

Jel-zaj arány

Balázs Boldizsár*
Eötvös Loránd Tudományegyetem

2015. június 10.

A gravitációs hullámok detektálása azért rendkívül komoly kihívás, mert a beérkező hullám detektorra gyakorolt hatásából származó jel olyan kicsiny, hogy a minden fizikai mérésből kiküszöbölhetetlen zajhatásoktól (a LIGO esetében: sörétzaj, fénynyomás zaja, szeizmikus zaj, hőmérsékleti zaj) nehéz elkülöníteni. Ahhoz, hogy egy adatsorba regisztrált jelnek a zajhoz viszonyított arányáról kvantitíve tudjunk beszélni, bevezetjük a jel-zaj arány matematikai fogalmát (angolul: Signal-to-Noise Ratio, SNR).

1. Az jel-zaj arány elméleti származtatása

Most a gravitációshullám-asztrófizikában leggyakrabban használt módon vezetjük be a jel-zaj arányt, de egyes szerzők ettől eltérhetnek, ezért az adott munkában mindig meg kell vizsgálni, hogy a szerzők milyen definícióval dolgoznak. Tegyük fel, hogy forrásmodellből és a forrás és a detektor egymáshoz képesti elhelyezkedéséből ismerjük, hogy a detektor kimenetében a forrásból érkező gravitációshullám-jel milyen egyoldali teljesítményspektrum-sűrűséggel (angolul: Power Spectral Density, PSD) jelenik meg. a gravitációs hullám jelének egyoldali teljesítményspektrum-sűrűségét (angolul: Power Spectral Density, PSD). Ez jelölésben: $|h(f)|^2$, lásd: 12. tétel [2]

Emellett ismerjük a zaj egyoldali PSD-jét is, amit $S(f)$ -fel jelölünk. Ez a detektor sok órás nyers kimenetéből becsülhető.

Ekkor a jel-zaj arány definíciója:

$$\text{SNR} \equiv 2\sqrt{\int_0^\infty \frac{|h(f)|^2}{S(f)} df} \quad (1)$$

Ahogy a neve is sugallja, ez a mennyiség azt mondja meg, hogy átlagosan hányszorosa az adatsorban rögzített jel energiája a háttérzajéhoz képest. A mérés előtt megválasztott jel-zaj arány küszöbérték (SNR_0) - ami jellemzően 3, 5, 8 értékeket vesz fel - azt dönti el, hogy a $|h(f)|^2$ PSD-vel megadott jel az $S(f)$ PSD-jű háttérzajban detektálható-e. Ha a (1) alapján kiszámolt SNR a küszöbértékkel egyenlő vagy afeletti, a jelet detektálhatónak ítéljük. Ha az alatti, a

*balazs.boldizsar@gmail.com

jelet detektálhatatlannak ítéljük. A küszöbérték megválasztása az eltűrt hamis riasztási gyakoriságtól függ (angolul: False Alarm Rate, FAR).[5] Minél kevésbé toleráljuk a hamis riasztásokat - azokat az eseményeket, amikor valójában nincsen gravitációs hullám, és a jel csak a zajból származik - annál magasabbra kell választanunk a jel-zaj arány küszöbértéket. A detektor zajának ismeretében egy még eltűrt FAR-hoz megtalálható az adott SNR_0 érték. Azonban ezt az értéket nem érdemes túlságosan magasnak választani, mert ekkor elszalaszthatunk valódi gravitációs hullámokat, tehát megnő a hamis tévesztési gyakoriság (angolul: False Dismissal Rate, FDR).[5] A kettő kényes egyensúlyának megtalálása teszi nehezzé a helyes SNR_0 érték megválasztását.

2. Logaritmikus skálán

A PSD-t gyakran logaritmikus skálán ábrázoljuk, ezért ilyen esetekben érdemes áttérni az ezen való integrálásra. Ekkor:

$$df = d(e^{\log(f)}) = e^{\log(f)} d(\log(f)) = f d(\log(f)) \quad (2)$$

A (1) egyenlet integrandusában megjelenő f szorzófaktort beolvaszthatjuk a számlálóba, és átírhatjuk a (1) egyenletet az alábbi alakba:

$$\text{SNR} = 2 \sqrt{\int_0^\infty \frac{|h(f)\sqrt{f}|^2}{S(f)} d(\log f)} = 2 \sqrt{\int_0^\infty \frac{H(f)}{S(f)} d(\log f)} \quad (3)$$

Ahol $H(f) \equiv |h(f)\sqrt{f}|^2$. Kihhasználva, hogy $H(f) = e^{\log(H(f))}$ és $S(f)^{-1} = e^{-\log(S(f))}$:

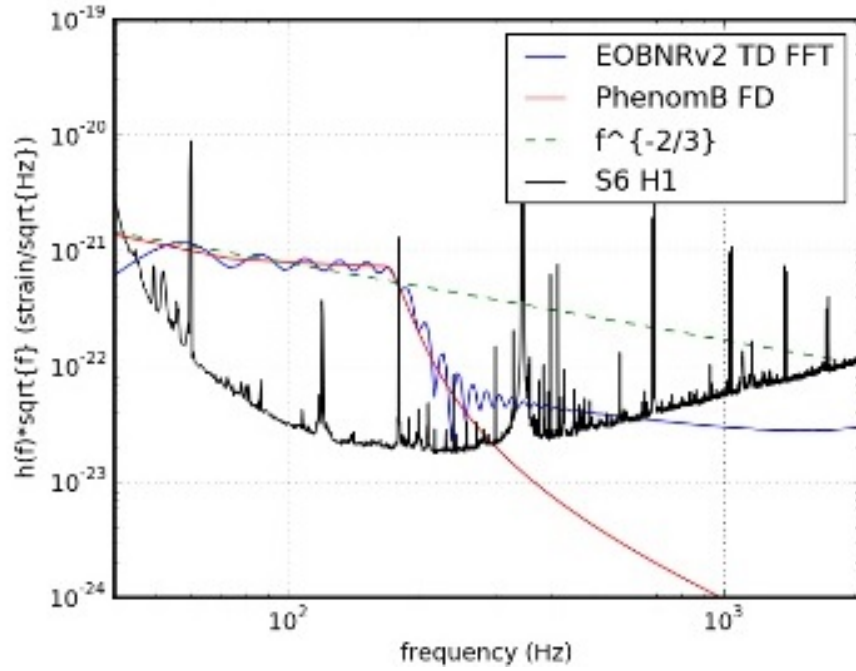
$$\text{SNR} = 2 \sqrt{\int_0^\infty \exp(\log(H(\log(f))) - \log(S(\log(f)))) d\log(f)} \quad (4)$$

Logaritmikus skálán ábrázolt $H(f)$ és $S(f)$ függvénygörbék mellett az exponens a függvénygörbék közötti különbséget adja meg. Ez minél nagyobb, azaz minél magasabban van a jel a zajszint felett, annál jobb a jel-zaj arányunk, ahogy azt a 1. ábra szemlélteti.

3. Az jel-zaj arány egy alternatív számítási módja

A valóságban nem áll rendelkezésünkre a zaj spektruma, csak a detektor kimenete, $v(t)$. Emellett elméleti modellekből ismerjük a várt jelalakot, $s(t)$ -t, de ez lehet az összegyűjtött adat egy konkrét szakasza is (egy olyan adatszakas, amiről az feltételezzük, hogy a detektálni kívánt gravitációshullám-jelet tartalmazza).

A zaj 0 várható értékű sztochasztikus adatsor, és a gyakorlatban Gauss-eloszlásúnak vehető. A normális eloszlás szórása ekkor az egyetlen szabad paraméter, ami jól jellemzi a zajt. A rendelkezésünkre álló függvényekből az alábbi módon tudjuk a zajt jellemezni:



1. ábra. A zaj amplitúdóspektrum-sűrűsége (angolul: Amplitude Spectral Density, vagy ASD) a hanfordi LIGO 4 km-es karhosszúságú detektorának S6 adatgyűjtési időszaka alatt és különböző jelek ASD-je . A zaj spektrumát látjuk feketével. Zölddel egymás felé spirálózva közelítő kettős jele (ez már $H(f)$), pirossal egy 50-50 naptömegű feketelyuk páré. A kék görbe ugyanennek a rendszernek a jele, ami a véges integrációs idő miatt ugrál (lásd [2]). A zaj ASD és görbék közötti területet számoljuk, amikor a jel-zaj arányt keressük. Az ábra a jel és a zaj viszonyát hivatott szemléltetni, további részletekért lásd az ábra forrását: [3]

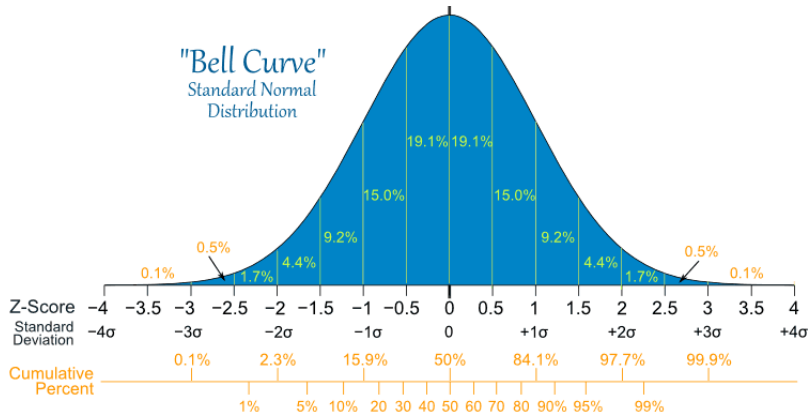
$$N \equiv \sqrt{\langle [v \star s]^2 \rangle} \quad (5)$$

Ahol a \star -gal jelölt művelet a keresztkorreláció művelete, ami a két függvény szorzatát különböző időeltolások mellett kiintegrálja:

$$(f \star g)(\tau) \equiv \int_{-\infty}^{\infty} f(t)g(\tau + t)dt \quad (6)$$

A várt jel és a detektor kimenetének az egyezésének mértékének mérésére a keresztkorrelációs függvényt használjuk. Bővebben lásd: 11. tétel, [1]

Ha a rendszerben nincsen jelen az $s(t)$ jel, akkor a $[v \star s]$ függvény értékei is zérus átlagú Gauss-eloszlást követnek, tehát a négyzetének a várható értéke (vagyis N) ennek az eloszlásnak a szórását adja.[6]



2. ábra. Gauss-eloszlás: Az ábrán látható, hogy a σ szóráshoz képest viszonyított eltérések észlelése mennyire valószínű. Az $\text{SNR}_0 = 3, 5, 8$ értékek a $[v \star s]$ eloszlás σ szórájának a három-, öt-, vagy nyolcszorosának felenek meg. Például a 3-as jel-zaj küszöbindex 3σ eltérést jelent, tehát egy ennél magasabb jel-zaj arányú eseményt rendkívül kicsiny valószínűséggel (0,1%) lehet a sztochasztikus jelből kapni.

Ahogy korábban írtuk, a várt jel és a detektor kimenet hasonlóságának mérésére a keresztkorrelációs függvényt használjuk. Tehát a t_0 időpontban érkező $s(t)$ alakú jel jellemzésére az alábbi mennyiséget definiáljuk:

$$S = | (v \star s)(t_0) | \quad (7)$$

Ennek a jelet jól jellemző S mennyiségnek, tehát a keresztkorrelációs függvény egy adott időeltolásnál felvett értékének, és a zajt jól becsülő N mennyiségnek hányadosával mérjük a jel-zaj arányt:

$$\text{SNR}(t_0) = \sqrt{\frac{S}{N}} \quad (8)$$

A korábban már példaként emlegetett $\text{SNR}_0 = 3, 5, 8$ értékek itt a $[v \star s]$ eloszlás σ szórájának a három-, öt-, vagy nyolcszorosának felenek meg. Így például a 3-as jel-zaj küszöbindex 3σ eltérést jelent, tehát egy ennél magasabb jel-zaj arányú eseményt rendkívül kicsiny valószínűséggel (0,1%) lehet a sztochasztikus jelből kapni (lásd 2. ábra). Ez tehát a hamis riasztási valószínűség (angolul: False Alarm Probability, FAP). Ebből kaphatjuk az 1. fejezetben már taglalt FAR-t, ha a FAP-ot leosztjuk a az adatgyűjtési időszak időtartamával. Megfordítva pedig egy tolerált FAR értékhez kiszámolhatjuk az SNR_0 értékét.

4. Periodikus jel jel-zaj aránya

Nézzünk egy konkrét példát: most tekintsünk egy véges T ideig megfigyelt periodikus jelet, ahol feltesszük, hogy a megfigyelés sokkal hosszabb, mint a periódusidő, azaz $T \gg f_0^{-1}$.

$$h(t) = \begin{cases} h_0 \cos(2\pi f_0 t) & 0 \leq t \leq T \\ 0 & \text{egyébként} \end{cases} \quad (9)$$

Itt h_0 a jel amplitúdója, f_0 a frekvenciája, t pedig a mérés kezdete óta eltelt idő.

A $h(t)$ függvényről feltételezzük, hogy ez az idősor, ahogy a gravitációshullám-jel a detektor adatsorában rögzítésre kerül. Ehhez adódik hozzá a detektor zajának idősora a kimeneten.

Mivel a számolásakor a zaj PSD-jének és a (9) egyenlet által megadott jelalak Fourier-transzformáltjának az arányát számoljuk, ami egy szűk frekvenciasávon kívül zérus, ezért ezen, a jel frekvenciáján felvett értékevel közelítjük a zaj PSD-jét. Így bevezetjük az alábbi jelölést: $\tilde{n} \equiv \sqrt{S(f_0)}$.

Így az (1) egyenletet átírva kapjuk

$$\text{SNR} = \frac{2}{\tilde{n}} \sqrt{\int_0^\infty |h(f)|^2 df} \quad (10)$$

Mivel a kétoldali PSD minden valós függvényre szimmetrikus, ezért erre áttérve (azaz negatív frekvenciákra is értelmezve a spektrumot) ezt írhatjuk:

$$\sqrt{\int_0^\infty |h(f)|^2 df} = \sqrt{\frac{1}{2} \int_{-\infty}^\infty |h(f)|^2 df} \quad (11)$$

A Parseval tétel[4] kimondja, hogy egy négyzetesen integrálható függvény esetén a függvény abszolútérték négyzetének idő- és frekvenciatérben vett integrálja megegyezik egymással. Általánosan írva:

$$\int_{-\infty}^\infty |X(f)|^2 df = \int_{-\infty}^\infty |x(t)|^2 dt, \quad (12)$$

ahol $x(t)$ Fourier-transzformáltja $X(f)$.

Így (11) egyenletben kihasználva (12)-t, és hogy $h(t)$ valós értékű:

$$\sqrt{\frac{1}{2} \int_{-\infty}^\infty |h(f)|^2 df} = \sqrt{\frac{1}{2} \int_{-\infty}^\infty h(t)^2 dt} \quad (13)$$

Ide már beírhatjuk a jel konkrét alakját, (9)-t:

$$\sqrt{\frac{1}{2} \int_{-\infty}^\infty h(t)^2 dt} = \sqrt{\frac{1}{2} \int_0^T h_0^2 \cos^2(2\pi f_0 t) dt} = \sqrt{\frac{T}{2}} h_{\text{rms}} \quad (14)$$

Ahol h_{rms} a jel négyzetes közepe, definíció szerint:

$$h_{\text{rms}} = \sqrt{\frac{1}{T} \int_0^T h(t)^2 dt} \quad (15)$$

Ez egy periodikus harmonikus jelre épp a h_0 amplitúdó $\sqrt{2}$ -ed része, azaz $h_{\text{rms}} = \frac{h_0}{\sqrt{2}}$, mivel a T integrálási időbe a függvény sok periódusa belefér, mivel $T \gg f_0^{-1}$, egy periódusra pedig:

$$\sqrt{f_0 \int_0^{\frac{1}{f_0}} \cos^2(t2\pi f_0) dt} = \frac{1}{\sqrt{2}} \quad (16)$$

Végül pedig (14)-ot behelyettesítve (10)-be:

$$\text{SNR} = \frac{h_0 \sqrt{T}}{\tilde{n}} \quad (17)$$

Tehát minél hosszabb T ideig tudok megfigyelni egy minél nagyobb h_0 amplitúdójú jelet, minél alacsonyabb \tilde{n} zaj-PSD mellett az adott jelfrekvencia körül, annál nagyobb jel-zaj arányt kapok. Mivel bármely jel felírható periodikus jelek összegeként, és a LIGO zaj PSD-je a néhány száz hertzes legérzékenyebb tartományban közel konstans, ezért a végeredményt egy gyors becslésre alkalmas ökölszabálynak tekinthetjük. Ezek szerint nem kiküszöbölhetetlen probléma, ha egy beérkező gravitációshullám-jel strain amplitúdója nem nagy, az teheti mégis a jelet detektálhatóvá, ha annak észlelése hosszabb T ideig történhet. Az ismert gravitációshullám-detektorok érzékenységgörbéje közvetlenül értelmezhető az adott detektorok által a választott szűk frekvenciatartományokon észlelhető jelamplitúdókként, ha azokhoz $\text{SNR}_0 = 1$ észlelhetőségi küszöbértéket és $T = 1$ s észlelési időhosszt feltételezünk. Mivel a keresett gravitációshullám-jelek forrásai rendszerint közel pillanatszerű összeomlások, összeolvadások, ezért az integrációs idő sem növelhető tetszőlegesen hosszúúra, így nincs más mód, mint a zajspektrumot a berendezés javításával csökkenteni.

Hivatkozások

- [1] <http://mathworld.wolfram.com/cross-correlation.html>.
- [2] Angler Gábor. A teljesítményspektrum és becslési módszerei. 2011.
- [3] Kari Alison Hodge. *The search for gravitational waves from the coalescence of black hole binary systems in data from the LIGO and VIRGO detectors*. PhD thesis, California Institute of Technology, 2014.
- [4] Marc-Antoine Parseval des Chenes. Memoire sur les series et sur l'integration complete d'une equation aux differences partielles lineaire du second ordre, a coefficients constants. 1799.
- [5] P. Raffai. *Az interferometrikus gravitációshullám-detektorok alkalmazása az asztrofizikában és a gravitációkutatásban*. PhD thesis, Eötvös Loránd Tudományegyetem, 2011.
- [6] Peter R. Saulson. *Fundamentals of Interferometric Gravitational Wave Detectors*. World Scientific Pub Co Inc, 1994.