Alignment of the CMS Muon System with Tracks using the MuonHIP Algorithm

Jim Pivarski, Alexei Safonov

June 4, 2008

1 Introduction: Why align the muon system?

CMS has two main tracking systems: a silicon detector for all-purpose track measurement and a system of muon chambers behind layers of iron, for identifying and tracking muons. The silicon tracker, hereafter simply called the "tracker," is a cylindrical volume around the interaction point with a 1 m radius, capable of resolving track momenta with high precision up to several hundred GeV. The muon system is a full tracking system on its own, with 18–44 layers of measurements (depending on polar angle) to robustly identify muons. However, it also extends the tracker's momentum reach for muons beyond a TeV, because even the innermost measurements are 4 m from the interaction point, thereby quadrupling each track's lever arm.

The impact of the muon system on momentum resolution is most significant above a TeV (Figure) because of the relative precision of the tracker and muon sensors, and the difficulty of tracking particles through a scattering medium. Hits on silicon strips in the tracker have an intrinsic resolution of x–y μm, while drift tube hits in the muon barrel and cathode strip hits in the muon endcap have intrinsic resolutions of about 200 μm. Also, low-momentum muons scatter in the iron, making their direction between measurements uncertain. Only very energetic muons with nearly straight trajectories get an additional benefit from the lever arm the muon system provides.

However, high-energy muons like these are key to many clean signatures of physics beyond the Standard Model— in particular, a dimuon resonance with a mass above 900 GeV would be striking evidence of new physics, requiring only about x pb^{-1} of 14 TeV collisions [ref]. The discovery potential of such a resonance is directly proportional to the real resolution of track momentum, which is a combination of errors from the intrinsic resolution of hits and incorrect assumptions about the sensors' positions when reconstructing tracks. The latter error is known as misalignment, and the misalignment of each detector element must be minimized to ensure early discovery of high-energy resonances (Figure).

From the point of view of momentum resolution, some chambers are more important than others. The exaggerated track in Figure Q illustrates the effect of the non-uniform magnetic field observed by muons passing through the muon system— the field strength even changes sign, reversing the curvature of the track. This places a limit on the advantage of extra lever arm: beyond about 5 m, muon system hits no longer contribute to the shared curvature measurement made by the tracker and muon system. Therefore, it is essential to determine the alignment of the inner chambers well, particularly in the $r\phi$ direction (perpendicular to the beamline), which constrains the curvature of all muon tracks.

Therefore, a high-quality alignment of the muon system would require the innermost muon chambers to be better aligned in $r\phi$ than their intrinsic resolution of 200 μm, relative to the tracker. The other chambers and alignment parameters should be better aligned than a few times this, about a millimeter. Rotation angles must also be aligned to such a degree that they don't displace hits on the far ends of the chambers, with the same resolution goals. This document describes a procedure to align the whole muon system using tracks measured by the chambers themselves, reaching the desired alignment goals with N pb^{-1} of collisions data.

1.1 Geometry of the muon system

The muon system has two main parts, a barrel circling the interaction point with drift tubes (DT) and an endcap of cathode strip chambers (CSC) on each side of the beamline (Figure). Resistive plate chambers (RPC) are interleaved between DT and CSC chambers, but these are low-resolution devices meant for extra redundancy in triggering and muon identification, not momentum resolution. DT chambers are mounted on wheels and CSCs on disks for modular assembly, and within these structures they are grouped into single-layered rings of chambers known as stations (labelled in Figure). Each station has a unique response to tracks due to its orientation and the quantity of material that the muons must penetrate to reach it, but within each station, the response should be identical because of the azimuthal symmetry of CMS. We therefore discuss alignment of chambers in each station individually.

Inside the chambers are 6, 8, or 12 sensitive layers, depending on the chamber type. CSCs have 6 layers of cathode strips and wires, each of which is sensitive to the two-dimensional position of passing muons. DT chambers in the inner three barrel stations (MB1–3) have 12 layers of one-dimensional drift tubes, the middle four at a 90-degree angle with respect to the top and bottom four, to provide measurements in both directions. Only the outermost barrel station (MB4) is a one-dimensional chamber, with 8 layers measuring $r\phi$.

Both chambers and layers are expected to be misaligned, though layer misalignments are smaller and do not change with time because the layers have been tightly fastened to the chamber frame. Layer misalignments have been under study since the chambers were assembled, and will be corrected before first collisions data arrive. We consider them outside the scope of this document.

Chambers, on the other hand, will physically move every time the CMS magnetic field increases or decreases, and do not return to the same positions for a given field strength. The most extreme case is ME1/1, the innermost endcap station, where chambers have been observed to move 1.4 cm when the CMS magnetic field bends the disk that supports them. Chambers are mounted on ball-and-socket joints, to ensure that they rotate rather than distort the shape of the layers inside, so we can treat them as rigid bodies, floating in 3 translational and 3 rotational degrees of freedom.

1.2 Methods of alignment

A system for measuring the geometry of the muon detector was integrated into its design, such that key positions are continuously observed by lasers and calipers. This system, known as the Muon Hardware Alignment System (MHAS), will determine the geometry of the muon chambers before any tracks are observed. Some degrees of freedom measured by this system are so-called "weak modes" of track-based alignment, and will therefore be dominated by hardware measurements until very large datasets become available. Also, only the hardware system can identify sudden changes in alignment due to slipping or settling into the magnetic field. Track-based alignment methods are limited to large, homogeneous datasets.

However, there are some advantages to track-based alignment, mainly that the tracks observe parameters which are directly relevant for reconstruction. For instance, tracks directly connect sensors in the tracker to the muon chambers, coaligning the active detector elements of the two systems, rather than their supporting structures. To achieve the same goal, the hardware measurements must be propagated through a series of mechanical devices, correcting for offsets from the active elements. Though track-based alignment is not equally sensitive to all degrees of freedom, it is most sensitive to degrees of freedom perpendicular to the tracks, which are exactly the ones that must be best aligned when reconstructing tracks.

Tracks in the muon system are stochastically deflected due to multiple scattering of the muons in material, particularly the iron yoke between stations. This can lead to deflections on the order of centimeters for individual tracks, particularly in the outermost stations and particularly for low-momentum muons. However, their distributions are symmetrically broadened without skewing the central value, reducing the statistical power of track-based alignment but not biasing it. More serious are systematic distortions, such as a mistaken material budget or incorrect map of the magnetic field, which we study in detail.

2 Overview of the MuonHIP algorithm

Alignment using tracks is usually complicated by the fact that it is coupled with track-fitting: the optimal alignment is given by the mean distance between tracks and their associated hits, but the best-fit of the tracks to those hits depends on the assumed detector geometry. This circular dependence is a major issue when aligning a self-contained system like the silicon tracker, but the coupling can be broken when aligning a muon tracking system next to the high-precision silicon tracker. Assuming that the silicon tracker has been correctly aligned, we can simply extend tracker-fitted tracks into the muon system and align muon chambers to them. This would accomplish two goals at once: relative alignment of the muon chambers, particularly along the lines of sight of tracks, and alignment of the whole muon system relative to the tracker.

This procedure is a special case of one of the algorithms used to align the tracker, the Hits and Impact Points (HIP) algorithm, which aligns components so as to minimize the distance between detector hits and the intersection of a track on the detector surface, known as an impact point. To reduce sensitivity to the initial misalignment, tracks are refit with Alignment Parameter Errors (APE), which is an alignment uncertainty added in quadrature with the intrinsic uncertainty of each hit in the track-refit. Typically, the HIP algorithm is used iteratively, refitting tracks with smaller APEs and correcting the sensor positions at each step, slowly converging to a solution that optimizes track-fits and aligned positions simultaneously. In our case, we refit tracks with a very large APE in the muon system (10 m), no APE in the tracker, and align the muon chambers to these "tracker-fitted" tracks in one iteration. Thus, we can accomplish our decoupled alignment using the same software tools as the tracker alignment community, because our procedure is implemented as a special case.

To summarize, our variant of the HIP procedure (MuonHIP) is as follows:

1. Collect a set of tracks with their associated hits in the tracker and the muon system.

2. Set APEs in the tracker to 0 and APEs in the muon system to ∞ (10 m).

3. Refit the tracks with these artificial hit uncertainties; track parameters are dominated by tracker measurements, completely insensitive to muon chamber positions, yet muon hits are included on tracks and residuals may be calculated.

4. Align muon chambers to these tracks, shifting and rotating them to minimize the residual (track position minus hit position) of muon hits.

A MuonHIP alignment is completed in one pass over the dataset because the resulting geometry is only a function of the tracker geometry and the set of tracks, not the muon geometry. We check our procedure by running it a second time and verifying that the second-iteration corrections are nearly zero.

2.1 Coordinate transformations and corrections

The static position and orientation of each muon chamber can be viewed as an active transformation from a single global coordinate system to local chamber coordinates. Hits are reconstructed relative to a reference point in the chamber from charge and timing measurements on hits and wires, knowing nothing of how the whole chamber is oriented in space. Tracks are fitted in a global coordinate system defined by the tracker, and must be transformed to the chamber's local coordinates for comparison with the hit. If the chamber is misaligned, the assumed coordinate transformation will lead to biases in the residuals between track impact points and hits, for instance, a millimeter misalignment in the chamber's local x direction would cause the x residuals to be systematically offset by one millimeter.

To talk about these parameters, we'll need a coordinate system: define the x and y directions to be the layer's measurement plane, with x at the center of the chamber corresponding to $r\phi$, sensitive to track curvature. The y direction is parallel to the beamline for DT chambers in the barrel and is pointing radially away from the beamline at the center of CSC chambers in the endcaps. The z direction, completing a right-handed coordinate system, is always perpendicular to the layer's measurement plane, radial for DT chambers and parallel to the beamline for CSCs.

Rotations are described by three angles, ϕ_x, ϕ_y, and ϕ_z, which are rotations around the local x, y, and z axes, respectively. The order of operations and signs of the angles are defined as follows: the active rotation matrix from local to global coordinates is

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_x & \sin\phi_x \\ 0 & -\sin\phi_x & \cos\phi_x \end{pmatrix} \begin{pmatrix} \cos\phi_y & 0 & -\sin\phi_y \\ 0 & 1 & 0 \\ \sin\phi_y & 0 & \cos\phi_y \end{pmatrix} \begin{pmatrix} \cos\phi_z & \sin\phi_z & 0 \\ -\sin\phi_z & \cos\phi_z & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1}$$

All of these definitions are illustrated in Figure Q.

The relationships between these parameters and global coordinates (such as identifying x with $r\phi$) apply only to chambers starting in the design geometry— in general, they are approximate because the parameters describe a correction from one assumed geometry to another. Discussing the geometry in terms of corrections gives us the convenience of working with small numbers in which zero always means no correction. Sometimes, however, we'll need to know the absolute position in space. Defining the initial position and rotation as $\vec{P}_0$ and $\mathbf{R_0}$, respectively, and the alignment corrections as $\vec{P}_c$ and $\mathbf{R_c}$, the corrected position and rotation of a chamber is

$$\vec{P} = (\mathbf{R_0})^{-1} \cdot \vec{P}_c + \vec{P}_0 \quad \text{and} \quad \mathbf{R} = \mathbf{R_c} \cdot \mathbf{R_0}. \tag{2}$$

Rotations are always applied first, so that absolute chamber positions are specified in global coordinates.

2.2 How to align chambers to tracks

Though we have said that the objective of the HIP algorithm is to minimize residuals between tracks and hits, we have not yet explained how this can be done, given a collection of residuals on a chamber. If a chamber were misaligned purely in one direction, x for example, the x residual of all hits would be offset from zero by the same amount, so the alignment correction should be the weighted mean of the x residuals distribution, negated.

$$\text{Alignment correction} = -\frac{b}{A} \pm \frac{1}{A} \text{ with } A = \sum_{\text{all hits}} \frac{1}{{\sigma_r}^2} \text{ and } b = \sum_{\text{all hits}} \frac{r}{{\sigma_r}^2}, \tag{3}$$

where r is the residual (track minus hit) and σ_r is its uncertainty, for each hit. It is important to propagate hit uncertainties because they are very inhomogeneous in the muon system: intrinsic uncertainties can vary by factors of several hundred percent, depending on whether the hit passed through or between cathode strips or wire groups in the CSCs, and how close the track approached each DT wire. Performing a weighted mean, rather than a simple mean, sharpens alignment resolution by more than a factor of two.

All 6 rigid body degrees of freedom can be determined from two residuals measurements, $\vec{r} = r_x$ and r_y, by exploiting trends in the residuals distribution as a function of the intersection between the track and the layer plane. For instance, misalignments in ϕ_z introduce linear trends in $r_x(y)$ and $r_y(x)$, which are orthogonal to the constant-offset effects of simple translations in x and y. The alignment corrections, which we'll describe with a 6-component vector $\vec{q}$, can all be determined at once through the 6×2 Jacobian matrix $(\partial q/\partial r)$. The alignment corrections still look like a weighted mean:

$$\vec{q} = -\mathbf{A}^{-1} \cdot \vec{b} \ \pm \ \mathbf{A}^{-1} \tag{4}$$

where

$$\mathbf{A} = \sum_{\text{all hits}} \left(\frac{\partial q}{\partial r}\right) \left({\sigma_r}^2\right)^{-1} \left(\frac{\partial q}{\partial r}\right)^T \tag{5}$$

and

$$\vec{b} = \sum_{\text{all hits}} \left(\frac{\partial q}{\partial r}\right) \left({\sigma_r}^2\right)^{-1} \vec{r}. \tag{6}$$

If we wish to fix some degrees of freedom, for reasons described in the next subsection, we would simply lower the dimensionality of $\vec{q}$ (or, equivalently, set rows of $(\partial q/\partial r)$ to zero).

Since there is no circular dependence between track-fitting and alignment in the MuonHIP procedure, this one correction is nearly sufficient to optimize the placement of all muon chambers. The only remaining distortion due to the method derives from the fact that rotating a layer plane changes the point where a track intersects it in global 3-space. This means that there's a tiny curvature in the $\vec{r} \times \vec{t}$ manifold which we project onto parameter space ($\vec{t}$ are the track parameters), so multiplication by $(\partial q/\partial r)$ is just a linear approximation to the full transformation from $\vec{r} \times \vec{t}$ to $\vec{q}$. We can easily correct this by iterating a second time, recomputing the track impact points with the first-iteration geometry, obtaining a set of corrections $\vec{q'}$ on top of the first iteration $\vec{q}$. However, these second order corrections are negligible and errors from detector effects are more of a concern.

2.3 Degrees of freedom and weak modes

Though all 6 degrees of freedom can in principle be derived from a system of 2 residuals and 5 track parameters, some of these are statistically limited. The most obvious example of a limited degree of freedom, or weak mode, is alignment parallel to the line of sight of tracks: if all tracks that pass through a chamber are straight and parallel, nothing can be said about the position of that chamber along the line parallel to the tracks because motion along that line leaves the residuals and impact points invariant (Figure?). Though there is a variation in track angles across each chamber's surface, this variation is small for tracks that come from the interaction point and tends to be confused with ϕ_x and ϕ_z because it's strongly correlated with y (see next paragraph). Unless we exploit a sufficiently large sample of non-interaction tracks, we will need to fix at least one coordinate, meaning that we would obtain the value from hardware measurements and not update it in the track-based alignment process.

The line of sight of tracks from the interaction point is a linear combination of the chambers' local y and z directions (with a tiny x component if misaligned), though mostly in the z direction, especially in the central barrel wheel. To simplify interpretation of the results, we would lock the z position of chambers if we had to fix one translational degree of freedom, allowing them to float in x and y. If the initial z positions are incorrect, chambers would align to y positions that preserve the straightness of straight tracks, as illustrated in Figure Q. These are not the true y positions, but they minimize error in momentum reconstruction, especially for the very straight tracks that motivate muon alignment.

The ϕ_x and ϕ_y angles are also weakly determined because they are second-order in residuals trends. Each rotation angle introduces an approximately-linear dependence in residuals across the layer surface because it rotates impact points around an arc whose length is proportional to the distance from the rotation center (see Figure Q). In the case of ϕ_z, rotation in the plane of the layer, the slope of $r_x(y)$ and $r_y(x)$ near the center is proportional to $\sin\phi_z$ ($\partial r_x/\partial y|_{y=0} \propto \partial r_y/\partial x|_{x=0} \propto \sin\phi_z$, Figure Q-c), so knowledge of this slope gives us first-order knowledge of ϕ_z. The first-order effects of ϕ_x and ϕ_y rotations are out of the plane of the layer, and are therefore not measured by tracks perpendicular to the layer plane (a reasonable approximation for this argument). Therefore, our access to ϕ_x and ϕ_y is limited to $\partial r_y/\partial y|_{y=0} \propto 1 - \cos\phi_x$ (Figure Q-a) and $\partial r_x/\partial x|_{x=0} \propto 1 - \cos\phi_y$ (Figure Q-b), which are second-order.

We should also consider that the muon chambers were designed to measure the p_T and ϕ of tracks with higher precision than η, because p_T is essential for setting the momentum scale of energetic muons. Therefore, local x is always the best measurement, especially in the CSCs where x is determined with high-precision cathode strips and y with coarse groups of ganged wires. Thus, the x positions will be better determined than y, and the ϕ_y angles will be better determined than ϕ_x.

3 Nominal results

Plots from baseline studies using 10 and 100 pb^{-1}, with and without some cosmic rays.

4 Error analysis and systematics studies

Statistical, systematic (tracker misalignments), both shrink with $\sqrt{N}$, assuming tracker alignment procedure goes well

Full explanation of the dependency of muon alignment on tracker alignment: it's a problem with curvature measurement, not pointing

incorrect magnetic field map, material budget, stuff like that (maybe miscalibration, though I think that's a small effect)

5 Identifying alignment quality with data

6 Concluding remarks