Non-parametric uncertainties in the dark matter velocity distribution

A. Fowlie, (2018), arXiv:1809.02323 [hep-ph]

Andrew Fowlie

December 10, 2018

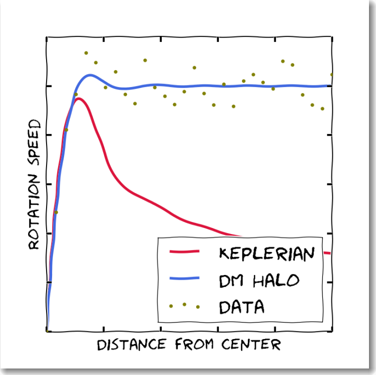
Nanjing Normal University

Nanjing

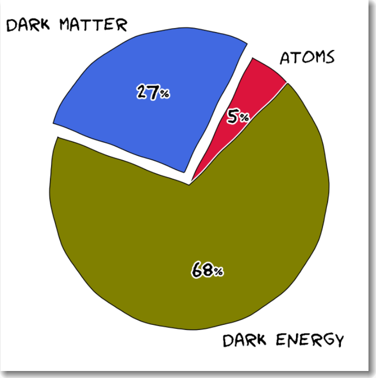
Table of contents

Dark matter

We all know the evidence for dark matter (DM) in gravitational interactions, e.g.



(I) Rotation curves



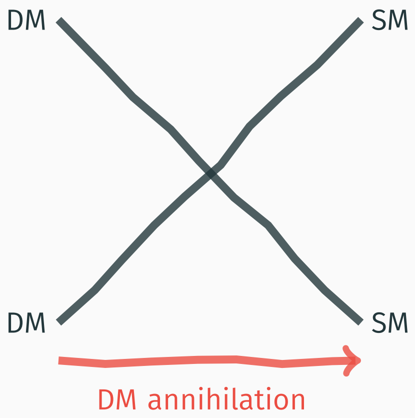
(II) CMB

WIMP miracle

Freeze-out of thermal equilibrium with bath of Standard Model (SM) particles sets relic density.

Freeze-out of thermal equilibrium with bath of Standard Model (SM) particles sets relic density.

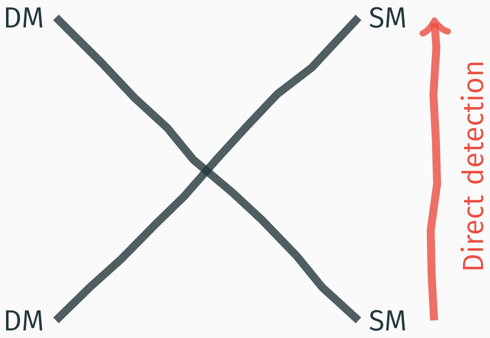
Correct prediction for weak interaction!



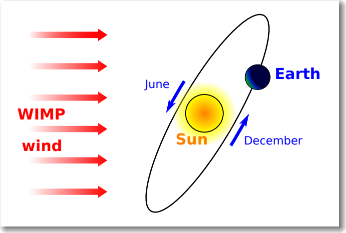
WIMP miracle

Freeze-out of thermal equilibrium with bath of Standard Model (SM) particles sets relic density.

Also predicts elastic scattering with SM!



We can search for DM in direct detection experiments. DM elastic scatters with nucleons in a detector on Earth.



In the laboratory frame there is a wind of DM because the Milky Way rotates inside a stationary DM halo.

Gran Sasso

There are major DD experiments in the Gran Sasso laboratory below the mountain

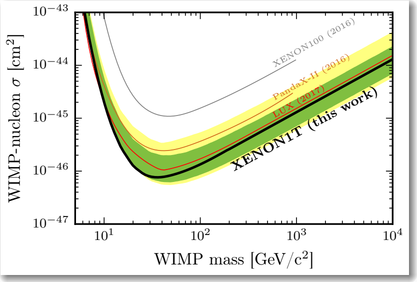
They are counting experiments — Poisson background from radioactive sources and neutrinos.

XENON1T

XENON1T uses one-tonne of liquid xenon — high atomic number, small radioactive contamination but expensive

With shielding and active background rejection techniques, backgrounds are under control — expect $\mathcal{O}(1)$ background event per tonne-year of exposure.

It set world-leading limit on spin-independent scattering cross-section between DM and nucleons



Two observed events, 1.62 expected background events. The number of expected signal events depends on the WIMP mass and scattering cross section, $s \propto \sigma$.

At high mass, the number density of DM $n = \rho / m$ and thus the flux falls and the limit weakens.

At low mass, the experimental efficiency falls as the energy recoil energy is tiny.

The flux of DM and amplitude of scattering in detector depend on DM velocity. The expected number of signal events is a moment of the velocity distribution,

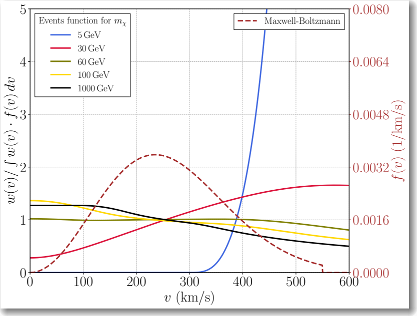
$$s = \int w(v) \cdot f(\vec{v})\, d^3v$$

The events function depends on the experiment, DM mass and scattering cross section. As detector not sensitive to direction, events function is isotropic in the laboratory frame.

The velocity distribution could be anisotropic in the galactic frame, but isotropic in simplest models.

Events function

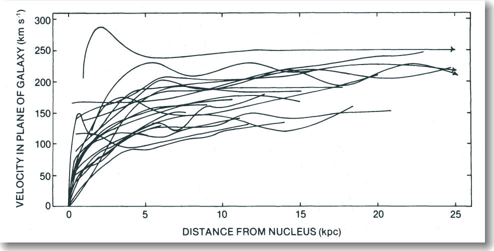
The shape of the events function depends on the DM mass.



Lighter masses need a greater velocity for detectable recoils.

Dark matter velocity profile

We don't know the identity of DM, but we know something about its density and velocity from e.g., rotation curves.



Velocity curves. Rubin et al [2].

Dark matter velocity profile

From $F_G = mv^2/r$, we find $\rho \propto 1/r^2$.

By the collisionless Boltzmann equation, this density corresponds to Maxwell-Boltzmann

$$f(v) \propto \begin{cases} v^2 e^{-\left(\frac{v}{v_0}\right)^2} & v < v_{\text{esc}} \\ 0 & v \geq v_{\text{esc}} \end{cases}$$

We truncate it at the escape velocity of our galaxy (though don't use $\rho \propto 1/r^2$ as v_{esc} and mass would be infinite).

This neglects non steady-state effects: clumps, streams and a possible dark disk.

Reasonable agreement with simulations.

Empirical

Is the DM velocity correlated with that of old, metal-poor stars?

- Determine $f(v)$ from correlation with metal-poor stars [3]. Significant departure from Maxwellian with anisotropic component
- Correlation is weak or non-existent [4]

I don't know the truth but techniques for incorporating that information could be challenging.

Modelling uncertainty in $f(v)$

Modelling uncertainty in profile

What if we want to reflect our uncertainty in the DM profile?

- Parametric approach: permit variation in v_0 and v_{esc} parameters in Maxwellian profile or shape parameters in another distribution.

What if we want uncertainty about distribution not just shape parameters?

- Non-parametric approach: permit all possible profiles.

How should we handle an infinite set of profiles?

Frequentist approach

1. Throw away all prior knowledge about DM in galaxy.
2. Profile infinite set of profiles by e.g., minimising chi-squared or maximising likelihood. Provable that the ansatz

$$f(v) = \sum_i \kappa_i \delta(v - v_i)$$

   is sufficient for minimising chi-squared and finding confidence intervals for signal rates [5, 6].
3. "Best-fit" profile not unique. One found from above procedure is an unphysical sum of delta-functions.

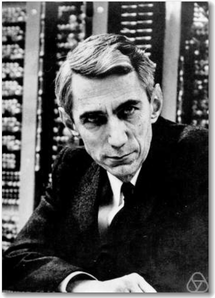
Bayesian approach

What if we could combine, in a coherent manner, experimental data and our background knowledge about the profile?

Maybe it isn’t exactly Maxwellian, but perhaps it’s something similar?

What can we do?

A possible solution?

Shannon's [7] information theory — Jaynes' [8] principle of maximum entropy — Skilling's [8] quantified maximum entropy.

Shannon entropy

Construct a measure of information learnt by receiving a message m_i that you expected with belief p_i. Requirements

- Anti-monotonic — more learnt from unexpected message
- $I \geq 0$ — information positive
- $I[p = 1] = 0$ — no new information if already certain about message
- $I[pq] = I[p] + I[q]$ — information additive for independent messages

imply that $I = -\ln p$.

Shannon entropy for discrete distributions is the expected information in a message [9]:

$$H = E[I] = -\sum_i p_i \ln p_i$$

Who learns the information?!

In the early days, this was a point of confusion:

- The sender: but he knows what he sent!
- The "pipe"/communication channel: this is strange, but was Shannon's thought and lead to ideas about channel capacity.
- The receiver: with this interpretation, the Shannon entropy measures ignorance of receiver/how much he expects to learn from message.

Relative Shannon entropy

Shannon's naive generalisation to continuous distributions

$$H = -\int p(x) \ln p(x) dx$$

violates Shannon's axioms and has other undesirable properties, e.g., it is not invariant under reparameterisations.

Correct expression found by limiting density of discrete points [10]

$$H = -\int p(y) \ln \frac{p(y)}{m(y)} dy$$

Meaning of $m(x)$

In the discrete case, the distribution with maximum Shannon information is uniform. This represents maximum ignorance.

In the continuous case, there is no such unique distribution because of covariance under changes of variable.

It is the age old question, which distribution represents ignorance? This is not solved; you must pick one, $m(x)$.

Which prior represents ignorance? Jaynes' used Shannon entropy to make his famous MaxEnt principle [11]

> *The prior that represents ignorance, subject to constraints, is the maximum entropy one.*

For example, if you know only $\langle x \rangle$, the MaxEnt distribution is the exponential. If you know $\langle x \rangle$ and $\langle x^2 \rangle$, the MaxEnt distribution is the Gaussian.

Skilling's quantified MaxEnt

Jaynes' MaxEnt principle worked for constraints on moments, but not not noisy data, and failed to provide a measure of reliablility.

Skilling's quantified MaxEnt

Jaynes' MaxEnt principle worked for constraints on moments, but not not noisy data, and failed to provide a measure of reliablility.

We in fact want $p(f)$ — a distribution upon possible choices of f.

Skilling's quantified MaxEnt

Jaynes' MaxEnt principle worked for constraints on moments, but not not noisy data, and failed to provide a measure of reliablility.

Skilling [12] demonstrated that if a general rule of assigning $p(f)$ exists, it must depend on the Shannon entropy by

$$p(f) \propto e^{\beta H[f,m]}$$

where β represents the strength of our prior conviction that $f = m$. He found this by assuming it must agree with a limiting case of a multinomial process.

Undesirable properties

Skilling's prior, however, suffers from problems as it in fact depends on the bin size, Δv. They relate to the fact that it is not divisible.

In the continuum limit $\Delta v \to 0$, the law of large numbers means that the prior overwhelmingly favors $f = m$ regardless of β.

MaxEnt inspired approach

We consider a multinomial process — throw a finite number, β, of lumps of probability to construct a distribution. The lumps falls into bins with probabilities from the default model, m.

This avoids the law of large numbers as we throw a finite number of finite lumps. We may still take the continuum limit $\Delta v \to 0$. It means, however, that $f(v)$ must, on small enough scales, appear spiky.

For $\beta \to \infty$, $f \to m$, i.e., the default model is selected.

Marginalization

The likelihood in a DD counting experiment is Poisson,

$$\mathcal{L} \equiv P(o \,|\, m_\chi, \sigma, f) = \frac{e^{-\lambda}\lambda^o}{o!}$$

with

$$\lambda = \int w(\vec{v}) \cdot f(\vec{v})\, d^3v + b$$

The events functions depends upon the DM mass, cross section and velocity distribution.

Marginalization

We want to marginalise uncertainty in $f(v)$,

$$\begin{aligned} \langle \mathcal{L} \rangle &= \sum P(o \mid m_\chi, \sigma, f) \cdot P(f \mid m) \\ &= \sum \frac{e^{-\lambda} \lambda^o}{o!} \cdot P(f \mid m) \end{aligned}$$

We are averaging upon our multinomial prior for the velocity distirbution, given our default one, a Maxwellian.

Marginalization

We may perform the sum! A general expression is in the paper. For two observed events,

$$\langle \mathcal{L} \rangle = \frac{1}{2} \left\langle e^{-w/\beta} \right\rangle_m^\beta \left(\frac{\beta - 1}{\beta} \frac{\left\langle w e^{-w/\beta} \right\rangle_m^2}{\left\langle e^{-w/\beta} \right\rangle_m^2} + \frac{1}{\beta} \frac{\left\langle w^2 e^{-w/\beta} \right\rangle_m}{\left\langle e^{-w/\beta} \right\rangle_m} \right)$$

where $\langle w \rangle_m \equiv \int w(\vec{v}) \cdot m(\vec{v})\, d^3 v$ etc.

Thus we can marginalize over uncertainty in the velocity profile by computing a few integrals.

Limits from XENON1T with uncertainty in the velocity profile

Treatment of XENON1T

We now have the ingredients that we need:

- Result of Poisson counting experiment from XENON1T. The experiment saw 2 events with 1.66 expected background events. We know the events function.

Treatment of XENON1T

We now have the ingredients that we need:

- Result of Poisson counting experiment from XENON1T. The experiment saw 2 events with 1.66 expected background events. We know the events function.
- Background knowledge about the velocity profile. The prior depends upon a parameter β describing our conviction that the profile is Maxwellian.

Treatment of XENON1T

We now have the ingredients that we need:

- Result of Poisson counting experiment from XENON1T. The experiment saw 2 events with 1.66 expected background events. We know the events function.
- Background knowledge about the velocity profile. The prior depends upon a parameter β describing our conviction that the profile is Maxwellian.
- A formalism for marginalising our uncertainty in the velocity profile. We can perform the marginalisation analytically, reducing it to a few integrals.

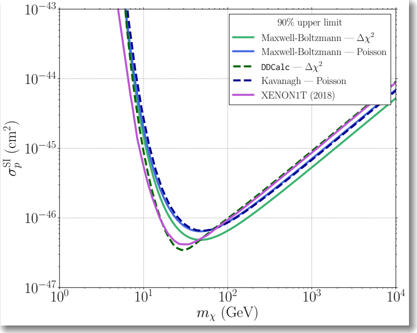
Validation — reproduce limit with Maxwellian

From a simple Poisson likelihood,

$$\mathcal{L} = \frac{e^{-\lambda}\lambda^{o}}{o!}$$

with $\lambda = s + b$, $b = 1.66$ and $o = 2$ we find our upper limit on the scattering cross section, σ, assuming a Maxwellian. The number of signal events, $s \propto \sigma$

Validation — reproduce limit with Maxwellian

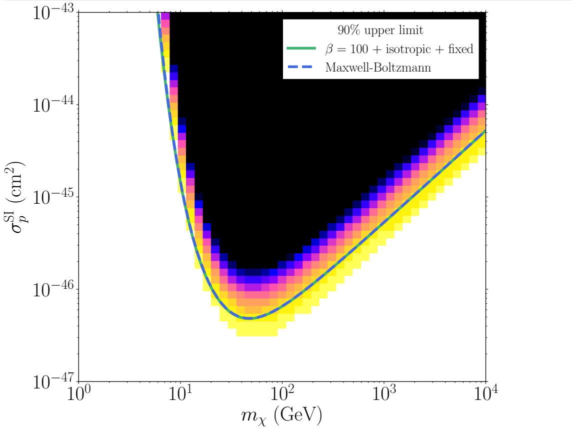


Reasonable agreement. The XENON1T official analysis used an unbinned analysis and achieved greater exclusion. This validates our description of the experiment.

Isotropic profiles

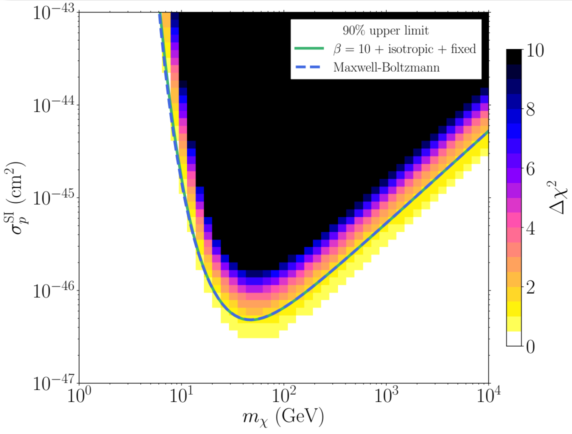
Assume that the $f(v)$ is isotropic in the galactic frame. Let's see how the observed limit changes as we marginalise uncertainty in the velocity distribution.

Isotropic profiles — varying β



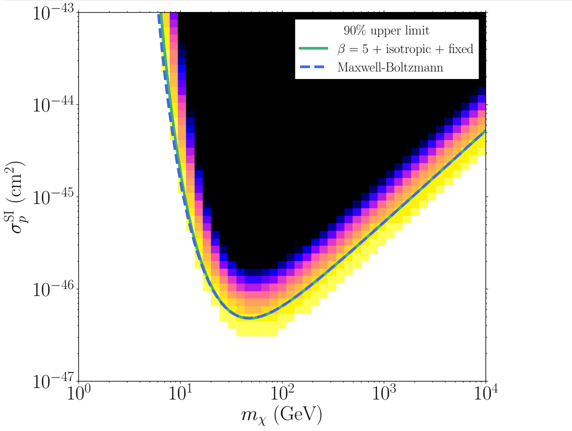
$\beta = 100$. The limit is indistinguishable from that from a Maxwellian.

Isotropic profiles — varying β

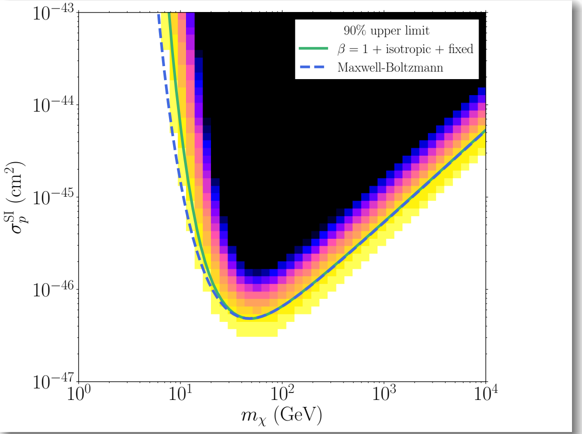


$\beta = 10$. Still no change!

Isotropic profiles — varying β



$\beta = 5$. Tiny weakening of the limit for light masses.

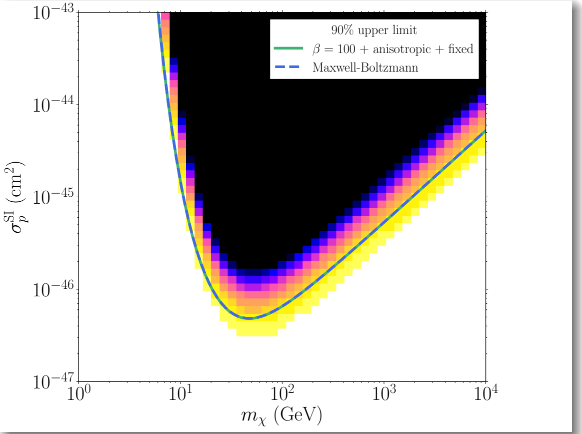


$\beta = 1$. Weakening visible for light masses, $m \lesssim 60$ GeV

Anisotropic profiles

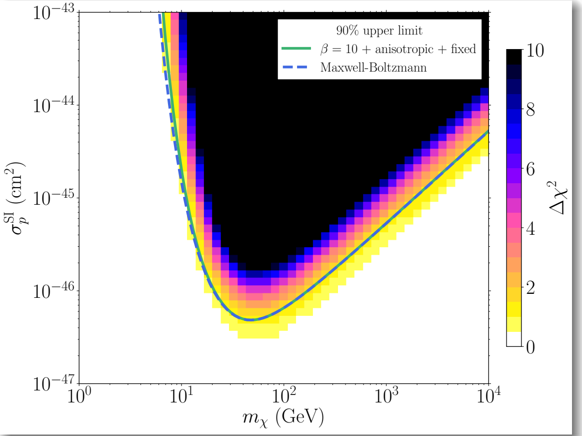
Now permit anisotropy. This is important — it could be tuned to counter the Earth's motion and result in a vanishing flux and signal.

Anisotropic profiles — varying β



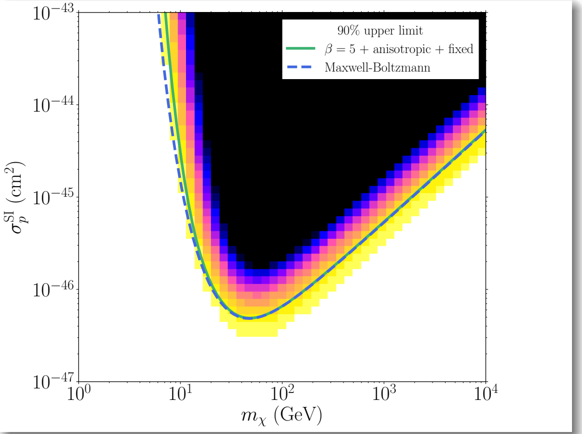
$\beta = 100$. The limit is indistinguishable from that from a Maxwellian.

Anisotropic profiles — varying β



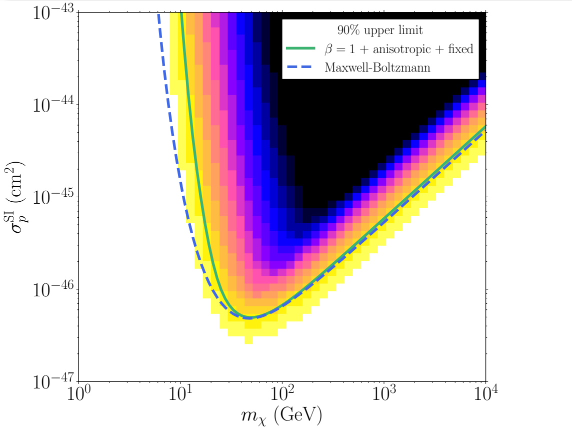
$\beta = 10$. Unlike the isotropic case, a change visible.

Anisotropic profiles — varying β



$\beta = 5$. Weakening of the limit for light masses.

Anisotropic profiles — varying β

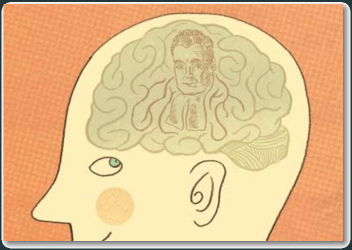


$\beta = 1$. Weakening visible for light and heavy masses, though limit at $m \approx 60\,\mathrm{GeV}$ quite robust.

Conclusions

- Technique for marginalising uncertainty in the DM velocity profile
- Strength of conviction that profile Maxwellian parameterised by β
- We find that once the uncertainty in marginalised, the XENON1T limit is quite robust with respect to the velocity profile
- Substantial impact noticeable in anisotropic case
- The multinomial prior was not perfect. Discrete lumps of probability. Can we do better?
- How can we incorporate information from metal-poor stars?

Questions?



References

[1] A. Fowlie, “Non-parametric uncertainties in the dark matter velocity distribution,” (2018), arXiv:1809.02323 [hep-ph].

[2] V. C. Rubin, W. K. Ford Jr., and N. Thonnard, “Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/,” Ap.J. 238, 471–487 (1980).

[3] L. Necib, M. Lisanti, and V. Belokurov, “Dark Matter in Disequilibrium: The Local Velocity Distribution from SDSS-Gaia,” (2018), arXiv:1807.02519 [astro-ph.GA].

References ii

[4] N. Bozorgnia, A. Fattahi, D. G. Cerdeno, C. S. Frenk, F. A. Gómez, R. J. J. Grand, F. Marinacci, and R. Pakmor, "On the correlation between the local dark matter and stellar velocities," (2018), arXiv:1811.11763 [astro-ph.GA].

[5] A. Ibarra and A. Rappelt, "Optimized velocity distributions for direct dark matter detection," (2017), arXiv:1703.09168 [hep-ph].

[6] G. B. Gelmini, J.-H. Huh, and S. J. Witte, "Unified Halo-Independent Formalism Derived From Convex Hulls," (2017), arXiv:1707.07019 [hep-ph].

[7] https://commons.wikimedia.org/wiki/File:ClaudeShannon_MFO3807.jpg.

[8] http://bayes.wustl.edu/.

[9] C. E. Shannon, "A Mathematical Theory of Communication," Bell System Technical Journal 27, 379–423 (1948).

[10] E. T. Jaynes, "Information Theory and Statistical Mechanics (Notes by the lecturer)," in Statistical physics 3 (1963), p. 181.

[11] E. T. Jaynes, "Information Theory and Statistical Mechanics," Phys. Rev. 106, 620–630 (1957).

[12] J. Skilling, "The Axioms of Maximum Entropy," in *Maximum-Entropy and Bayesian Methods in Science and Engineering: Foundations*, edited by G. J. Erickson and C. R. Smith (Springer Netherlands, Dordrecht, 1988), pp. 173–187.