DEAR Technical Note IR-3

Statistical and Systematic Errors in the DEAR Experiment

Ross King

Institute for Medium Energy Physics
Austrian Academy of Sciences, Vienna, Austria

April 15, 1998

1 Introduction

The purpose of this report is three-fold: First, in a section entitled **Statistical Errors**, to define the best approach to the fitting of the kaonic hydrogen spectrum, in the case of normal (greater than 2000 events in the K_α peak) statistics, and to give Monte Carlo results that predict the statistical power of the DEAR experiment under a variety of conditions. A future note will discuss the case of low statistics and the iterative fit algorithm. Second, in a section entitled **Systematic Errors**, to discuss the sources of all systematic uncertainties in the experiment and estimate their magnitudes based on Monte Carlo studies. Finally, in a section entitled **Conclusions**, to give a final estimate of the obtainable overall precision of the experiment.

2 Statistical Errors

2.1 Disentanglement

This word, though often used, is ill-defined. When we speak of the K-complex, there are three possible questions of interest. First, can we resolve the individual peaks? Second, can we determine the relative intensities of the individual peaks? Finally, can the K-complex be used to increase our knowledge of the K_α peak position?

The answer to the first question is "probably not." Given the most likely Γ_{1s} of around 400 eV, the K_β line is nothing more than a slight shoulder on the K-complex.

The second question is a matter of statistics. At our present goal of 10000 K_α X rays, the relative intensities of the various lines in the K-complex can only be determined to within 5-20%. There are simpy too many lines, too close together.

Nevertheless, this will be the most accurate measurement of these intensities yet performed, and will no doubt bring joy to the various producers of cascade code.

Concerning the final question, yes, of course we can independently resolve the peaks of the K-complex with sufficient statistics. This could be considered an independent proof of the identity of the K_α peak. Note however, that if one has sufficient statistics to fit the K-complex, the K_α peak will already be unambiguously identified. Furthermore, it is not necessary to fit the individual positions of the K-complex peaks; rather, the relative postions are fixed by QED and the absolute positions are fit by a single global shift. This means that these peaks can contribute to the statistical precision of the K_α peak position, even if the individual peaks will be poorly resolved.

2.2 Quantifying the Background

In order to compare with previous analyses, we adopt the following definition of the background percentage (BG):

$$\mathrm{BG}(\%) = \frac{\int_{-\Gamma/2}^{+\Gamma/2} B}{\int_{-\Gamma/2}^{+\Gamma/2} B+S} \cdot 100$$

where B is the background and S is the K_α signal, which are integrated over one Lorentzian width Γ of the kaonic hydrogen K_α peak.

Note that the standard signal-to-noise ratio (S/N) over this region is related to the background percentage BG in the following manner:

$$\mathrm{S/N} = \frac{100}{\mathrm{BG}(\%)} - 1$$

2.3 Fitting Function

In general, we use a fitting function with the form

$$F = S + B$$

where S is the signal function and B is the background as a function of energy. We assume that B will be fit by a low-order polynomial such that

$$B = B_0 + (B_1 \cdot E) + (B_2 \cdot E^2) + ... + (B_p \cdot E^p)$$

but it may be the case that a more complex function is required.

One might guess a signal function S with the following form:

$$S_1 = S(E_1...E_n, A_1...A_n, E_1^{cal}...E_m^{cal}, C_1...C_m, \sigma(E), \Gamma)$$

where E_n^{em} is the energy of the nth kaonic hydrogen peak, A_n is the amplitude of the nth kaonic hydrogen peak, $\sigma(E)$ is the energy-dependent detector resolution, and Γ is the width of the 1s level of kaonic hydrogen. The amplitude C_m and the energy E_m^{cal} are the parameters of the mth calibration peak.

However, this signal function is not optimal, because it treats each kaonic hydrogen peak independently, when in fact the relative positions are fixed. To use this function is to use too many parameters, which degrades the statistical power of the fit. In addition, the energy resolution of our detectors should be determined well in advance of the fit, and should only be allowed as a free parameter in a cross-check and not in the final fit.

The function using the minimum number of parameters is given by

$$S_2 = S(A_1 ... A_n, C_1 ... C_m, E_1^{cal} ... E_m^{cal}, \epsilon, \Gamma)$$

where ϵ is the shift of the 1s level in kaonic hydrogen due to the strong interaction. By using this fitting function, all the statistical power of the K-complex is included in the determination of ϵ and Γ_{1s}.

Written explicitly, the signal function is

$$S = \sum_{i=0}^{p} B_i \cdot E^i + \sum_{k=1}^{m} C_k \cdot G(E_k^{cal}, \sigma(E)) + \sum_{j=1}^{n} A_j \cdot V(E_j^{em} + \epsilon, \sigma(E), \Gamma)$$

where E^i is the average energy of channel i, G is a Gaussian, and V is a Voightian function. The value E_j^{em} is the QED prediction of the purely electromagnetic transition energy of the jth K-series transition in kaonic hydrogen. These values are constants rather than parameters of the fit. Similarly, the energy-dependent energy resolution $\sigma(E)$ is also a constant pre-determined function and not a parameter of the fit.

2.4 Implementaion

A Monte Carlo generator for the DEAR experiment requires the following:

- Gaussian peak generator
- Lorentzian peak generator
- Voightian peak fitter
- Chi-squared minimization routine

The first two items are obvious, as the detector response is modelled by a Gaussian, and the line-shape of the 1s level of kaonic hydrogen is Lorentzian. However, the resulting convolution of these two functions is the so-called *Voightian*, a function for which one requires a numerical approximation in order to fit. Such a routine was taken from the commercial software package **IGOR**.

The CERN package MINUIT was used as the chi-squared minimization routine. In order to generate random events with a Lorentzian probability distribution, the CERN routines FUNPRE and FUNRAN were used.

The Monte Carlo operation was guided by a control file, in which the following parameters were set (default values in brackets):

- *npeaks*: number of peaks in the kaonic hydrogen cascade [5]
- *nkas*: number of K_α events to generate [10000]
- *ntrials*: number of spectra to generate and fit [1000]
- *shift*: nominal value of ϵ [323 eV]
- *gamma*: nominal value of Γ_{1s} [400 eV]
- *back*: background percentage [30%]
- *sigma*: detector resolution at 6.5 keV [150 eV]

Typically ten-thousand K_α events were generated from a Lorentzian distribution characterized by a width *gamma*, at a position specified by *shift*. The remaining *npeaks*-1 peaks were generated with relative intensities suggested by the KEK kaonic hydrogen spectrum [2].

The CCD energy resolution as a function of energy is well-modelled by the equation

$$\Delta E(\sigma) = \omega\sqrt{N^2 + \frac{F \cdot E}{\omega}}(eV)$$

where E is the x-ray Energy, F is the Fano factor (≈ 0.12), ω is the conversion energy for an electron-hole pair in silicon (=3.68 eV), and N is the factor associated with noise in the charge transfer and readout of the chips. The above equation has been used to fit data from our CCDs, obtaining a value for N of 15 ± 0.6.

The generated Lorentzian peaks are convoluted with the resulting function, scaled by the factor *sigma* at 6.5 keV, and the energy-dependent CCD efficiency, taken from [4]. The resulting point is then binned in a 100-channel histogram with an energy range from 0 to 10 keV.

The Titanium K_α and K_β calibration peaks are also generated in a 100-channel histogram, convoluted with the detector resolution, with normalizations scaled by $1.5 \cdot nkas$ and $0.15 \cdot nkas$, respectively.

The background was similarly generated, again using a polynomial function and the CERN routines FUNPRE and FUNRAN, scaling to the number of K_α events by the factor *back*, and similarly convoluting with the energy resolution and efficiency, and binned in to a histogram of 100 channels.

The final 100-channel spectrum is the sum of the signal, calibration, and background histograms. An example spectrum, with 5000 generated K_α events and a 60% background, is shown in figure 1.

Once a spectrum had been generated, MINUIT was called, with initial parameters that were specified in the control file. Good convergence was, as always, contingent upon choosing a good set of starting values for the fit parameters.

In order to properly fit a histogram, the fit functions (Gaussians for calibration peaks, Voightians for kaonic hydrogen peaks) were integrated over a given channel width using the CERN routine DGAUSS.

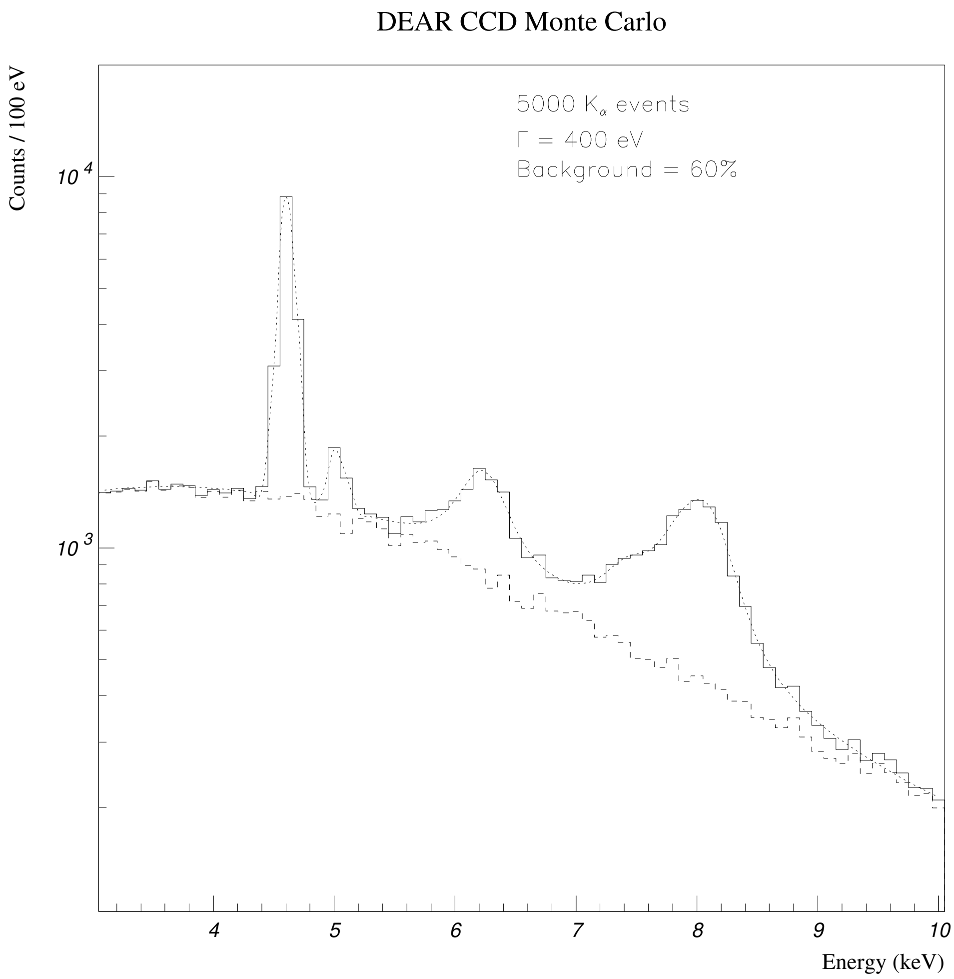

Figure 1: A sample generated kaonic hydrogen spectrum.

The procedure described above was repeated *ntrials* times, and the results of the fit (the fit parameters and the resulting chi-squared value) were written to a file, later to be analyzed in PAW.

2.5 Results

Histograms of one thousand fit results for a specific case (5000 K_α events, 60% background) are shown in figure 2. One can observe that the fits of the shift and width are unbiased relative to the true values (the vertical lines). Furthermore, the χ^2 per degree of freedom for the fits is nicely distributed around one.

For the sake of comparison, we also ran the same set of fits, with one difference: the shift of the K_α peak and the shift of the K-complex were separate parameters. The results are shown in figure 3. One can see that the K-complex yields relatively little information (hence the large sigma of the fit distribution) about the shift; nevertheless, the expected error on the K_α shift is 8.7 eV in this case, compared to 8.2 eV for the case of a single fit parameter.

A summary of the Monte Carlo results (using the fitting function $B + S_2$) is given in table 1.

K_α events	Background (%)	$\Delta\epsilon$ (eV)	$\Delta\Gamma_{1s}$ (eV)
10000	0	3.5	5.9
	30	4.3	9.0
	60	6.2	13.7
	80	8.3	21.8
	90	12.6	28.0
5000	0	4.5	7.0
	30	6.1	14.5
	60	8.2	20.5
	80	11.6	30.8
	90	16.6	44.9
2000	0	7.2	11.1
	30	9.4	22.5
	60	12.8	32.7
	80	18.4	48.2
	90	25.7	68.5

Table 1: Statistical errors on shift and width parameters for various combinations of background and signal levels.

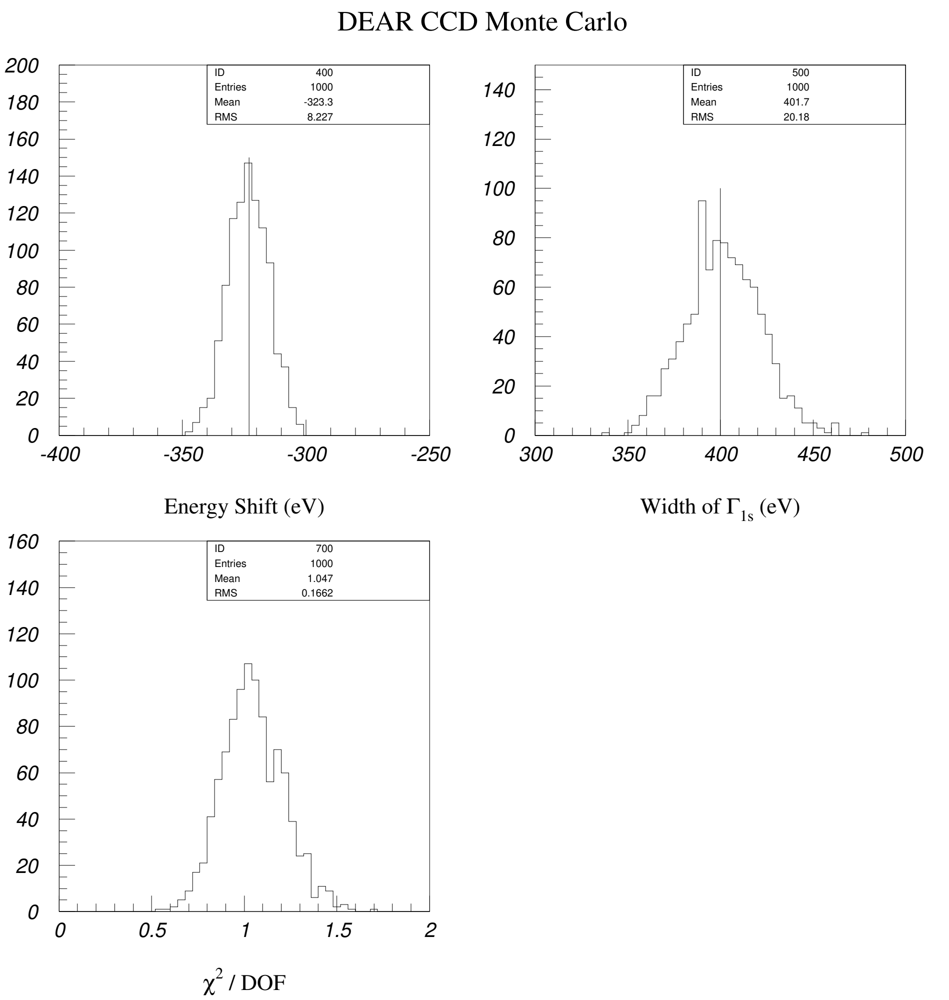

Figure 2: Sample fit results, 5000 K_α events, 60% background.

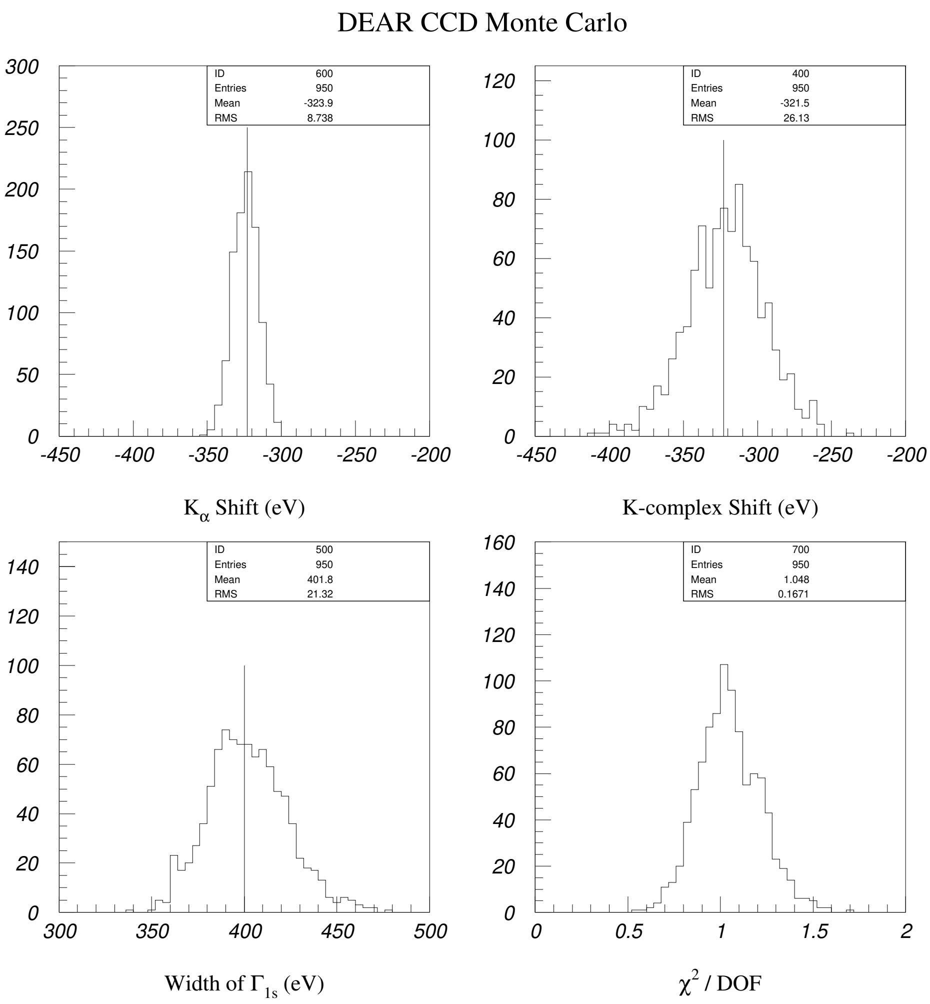

Figure 3: Fit results when the K_α and K-complex are fit separately, 5000 K_α events, 60% background.

3 Systematic Errors

In a previous section of this report, we have shown that we can expect a statistical error on the position of the K_α peak to be on the order of 4 eV for 10000 K_α events (assuming Γ_{1s} = 400 eV; this error scales with Γ_{1s}). In this chapter we discuss the systematic errors associated with the measurement, explain the strategy developed to minimize these errors, and give an estimate of the systematic error associated with such a measurement.

3.1 Sources of Systematic Error

3.1.1 Electromagnetic transition energies

In order to measure the component of the kaon-nucleon interaction due to the strong force, we must measure the shift of position of the K_α line, $\Delta E_{2p\rightarrow 1s}$, where

$$\Delta E_{2p\rightarrow 1s} = E^{measured}_{2p\rightarrow 1s} - E^{EM only}_{2p\rightarrow 1s}$$

that is, we must measure the position of the $K_\alpha$line and then subtract the component due only to the electromagnetic force, $E^{EM only}_{2p\rightarrow 1s}$. The purely electromagnetic $2p \rightarrow 1s$ transition energy in kaonic hydrogen is calculated to be $6480 \pm 1eV$ [1], and this 1 eV uncertainty directly contributes to the uncertainty on $\Delta E_{2p\rightarrow 1s}$.

The calibration of our detectors also relies on the experimental determination of the energies of this electronic X-rays from the elements in our calibration targets. These energies are known to within 0.03 eV [3], which has a negligible effect on our determination of the kaonic hydrogen parameters.

3.1.2 The linearity of the CCD detectors

From the standpoint of the physical processes leading to charge deposition in the CCD material, the response must be linear. The response of the detection system as a whole, including amplifiers and ADCs, must be characterized experimentally.

There are eight CCD chips in the cryogenic target system, and each chip is divided into two separate areas, which are driven and read out by independent detector electronics. This means that there are effectively sixteen independent detectors, all of which must be independently calibrated and characterized.

The chips and their corresponding readout electronics packages (so-called Readout Boxes) will be characterized by long exposures to multi-material targets which are in turn activated by a strong radioactive source. At the time of this writing, we have made preliminary measurements of this type. The prototype target consisted of Titanium (K_α = 4.51 keV), Copper (K_α = 8.05 keV), and Zirconium (K_α = 15.78 keV). Exposures on the order of 20 hours yielded approximately 5000 counts per peak per detector ($\frac{1}{2}$ CCD). The resulting spectra were fit by multi-Guassian peaks to yield the mean channel of each calibration peak. We then fit the channel versus energy data with a quadratic function, and found that the fit quadratic term was consistent with zero (indicating good linearity). As an estimate of the systematic uncertainty, we took the error on the quadratic term, which suggested an error on the order of 2 eV near the kaonic hydrogen K_α peak.

Before the DEAR system is installed on $DA\Phi NE$, we intend to carry out a more precise analysis, including more target materials and higher statistics. Therefore this estimated 2 eV systematic error is considered an upper limit only.

3.1.3 The energy calibration of the CCD detectors

Once the linearity test described above has been accomplished, one might be tempted to use the results of the linearity fit to determine the energy of the channel of the final fit K_α peak. This would be ill-advised, because the energy scale of a given detector could change over long periods of time, due to temperature, radiation, or other unknown effects.

We will implement our so-called *long-term calibration scheme* to partially eliminated this problem. We intend to line the walls of the cryogenic target with Titanium and Zirconium foils. Electronic X rays will be excited in these foils by kaons, pions, and other high-energy beam-related backgrounds, and these X rays will be measured simultaneously with the kaonic hydrogen X rays. Thus the final fit of the kaonic hydrogen spectra will automatically include the precisely-known peaks from Titanium and Zirconium. Random shifts in the detector gain will be convoluted with the natural Gaussian width of the detector response function, and will be measured and included in the fit of these long-term calibration peaks, which will then be used to determine the energy of the mean channel of the kaonic hydrogen K_α peak.

However, this is not a complete solution. It could happen, for example in the case of a steady degradation of the gain over the course of the experiment, that the additional width introduced by the gain shift is large compared to our detector resolution, leading to badly degraded statistical performance. Even more dangerous would be sudden jumps in the gain of a channel, which would lead to convoluted multiple peaks in the final spectra.

In order to guard against these potential problems, we also intend to implement the so-called *short-term calibration scheme*. This will involve one hour of data-taking per day, during which we do not try to measure kaonic X rays, but rather X rays excited in our Ti and Zr calibration foils by a strong radiation source inserted between the beam pipe and the target vacuum entrance window. Test measurements indicate that we could induce 600 counts/hour/$\frac{1}{2}$ CCD with a 4 MBq ^{241}Am source. Monte Carlo studies show that just 100 counts per peak is enough to track any shifts larger than 12 eV, which are negligible compared the expected 150 eV intrinsic energy resolution of the detectors. Using the short-term calibration information, data from each short-term calibration period can be scaled appropriately before being added to the final spectra.

By implementing the short-term and long-term calibration schemes as described, we will be able to compensate for any sudden systematic gain shifts, which could severely compromise our ability to fit the final kaonic hydrogen spectra. Instead, we will have a set of well-behaved spectra with the narrowest possible calibration peaks from Titanium and Zirconium built-in.

In this case, the systematic error in the energy determination of the kaonic hydrogen K_α peak will be determined by the precision to which the positions of the calibration peaks are known, which is in turn a question of statistics.

We have made two estimates of the expected number of excited Titanium X rays. One very rough estimate is based on the number of X rays seen in the KEK KpX experiment, which also used a Titanium calibration target. Under the assumption that the energy spectra of kaons entering the KpX target was similar to that which should enter the DEAR target, and correcting for differing geometric acceptances of the two experiments, we estimate a ratio of 1:1 for Titanium:Kaonic Hydrogen K_α X rays.

A more precise estimate has been made, based on the flux of high-energy background photons in the DEAR target projected by Monte Carlo. This study, combined with a simulation to determine the optimal arrangement of calibration foils within our target, suggests that we can expect 15000 Ti K_α X rays along with our 10000 kaonic hydrogen K_α X rays, and in addition 10000 Zr K_α X rays, in our final kaonic hydrogen spectrum. With 15000 Ti events and 10000 Zr events, we can determine the energy of the mean channel of the fit kaonic hydrogen K_α to within 1.0 eV. We also find, assuming that our projected ratios of calibration X rays to kaonic hydrogen X rays are correct, that the systematic error associated with the CCD energy scale will always be smaller than the statistical error.

3.1.4 Uncertainty in the fitting procedure

We have done extensive Monte Carlo studies of the fitting procedure for the final kaonic hydrogen spectra. We have demonstrated that we can fit multiple kaonic hydrogen peaks and calibration peaks over a large but smooth background without introducing any systematic biases. However, in order to achieve these unbiased fits, we require three important pieces of information: the efficiency of the CCDs as a function of energy, the energy resolution of the CCDs as a function of energy, and the background shape.

The CCD efficiency has been successfully modelled and measured, including the effects of window absorption [4]. Our Monte Carlo studies indicate that our present level of uncertainty in the efficiency will not introduce any significant systematic bias in our fit of the final spectra.

As was previously mentioned, the CCD energy resolution as a function of energy is well-modelled by the Fano equation. The factor N must be measured for each detector, and in fact this data is obtained as a bonus when we make the linearity characterization for each detector.

By allowing N to vary as a free parameter in our Monte Carlo fits, we determined that the 5% uncertainty in the value of N has no effect on the fit error of the K_α peak position (as expected; this only depends on the symmetry of the fitting functions). On the other hand, no knowledge of the detector resolution increases the error on the K_α width from 5 to 10 eV. As the final energy resolution function will be known to better than 5%, the actual systematic uncertainty should be negligible.

One additional source of uncertainty related to the energy resolution is the fact that each fit peak is assumed to be a Gaussian (in the case of calibration peaks) or Voightian (a convolution of a Guassian with a Lorentzian, in the case of kaonic hydrogen peaks). These functions are only defined for for a constant Gaussian width (σ) over the range of the peaks, which is not the case in a real detector. However, our simulations have shown that making the approximation that the energy resolution is constant over the range of a given peak does not introduce any measurable systematic bias.

The final source of systematic error is our lack of knowledge concerning the background shape. For there is certain to be a flux of real X rays illuminating the CCD detectors with an energy distribution that can only be very roughly estimated. However, the rate of these X rays is estimated to be on the same order as that of the kaonic hydrogen X rays. This means that to measure the the background directly using an empty target could take an appreciable fraction of our running time. Or, in the unlikely case that the experimental setup is the same for the kaonic deuterium phase of the experiment, the background in the kaonic hydrogen energy range will be directly measured during this second phase. However, for an initial estimate, we are forced to consider a worst-case scenario.

For example, K_α X rays from Iron would be a particularly dangerous background peak, found at 6.4 keV, in a position that could be masked by the kaonic K_α peak - and indeed, sufficiently close to introduce a significant shift-bias with relatively few events. Of course the greatest possible care has been taken to exclude iron (and the equally dangerous manganese, $K_\alpha = 5.9$ keV) from the DEAR target environment. The target cell itself was machined from a special low manganese content ($\leq 0.15\%$) alloy of aluminum (AlMgSi 0.5). The same is true of the target vacuum chamber. All screws are either made of titanium or are titanium-coated. Thus to first order, there is no iron (or manganese, or chromium) within the solid-angle of the CCD detectors. However, with the very high flux of potential excitation background and the very long exposures of the CCDs (perhaps hundreds of days), one must consider the potential for such peaks from very small sources. In addition, the assumption of some unknown peak under the kaonic hydrogen K_α peak is the most pessimistic approach, since counts concentrated in a peak would bias the mean much more effectively than counts spread over some unusual (but smooth compared to a peak) background shape.

In order to quantify the effect of such unknown peaks, we have simulated a small (0.2% of the kaonic hydrogen peak) iron peak that is entirely uncompensated (i.e. not fit) by our Monte Carlo. Such a peak will bias our fits and lead to a systematic shift of $\Delta E_{2p\rightarrow 1s}$ on the order of 1 eV. However, a completely unknown peak (one not included in the background fitting function) is a worst-case scenario. A *known* peak under the kaonic hydrogen K_α peak can be effectively fit without introducing any systematic biases. However, the necessary inclusion of additional free parameters will increase the resulting statistical error of the fit.

Another source of uncertainty is the background shape itself. In order to estimate this, we generated a quadratic background with a very weak quadratic term, but fit only a linear background. The size of this background was chosen so that the mean of the resulting chi-squared distribution was about one sigma away from the mean of the nominal distribution (that in which a true linear background is fit with a first-order polynomial). This is the one sigma limit for a detectable effect, that is, with a chi-squared value in this range, we cannot distinguish between a poor background fit and a statistical fluctuation of the chi-squared value itself. The results of a series of 1000 fits, with 10000 generated K_αX rays and a quadratic background, are shown in figure 4. We see that, under these conditions, the systematic effect on the shift is small, less than 2 eV. However, the effect on the width is dramatic, on the order of 15 eV. Note also that this estimate is based on an assumed background of 60%; in the case of higher backgrounds, the effect will be even more significant.

Fortunately, both of these uncertanties will be addressed when we use a deuterium target. By using the same experimental setup, we will directly measure the background shape under the kaonic hydrogen K_α peak up to the kaonic deuterium K_α energy (around 7.8 keV). This measurement will allow us to constrain the background term during the fitting of the kaonic hydrogen spectrum. We therefore assign a maximum combined uncertainty of 1.5 eV to the background fitting procedure.

4 Conclusions

The statistical error on the determination of $\Delta E_{2p\to 1s}$ is strongly dependent on the machine background and the final luminosity. Assuming optimal conditions (at least 500 pb^{-1} and 30% background), we can hope to achieve a statistical precision of 4 eV.

The estimated contributions to systematic errors on $\Delta E_{2p\to 1s}$ are summarized in the following table:

Source	**Error (eV)**
$E^{EM\,only}_{2p\to 1s}$	1.0
Energy Calibration of CCDs	1.0
Non-linearity of CCDs	≤ 2.0
Background Uncertainty	≤ 1.5
TOTAL:	$\leq$ **3.0**

Thus, we estimate a systematic error on the shift of the K_α line in kaonic hydrogen on the order of 4 eV, that is, on the same level as the statistical error. The optimistic final combined error estimate is approximately **5 eV**, or a relative error of approximately 1-2%.

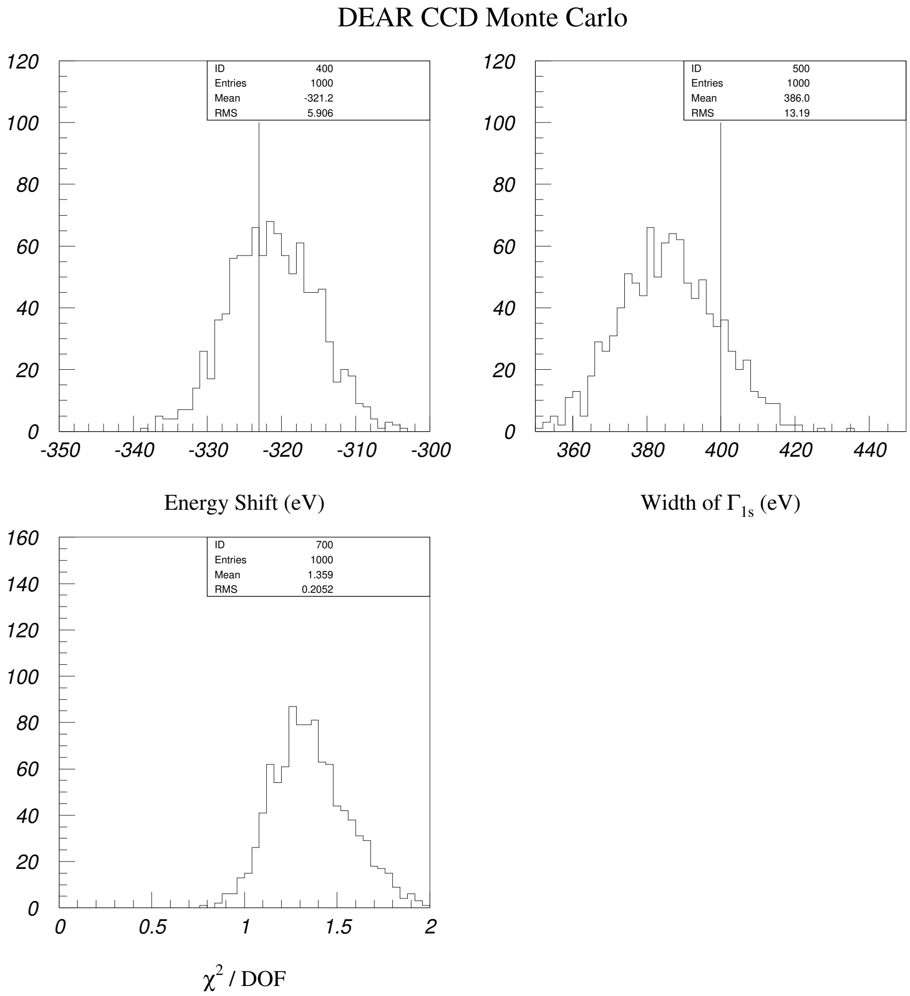

Figure 4: Results of a linear fit to a generated quadratic background, 10000 K_α events, 60% background.

References

[1] M. Iwasaki *et al.*, Phys. Rev. Lett. 78 (1997) 3067.

[2] T. Ito, *Observation of Kaonic Hydrogen Atom X Rays*, Ph.D. Thesis, University of Tokyo, UTPN-227, 1997.

[3] J.A. Bearden, Rev. Mod. Phys. 39 (1967) 78.
J.A. Bearden and A.F. Burr, Rev. Mod. Phys. 39 (1967) 125.

[4] B. Lauss, *Kaskade und Transfer von Myonen in Wasserstoff*, Ph.D. Thesis, University of Vienna.