] via

$$x_k := t_k / w$$

then the normalised time of the k-th event follows $f(x_k \mid m, w) = \mathrm{Beta}(x_k; k, m+1-k)$
i.e. the **k-th order statistic** of **uniform distribution**

Transforming back in time domain

$$f(t_k \mid m, w) = \frac{1}{w} \times \mathrm{Beta}\left(\frac{t_k}{w}; k, m+1-k\right)$$

The statistical basis behind the newly proposed β method

Hypothesis: **stationary Poisson process** with constant rate r

$$P_{\text{survival}}(m \mid rw) = 1 - \sum_{k=0}^{m-1} \frac{(rt)^k \, e^{-rt}}{k!} = \sum_{k=m}^{\infty} \text{Pois}\,(k, rt)$$

probability of observing at least m events in a given time interval $[0, t]$

Key Idea: Order Statistics

In a fixed time window $W = [0, w]$, the m events are uniformly distributed with

$$0 < t_1 < t_2 < \cdots < t_m < w$$

Normalising to [0,1] via

$$x_k := t_k / w$$

then the normalised time of the k-th event follows $f(x_k \mid m, w) = \text{Beta}(x_k; k, m+1-k)$
i.e. the **k-th order statistic** of **uniform distribution**

Transforming back in time domain

$$f(t_k \mid m, w) = \frac{1}{w} \times \text{Beta}\left(\frac{t_k}{w}; k, m+1-k\right)$$

Order statistics: Theory

Definition

Let $X_1, X_2, \cdots, X_N$ be N random variables *i.i.d* with $F(x)$ distribution, **sorting** in increasing order

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(N)}$$

are r.v. as well, called **order statistics**

Cumulative distribution

- **largest order statistic** $X_{(N)} = \max\{X_1, X_2, \cdots, X_N\}$

$$F_{X_{(N)}}(x) = \Pr(X_{(N)} \leq x) = [F(x)]^n$$

- **first order statistic** $X_{(1)} = \min\{X_1, X_2, \cdots, X_N\}$

$$F_{X_{(1)}}(x) = \Pr(X_{(1)} \leq x) = 1 - [1 - F(x)]^n$$

- **k-th order statistic**

$$F_{X_{(k)}}(x) = \Pr(\text{at least } k \text{ of } X\text{'s are} \leq x) = \sum_{j=k}^{n} \binom{n}{j} [F(x)]^j [1 - F(x)]^{n-j}$$

Order statistics: Theory

Definition

Let $X_1, X_2, \cdots, X_N$ be N random variables *i.i.d* with $F(x)$ distribution, **sorting** in increasing order

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(N)}$$

are r.v. as well, called **order statistics**

Cumulative distribution

- **largest order statistic** $X_{(N)} = \max\{\, X_1, X_2, \cdots, X_N \,\}$

$$F_{X_{(N)}}(x) = \Pr(X_{(N)} \leq x) = [F(x)]^n$$

- **first order statistic** $X_{(1)} = \min\{\, X_1, X_2, \cdots, X_N \,\}$

$$F_{X_{(1)}}(x) = \Pr(X_{(1)} \leq x) = 1 - [1 - F(x)]^n$$

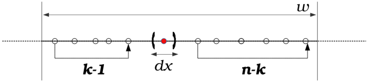
- **k-th order statistic**

$$F_{X_{(k)}}(x) = \Pr(\text{ at least } k \text{ of } X\text{'s are} \leq x) = \sum_{j=k}^{n} \binom{n}{j} [F(x)]^j \, [1 - F(x)]^{n-j}$$

Order statistics: Theory

PDF k-th order statistic

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} f(x)[F(x)]^{k-1}[1-F(x)]^{n-k}$$



Heuristic Proof

$$\begin{aligned}
\Pr(X_{(k)} \in [x, x+dx]) &= \Pr(\text{one of the } X\text{'s } \in [x, x+dx] \quad \text{and} \quad k-1 \quad \text{of the others } X\text{'s} < x) \\
&= [\, n \Pr(X \in [x, x+dx]) \,] \Pr(k-1 \text{ of other } X\text{'s} < x) \\
&= n\, f(x)\, dx \quad \binom{n-1}{k-1} [\Pr(X < x)]^{k-1} \, [\Pr(X > x)]^{n-k} \\
&= n \binom{n-1}{k-1} f(x)\, dx \quad [F(x)]^{k-1} \, [1 - F(x)]^{n-k}
\end{aligned}$$

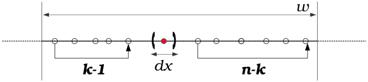
Recalling Gamma function definition for integers, $\Gamma(n+1) = n!$

$$f_{X_{(k)}}(x) = \frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k)} f(x)[F(x)]^{k-1}[1-F(x)]^{n-k} \qquad \square$$

Order statistics: Theory

PDF k-th order statistic

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} f(x)[F(x)]^{k-1}[1-F(x)]^{n-k}$$



Heuristic Proof

$$\begin{aligned}
\Pr(X_{(k)} \in [x, x+dx]) &= \Pr(\text{one of the } X\text{'s } \in [x, x+dx] \quad \text{and} \quad k-1 \quad \text{of the others } X\text{'s} < x) \\
&= [\, n \Pr(X \in [x, x+dx]) \,] \Pr(k-1 \text{ of other } X\text{'s} < x) \\
&= n\, f(x)\, dx \quad \binom{n-1}{k-1} [\Pr(X < x)]^{k-1} [\Pr(X > x)]^{n-k} \\
&= n \binom{n-1}{k-1} f(x)\, dx \quad [F(x)]^{k-1} [1 - F(x)]^{n-k}
\end{aligned}$$

Recalling Gamma function definition for integers, $\Gamma(n+1) = n!$

$$f_{X_{(k)}}(x) = \frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k)} f(x)[F(x)]^{k-1}[1-F(x)]^{n-k} \qquad \square$$

Order statistics: Theory

Beta distribution

$$\text{Beta}(u;\alpha,\beta) = \frac{u^{\alpha-1}(1-u)^{\beta-1}}{\int_0^1 v^{\alpha-1}(1-v)^{\beta-1}\,dv} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\,u^{\alpha-1}(1-u)^{\beta-1} \qquad u\in[0,1]$$

Order statistics of Uniform distribution on Unit interval

Let $X_1, X_2, \cdots, X_N$ be N random variables *i.i.d* with $X \sim \text{Uniform}(0,1)$, the **k-th order statistic**

$$\begin{aligned} f_{X_{(k)}}(u) &= \frac{n!}{(k-1)!(n-k)!}u^{k-1}(1-u)^{n-k} \qquad \text{with } 0 \le u \le 1 \\ &= \frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k)}u^{k-1}(1-u)^{n-k} \\ &= \text{Beta}(u;k,n-k+1) \end{aligned}$$

Thus

$$X_{(k)} \sim \text{Beta}(k, n-k+1)$$

Order statistics: Theory

Beta distribution

$$\text{Beta}(u; \alpha, \beta) = \frac{u^{\alpha-1}(1-u)^{\beta-1}}{\int_0^1 v^{\alpha-1}(1-v)^{\beta-1}\, dv} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, u^{\alpha-1}(1-u)^{\beta-1} \qquad u \in [0,1]$$

Order statistics of Uniform distribution on Unit interval

Let $X_1, X_2, \cdots, X_N$ be N random variables *i.i.d* with $X \sim \text{Uniform}(0,1)$, the **k-th order statistic**

$$\begin{aligned} f_{X_{(k)}}(u) &= \frac{n!}{(k-1)!(n-k)!} u^{k-1}(1-u)^{n-k} \qquad \text{with } 0 \le u \le 1 \\ &= \frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k)} u^{k-1}(1-u)^{n-k} \\ &= \text{Beta}(u; k, n-k+1) \end{aligned}$$

Thus

$$X_{(k)} \sim \text{Beta}(k, n-k+1)$$

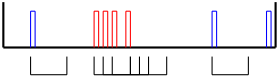
Dynamic clustering

PDF: by construction $t_1 = 0 \longrightarrow \Delta t = t_m$

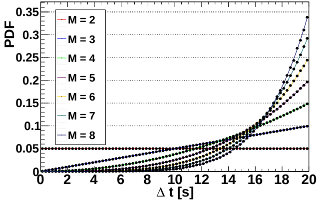
$$f_{\rm D}(\Delta t \mid m, w) = f_\beta(t_m \mid m-1, w)$$
$$= \frac{1}{w} \times \mathrm{Beta}\left(\frac{\Delta t}{w}; m-1, 1\right)$$
$$= \frac{m-1}{w^{m-1}} \times \Delta t^{m-2}$$

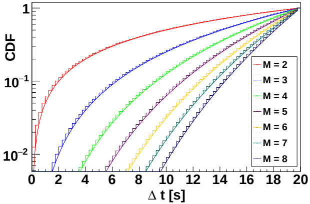
CDF: fixed multiplicity m

$$F_{\rm D}(\Delta t \mid m) = \mathrm{Prob}(t \leq \Delta t \mid m, w)$$
$$= \int_0^{\Delta t} dt\, f_{\rm D}(t|m, w)$$
$$= \frac{1}{m}\mathrm{Beta}\left(\frac{\Delta t}{w}; m, 1\right)$$
$$= \left(\frac{\Delta t}{w}\right)^{m-1}$$



Numerical results + theoretical distributions





Dynamic clustering

Discrete multiplicity density distribution $\longrightarrow$ truncated Poisson pdf

$$g_{\mathrm{D}}(m \mid r, w) = \frac{(rw)^{m-1}\, e^{-rw}}{(m-1)!}\, \frac{1}{1-e^{-rw}} = \frac{\mathrm{Pois}(m-1, \mu = rw)}{1-e^{-\mu}}$$

Joint probability density distribution

$$j_{\mathrm{D}}(m, \Delta t \mid r, w) = f_{\mathrm{D}}(\Delta t \mid m, w) \times g_{\mathrm{D}}(m \mid r, w) = \frac{re^{-rw}}{1-e^{-rw}}\, \frac{(r\Delta t)^{m-2}}{(m-2)!}$$

Key Idea: Decision Boundary

Build decision boundary, $\{\Delta t_k\}$ with $k \geq 2$, on the $m - \Delta t$ plane such that the required **false alert rate** is ensured in analytical way $\longrightarrow$ **no MC-based thresholds**

Dynamic clustering

Discrete multiplicity density distribution $\longrightarrow$ truncated Poisson pdf

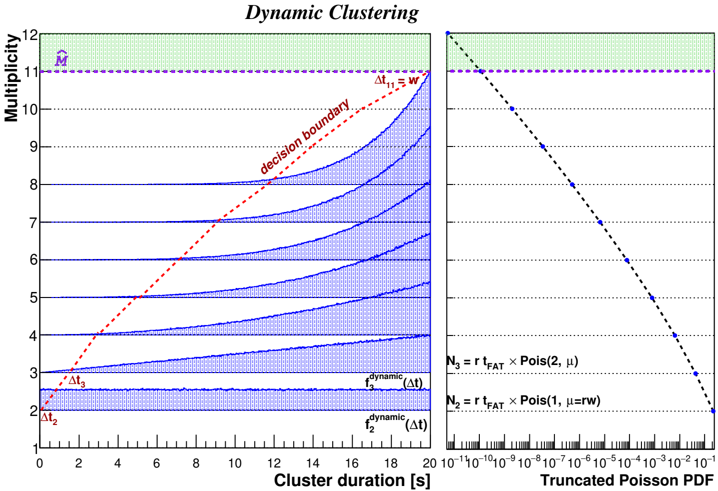
$$g_{\mathrm{D}}(m \mid r, w) = \frac{(rw)^{m-1}\, e^{-rw}}{(m-1)!} \, \frac{1}{1 - e^{-rw}}$$
$$= \frac{\mathrm{Pois}(m-1, \mu = rw)}{1 - e^{-\mu}}$$

Joint probability density distribution

$$j_{\mathrm{D}}(m, \Delta t \mid r, w) = f_{\mathrm{D}}(\Delta t \mid m, w) \times g_{\mathrm{D}}(m | \, r, w)$$
$$= \frac{re^{-rw}}{1 - e^{-rw}} \, \frac{(r\Delta t)^{m-2}}{(m-2)!}$$

Key Idea: Decision Boundary

Build decision boundary, $\{\Delta t_k\}$ with $k \geq 2$, on the $m - \Delta t$ plane such that the required **false alert rate** is ensured in analytical way $\longrightarrow$ **no MC-based thresholds**



Key Idea: Decision Boundary

Build decision boundary, $\{\Delta t_k\}$ with $k \geq 2$, on the $m - \Delta t$ plane such that the required **false alert rate** is ensured in analytical way $\longrightarrow$ **no MC-based thresholds**

Dynamic clustering: Decision Boundary Building

Constraint

$$\sum_{k=2}^{\infty} N_k \times F_{\mathrm{D}}(\Delta t_k \mid k) \leq 1$$

Expected number of clusters with multiplicity k over t_{FAT}

$$N_k = r t_{FAT} \times \mathrm{Pois}\left(k-1, \mu = rw\right)$$

Find $\hat{M}$ such that

$$\sum_{j=\hat{M}}^{\infty} N_j < 1$$

e.g. the standard imitation frequency threshold

Set $\Delta t_j = w$ for $j \geq \hat{M}$, the constraint can be rearranged as

$$\sum_{k=2}^{\hat{M}-1} \mathrm{Pois}\left(k-1, \mu\right) \times \left(\frac{\Delta t_k}{w}\right)^{k-1} \leq \frac{1 - \sum_{j=\hat{M}}^{\infty} N_j}{r t_{FAT}}$$

Dynamic clustering: Decision Boundary Building

Constraint

$$\sum_{k=2}^{\infty} N_k \times F_{\mathrm{D}}(\Delta t_k \mid k) \leq 1$$

Expected number of clusters with multiplicity k over t_{FAT}

$$N_k = rt_{FAT} \times \mathrm{Pois}\left(k-1, \mu = rw\right)$$

Find $\hat{M}$ such that

$$\sum_{j=\hat{M}}^{\infty} N_j < 1$$

e.g. the standard imitation frequency threshold

Set $\Delta t_j = w$ for $j \geq \hat{M}$, the constraint can be rearranged as

$$\sum_{k=2}^{\hat{M}-1} \mathrm{Pois}\left(k-1, \mu\right) \times \left(\frac{\Delta t_k}{w}\right)^{k-1} \leq \frac{1 - \sum_{j=\hat{M}}^{\infty} N_j}{rt_{FAT}}$$

Dynamic clustering: Decision Boundary Building

Constraint

$$\sum_{k=2}^{\infty} N_k \times F_{\mathrm{D}}(\Delta t_k \mid k) \leq 1$$

Expected number of clusters with multiplicity k over t_{FAT}

$$N_k = r t_{FAT} \times \mathrm{Pois}\left(k-1, \mu = rw\right)$$

Find $\hat{M}$ such that

$$\sum_{j=\hat{M}}^{\infty} N_j < 1$$

e.g. the standard imitation frequency threshold

Set $\Delta t_j = w$ for $j \geq \hat{M}$, the constraint can be rearranged as

$$\sum_{k=2}^{\hat{M}-1} \mathrm{Pois}\left(k-1, \mu\right) \times \left(\frac{\Delta t_k}{w}\right)^{k-1} \leq \frac{1 - \sum_{j=\hat{M}}^{\infty} N_j}{r t_{FAT}}$$

Dynamic clustering: Decision Boundary Building

β discrete filter

Introducing the $\{\beta_k\}_{k=2}^{\hat{M}-1}$ discrete filter, as follows

$$\mathrm{Pois}\,(k-1,\mu) \times \left(\frac{\Delta t_k}{w}\right)^{k-1} = \beta_k \times \frac{1-\sum_{j=\hat{M}}^{\infty} N_j}{rt_{FAT}}$$

$$\sum_{k=2}^{\hat{M}-1} \beta_k = 1$$

Such set of equations allow **exploring the region $m < \hat{M}$**

Decision boundary : Analytical Formula

Finally we get

$$\Delta t_k = \begin{cases} w\left[\beta_k \cdot \frac{1-\sum_{j=\hat{M}}^{\infty} N_j}{N_k}\right]^{1/(k-1)} & \text{if } k < \hat{M} \\ w & \text{if } k \geq \hat{M} \end{cases}$$

Dynamic clustering: Decision Boundary Building

β discrete filter

Introducing the $\{\beta_k\}_{k=2}^{\hat{M}-1}$ discrete filter, as follows

$$\mathrm{Pois}\,(k-1,\mu) \times \left(\frac{\Delta t_k}{w}\right)^{k-1} = \beta_k \times \frac{1 - \sum_{j=\hat{M}}^{\infty} N_j}{rt_{FAT}}$$

$$\sum_{k=2}^{\hat{M}-1} \beta_k = 1$$

Such set of equations allow **exploring the region $m < \hat{M}$**

Decision boundary : Analytical Formula

Finally we get

$$\Delta t_k = \begin{cases} w \left[\beta_k \cdot \frac{1 - \sum_{j=\hat{M}}^{\infty} N_j}{N_k}\right]^{1/(k-1)} & \text{if } k < \hat{M} \\ w & \text{if } k \geq \hat{M} \end{cases}$$

Dynamic clustering: β filter application

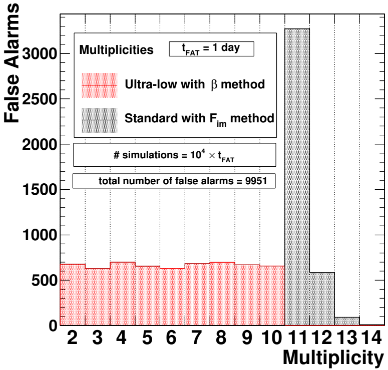
Uniform filter $\beta_k = \beta$

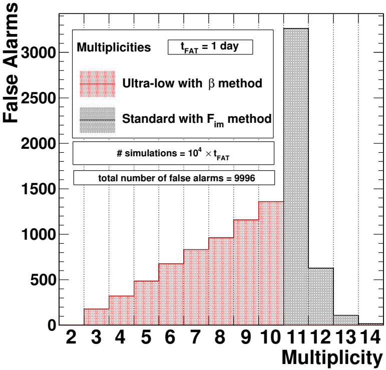
$\sum_{k=2}^{\hat{M}-1} \beta = 1 \longrightarrow \beta = \frac{1}{\hat{M}-2}$

Ramp filter $\beta_k = \beta \times (k-2)$

$\sum_{k=2}^{\hat{M}-1} \beta_k = 1 \longrightarrow \beta = \frac{2}{(\hat{M}-2)(\hat{M}-3)}$

Example: *dynamic* clustering, w =20 s, t_{FAT} = 1 day, r = 0.1 Hz, $N_{simu} = 10^4 \times t_{FAT}$





Remark on β filter method

High flexibility of β filter method $\longrightarrow$ set **additional constraints** based on *e.g.*

- penalising region where non-poissonian background occurs more frequently
- discarding the multiplicities with Δt_k below the timing resolution of the detector
- ...

General polynomial β filter

$$\beta_k = \beta \times \sum_{j=0}^{\infty} a_j\, k^j$$

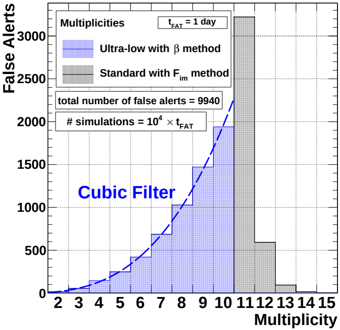
Reminder: β discrete filter

Dynamic clustering

$$\text{Pois}\,(k-1,\mu) \times \left(\frac{\Delta t_k}{w}\right)^{k-1} = \beta_k \times \frac{1 - \sum_{j=\hat{M}}^{\infty} N_j}{r t_{FAT}}$$

$$\sum_{k=2}^{\hat{M}-1} \beta_k = 1$$

$a_3 = 1,\ a_2 = a_1 = 0,\ a_0 = -1$



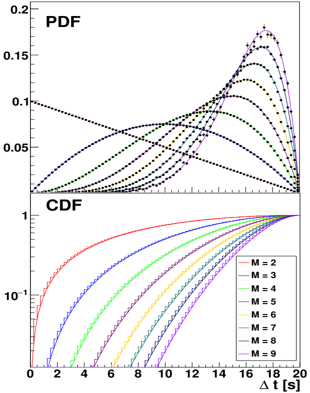
Shifted clustering

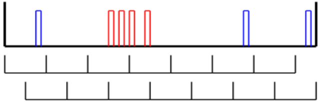
PDF: $\Delta t = t_m - t_1 \longrightarrow$ **sample range** of order statistics of uniform distribution

$$f_S(\Delta t|m,w) = \frac{m\,(m-1)}{w}\left(\frac{\Delta t}{w}\right)^{m-2}\left(1-\frac{\Delta t}{w}\right)$$
$$= \frac{1}{w} \times \mathrm{Beta}\left(\frac{\Delta t}{w}; m-1, 2\right)$$

CDF: fixed multiplicity m

$$F_S(\Delta t\,|\,m) = \mathrm{Prob}(t \leq \Delta t|m,w)$$
$$= \int_0^{\Delta t} dt\, f_S(t|m,w)$$
$$= \mathrm{Beta}\left(\frac{\Delta t}{w}; m, 1\right) \times$$
$$\left[1 - \left(\frac{m-1}{m}\right)\left(\frac{\Delta t}{w}\right)\right]$$





Shifted clustering

- discrete multiplicity pdf

$$g_S(m \mid r, w) = \frac{\mathrm{Pois}(m, rw)}{1 - e^{-rw}(1 + rw)}$$

- $\#$ clusters, multiplicity k over t_{FAT}

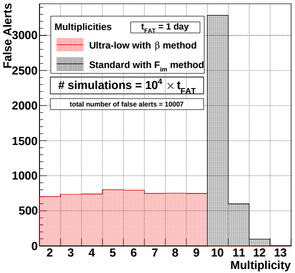
$$N_k = N_W \times \mathrm{Pois}\,(k, \mu = rw)$$

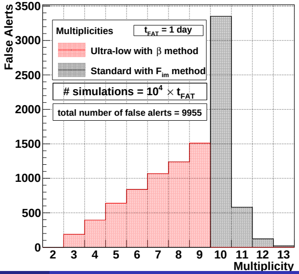
where $N_W = (2t_{FAT})/w - 1$

Likewise the dynamic case

$$N_k \times F_S(\Delta t_k \mid k) = \beta_k \times [1 - \sum_{j=\hat{M}}^{\infty} N_j] \quad \text{for } k < \hat{M}$$

solvable for $\{\Delta t_k\}_{k=2}^{\hat{M}-1} \longrightarrow$ root between $[0, w]$ of the associated k-th degree polynomials





Shifted clustering

- discrete multiplicity pdf

$$g_S(m \mid r, w) = \frac{\mathrm{Pois}(m, rw)}{1 - e^{-rw}(1 + rw)}$$

- $\#$ clusters, multiplicity k over t_{FAT}

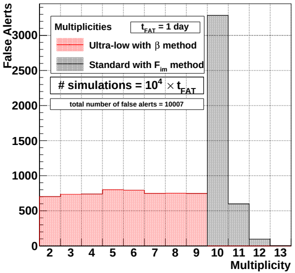
$$N_k = N_W \times \mathrm{Pois}\left(k, \mu = rw\right)$$

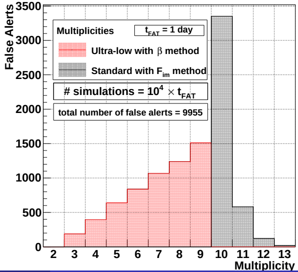
where $N_W = (2t_{FAT})/w - 1$

Likewise the dynamic case

$$N_k \times F_S(\Delta t_k \mid k) = \beta_k \times [1 - \sum_{j=\hat{M}}^{\infty} N_j] \quad \text{for } k < \hat{M}$$

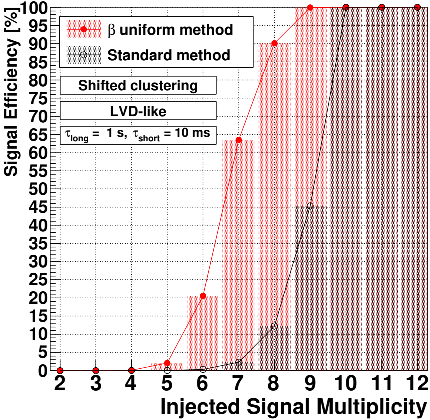
solvable for $\{\Delta t_k\}_{k=2}^{\hat{M}-1} \longrightarrow$ root between $[0, w]$ of the associated k-th degree polynomials





Performance Enhancement Evaluation

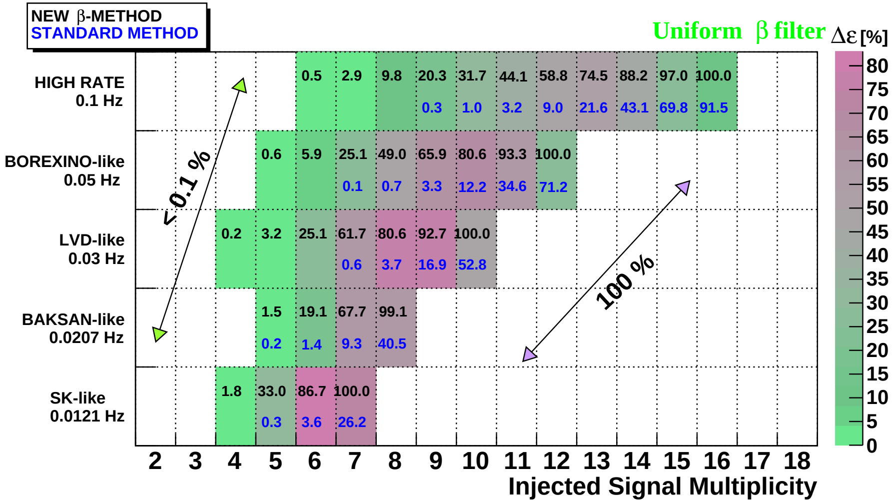
Shifted clustering, r =0.03 Hz, w =20 s, $\tau_{\mathrm{short}} = 10$ ms and $\tau_{\mathrm{long}} = 1$ s



SN-like ν signal $$f_{\mathrm{signal}}(t) = e^{-\frac{t}{\tau_{\mathrm{long}}}}\left(1 - e^{-\frac{t}{\tau_{\mathrm{short}}}}\right)$$

Performance Enhancement Evaluation

Dynamic clustering, w =20 s, τ_{short} = 10 ms and τ_{long} = 1 s



Summary

New β filter method

Main Result

- important **performance improvement** (up to $\sim 80\%$) for **real-time SN-ν monitor** with respect to standard technique

Pro's

- exploit **timing information** $\longrightarrow$ bi-dimensional cut on $m - \Delta t$ plane
- **no MC-based thresholds** $\longrightarrow$ only 1D root-finding algorithm required
- **flexibility** in the construction of $\{\Delta t_k\}_{k=2}^{\hat{M}-1}$ decision boundary $\longrightarrow$ allow setting detector-based thresholds
- **independent** on the nature of clusters $\longrightarrow$ applicable on wide class of real-time discrimination processes where **sliding-time windows** are involved + **fixed false discovery rate** is required

Con's

- careful **identification** of **non-poissonian background** component

Summary

New β filter method

Main Result

- important **performance improvement** (up to $\sim 80\%$) for **real-time SN-ν monitor** with respect to standard technique

Pro's

- exploit **timing information** $\longrightarrow$ bi-dimensional cut on $m - \Delta t$ plane
- **no MC-based thresholds** $\longrightarrow$ only 1D root-finding algorithm required
- **flexibility** in the construction of $\{\Delta t_k\}_{k=2}^{\hat{M}-1}$ decision boundary $\longrightarrow$ allow setting detector-based thresholds
- **independent** on the nature of clusters $\longrightarrow$ applicable on wide class of real-time discrimination processes where **sliding-time windows** are involved + **fixed false discovery rate** is required

Con's

- careful **identification** of **non-poissonian background** component

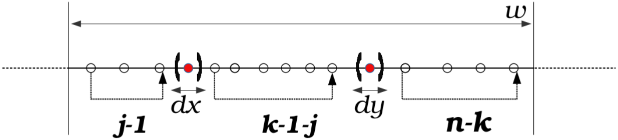
References

- M. Mattiazzi, M. Lamoureux and G. Collazuol, *On a new statistical technique for the real-time recognition of ultra-low multiplicity astrophysical neutrino burst*, arXiv:2106.12345
- O. Halim, C. Vigorito, C. Casentini, G. Pagliaroli, M. Drago and V. Fafone, *Expanding Core-Collapse Supernova Search Horizon of Neutrino Detectors*, J. Phys. Conf. Ser. **1468** (2020) 012154 [arXiv:1911.11450]
- C. Casentini, G. Pagliaroli, C. Vigorito and V. Fafone, *Pinpointing astrophysical bursts of low-energy neutrinos embedded into the noise*, JCAP **08** (2018) 010 [arXiv:1801.09062]
- N. Yu. Agafonova et al., *On-line recognition of supernova neutrino bursts in the LVD detector*, Astropart. Phys. **28** (2008) 516 [arXiv:0710.0259]
- SNEWS collaboration, *SNEWS 2.0: a next-generation supernova early warning system for multi-messenger astronomy*, New J. Phys. **23** (2021) 031201 [arXiv:2011.00035]
- P. Antonioli et al., *SNEWS: The Supernova Early Warning System*, New J. Phys. **6** (2004) 114 [astro-ph/0406214]

Thanks for your attention

Backup

[Backup] Joint PDF of two order statistics



Likewise the pdf of k-th order statistic's case

$$f_{X_{(j)},X_{(k)}}(x,y)dx\,dy = n\,(n-1) \qquad \cdots \qquad f(x)f(y)\,dx\,dy \quad \text{where } x \leq y$$

$$f_{X_{(j)},X_{(k)}}(x,y) = \frac{n(n-1)(n-2)!}{(j-1)!(k-j-1)!(n-k)!}[F(x)]^{j-1}[F(y)-F(x)]^{k-1-j}[1-F(y)]^{n-k}f(x)f(y)$$

$$f_{X_{(j)},X_{(k)}}(x,y) = \frac{n!}{(j-1)!(k-j-1)!(n-k)!}[F(x)]^{j-1}[F(y)-F(x)]^{k-1-j}[1-F(y)]^{n-k}f(x)f(y)$$

[**Backup**] Sample range of o.s. of uniform distribution

$$f_{X_{(1)},X_{(n)}}(x,y) = \frac{n!}{(n-2)!}[F(y) - F(x)]^{n-2} f(x) f(y)$$

Let $X_1, X_2, \cdots, X_N$ be N random variables *i.i.d* with $X \sim \text{Uniform}(0,1)$

$$f_{X_{(1)},X_{(n)}}(x,y) = n(n-1)[y-x]^{n-2} \quad \text{where} \quad 0 \leq x \leq y \leq 1$$

change of variable $w = y - x$ and $z = x$

$$f_{X_{(1)},X_{(n)}}(w,z) = n(n-1)w^{n-2}$$

marginalising, $y \leq 1 \longrightarrow w + z \leq 1$

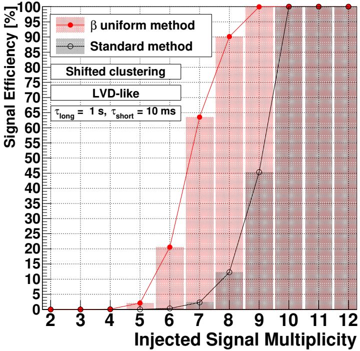
$$\begin{aligned} f_{X_{(1)},X_{(n)}}(w) &= \int_0^{1-w} f_{X_{(1)},X_{(n)}}(w,z) dz \\ &= n(n-1)w^{n-2}(1-w) \\ &= \text{Beta}(w; n-1, 2) \end{aligned}$$

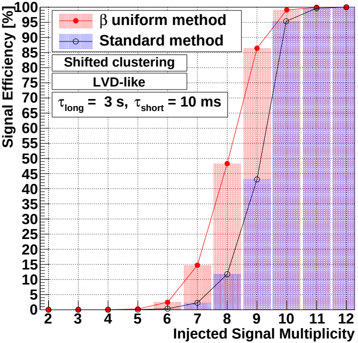
Thus **sample range** of order statistics of Uniform distribution

$$\boxed{f_{1,n}(w) \sim \text{Beta}(w; n-1, 2)}$$

[Backup] Performance Enhancement Evaluation

Shifted clustering, r =0.03 Hz, w =20 s, $N_{\text{simulations}} = 10^7$





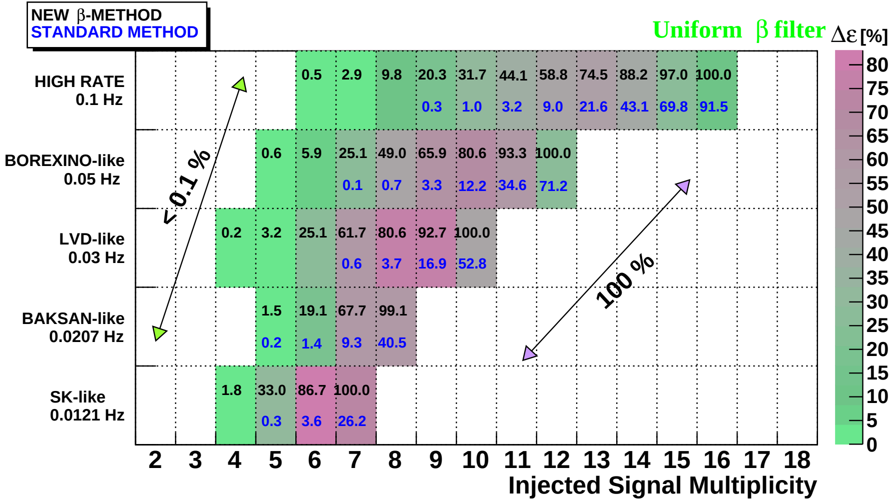
$\tau_{\text{short}} = 10$ ms and $\tau_{\text{long}} = \mathbf{1}$ **s**

$\tau_{\text{short}} = 10$ ms and $\tau_{\text{long}} = \mathbf{3}$ **s**

SN-like ν signal $\quad f_{\text{signal}}(t) = e^{-\lambda_{\text{long}} t}(1 - e^{-\lambda_{\text{short}} t}) \quad \lambda = 1/\tau$

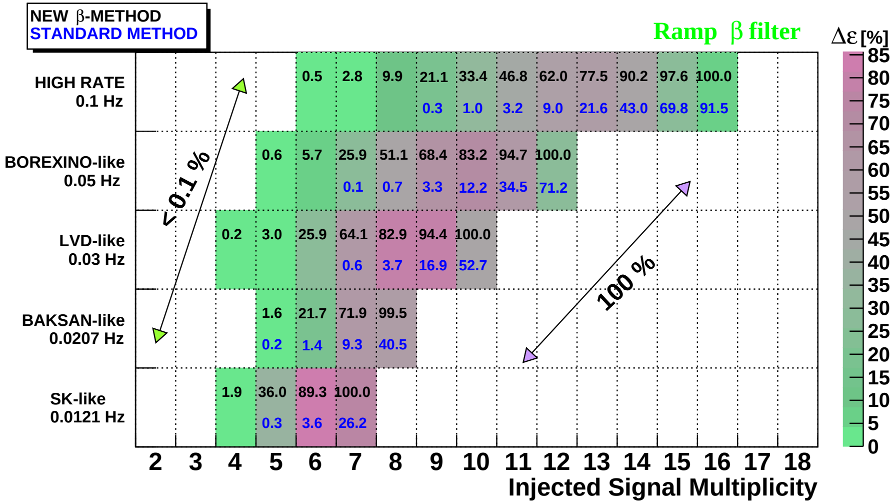
[Backup] Performance Enhancement Evaluation

Dynamic Uniform clustering, w =20 s, $\tau_{\rm short}$ = 10 ms and $\tau_{\rm long}$ = 1 s



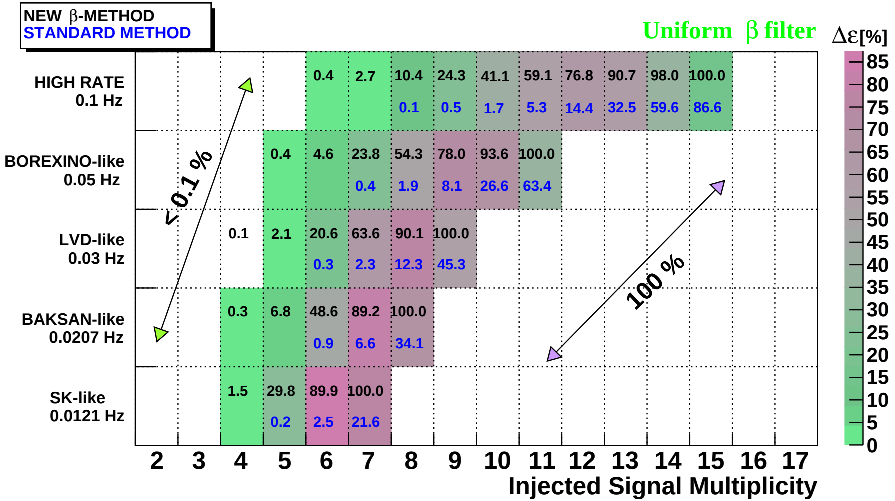
[Backup] Performance Enhancement Evaluation

Dynamic Ramp clustering, $w = 20$ s, $\tau_{\rm short} = 10$ ms and $\tau_{\rm long} = 1$ s



[Backup] Performance Enhancement Evaluation

Shifted Uniform clustering, w =20 s, $\tau_{\rm short}$ = 10 ms and $\tau_{\rm long}$ = 1 s



Shifted Ramp clustering, w =20 s, τ_{short} = 10 ms and τ_{long} = 1 s