Draft version 0.1

# ATLAS NOTE

ATL-PHYS-PUB-2008-000

July 29, 2008

## Tag and probe with background

David L. Adams

### Abstract

Tag and probe.

# 1 Introduction

Tag and probe is useful.

# 2 No background

We begin by assuming there is no background under our signal, i.e. all probes are correctly identified as true muons. The number of found probes, $F$, may be expressed in terms of the number of candidate probes $N$ and the efficiency $\varepsilon$:

$$F = \varepsilon N$$

and so the efficiency may be estimated from measurement

$$\hat{\varepsilon} = F/N \tag{1}$$

The error in the estimator is the usual binomial result which we derive by expressing the total number of probes in terms of two independent measurements: the number found ($F$) and the number lost ($L$):

$$N = F + L$$

The estimate for the efficiency is expressed in terms of these independent variables

$$\hat{\varepsilon} = F/(F+L)$$

and the error in this estimate is

$$\sigma(\hat{\varepsilon}) = \frac{\partial\hat{\varepsilon}}{\partial F}\sigma(F) \oplus \frac{\partial\hat{\varepsilon}}{\partial F}\sigma(L) = \frac{L\sigma(F)\oplus F\sigma(L)}{N^2} = \sqrt{\frac{\varepsilon(1-\varepsilon)}{N}} \tag{2}$$

where the symbol $\oplus$ indicates addition in quadrature and the square root is used as the estimate of the error in the count of found or lost probes, i.e. $\sigma(F) = \sqrt{F}$.

# 3 Background

Next we consider the case where our $N$ probes includes $N_\mu$ muons and $N_h$ candidates that are not muons, e.g. pions or kaons:

$$N = N_\mu + N_h$$

Again using $\varepsilon$ to denote the efficiency to identify a muon and introducing a misidentification probability $\gamma$ for the background, we obtain the number of found probes:

$$F = \varepsilon N_\mu + \gamma N_h$$

If we use our previous estimator for the efficiency, then we have a systematic error

$$\hat{\varepsilon} - \varepsilon = \frac{\varepsilon N_\mu + \gamma N_h}{N} = \frac{N_h}{N_\mu}(\varepsilon - \gamma) \tag{3}$$

We underestimate the efficiency (in the typical case where $\gamma$ is less than $\varepsilon$). If we know the impurity (background fraction), this expression can be used to obtain an unbiased estimate of the efficiency and the uncertainty arising from our uncertainty in the impurity.

The purity is typically enhanced by selecting probes for which the tag-probe mass is at a resonance, e.g. $Z$, $J/\psi$ or $\Upsilon$. In this case the purity may be estimated by fitting the peak and assuming a smooth continuum for the background. However there is still potential for systematic error arising from the fit procedure and the implicit assumption that continuum background does not include real muons.

The last danger can be circumvented by independently fitting both the candidate and found spectra (or the found and lost) to eliminate the continuum. This is discussed in the section ”Alternative approach” in reference [1]. Here we examine techniques to estimate or remove the background contamination without spectral fits.

# 4 Same-sign background estimate

So far we have ignored the relative charge signs of the the tag and probe. We can separately count same-sign (SS) and opposite-sign (OS) probes to obtain four measurements in terms of eight unknowns:

$$N^{OS} = N_\mu^{OS} + N_h^{OS}$$

$$N^{SS} = N_\mu^{SS} + N_h^{SS}$$

$$F^{OS} = F_h^{OS} + F_h^{OS}$$

$$F^{SS} = F_h^{SS} + F_h^{SS}$$

We eliminate one unknown by assumming the same efficiency for OS and SS probes:

$$F_\mu^{OS} = \varepsilon N_\mu^{OS}$$

$$F_\mu^{SS} = \varepsilon N_\mu^{SS}$$

and another by assuming the same misidentification probability for both:

$$F_h^{OS} = \gamma N_h^{OS}$$

$$F_h^{SS} = \gamma N_h^{SS}$$

Define the same-sign impurity in the signal by

$$\delta_\mu = \frac{N_\mu^{SS}}{N_\mu^{OS} + N_\mu^{SS}} = \frac{N_\mu^{SS}}{N_\mu} \tag{4}$$

and the background asymmetry

$$\alpha_h = \frac{N_h^{OS} - N_h^{SS}}{N_h^{OS} + N_h^{SS}} \tag{5}$$

Our critical assumption will be that this parameter is small.

Our four measurements can now be expressed in terms of six unknowns:

$$N^{OS} = (1-\delta_\mu)N_\mu + \frac{1}{2}(1+\alpha_h)N_h$$

$$N^{SS} = \delta_\mu N_\mu + \frac{1}{2}(1-\alpha_h)N_h$$

$$F^{OS} = (1-\delta_\mu)\varepsilon N_\mu + \frac{1}{2}(1+\alpha_h)\gamma N_h$$

$$F^{SS} = \delta_\mu \varepsilon N_\mu + \frac{1}{2}(1-\alpha_h)\gamma N_h$$

We combine these to obtain an estimator for the efficiency

$$\begin{aligned}\hat{\varepsilon} &= \frac{F^{OS} - F^{SS}}{N^{OS} - N^{SS}} \qquad (6)\\ &= \frac{(1-2\delta_\mu)\varepsilon N_\mu + \alpha_h \gamma N_h}{(1-2\delta_\mu) N_\mu + \alpha_h N_h}\\ &= \varepsilon - \frac{\varepsilon - \gamma}{(1-2\delta_\mu)} \frac{N_h}{N_\mu} \alpha_h + O(\alpha_h^2) \qquad (7)\end{aligned}$$

which is unbiased if we assume $\alpha_h = 0$.

Again we evaluate the uncertainty in the estimator by introducing lost counts, $L^{OS}$ and $L^{SS}$:

$$N_k^A = F_k^A + L^A \quad \{A = OS, SS\}$$

and use independent counting statistics for these and the found values to obtain the uncertainty in the estimator:

$$\begin{aligned}\sigma(\hat{\varepsilon}) &= \frac{\partial \hat{\varepsilon}}{\partial F^{OS}} \sigma(F^{OS}) \oplus \frac{\partial \hat{\varepsilon}}{\partial L^{OS}} \sigma(L^{OS}) \oplus \frac{\partial \hat{\varepsilon}}{\partial F^{SS}} \sigma(F^{SS}) \oplus \frac{\partial \hat{\varepsilon}}{\partial L^{SS}} \sigma(L^{SS})\\ &= \frac{(L^{OS} - L^{SS})\sqrt{F^{OS}} \oplus (F^{OS} - F^{SS})\sqrt{L^{OS}} \oplus (L^{OS} - L^{SS})\sqrt{F^{SS}} \oplus (F^{OS} - F^{SS})\sqrt{L^{SS}}}{(N^{OS} - N^{SS})^2}\\ &= \frac{\sqrt{\hat{\varepsilon}^2 (N^{OS} + N^{SS}) + (1 - 2\hat{\varepsilon})(F^{OS} + F^{SS})}}{(N^{OS} - N^{SS})} \qquad (8)\end{aligned}$$

Our assumption that the background asymmetry vanishes follows from two other assumptions about the background: the signs of the tag and probe are not correlated and tags are equally likely to be found with either charge. The former is not exact should be fairly good in the typical case at a collider such as the LHC where many charged particles are produced. Both assumptions should be tested both with Monte Carlo and collision data to obtain an estimate for the consequent systematic error and possibly to obtain a correction for the efficiency estimate.

The LHC is a proton-proton collider and and so we expect to find more positively charged muons and hence a nonzero background asymmetry. The next section introduces a formalism to handle this.

## 5 Summary and conclusion

## 6 Acknowledgements

## References

[1] ATLAS collaboration, ”In-situ Determination of the Performance of the ATLAS Muon Spectromter”