YFY201-2025-PR1: Information Entropy in Atoms

Georgia Pavlina Tsoukaneri

July 2025

1 Introduction

This study in computational quantum mechanics focuses on the information content of elements with Z = 2 (He) to Z = 10 (Ne) and its dependence on size and scale. It replicates the results of the paper "Information entropy, information distances, and complexity in atoms" by Chatzisavvas et al[1]. The project starts from calculating position-space and momentum-space density using STO orbitals and uses these densities to compute:

- Shannon entropy in position and momentum spaces
- Onicescu information and how it compares to ionization potential
- Kullback-Leibler and Jensen-Shannon divergences as information distances between
  - the RHF density vs an asymptotic density depending on the ionization potential
  - the RHF electron configuration of the atom vs the Thomas Fermi model for multi-electron atoms

2 Theoretical Framework

Our approach focuses on the information content of different elements, measured through different measures and the similarity of different distributions based on the concept of information distance.

2.1 Representation of the wavefunction

In computational physics , the Hartree–Fock (HF) method is a method of approximation for the determination of the wave function and the energy of a quantum many-body system in a stationary state. The RHF orbitals need to be expanded in a basis. The basis of choice for our analysis is the Slater Type Orbitals.[2]
RHF orbitals will also be compared to the semiclassical Thomas-Fermi model for multi-electron atoms.

2.2 Information Entropy in Quantum Systems

The Shannon information entropy S_r in position space may be defined as

$$S_r = -\int \rho(r) \ln \rho(r) dr,$$

where $\rho(r)$ is the electron density distribution normalized to unity. The corresponding information entropy S_k in the momentum-space representation is

$$S_k = -\int n(k) \ln n(k) dk,$$

A high information entropy S_r in position space points to highly delocalized electrons with a flattened distribution. Similarly, a maximum in S_k points to great delocalization in momentum space.[3] Another measure of the information content of a quantum system is the concept of information energy, introduced by Onicescu. Defined as

$$E = \int \rho^2(x) dx.$$

for a continuous density distribution. E becomes minimum for equal probabilities (total disorder). This means we can define a measure of information content as

$$O = \frac{1}{E_r E_k}$$

The last measure of information that will be used in this analysis is the order parameter Ω, defined as

$$\Omega = 1 - \frac{S}{S_{max}}$$

where S is the information entropy of the system and Smax is the maximum entropy accessible to the system.

2.3 Kullback-Leibler and Shannon-Jensen information distances

In mathematical statistics, the Kullback-Leibler (KL) divergence, is a type of statistical distance. It shows how much a model probability distribution Q is different from a true probability distribution P. Mathematically, it is defined as

$$KL(P \parallel Q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} dx.$$

KL is not symmetrical, in the sense that $KL(P \parallel Q) \neq KL(Q \parallel P)$, so we use a "symmetrized" version of it, SK, defined as

$$SK = \int \rho_A(r) \ln \frac{\rho_A(r)}{\rho_B(r)} dr + \int \rho_B(r) \ln \frac{\rho_B(r)}{\rho_A(r)} dr.[4]$$

The Jensen-Shannon divergence entropy is defined as follows:

$$J = -\int \left(\frac{\rho_A(\mathbf{r}) + \rho_B(\mathbf{r})}{2} \right) \ln \left(\frac{\rho_A(\mathbf{r}) + \rho_B(\mathbf{r})}{2} \right) d\mathbf{r} + \frac{1}{2} \int \rho_A(\mathbf{r}) \ln \rho_A(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \rho_B(\mathbf{r}) \ln \rho_B(\mathbf{r}) d\mathbf{r}.$$

and has a similar physical meaning to the kullback distance.

3 Methodology

3.1 Implementation Overview

In order to carry out our analysis we followed the following workflow:

1. Extraction of the RHF coefficients for the STO basis from the relative table in the paper by Carlos F. Bunge and Jose A. Barrientos.[5]

2. Calculation of position-space density $\rho(r)$ and momentum-space density $n(k)$.

3. Calculation of the information content based on the Shannon and Onicescu metrics.

4. Evaluation of asymptotic density distributions using the ionization potential and comparison to the ground state using Kullback and Jensen distances.

5. Evaluation of density distribution based on the Thomas-Fermi model and comparison to RHF using KL and J.

3.2 Computational Details

Image to text recognition: Extraction of the RHF table

In order to digitize the RHF coefficients used for the expansion in STO orbitals, we provided the AI model Deepseek with the table from the paper as an image. We then instructed it to generate Python code to create a csv file containing the values included in the table. We specified the desired format of the table to be so that the exponents and coefficients for each orbital basis (1s, 2s, 2p) are contained in a different column of the csv file, as showcased in the mock table below.

Table 1: Gaussian-type orbital basis set coefficients with s- and p-orbital separation.

Element	**Z**	**Orbital**	**Exponent**	$\mathbf{C_{1s}}$	$\mathbf{C_{2s}}$	$\mathbf{C_{2p}}$
He	2	1S	1.4595	1.3479	0.0	0.0
He	2	3S	5.3244	-0.001613	0.0	0.0
Li	3	1S	4.3069	0.141279	-0.022416	0.0
Li	3	2S	0.7667	-0.000305	0.340542	0.0
⋮	⋮	⋮	⋮	⋮	⋮	
B	5	2P	5.7416	0.0	0.0	0.0076
B	5	2P	0.8494	0.0	0.0	0.432795
C	6	2P	1.0242	0.0	0.0	0.350701
O	8	2P	1.2473	0.0	0.0	0.231788
Ne	10	2P	1.4948	0.0	0.0	0.221831

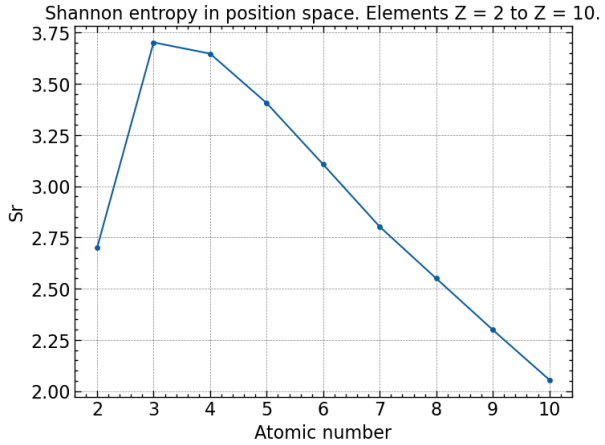
The csv file created this way was overally correct and needed small manual adjustments and corrections.

Calculation of electron densities and information metrics

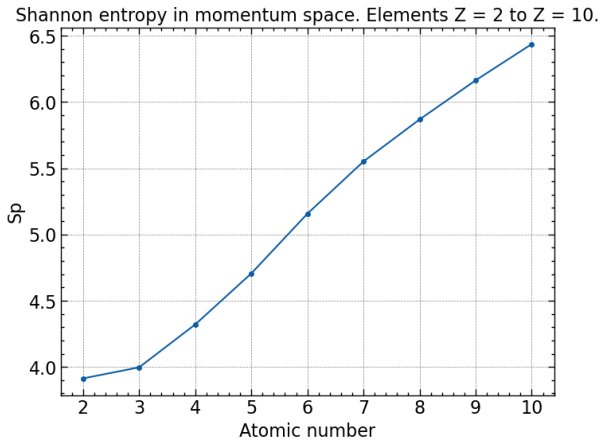
The electron denisites for each orbital, the Shannon and Onicescu entropies, the order coefficient Ω, the electron densities for the ionized elements and the Kullback and Jensen distance metrics were calculated using a reference file for all formulas for Neon, written in Mathematica. The file was personally provided to us by K. Chatzisavvas, one of the writers of the original paper. Using the reference formulas, we generalized the calculations for all elements with Z = 2 to Z = 10 in Python. We also used the documented ionization potentials for those elements, as found in Wikipedia[6]. All calculations performed can be found in the accompanying Jupyter notebook. The implemenation used Python with scientific libraries: numpy for numerical calculation, scipy for numerical integration, pandas for data handling and matplotlib for visualization.

4 Results and Analysis

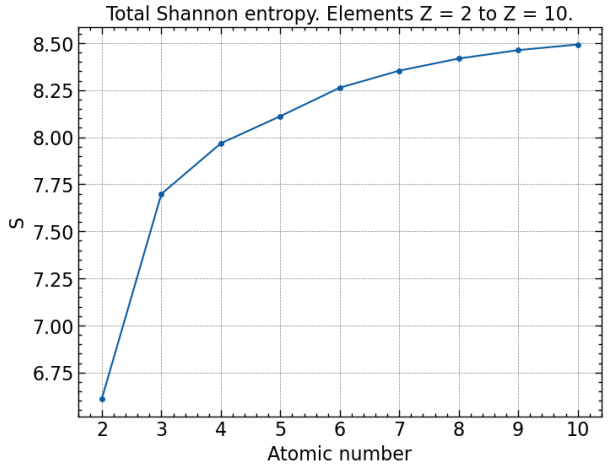
4.1 Shannon Entropy and order parameter



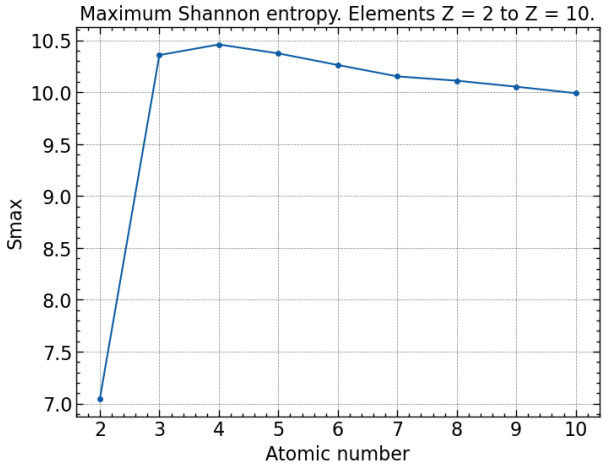
(a) S_r (Position space entropy)



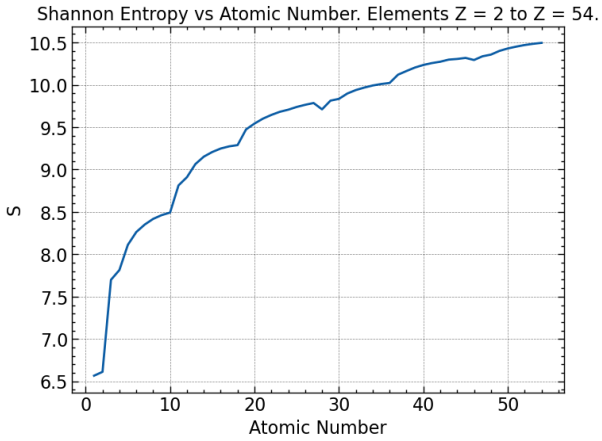
(b) S_k (Momentum space entropy)



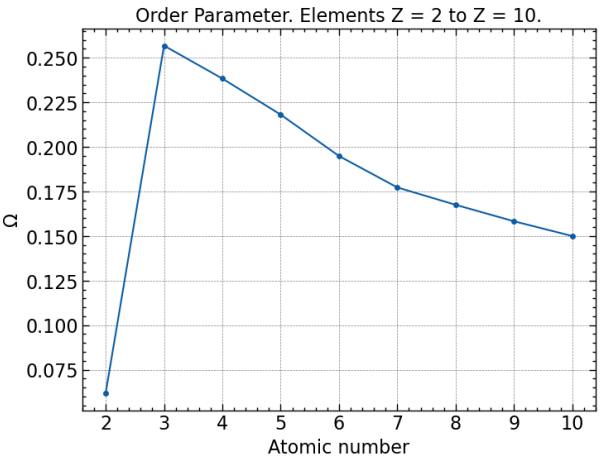
(c) S (Total entropy)



(d) S_{max} (Maximum possible entropy)



(e) S for Z = 2 to Z - 54



(f) Ω (Order parameter)

Figure 1: 1-5: Information entropy measures in different spaces: (a) position space, (b) momentum space, (c) total entropy, and (d) maximum possible entropy (e) total entropy for Z = 2 to Z = 54. 6: (f) order parameter

Shannon entropy and order parameter for Z = 2 to Z = 10

An overall decrease is noted in the Shannon entropy in position space while a monotonic increase is noted in the Shannon entropy in momentum space. The total Shannon entropy also increases, although less than linearly. The maximum entropy available to the system increases for the first few elements and then gradually falls.
The order parameter sharply increases from He to Li and then gradually falls hinting at closed shell effects taking place.
The scope of the analysis we can carry out is rather limited, since we only treated the elements with Z = 2 to Z = 10, so no comments on the large scale behavior of the system or the potential existence of closed shell effects can be made for now.
In order to infer more about the large scale dependence of entropy on the atomic number, we will also analyze S for elements Z = 1 to Z = 54, without doing the calculations ourselves, based on the data found in the paper by Chatzisavvas et al.

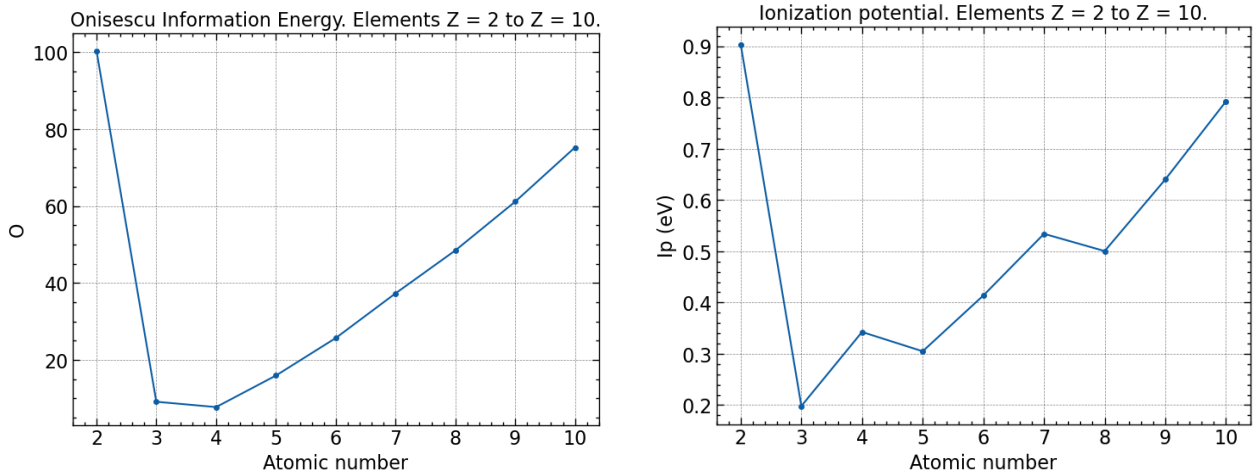
Shannon entropy for Z = 1 to Z = 54

Plotting the Shannon entropy over atomic number for elements with Z = 1 to Z = 54, we take notice of the fact that the entropy has a logarithmic dependence on the atomic number. Fitting a logarithmic curve to the our data produces a relationship of the format $S = a \cdot Z + b$ with parameters

Parameter	Value
a	1.05 ± 0.015
b	6.33 ± 0.047

Table 2: Fitted parameters for exponential relationship

4.2 Onicescu Energy and comparison to Ionization Potential



(a) Onicescu's energy $\mathcal{O}$

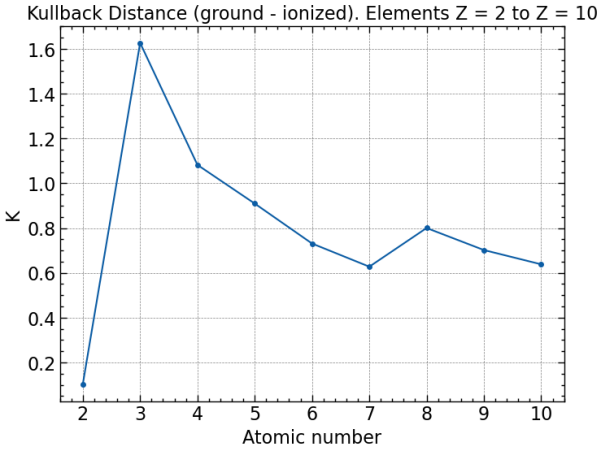
(b) Ionization potential I_p

Figure 2: Comparative analysis of (a) Onicescu's information energy and (b) ionization potential for the studied system.

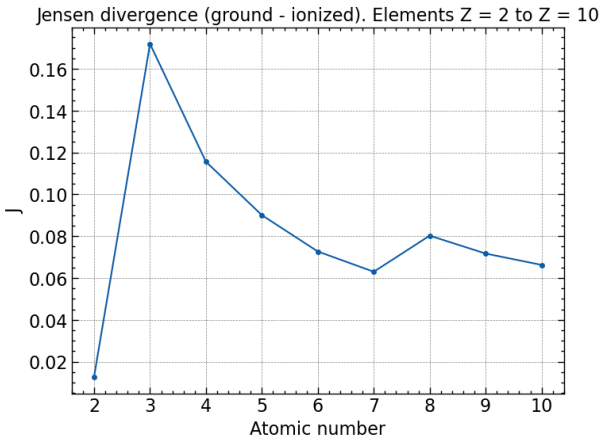
The Onicescu energy curve and the ionization potential curve share a similar shape. The reason for this might not be immidiately obvious, as they are completely diferrent quantities. A conceptual

understanding can be achieved taking into account that systems with a high Onicescu energy are more localized and therefore need more energy to become ionized.

4.3 Distribution distances: Ionized vs ground state



(a) Kullback-Leibler divergence
$Kl(\rho_0 \parallel \rho_1)$

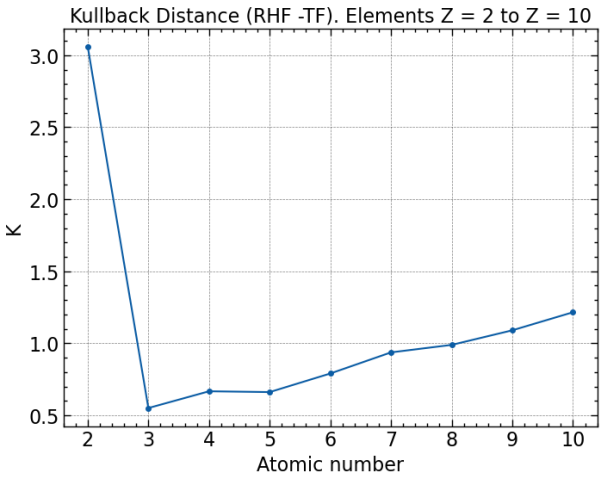


(b) Jensen-Shannon divergence
$J(\rho_0, \rho_1)$

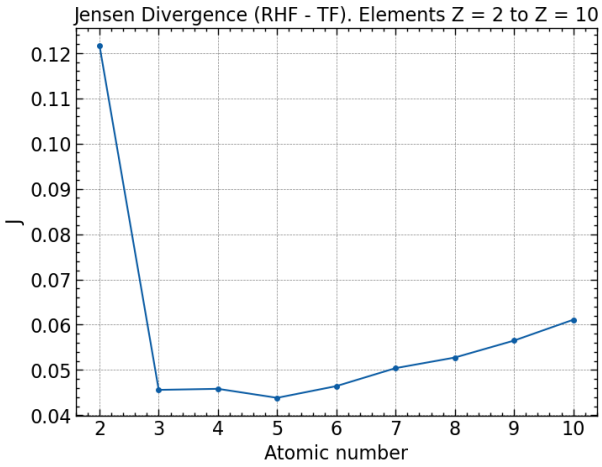
Figure 3: Information distances between ground (ρ_0) and ionized (ρ_1) states: (a) asymmetric Kullback-Leibler measure and (b) symmetric Jensen-Shannon measure.

As exhibited by both metrics, the distance between the asymptotic and ground densities rises sharply from He to Li and then gradually drops again. This possibly hints to closed shell effects, but the scope of the analysis would have to be extended beyond Z = 10 to confirm that.

4.4 Distribution distances: RHF vs Thomas-Fermi



(a) Kullback-Leibler divergence
$KL(P_{\mathrm{RHF}} \parallel P_{\mathrm{TF}})$



(b) Jensen-Shannon divergence
$J(P_{\mathrm{RHF}}, P_{\mathrm{TF}})$

Figure 4: Comparison between RHF orbitals and Thomas-Fermi model

As exhibited by both metrics, the distance between the RHF orbital density distribution and the Thomas-Fermi model is very big for He and then sharply drops. It then rises gradually again. Given

the fact that Thomas-Fermi excels in describing larger atoms with more electrons, we expect it to become more accurate as Z grows large and therefore closer to the RHF interpretation. Therefore, the gradual increase in distance from Li and on is probably attributed to closed shell effects. This cannot be verified without extending the scope beyond Z = 10.

5 Discussion

5.1 Entropy Trends and Shell Effects

Our calculations reproduce the characteristic entropy behavior reported by Chatzisavvas et al, showing:

- An overall decrease in position-space entropy (S_r) with increasing Z
- A monotonic increase in momentum-space entropy (S_k)
- Non-linear growth of total entropy ($S = S_r + S_k$)

The opposing trends in S_r and S_k reflect the position-momentum uncertainty principle, while the total entropy's sublinear growth suggests increasing electron correlation effects. The maximum available entropy peaks at intermediate Z values (Li-Be), hint at closed shell effects as analyzed in the original paper.
Expanding the analysis beyond Z = 10 confirms a logarithmic increase in information entropy as an element grows larger. An atom cannot get infinitely complex just by adding more matter to it.

5.2 Divergence between different states

The divergence metrics between rhf and asymptotic (Fig. 3) reveal:

- Sharp Li-He contrast in $KL(\rho_0 \| \rho_1)$ and $J(\rho_0, \rho_1)$
- Subsequent gradual decrease for heavier elements

This pattern potentially mirrors the paper's finding that "that closed-shell atoms are less complex than neighboring ones."
The Thomas-Fermi comparisons (Fig. 4) further show:

- Exceptionally poor TF performance for He ($Z = 2$)
- Potential non-monotonic accuracy improvements with Z

confirming that semiclassical models require $Z \gg 10$ to match quantum calculations.

References and Resources

[1] K. Ch. Chatzisavvas, Ch. C. Moustakidis, and C. P. Panos. "Information entropy, information distances, and complexity in atoms". In: *J. Chem. Phys.* 123 (2005). Primary reference for methodology and results, p. 174111.

[2] Wikipedia. *Roothaan Equations*. Accessed July 2025; used for RHF overview. 2025. URL: `https://en.wikipedia.org/wiki/Roothaan_equations`.

[3] Wikipedia. *Shannon entropy*. Accessed July 2025; used for introductory concepts. 2025. URL: `https://en.wikipedia.org/wiki/Entropy_(information_theory)`.

[4] Wikipedia. *Kullback-Leibler divergence*. Accessed July 2025; used for definitions. 2025. URL: `https://en.wikipedia.org/wiki/Kullback-Leibler_divergence`.

[5] Carlos F. Bunge, José A. Barrientos, and Annik Vivier Bunge. "Roothaan-Hartree-Fock ground-state atomic wave functions". In: *At. Data Nucl. Data Tables* 53 (1993). Source for STO coefficients, pp. 113–162.

[6] Wikipedia. *Ionization energies of the elements (data page)*. Accessed July 2025; source for ionization potentials. 2025. URL: `https://en.wikipedia.org/wiki/Ionization_energies_of_the_elements_(data_page)`.

To carry out our analysis we used formal academic resources for extracting data as well as non-formal sources like Wikipedia to gain conceptual understanding and reference definitions in our report. All computational effort was carried out using local resources to run Python code in Jupyter Notebooks. The LLM model Deepseek was used for image to text recognition. All code and data is uploaded to Github.