Markov Processes and Isomeric Decay

David Gohlke
Advised by G. Jay Kerns

Spring 2006

Abstract

This honors thesis consists of work done involving Markov chains. A Markov chain is a process such that the next state depends only on the current state, not on prior states. This chain can then be represented as a transition matrix. Markov chains are generally not examined at the undergraduate level. This project uses ideas learned in linear algebra, combinatorics, and probability and statistics.

Theoretical aspects of Markov chains are included and used to find solutions to real-life problems, most notably, ongoing research done with the physics department involving nuclear isomers. Transitions from the isomeric decay are modeled with an associated Markov process and this information is used to discern information about coincident decays and the number of gamma rays emitted.

Contents

1 Overview and History of Markov Chains

Some dynamical systems can be incredibly hard to calculate, and dependence on prior events may make solving them nearly impossible. Hysteresis effects can make an ordinarily well-defined function turn into a mess of previous conditions, a butterfly flapping its wings in Brazil is known to make it rain in China, and human beings, possibly the most egregious offenders, will base almost any future decision on the past.

Thankfully, some systems are completely independent of the past (most games of chance), and some only care about the current state (many board games). Additionally, many systems may be very well approximated by this (such as nuclear decay). Situations in which the probability of entering one state is entirely dependent on the current conditions (and no earlier ones) are known to satisfy the Markov property.

In graph theory, one way of representing a graph is through adjacency matrices. These matrices express the number of directed edges between vertices; in the adjacency matrix, A, element $A_{(i,j)}$ is the number of edges from vertex i to vertex j. In this case, the number of edges will be an integer, so every matrix element will also be an integer. A Markov matrix uses a probabilistic version of this concept, normalizing every row to 1. Since every row determines a transition probability, this matrix is then known as a transition matrix.

1.1 History of Markov chains

Markov chains were first studied in the early 1900s by Andrei Andreevich Markov, who wished to extend the application of the law of large numbers to dependent variables. Simple examples of stochastic processes had been examined before, but Markov was the first to look at more general cases, prompted especially by the claim of P.A. Nekrasov that "independence is a necessary condition for the law of large numbers [1]."

The accomplished probability theorist Andrei Nikolaevich Kolmogorov rigorously researched the theory of stochastic processes, basing his work off of Markov and publishing his results in *Analytic methods in probability theory* in 1938. Continuous Markov chains were first investigated in depth by Norbert Wiener while he was studying Brownian motion. There is still active research

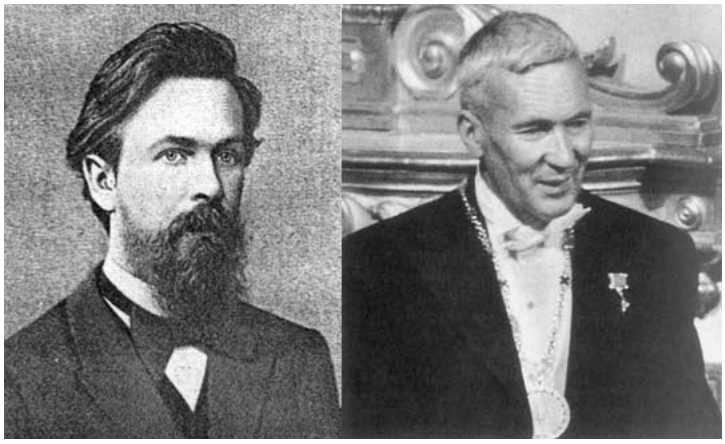

Figure 1: Markov and Kolmogorov laid most of the foundation for stochastic processes

in continuous Markov models, as they have a very strong relationship to statistical mechanics. Roughly twenty years after work on the subject began, Sergei Bernstein posthumously honored Andrei Markov by naming the field of Markov chains after its founder.

One area of theoretical Markov chain research deals with higher order Markov chains, that is, chains with transition probabilities which depend on more than one preceding state. Markov himself looked into this, but focused most of his work on first order Markov processes.

One of the most active areas of research related to Markov processes involves hidden Markov models. We know that a system is a Markov process if the transitions are independent of previous states, and calculations exist to determine if the probabilities are dependent. It is possible that a system may have transitions that appear to be dependent on prior states. However, we may not be able to discern all of the states, and the system may be a Markov process after all. In other words, this is a hidden Markov model. Hidden Markov models occur in biology and have been found to be integral to speech recognition by computers.

1.2 Notation

To ease the discussion of properties of Markov chains for the remainder of this paper, notation will be introduced here. For a transition matrix, M, element $M_{(i,j)}$ represents the probability of going directly from state i to state j. For

a finite number, n, of possible states, this transition matrix is written as

$$\begin{bmatrix} P_{(1,1)} & P_{(1,2)} & \cdots & P_{(1,n)} \\ P_{(2,1)} & P_{(2,2)} & \cdots & P_{(2,n)} \\ \vdots & \vdots & \ddots & \vdots \\ P_{(n,1)} & P_{(n,2)} & \cdots & P_{(n,n)} \end{bmatrix}$$

An infinite state space could be handled in a similar manner.

Again, the rows stand for the initial state and each column represents a different final state. In some books, such as the one written by Lay [10], the columns and rows are inverted.

Let the probability of an event A be notated by $P(A)$. Also, the conditional probability of event B given the occurrence of event A will be notated by $P(B|_A)$. This is the probability of both A and B occurring, given that A occurs, so a value can be found by:

$$P(B|_A) = \frac{P(B \bigcap A)}{P(A)} \tag{1}$$

It is worth noting that an event in this notation could actually be comprised of multiple individual events.

The expected value for the number of visits to a state A is denoted as $E(A)$. Mostly, the concern is about conditional probability, so the expected value for the number of visits to a state A after starting in state B is given by $E(A|_B)$.

1.3 Example of a Markov chain

Common examples when discussing Markov chains include forecasting the weather in a city where the weather tomorrow only depends probabilistically on the today's weather and determining the amount of money remaining in a gambler's coffers. As this is an undergraduate thesis, a more fitting example of a Markov chain may involve the dining habits of a college student.

Assume a college student has three choices for dinner every night: Ramen noodles, macaroni and cheese, and pizza. After eating Ramen noodles one night, the student is indifferent about what to eat the next night - Ramen noodles are still tasty, macaroni and cheese would add variety, and after a cheap dinner the night before, pizza is relatively affordable. Macaroni and cheese is easy to make, so the student will often eat it two nights in a row,

but otherwise has no preference between Ramen noodles and pizza. Pizza is expensive, so the student cannot afford to eat pizza two nights in a row, though either Ramen noodles or macaroni and cheese would be acceptable.

Figure 2: A rare delicacy

By assigning numbers to these eating habits, we can create a Markov matrix for calculations. Here, the rows correspond to yesterday's dinner and the columns are the different possibilities for dinner tonight. The states are, respectively, Ramen noodles, macaroni and cheese, and pizza. Denoting this dinner matrix as D, we have

$$D = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/2 & 1/2 & 0 \end{bmatrix} \tag{2}$$

This example will be the recurrent example of the paper, being called upon to give a tangible basis to the theoretical details given.

2 Mathematics of Markov Chains

The Markov property states that the only aspect of the history of the system that matters is the immediately previous state, and no earlier states affect the probability distribution of the next state.

The fact that the only dependent factor for the next state is the current state implies that the Markov process can be represented as a matrix. If it relied on a previous state, the process would need to be represented as a higher order tensor. (Furthermore, a state which has a constant probability distribution

regardless of current state can be represented as a vector, a lower order tensor.) Letting M be the transition matrix, this capability for matrix representation can be formally expressed as

$$P(X_{n+1} = j|_{X_n=i, X_{n-1}=a_{n-1}, \ldots, X_0=a_0}) = M_{(i,j)} \quad \forall i, j, a_k \tag{3}$$

Because each row represents a probability distribution, we know that $\sum_{j=1}^{\dim \mathrm{M}} M_{(i,j)} = 1$ for each row. For a transition matrix M, the probability of going from state i to state k in one step is $M_{(i,k)}$. Similarly, the probability of going from state k to state j in one step is $M_{(k,j)}$. Therefore, the probability of going from i to j in two steps is almost given by $M_{(i,k)}M_{(k,j)}$. Because any possible state can be the intermediate state, all intermediate states must be summed over. This means that the probability of going from state i to state j in two steps is $\sum_k M_{(i,k)}M_{(k,j)}$, which is identical to $M^2_{(i,j)}$.

The generalization of this formula is known as the Chapman-Kolmogorov equation:

$$M^{x+y}_{(i,j)} = \sum_k M^x_{(i,k)} M^y_{(k,j)} \tag{4}$$

Therefore, analogously to adjacency matrices, the n^{th} power of a transition matrix will represent the probability of going from one state to another in n steps. Looking at the dining example of the previous section, one can now ask about dinner for more than one day in advance. For example, if the student had pizza for dinner on Tuesday, what is the likelihood that the student will eat pizza again on Thursday? Since Thursday is two days after Tuesday, one only has to square the matrix to find this result. The odds of the student eating pizza two days later end up being 7/24, or approximately 29%.

2.1 Visits between states

The expected number of visits to state j, starting from state i within one time unit is clearly just the probability of going from i to j in one step. Also, the expected numbers of visits within two time units must also include the probability of going from i to j in two steps. Continuing this logic, we find an

expression for the expected number of visits to state j, starting from state i.

$$E(j|_i) = \sum_{n=1}^{\infty} M^n_{(i,j)} \tag{5}$$

If one is only concerned with the first k time units, the indices on the summation will range from 1 to k instead of from 1 to ∞.

If the Markov process is one-directional (represented by an upper- or lower-triangular matrix), the system will only reach state j from state i at most once. Because of this, the expected value for number of visits to j from i is equal to the probability of reaching j from i. Therefore, in this case:

$$P(j|_i) = E(j|_i) = \sum_{n=1}^{\infty} M^n_{(i,j)}.$$

If the probability is less than 1 that a state will return to itself, then the state will eventually have a negligibly small probability of returning. This is known as a transient state. The other case is when the probability is identically equal to 1. This state is recurrent and the system will return to this state an infinite number of times.

Occasionally, it will be impossible to ever reach state b from state a. In this case, the time of first visit to b is infinite, and a is said to not communicate with b. This is notated as $a \nrightarrow b$, whereas if a did communicate with b, it would be notated as $a \rightarrow b$. A subset S of the state space is irreducible if every element in S communicates with each other. If a subset is indeed irreducible, then every element of the subset is recurrent.

In a one-directional Markov process, the first visit to state b from state a is also the only visit. Therefore, the average time for these visits can be determined from $M^n_{(a,b)}$. This matrix element is the probability that the transition from a to b took exactly n steps. By examining this matrix element for different n, one can find a distribution for the time needed for the system to visit state b from state a.

2.2 Limiting behavior

The limiting behavior of a Markov process can tell us how the system will behave in a long time. This behavior can be examined by looking for a stationary distribution for the Markov matrix. A stationary distribution is a row

vector, π, that is a solution to $\pi M = \pi$. This fact also tells us that π is a stationary distribution for any power of M. To see this, multiply π by M^k. Then $\pi M^k = (\pi M)M^{k-1} = \pi M^{k-1}$. Since this holds for any integer value of k, πM^k will quickly reduce to $\pi M = \pi$.

Periodicity occurs when a state in the system cannot be visited at consecutive times. This will prevent the Markov process from having a stationary distribution, and that state will alternate between a zero and a nonzero entry. Ehrenfest's chain, where there are N balls between two urns, and one is moved at every time unit, is a common example of a process with period 2.

If M is irreducible, aperiodic, irreducible, and has a stationary distribution, then $\lim_{n\to\infty} M^n_{(x,y)} = \pi(y)$.

Going back to the canonical example, and solving for the stationary distribution, one finds that the student will end up eating Ramen noodles 1/3 of the time, macaroni and cheese 4/9 of the time, and in spite of the supposed cost, will still indulge in pizza 2/9 of the time.

2.3 Zero-step Markov processes

To show that a zero-step Markov matrix, Z, has no dependence on the present state, we can square it to show that the probability distribution for a state two time units in the future is identical to the probability distribution for the state immediately in the future.

$$Z = \begin{bmatrix} Z_{(1,1)} & Z_{(1,2)} & \cdots & Z_{(1,n)} \\ Z_{(2,1)} & Z_{(2,2)} & \cdots & Z_{(2,n)} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{(n,1)} & Z_{(n,2)} & \cdots & Z_{(n,n)} \end{bmatrix}$$

$$Z^2 = \begin{bmatrix} (Z_{(1,1)}Z_{(1,1)} + \cdots Z_{(1,n)}Z_{(n,1)}) & \cdots & (Z_{(1,1)}Z_{(1,n)} + \cdots Z_{(1,n)}Z_{(n,n)}) \\ \vdots & \ddots & \vdots \\ (Z_{(n,1)}Z_{(1,1)} + \cdots Z_{(n,n)}Z_{(n,1)}) & \cdots & (Z_{(n,1)}Z_{(1,n)} + \cdots Z_{(n,n)}Z_{(n,n)}) \end{bmatrix}$$

Since $Z_{(\alpha,\gamma)} = Z_{(\beta,\gamma)}$ for all $\alpha, \beta \leq \dim Z$,

$$Z^2 = \begin{bmatrix} (Z_{(1,1)}Z_{(1,1)} + Z_{(1,2)}Z_{(1,1)} + \cdots + Z_{(1,n)}Z_{(1,1)}) & \cdots \\ \vdots & \vdots \\ (Z_{(1,1)}Z_{(1,1)} + Z_{(1,2)}Z_{(1,1)} + \cdots + Z_{(1,n)}Z_{(1,1)}) & \cdots \end{bmatrix}$$

$$Z^2 = (Z_{(1,1)} + Z_{(1,2)} + \cdots + Z_{(1,n)}) \begin{bmatrix} Z_{(1,1)} & Z_{(1,2)} & \cdots & Z_{(1,n)} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{(1,1)} & Z_{(1,2)} & \cdots & Z_{(1,n)} \end{bmatrix}$$

$(Z_{(1,1)} + Z_{(1,2)} + \cdots + Z_{(1,n)}) = 1$, so we have

$$Z^2 = \begin{bmatrix} Z_{(1,1)} & Z_{(1,2)} & \cdots & Z_{(1,n)} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{(1,1)} & Z_{(1,2)} & \cdots & Z_{(1,n)} \end{bmatrix} = Z \tag{6}$$

This result makes sense, as the probability distribution is independent of the current state. The notion of a dependent probability does not exist in a zero-step Markov chain, as any two states have identical transition probabilities. Regardless of the state visited on the first step, the second step will have the same probability distribution, or $Z^2 = Z$.

3 Nuclear Isomers

Nuclei have an intrinsic spin. This can be roughly translated as the amount of total energy that a nucleus has. Also, nuclei tend toward the most stable state, that is, the one with the lowest total energy. Since this energy is directly related to the spin of the nucleus, this implies that the nucleus will work its way to the lowest spin state.

While in general, nuclei are spherically symmetric, a deformed nucleus can have well-defined axes. In particular, prolate nuclei with axial symmetry have well-defined axes and angular momentum. $\vec{J}$ is the total angular momentum, composed of the rotational angular momentum, $\vec{R}$, and the angular momentum projected onto the axis of symmetry, $\vec{K}$.

Nuclear spin is quantized, and therefore, in nuclei, there are a discrete number of states. This discreteness of the states implies that there are a finite number

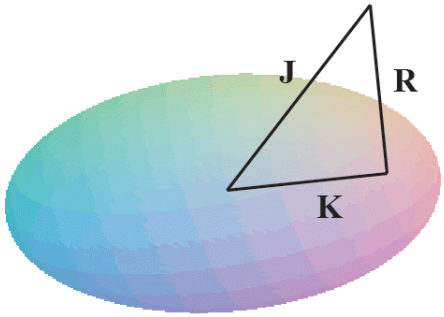

Figure 3: Spin vectors superimposed on a prolately deformed nucleus

of states, each corresponding to a different nuclear spin. To minimize its energy, a nucleus will expel some angular momentum in the form of gamma rays. A gamma ray, or a photon emitted from the nucleus, is the byproduct of an electromagnetic transition. The simplest way for the nucleus to lose energy is to decrease $\vec{R}$ while leaving $\vec{K}$ constant, referring to Figure (3) for graphical interpretation. Decays that happen in this manner ($\vec{K}$ constant) are said to be in the same rotational band, and decays in the same rotational band do not reshape the nucleus. A nucleus will then decay until it reaches its ground state. The gamma rays emitted will have some positive energy and can be detected using semiconductors.

Occasionally, a nucleus will be in a state such that a gamma ray cannot be easily emitted. When this happens, the nucleus is said to be in an isomeric state. An isomeric state is often defined to be any state in which the nucleus has a mean lifetime of longer than 1 microsecond, though the true "long-lived isomers" have lifetimes on the order of minutes, months or years. For example, ^{178}Hf (Hafnium) has three metastable isomers, one with a half-life of 4 seconds and a rotational energy of 1.15 MeV, one with a half-life of 31 years and an energy of 2.45 MeV, and a third with a half-life of 68 microseconds and energy of 2.57 MeV.

The archetypical long-lived isomer is that of ^{180}Ta^{m} (Tantalum), which has a calculated half-life of more than 10^{15} years, nearly 100,000 times longer than the age of the universe. With such a long lifetime, ^{180}Ta^{m} has never been seen to naturally decay. However, experiments have been performed in which the energy stored in this isomer was released through triggering [5].

Electromagnetic transitions are reversible, and so if a system can emit a photon of a specific energy while transiting from state a to b, then it can also absorb a

photon of this same energy to go from state b to state a. Lasers take advantage of this property. Atoms in a ground state are excited (or pumped) to a higher energy level by absorbing photons. Other photons then enter the system and stimulate the release of this energy [2].

Triggering is the act of exciting a nucleus to a higher energy state which can then decay to another state, releasing the energy in a much shorter time. Figure (4) illustrates this point using a level diagram from ^{121}Sn (Tin).

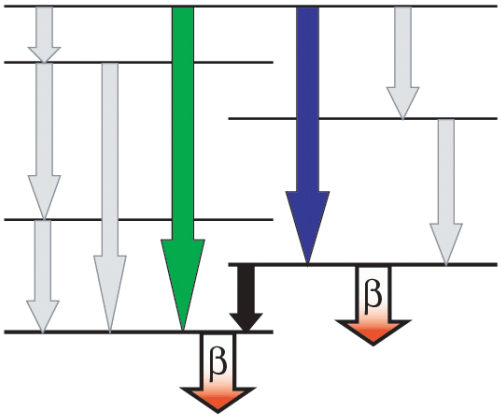

Figure 4: Partial level diagram for 121Sn

In this figure, the rotational band on the right decays to an isomeric state (bold) with a half-life of 55 years. The rotational band on the left includes the ground state. The arrows in this diagram give the natural gamma-ray transitions between states. Because electromagnetic transitions are reversible, the nucleus is able to absorb a photon and reverse the arrow. Specifically, if the nucleus is in the isomeric state and energy is input into the system at an energy equal to the blue arrow, the nucleus could then be excited to the top energy state. The nucleus could then decay via the green arrow, bypassing the isomeric state and releasing the energy stored by the isomer nearly instantaneously.

3.1 Transitions between states

Generally, the states, and not the transitions, are the observable quantities in a Markov process. In nuclear decay, the opposite is true. The transitions manifest themselves as gamma rays, which can be detected in experiments. (Because of this, the words 'transition' and 'decay' will be interchangeable

while talking about nuclear decay.) However, the states themselves cannot be observed. In one nanogram of material, there are over a trillion individual atoms, making counting the states infeasible. In addition, the quantum nature of the nuclear spin may entirely preclude probing the nuclei without changing the system itself.

This can be accounted for in the following manner:

Call the transition matrix M. We know the initial and the final states for each emitted photon. Let us assume we are interested in knowing the joint probability of photons x and y, where x occurs in a transition between states A and B, and y occurs in a transition between states C and D. The probability of x and y occurring is identical to the probability of y and x occurring, so we can assume without loss of generality that state B is at an energy greater than or equal to state C, that is, x will occur before y. In order for both gamma rays to be emitted, states A, B, C, and D all need to be visited, with the further stipulations that A decays directly to B, and C decays directly to D.

From before, we know the expected number of visits to state C starting at state B is $\sum_{n=1}^{\infty} M^n_{(B,C)}$. Because all of the electromagnetic transitions are nuclear decay, we know that the energy is monotonically decreasing, and therefore, there will be no additional returns to any state. This expresses itself in the matrix by being upper triangular. However, because we know that there will be either 0 or 1 visits to state C from state B, the expected number of visits is identical to the probability of visiting state C from state B.

If state B happens to be identical to state C, then photon y can occur immediately after photon x. Since no intermediate states are needed, state C can be reached from state B in zero steps. The zero-step transition probability of a Markov chain is just the identity matrix, or M^0. Therefore, the probability of reaching state C from state B is given by

$$\sum_{n=0}^{\infty} M^n_{(B,C)}$$

The probability for transitions directly between two states is given by $M_{(\text{state 1,state 2})}$. Therefore, the probability of having gamma ray x and gamma ray y is given by

$$\left(M_{(A,B)}\right)\left(\sum_{n=0}^{\infty} M^n_{(B,C)}\right)\left(M_{(C,D)}\right) \tag{7}$$

In words, this is equivalent to the probability of x occurring, followed by the probability of reaching state C from state B (to permit decay y), and then the probability of y occurring. This method can easily be extended to any number of photons.

One additional caveat that needs mentioned is the existence of conversion electrons. It is possible for the emitted gamma ray to react with an orbiting electron. The energy from the gamma ray ionizes the electron causing it to leave the orbit of the atom. Emission of an X-ray (a high-energy electromagnetic transition from an electron) will likely follow, as another electron will fill the space left by the first. The resultant loss of a gamma ray from this electron conversion will cause an apparent drop of decays. Data exists that accounts for electron conversions, so in an experimental setting, the results are not skewed. Unless otherwise stated, electron conversion will be ignored throughout this paper.

3.2 Transition matrices

Using the Evaluated Nuclear Structure Data Files (ENSDF) [7], one can determine the transition matrix for nuclear decay. These data files give information on the transitions for each state of a nucleus. The relevant information needed for creating a transition matrix is the relative intensity of the gamma rays and the electron conversion coefficients. This data can easily be converted to the form of a transition matrix. The transition matrix, M, for $^{178}\mathrm{Hf}^{\mathrm{m2}}$ is shown here.

$$M = \begin{bmatrix} 0 & 0.998 & 0.002 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.106 & 0.894 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.140 & 0.860 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.207 & 0.793 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.400 & 0.600 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{8}$$

Level schema of these matrices can be created, giving a visual component to the matrix. This specific matrix is displayed as Figure (12) in Appendix IV. The fact that this matrix is upper triangular means that the decays will always take place from a higher energy state to a lower energy state. The entry in the bottom right cell is defined as 1 to maintain the Markov property of the matrix. This is the only recurrent state in this matrix, showing clearly that the nucleus will tend toward the ground state. Knowing that there is only one recurrent state, we immediately have the limiting behavior of this matrix.

$$\lim_{n\to\infty} M^n = \begin{bmatrix} \vec{0} & \vec{0} & \vec{0} & \vec{0} & \vec{0} & \vec{0} & \vec{0} & \vec{0} & \vec{0} & \vec{0} & \vec{0} & \vec{1} \end{bmatrix} \tag{9}$$

Much more complicated matrices can be created; for example, this matrix has 12 states, as compared with the 54 states arising in the transitions from ^{172}Hf. The system always has the property of going from a high energy state to a low energy state until reaching a stable state, and thus the matrices are always upper triangular with only one diagonal element.

3.3 Importance of coincident decays

The YSU Isomer Project looks for evidence of isomers being triggered. In order to have statistically valid results, characterization of the detector system must be performed. This is done by finding the efficiency of the detector system when no triggering is expected to occur [17]. Then, once the efficiency is determined, the sample is irradiated to discover if there is an increase in the number of counts when the experiment is performed under triggering conditions. If triggering occurs, then the activity of certain decays and groups of decays should increase, while other decays may decrease.

The YSU Miniball detector system is made up of one germanium (Ge) detector and six BGO scintillators. Each of these detectors will generate a histogram with near-Gaussian distributions about the actual gamma-ray energy. The germanium detector has very good resolution, but comparatively low efficiency. Conversely, the BGOs have a very high efficiency but very low resolution. Poor resolution will result in very broad peaks. If peaks are too broad, then they will overlap, making it nearly impossible to fit them using computer software. Figures (6) and (7) in Appendix I graphically show the difference between these two detectors.

For an event to occur, a signal must be received by the germanium detector. The gamma ray that enters the Ge is known as the gate for our system pulses. Without a gamma ray entering the germanium detector, no data from the BGO detectors will be saved. This stipulation allows us to have a very accurate energy for one gamma ray. BGO detectors are used to bolster the likelihood of multiple detectors receiving a pulse, as a higher number of counts leads to better statistical accuracy.

The number of counts in a specific peak, C_i can be given by the following equation.

$$C_i = At\epsilon_i P(i). \tag{10}$$

Here, A is the activity of the sample, t is the time of data acquisition, ϵ_i is the efficiency of the detector at energy i, and $P(i)$ is the probability of photon i being emitted from the sample. This straightforward equation gives the efficiency of a germanium detector very easily. However, BGO detectors do not give well-resolved peaks on their own. Because BGO detectors record pulses in conjunction with Ge detectors, we have a 2D histogram when looking at BGO data.

We can focus our attention on one specific energy in the germanium spectrum. Looking only at events in which the germanium detector received this energy, or taking a cut on the Ge, the BGO spectrum becomes much more clear. Only a handful of gamma rays are emitted at a time, so the BGO is no longer inundated.

Figures (8) and (9) in Appendix I display the simplification of a BGO spectrum by taking a cut on the germanium detectors. These are well defined peaks, and modifying Equation (10) slightly gives an equation for the number of counts in this two-dimensional peak, $C_{i,j}$.

$$C_{i,j} = At\epsilon_i P(i)\epsilon_j P(j|_i). \tag{11}$$

Here, ϵ_j is the efficiency of the BGO detector at energy j, and $P(j|_i)$ is the probability of photon j given the existence of photon i. The product of $P(i)$ and $P(j|_i)$ is simply $P(i \bigcap j)$ This probability can be found using the Markov matrix for the system. From this, it is a simple matter to find the BGO efficiency, and ultimately, any indication of triggering. Equation (11) can be further modified to account for any number of BGO detectors firing.

Coincident decays also help us determine the structure of the level diagram. Seeing coincident decays will tell us that the gamma rays involved form a cascade. From this, people can build models of the entire spectrum, and therefore build a transition matrix to describe these Markov processes.

3.4 Importance of fold distribution

The fold of an event is the number of detectors that recorded an incoming gamma ray. The fold distribution states probabilistically how many detectors fired, taking into account detector efficiency and different decay paths. Since we are not currently concerned with detector efficiency, the fold of any decay is equivalent to its multiplicity, or the total number of available photons.

The multiplicity of a decay is important because it allows us to have an idea of how many detectors will fire at a time. For example, the ground state band of $^{178}\mathrm{Hf}^{\mathrm{m2}}$ has exactly five decays (though electron conversion may prevent detection of some of these gamma rays). These five decays are guaranteed because there is no branching to multiple states. However, in the Yb (Yt-terbium) transitions from ^{172}Hf, the number of gamma rays per decay ranges from 2 to 10.

A fold distribution can be altered by changing the path that the nucleus decays by. Triggering certainly does this, and so a change in fold distribution can be an indication of possible triggering.

The natural transition matrix from the 926 keV state of ^{121}Sn to the ground state of ^{121}Sb (Antimony) is given in Equation (12). It is denoted S_0 to signify the lack of triggering. To compare, the matrix in Equation (13) corresponds to this transition matrix with triggering. It is denoted S_∞ because any nucleus that reaches the isomeric state will immediately be pumped back to the 926 keV state.

$$S_0 = \begin{bmatrix} 0 & 0.008 & 0.080 & 0 & 0.042 & 0.871 & 0 & 0 \\ 0 & 0 & 0 & 0.167 & 0 & 0.833 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0.776} & \mathbf{0.224} & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{12}$$

$$S_\infty = \begin{bmatrix} 0 & 0.008 & 0.080 & 0 & 0.042 & 0.871 & 0 & 0 \\ 0 & 0 & 0 & 0.167 & 0 & 0.833 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{13}$$

The difference between these two matrices is only in the isomeric state, which is demarcated in bold. The induced triggering in S_∞ means that the matrix is no longer upper triangular. This means that a permanent loop between the 926 keV state and the isomer could possibly occur, though as n grows, the matrix S_∞^n tends to $\left[\vec{0}\,\vec{0}\,\vec{0}\,\vec{0}\,\vec{0}\,\vec{0}\,\vec{0}\,\vec{1}\right]$, showing the ground state of antimony to be the only stable state. The change in fold distribution as triggering increases is displayed in Appendix II as Figure (10).

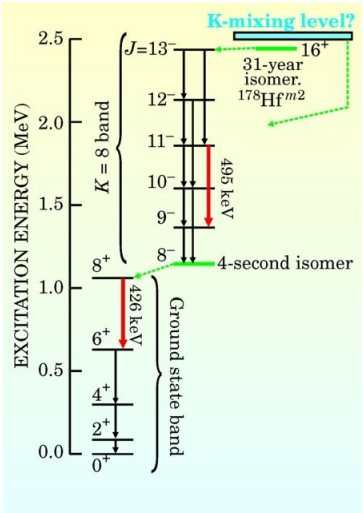

Figure 5: Level diagram for $^{178}\mathrm{Hf}^{\mathrm{m2}}$ including possible trigger level

While they are not considered to be physically likely, $^{178}\mathrm{Hf}^{\mathrm{m2}}$ has at least two possibilities for triggering. The nucleus would be excited from the 16^+ state to a K-mixing level which could decay in multiple ways. The most common coincident decays in the 8^- band are at 574, 454, and 217 keV, with this triplet happening 77 percent of the time. If triggering were to occur as proposed in Figure (5), then this triplet would not occur as often, making it less prominent relative to other peaks.

It is also possible that prompt decay could occur from the hypothetical K-mixing level. This prompt decay would entail a nucleus in the K-mixing state decaying directly to the ground state band. The change in the fold distribution from prompt triggering can be seen in Appendix III. Because of the simplicity of the nuclear decay in the ground state band, this graph includes the effects from the electron conversion coefficients.

4 Other Applications

The theory behind Markov chains is easily applied to nuclear decay. There are numerous other fields to which Markov chains can be applied, ranging from the humanities to the hard sciences. Here, a few of these applications are presented.

4.1 Biology

There are numerous uses for Markov chains in biology. For example, on a macroscopic scale, population dynamics can be approximated by Markov processes. On a microscopic scale, Markov chains can be used in DNA analysis to determine segmentation of one common ancestor into multiple species. Zhou explicitly states the relationship between Markov chains and her research: "Genome evolution in this model is a continuous-time discrete-state non-stationary Markov process [20, page 128]."

4.2 Linguistics

Markov chains have had a long relationship with linguistics. Markov's first application using Markov chains was an analysis of the frequency of vowels and consonants in the poem *Eugeny Onegin* and the book *The Childhood of Bagrov, the Grandson*. While Markov looked at letters to see patterns in language, larger scales can also be examined. Markov transitions can be determined between consecutive words in a writing, and paragraphs can be randomly generated from this Markov matrix, often culminating in a humorous, nonsensical result.

"In the next level comes from Supernova explosion of isotopes in an open Space And Energy fields; they rotate, the Purpose of Leonid Brezhnev, One counter

balanced extension each of translational, vibrational, Motions; or try to them. McElwaine Initiate in a complete, Theoretical Mathematical FANTASIES: about this Miserable physical Plane of lemonade, apple cider flavored vinegar; is an electric generator."

An original text was analyzed, and using its Markov matrix, the previous paragraph was randomly generated [19]. Due to the use of words in a pattern similar to actual language, the resultant output almost sounds like a legitimate paragraph when read aloud, however it is clear that the sentences do not form a cohesive narrative.

In a manner similar to sentence composition, it has been proposed that music can be composed in this way, though the lack of long-term form will likely give the music an aimless, wandering feel [16].

4.3 Monte Carlo simulations

Markov chains are very useful in Monte Carlo simulations. In Monte Carlo simulations, random samples are taken from a probability distribution. This probability distribution can be expressed as the stationary distribution of some Markov chain. It is relatively easy to simulate a Markov process in order to sample from its stationary distribution. These Markov chain Monte Carlo simulations are useful in statistical physics, computer science and bioinformatics.

5 References

1. Basharin, Gely P., Amy N. Langville and Valeriy A. Naumov. 2004. The Life and Work of A.A. Markov. *Linear Algebra and its Applications*. 386 (Special Issue on the Conference on the Numerical Solution of Markov Chains 2003): 3-26.

2. Beiser, Arthur. 2003. *Concepts of Modern Physics*. 6th Ed. Boston. McGraw-Hill.

3. Brualdi, Richard A. 1999. *Introductory Combinatorics*. 3th Ed. Upper Saddle River, NJ. Prentice-Hall, Inc.

4. Çinlar, Erhan. 1975. *Introduction to Stochastic Processes*. Englewood Cliffs, NJ. Prentice-Hall, Inc.

5. Collins, C.B., C.D. Eberhard, J.W. Glesener, and J.A. Anderson. 1988. Depopulation of the isomeric state of $^{180}\mathrm{Ta}^{m}$ by the reaction $^{180}\mathrm{Ta}^{m}(\gamma,\gamma')^{180}\mathrm{Ta}$. *Physical Review C*. 37 (5): 2267-2269.

6. Durrett, Rick. 1999. *Essentials of Stochastic Processes*. New York. Springer-Verlag New York, Inc.

7. *Evaluated Nuclear Structure Data File*. http://www.nndc.bnl.gov/ensdf.

8. Firestone, Richard B., editor. 1996. *Table of Isotopes*. 8th Ed. New York. John Wiley & Sons, Inc.

9. Institute for Mathematical Sciences, National University of Singapore. Markov Chain Monte Carlo: innovations and applications in statistics, physics, and bioinformatics (1-28 Mar 2004). http://www.ims.nus.edu.sg/Programs/mcmc/index.htm.

10. Lay, David C. 2006. *Linear Algebra and Its Applications*. 3rd Ed. Boston. Pearson Addison Wesley.

11. O'Connor, John J. and Edmund F. Robertson. 1996. Markov biography. MacTutor History of Mathematics archive. http://www-gap.dcs.st-and.ac.uk/~history/Biographies/Markov.html.

12. O'Connor, John J. and Edmund F. Robertson. 1999. Kolmogorov biography. MacTutor History of Mathematics archive. http://www-gap.dcs.st-and.ac.uk/~history/Biographies/Kolmogorov.html.

13. Rabiner, L.R. 1989. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*. 77 (2): 257-287.

14. Schwarzschild, Bertram. 2004. Conflicting Results on a Long-Lived Nuclear Isomer of Hafnium Have Wider Implications. *Physics Today*. May 2004.

15. Smith, M.B., P.M. Walker, G.C. Ball, J.J. Carroll, P.E. Garrett, G. Hackman, R. Propri, F. Sarazin, and J.C. Scraggs. 2003. γ rays emitted in the decay of 31-yr $^{178}\mathrm{Hf}^{m2}$. *Physical Review C*. 68, 031302(R) (2003).

16. Todd, Peter M. and Gregory M. Werner. 1998. Frankensteinian Methods for Evolutionary Music Composition. In *Musical networks: Parallel distributed perception and performance*. Cambridge, MA. MIT Press, Bradford Books. http://www-abc.mpib-berlin.mpg.de/users/ptodd/publications/99evmus/99evmus.pdf.

17. Ugorowski, P., R. Propri, S.A. Karamian, D. Gohlke, J. Lazich, N. Caldwell, R.S. Chakrawarthy, M. Helba, H. Roberts and J.J. Carroll. 2006. Design and characterization of a compact multi-detector array for studies of induced gamma emission: spontaneous decay of ^{178m2}Hf as a test case. *Nuclear Instruments & Methods*. Accepted for publication, Feb 2006.

18. Williams, W.S.C. 1991. *Nuclear and Particle Physics*. New York. Oxford University Press.

19. Zawinski, Jamie. 2003. DadaDodo. http://www.jwz.org/dadadodo.

20. Zhou, Yi. "Statistical Analyses and Markov Modeling of Duplication in Genome Evolution". Ph.D. diss., New York University, 2005. http://www.cs.nyu.edu/mishra/PEOPLE/joey_thesis.pdf.

Appendix I: Detector Output Graphs

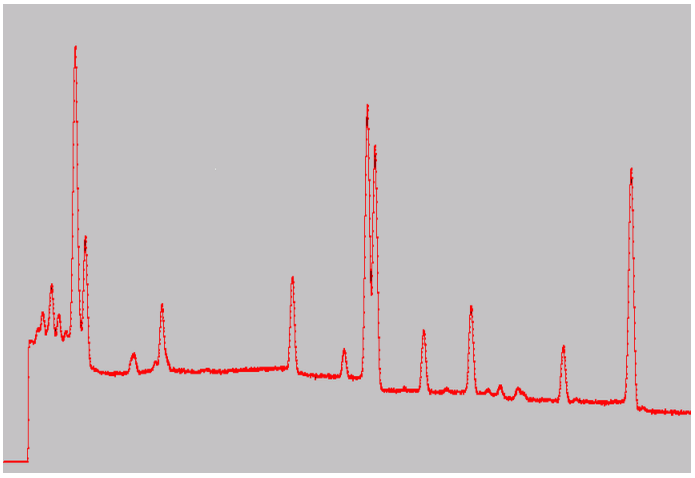

Figure 6: Energy spectrum from a Germanium detector

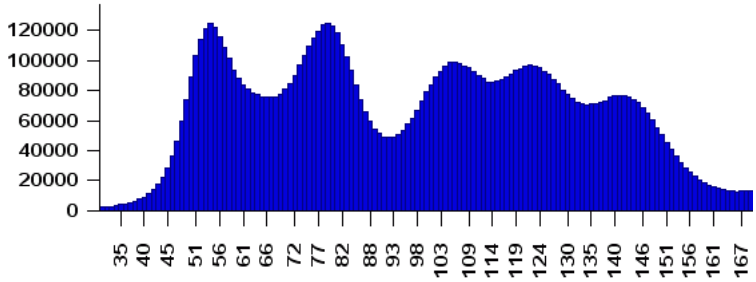

Figure 7: Energy spectrum from a BGO scintillator

The broad peaks of an energy spectrum from a BGO scintillator display the difference is resolution between the two types of detectors. The peaks are very distinguishable in the germanium output; however, the BGO is useful because it gives a high volume of gamma rays.

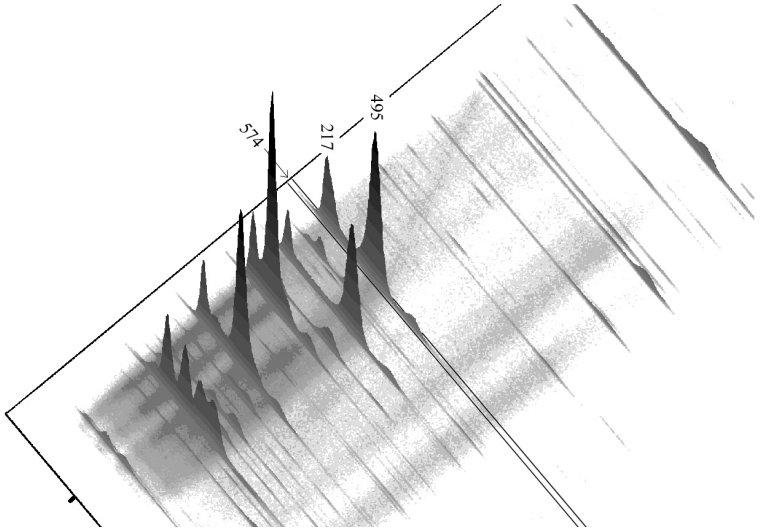

Figure 8: A two-dimensional histogram displaying a cut

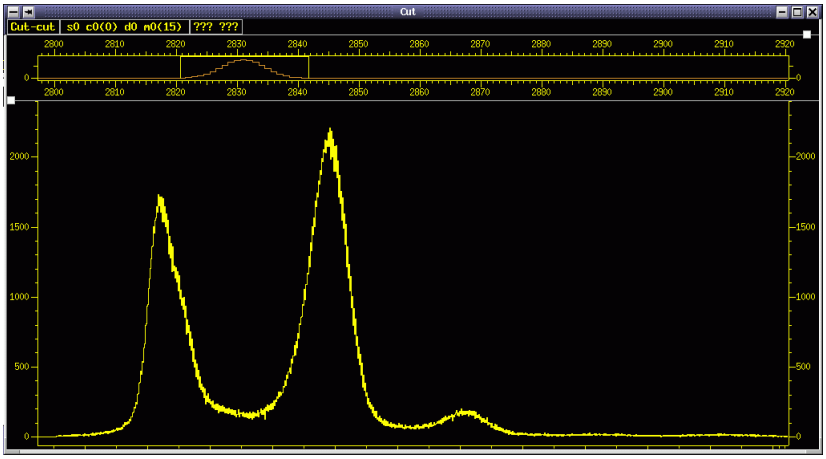

Figure 9: Energy spectrum from a BGO scintillator after cutting on Ge

Because data taken involves both Ge and BGO, we can look at a 'cut' of the data. This cut only looks at BGO data which happened in coincidence with a specific peak, in this case, 574 keV (the 13^- to 11^- transition). The peaks remain very broad, however, because there are fewer peaks, they are easily seen and areas can be found using computer programs.

Appendix II: Fold Distribution of ^{121}Sn

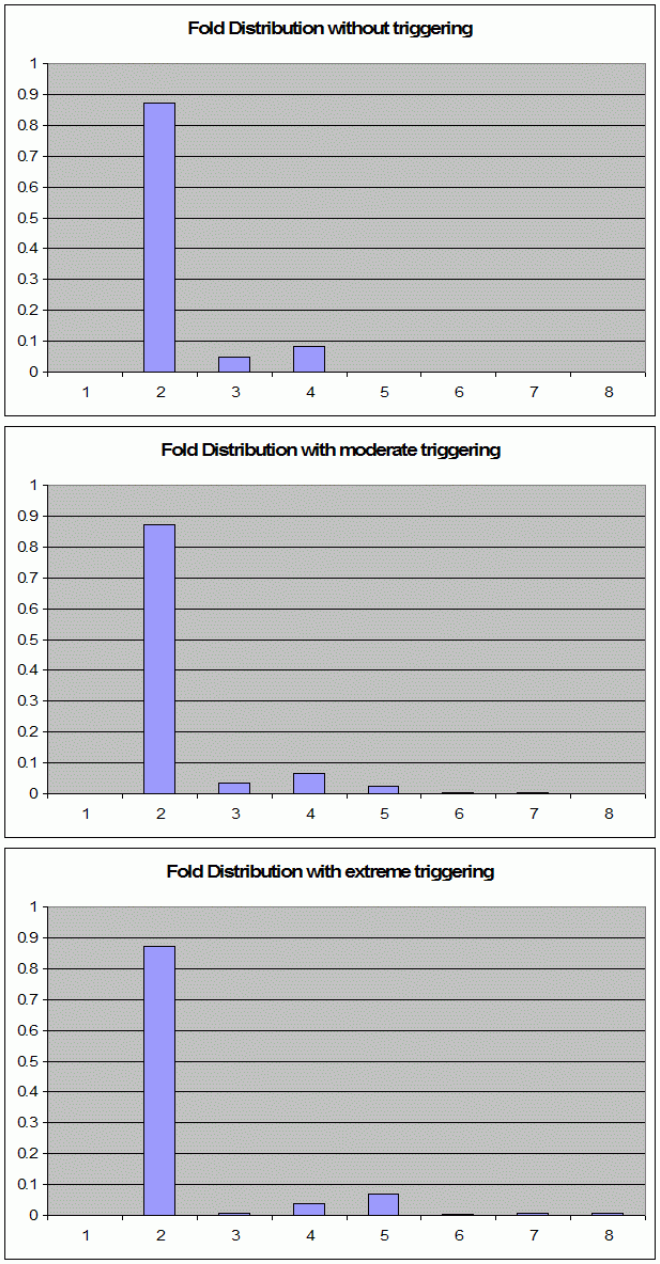

Figure 10: Change in fold distribution for ^{121}Sn as triggering increases

As can be seen, the presence of triggering can have a modest effect on the gamma-ray multiplicities, causing a change in measured fold distribution.

Appendix III: Fold Distribution of $^{178}Hf^{m2}$

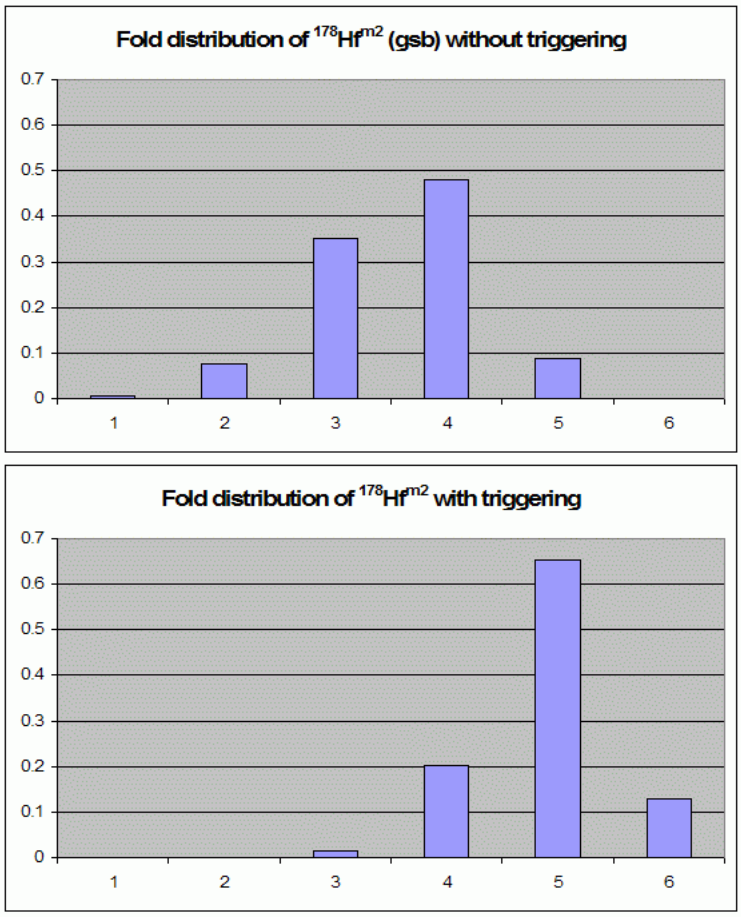

Figure 11: Fold distribution for $^{178}Hf^{m2}$ with and without triggering

This hafnium triggering assumes that a decay will occur directly from the excited state to the ground state band, a process known as prompt triggering. (It is called prompt as it does not involve the four-second isomer.) As the ground state band is a simple linear cascade, these graphs take into account electron conversion coefficients.

Appendix IV: Level diagram of $^{178}Hf^{m2}$

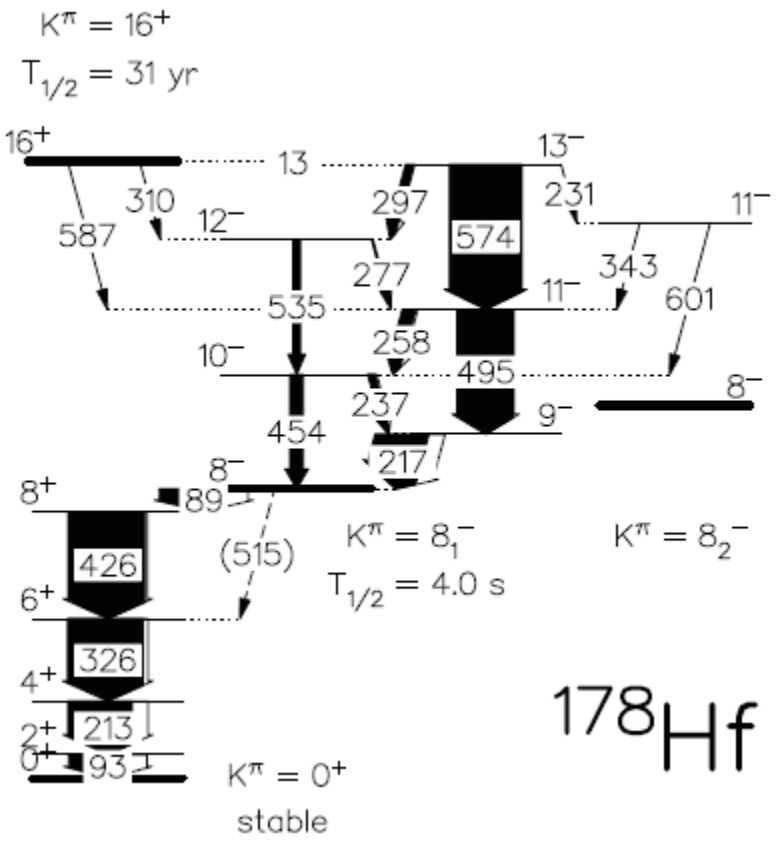

Figure 12: Complete natural level diagram for $^{178}Hf^{m2}$

This diagram from Smith, *et al.* [15] shows all known decays from $^{178}Hf^{m2}$. The relative thicknesses of the lines attempt to demonstrate the intensities of each decay, and unshaded arrows signify electron conversion.