4

Vertex Detector

4.1 Vertex Detector Requirements

The major motivation for building BaBar is to find and precisely measure *CP* violation in the decays of neutral B mesons. To this end, one needs to determine for each event the time interval between the two B-meson decays. This is accomplished by reconstructing the two primary decay vertices, which is the main task of the vertex detector. For particles with low transverse momenta ($p_t < 100\,\mathrm{MeV}/c$), which may not be reconstructed in the drift chamber, the vertex detector should also provide full tracking information. Due to this capability to perform stand-alone tracking, the BaBar vertex detector is known as the silicon vertex tracker (SVT).

Below, we will discuss the requirements for the SVT and give a concise overview, including performance studies. A detailed discussion of all design aspects completes this chapter.

4.1.1 Resolution

Without a measurement of the B decay vertex, no useful *CP* asymmetries can be extracted at the $\Upsilon(4S)$. Therefore, the most important function of the SVT is the determination of the B decay positions, especially along the beam direction. The measurement of *CP* asymmetries does not place a very stringent requirement on the intrinsic position resolution; there is less than a 10% loss in precision in the asymmetry measurement due to imperfect vertex resolution if the separation between the B vertices Δz is measured with a resolution of approximately one half the mean separation, which is $250\,\mu$m at PEP-II [Sny94]. This translates into a single vertex resolution of better than about $80\,\mu$m both for *CP* eigenstates and for tagging final states. Vertex resolution in this range is readily achieved with silicon strip detectors.

We can benefit from additional precision over that required for the measurement of Δz. Pattern recognition, vertex reconstruction, and, in particular, background rejection will all be improved by better intrinsic position resolution, which in turn will improve the efficiency and purity of the samples used for the *CP* asymmetry measurements. Therefore, our aim is to achieve the best precision practical. Constraints which limit the achievable resolution include the total number of readout electronics channels, mechanical constraints, and considerations of cost and schedule. The multiple scattering in the beam pipe and in the silicon itself sets a lower limit on the useful intrinsic resolution, corresponding to a point resolution of about 10–15 μm for measurements made close to the IP and 30–40 μm for the outer layers [For94].

The vertex detector also dominates in the determination of the track angles. The angle measurements should be good enough that they do not significantly contribute to the experimental uncertainty in the measurements of track impact parameters or the invariant mass of combinations of tracks. This requirement does not impose a more stringent limit on point resolution than that determined through consideration of the issues discussed above. The tracking resolution needed for matching tracks to the particle identification system and the calorimeter is mostly provided by the drift chamber.

4.1.2 Acceptance

The coverage of the SVT must be as complete as technically feasible, given the location of the B1 magnets below 17.2° (300 mr) with respect to the beam line in both the forward and backward directions. To maximize the space available for the SVT and its readout electronics in the boost direction, machine components such as cooling manifolds and vacuum flanges are located in the backward region. Therefore, it is only possible to extend the backward coverage to within 30° of the beam line. In the $\Upsilon(4S)$ center-of-mass, this corresponds to $-0.95 < \cos\theta_{cm} < 0.87$.

There should be as little material as possible within the active tracking volume. Special attention must be paid to minimizing the amount of material between the IP and the first measurement in order to reduce multiple scattering. The beam pipe itself contributes $0.006X_0$ at normal incidence. Material located beyond the inner layers does not significantly degrade the measurement of track impact parameters, but does affect the performance of the overall tracking system and leads to increased photon conversions in the active region.

4.1.3 Efficiency

Our goal is to achieve close-to-perfect track reconstruction efficiency within the active volume of the tracking detectors when information from both the drift chamber and the SVT is used. The pattern recognition capabilities of the combined tracking system must be robust enough to tolerate background levels up to 10 times nominal, where nominal is defined for a 1 nTorr vacuum pressure in the beam pipe within ±30 m of the IP. Low momentum particles that do not traverse many drift chamber planes, such as many of the charged pions from D^* decays, must be reconstructed in the SVT alone. For this category of tracks, with $p_{\rm t}$ less than 100 MeV/c, we want to achieve reconstruction efficiencies of at least 80–90%. The SVT must also be efficient for particles such as K^0_Ss that decay within the active volume. Together, these determine the number of measurements along a track and the necessary single-hit efficiency.

The impact parameter resolutions are determined by the precision of the measurement closest to the IP. The performance of the inner vertex layers must therefore be optimized for both high efficiency and good point resolution. It follows that redundancy for the first measurement is an important design requirement.

4.1.4 Radiation Tolerance

The expected machine-related backgrounds set the requirements for the data transmission bandwidth and radiation resistance of all components located close to the interaction region [Lev94]. As described in Chapter 12, the expected dose in the innermost layer of the SVT averages about 33 krad/yr. The radiation is highly nonuniform in azimuth, peaking in the bend plane of the accelerator with a local maximum of up to 240 krad/yr over a small region covering approximately 6° in azimuth. At the location of the front-end electronics for Layer 1 the maximum dose is 110 krad/yr, with an average value of 47 krad/yr. Detector and front-end electronics are specified to be able to withstand at least 10 times the annual radiation dose. The readout electronics must be fabricated with radiation-hard technologies, and special attention must be paid to the sensitivity of the detector performance to radiation.

4.1.5 Reliability

The SVT is mounted inside a support tube of radius 20 cm, which also supports and aligns the machine elements closest to the IP. Access to the SVT is not possible without a major shutdown involving removal of the support tube from the detector. The reliability requirements for the SVT are therefore more stringent than usual for such a device, with implications for engineering design at all levels. The detector layout must provide redundant measurements wherever possible; the electronic readout must be very robust and have redundancy built in for critical data and control lines; and the functionality of all components must not be compromised by exposure to the expected radiation levels. The detector monitoring and interlock system must serve as a safeguard against catastrophic failure in the event of a component malfunction or a simple human error.

4.2 Vertex Detector Overview

4.2.1 Choice of Technology

The SVT design is based on double-sided silicon microstrip detectors. The characteristics of this technology that make it attractive for the *BaBar* detector are: high precision for measuring the location of charged particles, tolerance to high background levels, and reduction in mass made possible through double-sided readout. The process for the fabrication of double-sided silicon detectors is now mature enough to be employed in a large-scale application and to meet the performance standards outlined above.

4.2.2 Detector Layout

The SVT will provide five measurements, in each of two orthogonal directions, of the positions of all charged particles with polar angles in the region $17.2^\circ < \theta < 150^\circ$. A three-dimensional cut-away view of the SVT is shown in Figure 4-1. Each of the three inner layers has six detector modules, arrayed azimuthally around the beam pipe, while the outer two layers consist of 16 and 18 detector modules, respectively. A side view of the detector is shown in Figure 4-2, and an end view is shown in Figure 4-3.

The inner detector modules are traditional barrel-style structures, while the outer detector modules employ a novel arch structure in which the detectors are electrically connected across an angle. The bends in the arch modules minimize the area of silicon required to cover the solid angle and also avoid very large track incident angles.

In order to satisfy the requirement of minimizing material in the detector acceptance region, one of the main features of the SVT design is the mounting of the readout electronics entirely outside the active detector volume. There is a 1 cm space between the 300 mr stay-clear and the B1 magnet in the forward direction; all of the forward electronics are mounted here. In the backward direction, there is space below about 500 mr. In both directions, space is very tight, and the electronic and mechanical designs are closely coupled in the narrow region available.

The layout specifications for this five-layer design are given in Table 4-1. The strips on the two sides of the rectangular detectors in the barrel regions are oriented parallel (ϕ strips) or perpendicular (z strips) to the beam line. In the forward and backward regions of the two outer layers, the angle between the strips on the two sides of the trapezoidal detectors is approximately 90°, and the ϕ strips are tapered. Floating strips are used to improve the position resolution for near-perpendicular angles of incidence; the capacitive coupling

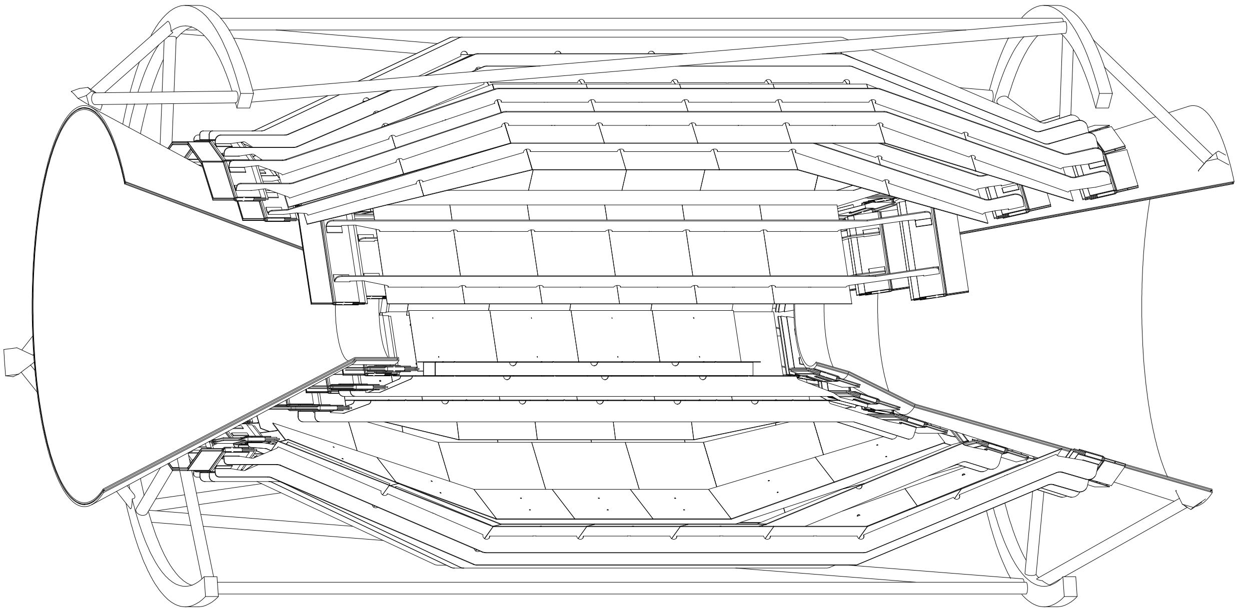

Figure 4-1. *Three-dimensional cutaway view of the SVT.*

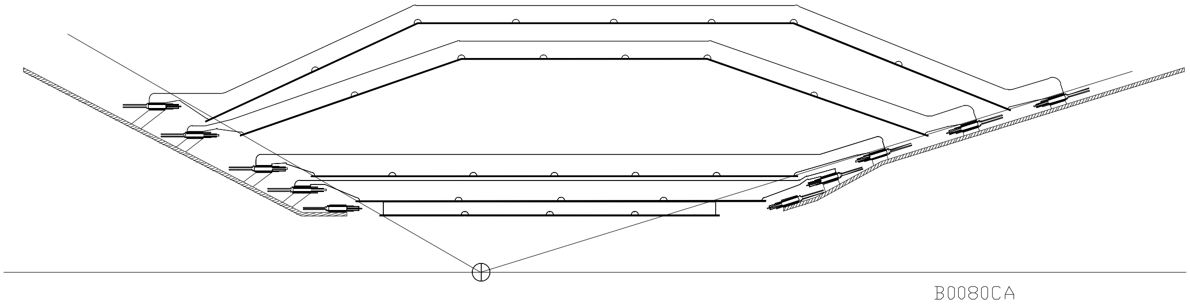

Figure 4-2. *Cross-sectional view of the SVT in a plane containing the beam axis.*

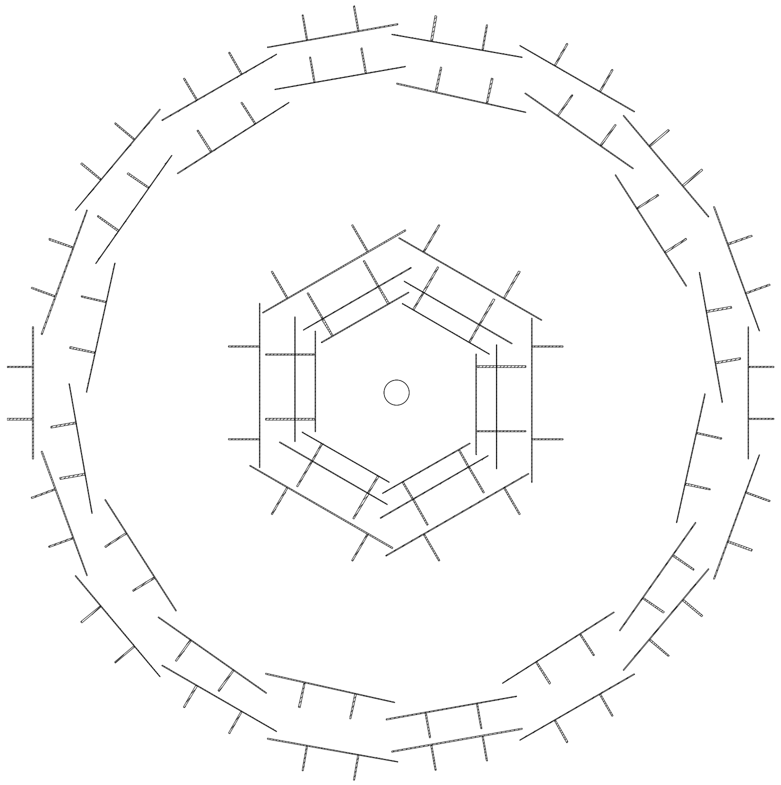

Figure 4-3. *Cross-sectional view of the SVT in a plane perpendicular to the beam axis. The lines perpendicular to the detectors represent structural beams.*

between the floating strip and the neighboring strips results in increased charge sharing and better interpolation. For larger incident angles, the wider readout pitch minimizes the degradation in resolution that occurs because of the limited track path length associated with each strip. These issues are discussed in more detail in Section 4.3.

The design has a total of 340 silicon detectors of seven different types. The total silicon area in the SVT is $0.94\,\mathrm{m}^2$, and the number of readout channels is $\sim$150,000.

4.2.3 Electronic Readout

As emphasized above, all readout electronics are located outside the active volume, below 300 mr in the forward direction and below about 500 mr in the backward region. To accomplish this, ϕ strips on the forward or backward half of a detector module are electrically connected with wire bonds. This results in total strip lengths associated with a single readout channel of up to $\sim$14 cm in the inner three layers and up to $\sim$24 cm in the outer two layers.

Quantity	Layer 1	Layer 2	Layer 3	Layer 4a	Layer 4b	Layer 5a	Layer 5b
Radius (mm)	32	40	54	120	127	140	144
Wafers/Module	4	4	6	7	7	8	8
Modules/Layer	6	6	6	8	8	9	9
Silicon Area (cm^2)	457	683	1072	1506	1582	2039	2082
Overlap in ϕ (%)	2.4	1.8	1.8	4.0	4.0	2.0	2.0
Readout pitch (μm):							
ϕ	50	50	50	65–100		65–100	
z	100	100	100	200		200	
Floating Strips:							
ϕ	—	—	—	1		1	
z	1	1	1	1		1	
Intrinsic Resolution(μm):							
ϕ	10	10	10	10–12		10–12	
z	12	12	12	25		25	
R.O. Section/Module	4	4	4	4		4	
ICs/R.O. Section	6	8	10	4		4	
Readout Channels	18432	24576	30720	32768		36864	
Strip Length (mm):							
ϕ	95	115	136	177/223	186/232	232/241	241
z	39	48	64	35–52	35–52	35–52	35–52
z Ganging:							
Forward $\times 2$	34%	18%	7%	82%	88%	71%	60%
Forward $\times 3$						29%	40%
Backward $\times 2$	34%	18%	7%	82%	71%	60%	60%
Backward $\times 3$				18%	29%	40%	40%

Table 4-1. *Parameters of the SVT layout. See text for more detail on the meanings of the different quantities. The intrinsic resolution is calculated at 90° track incidence assuming $S/N = 20:1$. The z-ganging numbers represent the percentage of detector strips connected to one other strip ($\times 2$) or two other strips ($\times 3$).*

The signals from the z strips are brought to the readout electronics using fanout circuits consisting of conductive traces on a thin flexible insulator (for example, copper traces on Kapton). The traces are wire-bonded to the ends of the z strips. To read out all the z strips with the same number of electronics channels as the ϕ strips, some z strips will be electrically connected, or ganged, together. The length of the z strips is much shorter, typically 5 cm in the inner layers and either 10 or 15 cm in the outer layers where there is either $\times 2$ or $\times 3$ ganging.

Front-end signal processing is performed by ICs mounted on hybrid circuits that distribute power and signals, and thermally interface the ICs to the cooling system. The signals from the readout strips, after amplification and shaping, are compared to a preset threshold. The time interval during which they exceed the threshold (time over threshold, or TOT) is an analog variable related to the charge induced on the strip. This time interval is digitally recorded prior to readout. Unlike the ordinary peak-amplitude measurement at the shaper output, the TOT technique has a nonlinear input-to-output relationship which is approximately logarithmic. This is an advantage since it compresses the dynamic range and allows one to achieve good position resolution and large dynamic range with a minimum number of bits.

TOT readout is less complicated than a linear analog readout system, resulting in less development time, lower cost, and smaller space and power requirements. At the same time, it is a significant improvement over a purely digital (hit/no-hit) readout, and provides sufficient analog resolution for position interpolation, time-walk correction, and background rejection. The readout IC is expected to be about 8–9 mm long and to dissipate no more than 2.0 mW per channel. The total power that will be generated by the SVT readout is $\sim$300 W.

For each channel with a signal above threshold, the strip number within the readout module, the time of arrival of the signal, and the digital value of the TOT will be read out. There are four readout sections per detector module, where the module is divided in half along z, and the ϕ and z strips are grouped together separately. The data from one-half of a detector module will be transmitted from the hybrid on a flexible cable to a transition card located approximately 40 cm away, where the signals are converted to transmission via conventional cables.

4.2.4 Mechanical Support

The silicon detectors and the associated readout electronics are assembled into mechanical units called detector modules. Each module contains several silicon detectors glued to low-mass beams constructed of carbon and Kevlar fiber-epoxy laminates. The beams are attached to the hybrid electronic circuits at each end. A ceramic or aluminum substrate for the hybrid provides precise mechanical mounting surfaces and is the heat sink for the electronics.

The inner layer is supported from the second layer; the detector modules from Layers 1 and 2 are glued together with rigid beams, forming sextants which are then mounted from the support cones in the forward and rear directions. Each detector module of the third, fourth, and fifth layer is mounted on the support cones independently of the other modules. In the fourth and fifth layers, there are two different types of modules in each layer, an inner one, labeled a, and an outer one, labeled b, occupying slightly different radial positions. Thus there are seven different types of detector modules.

The support cones are double-layered carbon-fiber structures which are mounted from the B1 magnets. Cooling water flows between the two carbon-fiber layers around aluminum mounts which protrude through the outer surface of the cone. Mounting pins in the hybrid structure provide the alignment between the modules and the aluminum mounts in the cone, and thermal contact is made to provide cooling for the front-end electronics located on the hybrid. The support cones are divided to allow the vertex detector to be assembled in two halves and then mounted on the B1 magnets by clam-shelling the pieces together.

The stiffness of the overall structure is provided by a very low mass space frame, constructed of carbon-fiber tubes, connecting the forward and backward support cones. It consists of rings at each end held rigid by 12 struts spanning the length of the detector. The rings are connected to the support cones by an additional series of 12 struts at each end. All material is carbon-fiber laminate. Preliminary finite element calculations show that this structure meets the tolerances for rigidity. The motivation for this space frame stems mainly from the possible relative motion of the two B1 magnets during the assembly procedure. Cooling water, power, and signal lines are routed along the B1 magnets to points outside the active region where manifolds for the cooling water and drivers for the electronics are located.

4.3 Detector Performance Studies

Cost, engineering complexity, space, and cooling requirements each constrain performance. In this section, we discuss how the SVT performance has been optimized within these constraints. Design and performance studies have considered the following:

- the impact of the number and locations of detector layers on track parameter resolution and pattern recognition;
- the effect of different readout schemes for recording charge information on intrinsic resolution, background rejection, and ultimately, on pattern recognition; and
- optimization of the readout pitch, taking into account the required resolution, the number of readout channels, and, in the case of the outer layers, the signal loss that occurs when the readout pitch becomes very large.

4.3.1 Resolution

Intrinsic Resolution

We have studied the effects on intrinsic position resolution of the strip pitch, readout pitch, and threshold levels, as well as various schemes for recording pulse information. The deposition of charge in a silicon strip detector was simulated by a Monte Carlo program, taking into account the effects of Landau fluctuations in energy loss, diffusion, channel-to-channel gain variations (5%), and noise. The simulation is in good agreement with experimental data on the energy loss in silicon, and with experimental measurements of position resolution in strip detectors.

The intrinsic position resolution was studied as a function of the track incident angle in the plane normal to the strips. Charged particles most often cross the detectors at close to normal incidence in the plane perpendicular to the ϕ strips; however, the angle of incidence of tracks in the plane perpendicular to the z strips extends to 300 mr from the beam line, or $\lambda = \pi/2 - 0.3 \approx 1.3$ rad, where λ is the dip angle measured relative to normal incidence. The resolution degrades significantly at large dip angles, especially for small readout pitch. This is due to inefficiency in the readout for strips with small signals deposited by tracks at grazing angles.

Figure 4-4(a) shows the intrinsic resolution for 300 μm-thick silicon as a function of the incident angle λ for four different strip and readout pitches. Strip and readout pitch coincide in the absence of floating strips; floating strips, when present, increase capacitive charge-sharing and thereby improve the point resolution. The simulation assumed a noise level of 1200 e^-, corresponding to a signal-to-noise ratio of about 20 for perpendicular tracks, and the threshold was set to 4 times the noise. The algorithm employed to determine the position uses the digital centroid for the central strips and applies a correction based on the relative amount of charge in the two edge strips of a cluster:

$$X = X_0 + 0.5 \cdot p \cdot \frac{Q_1 - Q_N}{Q_1 + Q_N},$$

where X_0 is the digital centroid (*i.e.*, geometrical center) of strips 2 through $N-1$; $Q_1(Q_N)$ is the charge deposited on the first(last) hit strip, and p is the readout pitch.

As seen in Figure 4-4(a), the resolution degrades significantly as the angle of incidence increases for 50 μm strip and readout pitch. With 100 μm strip and readout pitch, the resolution does not degrade as much with larger incident angles but, as expected, the resolution at normal incidence is about a factor of 2 worse than that for 50 μm pitch. On the other hand, 100 μm readout pitch with one floating strip gives a resolution equivalent to 50 μm readout pitch without floating strips for angles smaller than $45°(\lambda < 0.8)$. Above $45°$, it is equivalent to 100 μm strip pitch. Therefore, floating strips allow us to retain

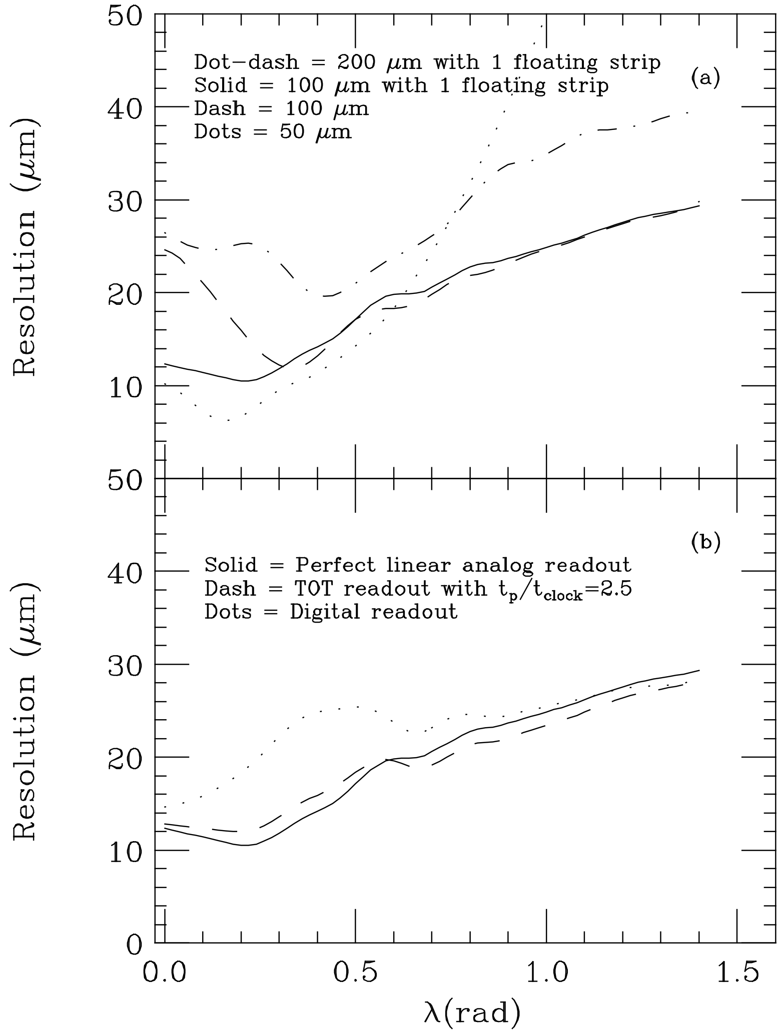

Figure 4-4. *(a) Intrinsic resolution for a 300 μm-thick silicon detector as a function of the dip angle for various strip configurations, assuming perfect analog readout. (b) Same as as (a), assuming 100 μm readout pitch and one floating strip, for various readout techniques.*

good resolution for small incident angles with fewer readout channels and less degradation in resolution for large incident angles. The upper curve in Figure 4-4(a) shows that for a 200 μm readout pitch with one floating strip, the resolution is about 25 μm at normal incidence. This configuration is used for the z strips in layers 4 and 5.

We have also studied the effect of the nonlinearity and quantization error of a TOT scheme compared both to an ideal analog readout and to a digital readout in which no information is recorded about the signal size [Roe94b]. The position is determined using the algorithm described above, but using the TOT, rather than the charge, to weight the edge strips.

The resolution is shown as a function of track incident angles in Figure 4-4(b). The signal shaping was assumed to have a peaking time 2.5 times longer than the period of the clock used in the digitization of the pulse length for this simulation. This value is consistent with that expected in the inner vertex layers, where fast shaping is necessary to cope with larger backgrounds. The resolution for a TOT readout scheme is practically the same as ideal analog readout. Digital readout, which provides hit/no-hit information only, shows much poorer resolution in comparison.

In addition to the advantage of improved resolution, analog readout may be useful for background rejection of non-MIP signals, for pattern recognition techniques which correlate pulse height with angle, and for correction of the time-walk which will affect the time-stamp information generated by the readout chip.

Track Parameter Resolution

The track parameter resolutions described in Chapter 2 were calculated using an analytical technique based on the method of Billoir [Bil84] which can be applied to arbitrary detector geometries in the program TRACKERR [Inn93]. Figure 4-5 shows the resolutions for the track impact parameter along the beam (σ_z) and in the plane perpendicular to the beam (σ_{xy}) for various momenta. The resolution is dominated by multiple scattering and not by the intrinsic position resolution of the silicon detectors, except for the highest momentum tracks from B decays.

A Monte Carlo analysis was performed to determine how the intrinsic resolution and the number of detector layers affect the track parameter resolution. The study included the effect of a catastrophic failure in Layer 1, the most important for vertex resolution. The intrinsic resolution was varied on the first two layers in the range 5–25 μm and on the outer two layers in the range 10–50 μm. The point resolution for Layer 3 was kept fixed at 10 μm. The quoted resolutions are for normally incident tracks; these were degraded as the dip angle λ increased, so that $\sigma = \sigma_0 \cdot (1 + 1.5\lambda)$.

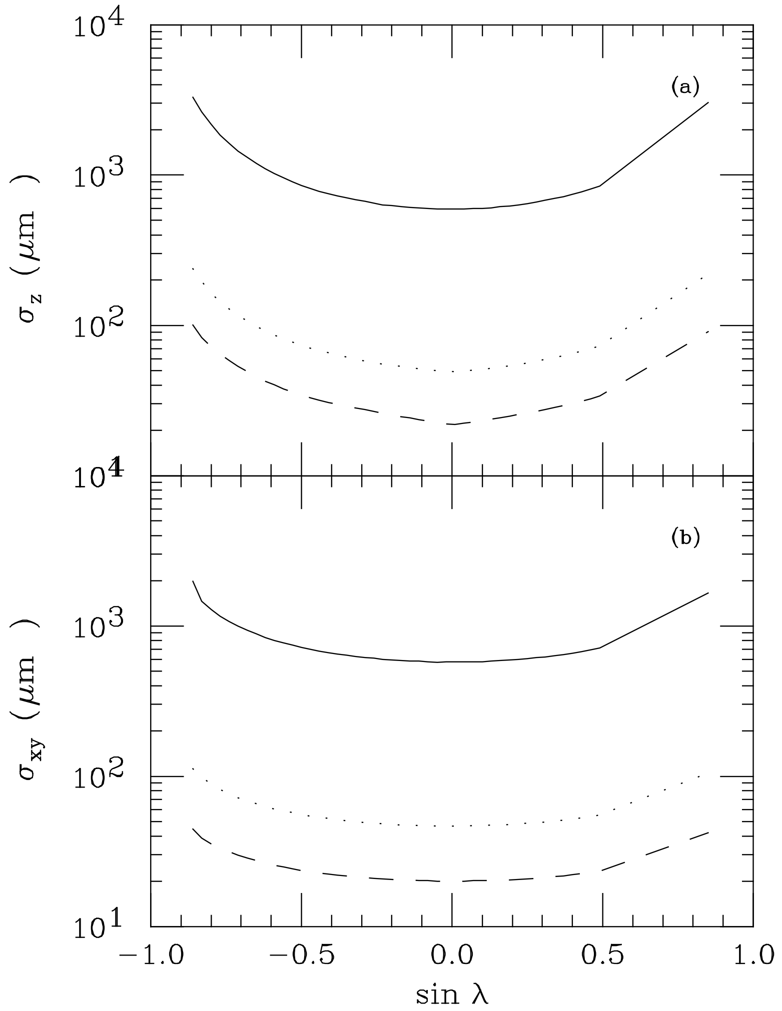

Figure 4-5. *(a) Impact parameter resolution along the beam (σ_z) and (b) in the plane perpendicular to the beam (σ_{xy}) vs. dip angle for three different momenta:* 100 MeV/c *(solid line),* 1.0 GeV/c *(dotted line), and* 3.0 GeV/c *(dashed line).*

In addition to the baseline geometry with five layers, we considered a four-layer layout which was obtained by removing the third layer. The basic conclusions of these studies are summarized as follows (see Reference [For94] for a full discussion):

- The angle and momentum resolution are essentially independent of the intrinsic resolution.
- The z impact parameter, on the other hand, does depend on the intrinsic resolution as shown in Figure 4-6(a) (the xy impact parameter has a similar behavior). In changing the intrinsic resolution on Layers 1 and 2 from 10 to 20 μm, respectively, σ_z degrades by about 15–20% for 1 GeV/c and by almost 40% for 3 GeV/c tracks. For soft tracks ($<$ 100 MeV/c), no dependence on the intrinsic resolution is observed because multiple scattering dominates. Most tracks produced in BaBar will be soft ($<$ 500 MeV/c); B vertices, however, will be determined by the hardest tracks in the event.

  The requirement that the track parameter resolution be dominated by multiple scattering for most tracks therefore leads us to choose the best intrinsic resolution technically feasible in Layers 1 and 2. Within the constraints imposed by the mechanical, electronic, and silicon detector performance, the best intrinsic resolution achievable in the inner layers is $\sim$10 μm in ϕ (employing 50 μm readout pitch) and $\sim$12 μm in z (employing 100 μm pitch with one floating strip). A larger pitch in z reduces the amount of ganging required and provides good resolution for tracks at small crossing angles.
- If for some reason Layer 1 is not usable, the tracking resolution is degraded, as shown in Figure 4-6(a). The degradation in the z-impact parameter resolution is 19% at 3 GeV/c and 22% at 1 GeV/c; thus, Layer 2 is an effective backup in case of a failure in Layer 1.
- The resolution on Layers 4 and 5 does not significantly affect any of the track parameters (Fig 4-6(b)), and the choice of the readout pitch in the outer layers is determined by other considerations (*i.e.,* noise, occupancy, and detector design considerations).
- Given the relatively soft momentum spectrum expected in BaBar, we were concerned that the material in the third layer (whose main motivation is pattern recognition) could degrade the tracking resolution. TRACKERR studies have shown [For94] that the effect of removing Layer 3 is in fact small (at most 5%) for high momentum tracks and negligible for soft tracks.

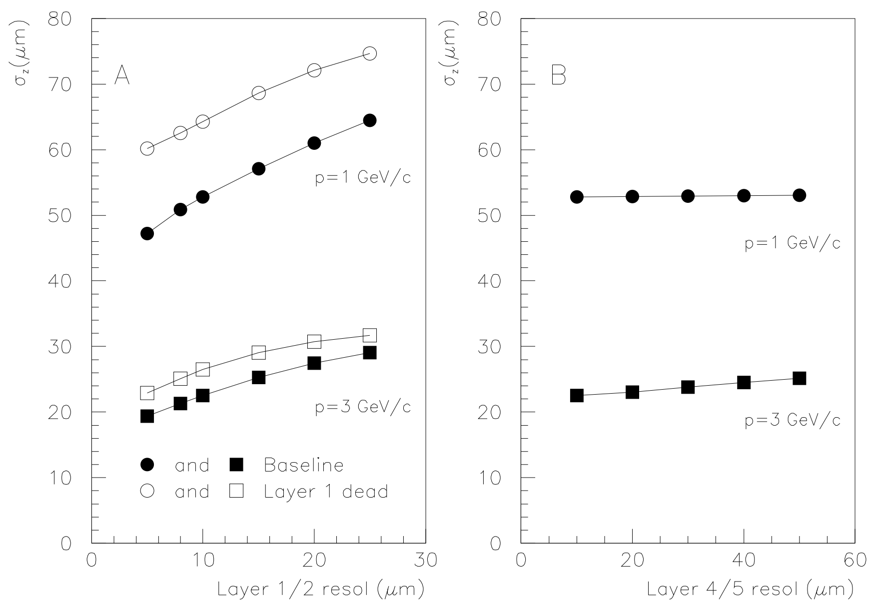

Figure 4-6. *z-impact parameter resolution as a function of the point resolution (a) in Layers 1 and 2 for the baseline design and for the case of a Layer 1 failure, and (b) in Layers 4 and 5. All values are at $\theta \approx 90\,\mathrm{deg}$ for 1 GeV/c and 3 GeV/c tracks.*

Vertex Resolution

The z vertex resolution has been calculated using a parameterized Monte Carlo simulation for three B decay modes: $B^0 \to \pi^+\pi^-$, $B^0 \to J/\psi\, K^0_S$, and $B^0 \to D^+D^-$. For each decay, the distribution of the difference between the reconstructed and generated B^0 vertex positions in the z direction is fit with a sum of two Gaussian functions. The results are shown in Table 4-2. The results for $B \to \pi^+\pi^-$ have been checked with a full GEANT-based Monte Carlo simulation, which gave reasonable agreement for both σ_n and σ_w. In all cases, the expected rms resolution (obtained by taking the weighted sum of σ_n and σ_w) is better than the $80\,\mu$m required for *CP* violation studies.

The z vertex resolution for the tagging B has also been studied with both parameterized and GEANT-based Monte Carlo simulations. The resolution depends upon which type of tag one employs. For direct leptons, we find $\sigma_n \approx 60\,\mu$m, while for kaon tags, $\sigma_n \approx 75\,\mu$m, with about 70–80% of the tags contained in the narrow Gaussian.

Good vertex resolution in the plane perpendicular to the beam pipe is also important. This information will help in associating tracks from B and D decays to the correct secondary or tertiary vertices, thus reducing combinatorial backgrounds. Figure 4-7(a) shows the distribution in the xy plane of the distance between two D vertices in the decay $B^0 \to D^+D^-$.

Mode	$\sigma_n(\mu\text{m})$	$\sigma_w(\mu\text{m})$	f_n (%)	rms
$\pi^+\pi^-$	27.5 ± 0.4	85.8 ± 3	84.0 ± 0.4	36.8 ± 0.6
ΨK_s	42.1 ± 0.7	168 ± 7	88.0 ± 0.2	57 ± 1.0
D^+D^-	46.5 ± 0.6	159 ± 10	94.0 ± 0.1	53.3 ± 0.8

Table 4-2. *A two-Gaussian fit to the z vertex resolution for various CP eigenstates; σ_n and σ_w are the rms of the narrow and wide Gaussians, respectively, and f_n is the fraction of the area contained in the narrow Gaussian. The rms is the weighted sum of σ_n and σ_w.*

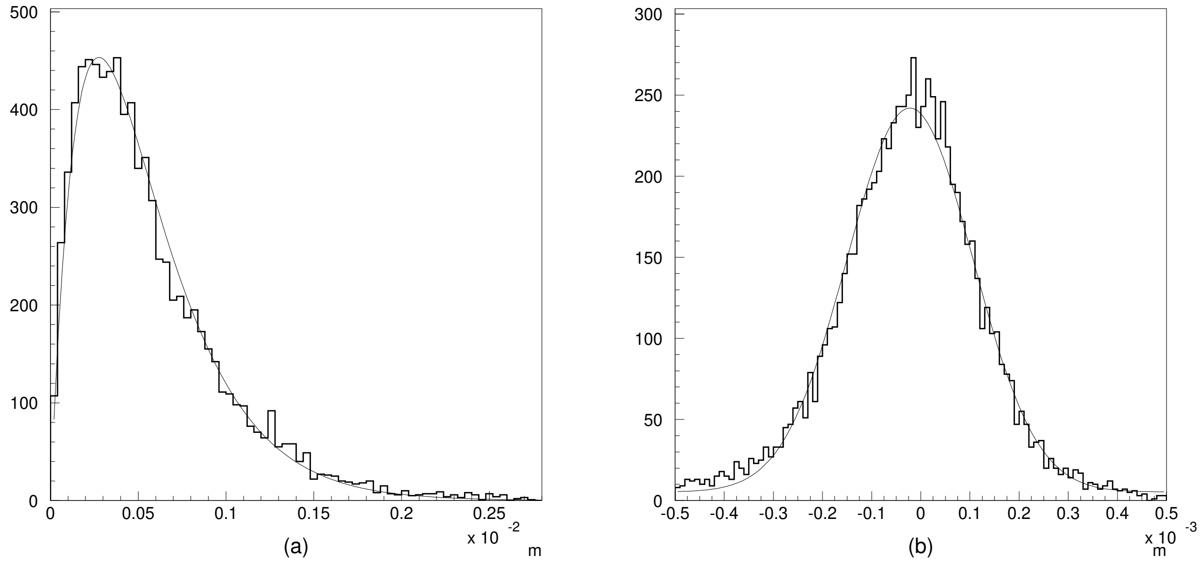

Figure 4-7. *(a) Separation between D vertices in the xy plane. (b) Resolution on the separation between the D vertices in the x-y plane. Note the different scales on the abscissa.*

The distribution is fit to a function of the form $A \cdot xe^{-x/\lambda}$ and yields a value of $\lambda = 276\,\mu$m. For comparison, the expected resolution on the separation between the two D decay points in the x-y plane is plotted in Figure 4-7(b) for the decay $D \to K\pi\pi$. The rms resolution is about $125\,\mu$m.

4.3.2 Pattern Recognition

The pattern recognition capabilities of the BABAR vertex detector have been studied using a GEANT simulation with a detailed model of the detector geometry. The simulation includes

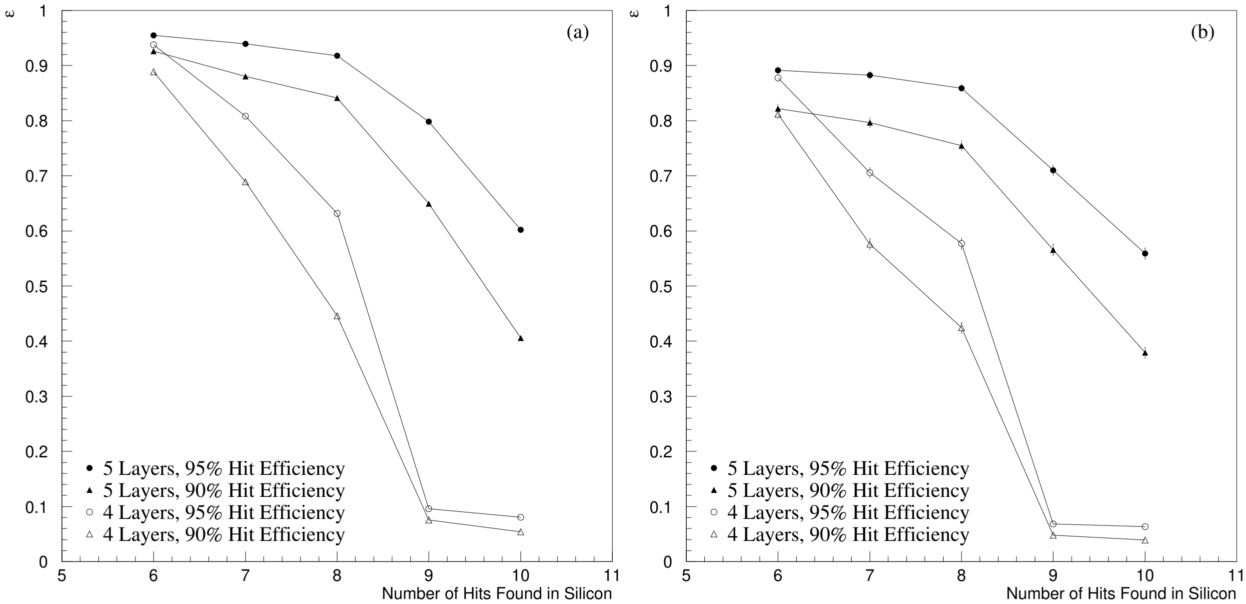

Figure 4-8. *Track finding efficiency for tracks from generic B decays: (a) all tracks, (b) tracks measured in the SVT only (p_{t} <∼ 90 MeV/c).*

detector overlaps, inactive regions near the edges of the wafers, strip inefficiency, background hits, and the effects of ganging the z strips. A method based on Kalman filtering techniques was used to locate tracks in the SVT. For tracks that went a sufficient distance into the drift chamber (corresponding to $p_{\mathrm{t}} > 90\,\mathrm{MeV}/c$), the results of a fit to the drift chamber track segment were used as a starting point for the silicon track finding. Hits in the SVT that were best associated with each of these track segments were added to the track. Once this was finished, these hits were removed from the SVT and a vertex-only pattern recognition algorithm was used on the remaining hits to find low-p_{t} tracks. This vertex-only algorithm was based on the same Kalman filtering technique but used combinations of hits in the outer layers of the SVT to generate the initial track parameters.

All tracks were required to cross five layers of silicon. The efficiency for finding tracks was computed by requiring a minimum number of silicon hits per track and by requiring each track fit to pass a fairly loose χ^2 cut ($\chi^2/N_{dof} < 10$). A single hit in the silicon was defined as either a ϕ or a z hit, so that a track going through five layers of silicon would generate ten hits. Six silicon hits was considered to be the minimum required in order for the number of degrees of freedom for a fit to the silicon-only tracks to be ≥ 1. The track finding efficiency for all tracks from generic B decays is plotted in Figure 4-8(a) as a function of the number of hits required. For this simulation, two geometries were studied: the baseline five-layer SVT, and a four-layer version in which Layer 3 was removed. The single hit efficiency was set to either 90% or 95%, and additional hits from beam-induced backgrounds at a level corresponding to 10 times nominal were overlayed on the hits from the $B^0\overline{B}^0$ event. The efficiency for the five-layer geometry remains high for up to eight required hits per track, while for the four-layer device, the efficiency is high only for the six-hit requirement.

In Figure 4-8(b), the efficiency is plotted for the same two geometries, but only for tracks which do not traverse enough layers in the drift chamber to be reliably found there. The overall efficiency is somewhat lower, and the same general behavior is observed, namely, a plateau for six, seven, or eight required hits in the five-layer detector and a sharply falling efficiency for more than six required hits in the four-layer version. The algorithm employed for finding tracks using the vertex detector only is quite sensitive to the detector efficiency, as seen in the figure. For comparison, in a perfect detector with 100% efficiency, the silicon-only track-finding efficiency would be 95% for the five-layer detector when requiring seven hits. Both the silicon-only and the combined silicon plus drift chamber track-finding efficiencies are quite insensitive to the background level.

The frequency at which fake tracks are found is also important to quantify. In this study, a fake track was defined as a track which had less than half of its hits from a single real track. A loose impact parameter cut was imposed, rejecting any tracks originating more than 5 cm from the nominal beamspot in z or more than 2 cm in xy. The average number of fake tracks per event is shown in Figure 4-9 as a function of the number of hits found in the silicon. This measurement was also performed assuming a background level that was 10 times nominal. In order to achieve reasonable efficiency for good tracks in the four-layer geometry, one cannot ask for more than six hits to be found per track. Therefore, the relevant rate in the four-layer geometry is 0.13 fake tracks per event. The five-layer geometry provides a relative reduction in the fake rate of a factor of $\sim$3 if one requires seven or eight hits on a track, without any significant reduction in track-finding efficiency. Approximately half of the fake tracks have $p_{\rm t}$ (as measured in the silicon) large enough to generate a corresponding track segment in the drift chamber. Most of these fakes should therefore disappear when attempts are made to match them to drift chamber hits.

4.3.3 Solid Angle Coverage

As discussed above, the SVT must have high efficiency over the region $17.2^\circ < \theta < 150^\circ$. The stay-clear at 17.2° (300 mr) in the forward region does not allow greater coverage, and beam-line components in the backward region prevent extension below $\sim$500 mr.

The geometry of the SVT in the inner layers is hexagonal, so the solid angle coverage varies as a function of azimuth [Roe94a]. To meet the solid angle coverage requirements, the first two layers are extended along $+z$ so that a track emitted at 300 mr from $z = 0$, at any azimuthal angle, is in the SVT acceptance. There is then some coverage down to 245 mr for tracks in certain azimuthal regions.

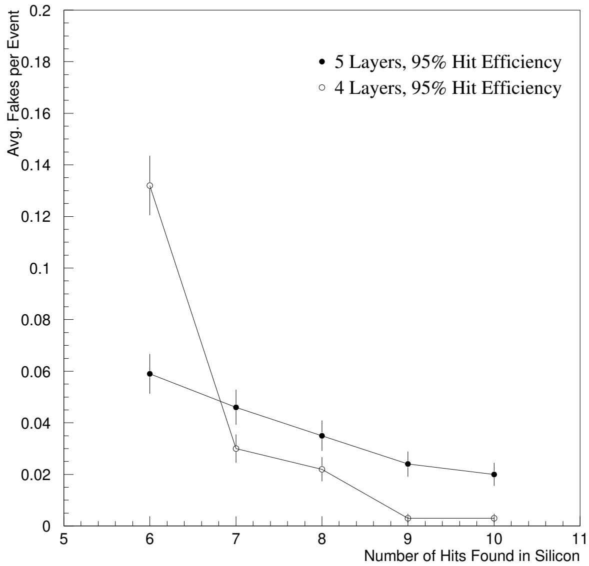

Figure 4-9. *Fake track rate in the SVT at 10 times nominal background rate.*

In the outer layers it is not possible to extend the detectors beyond the 300 mr stay clear. However there are many more staves per layer, and coverage extends to between 300 and 310 mr everywhere.

In the backward direction, there is a similar variation of coverage with azimuthal angles. At the minimum radii of the inner layers there is nominal coverage to within 30° of the beam line. In the center-of-mass, the SVT covers $-0.95 < \cos\theta_{cm} < 0.87$, so the acceptance in the backward direction is actually greater than that in the forward direction in this frame.

4.4 Silicon Detectors

The SVT will be constructed from double-sided, AC-coupled silicon strip detectors. These solid state devices are a technically mature solution to the requirements the SVT must meet to provide precise, highly segmented, robust tracking near the interaction point. The detailed requirements which the detectors must meet are discussed below.

4.4.1 Requirements

Readout Strip Efficiency. The silicon detectors must maintain high single-point efficiency in order to achieve the requirements given in Section 4.1 for high overall track reconstruction efficiency and good tracking resolution. Loss of efficiency can occur from intrinsic strip inefficiencies, from bad interconnections, or from faulty electronics channels. Intrinsic strip inefficiencies can occur due to production defects which result in strips with unacceptably large leakage currents, from accidents during assembly causing the strip to be physically damaged, or from a breakdown in the AC-coupling capacitor. The latter problem is referred to as a pinhole and is due to a small hole in the oxide separating the implant from the metal readout strip above it. Pinholes can occur during fabrication, or they can be generated later on. Understanding and controlling the level of pinholes is one of the primary concerns in our program of silicon detector R&D.

Our goal is to achieve an overall single detector strip failure rate of less than 1%. Data from a large production of double-sided DC-coupled detectors (ALEPH) show that 60–70% can be achieved with a maximum inefficiency of 1%. On this basis, we expect that a 50% yield can be achieved for double-sided AC-coupled detectors while maintaining similar standards.

Point Resolution. As described in Section 4.3, we have determined from Monte Carlo simulations [For94] that the intrinsic point resolution should be 15 μm or better in both z and ϕ for the inner layers. These are the point resolutions for tracks at near-normal incidence. As the angle between the track and the plane normal to the strip increases, the resolution degrades. We require the resolution to degrade by no more than a factor of approximately 3 for angles up to 75° ($\lambda \sim 1.3$) from normal.

Radiation Hardness. A further requirement is that the quoted resolution values hold up to an integrated dose of ~2 Mrad of ionizing radiation (electromagnetic in origin). This requirement leads to the use of AC-coupled detectors in order to avoid the problems associated with direct coupling of the large leakage currents which can occur at such large doses. It also has implications in the choice of the biasing scheme.

Minimum Mass. To achieve good vertex resolution, it is especially important to minimize the material up to and including the first measurement. This requirement, and the need to provide precise vertexing in both z and ϕ, leads to the choice of double-sided detectors. We plan to use 300 μm-thick silicon wafers, which are a standard choice and present acceptable handling properties. While it may be possible to go to 200 μm-thick silicon, the gain is only by the square-root of the total material before the first measurement (including the $0.006X_0$ in the beam pipe), while the fabrication and handling will be much more problematic.

4.4.2 Silicon Detector Design

From the above requirements and from the discussion in Sections 4.1–4.3, we have arrived at the detector specifications and design parameters which are described in this section. A more complete discussion can be found in Reference [Bat94b].

Substrate and implant type. The wafers will be n-type, with a resistivity in the range 4–8 kΩ cm, corresponding to a depletion voltage of 40 to 80 V. These values seem to be a reasonable compromise between the need to have a low depletion voltage and the need to avoid type inversion in the presence of radiation damage.

We will employ p^+ strips on the junction side and n^+ strips on the ohmic side, with p^+-blocking implants in between; see Figure 4-10 for a cross-sectional view. This choice has proven to be a reliable technology [Bat94a] which is directly available without extensive R&D.

Coupling to preamplifier. The strips are connected to the preamplifiers through a decoupling capacitor. AC coupling prevents the amplifier from integrating the leakage current with the signal; handling high leakage currents due to radiation damage imposes an additional burden on the preamplifier design and has other undesirable operational implications.

The value of the decoupling capacitance must be much larger than the total strip capacitance, which is as large as 35 pF. It is possible to use DC-coupled detectors with external capacitors on a separate chip [Bat92]; however, with an electronics pitch of 50 μm, we would need a ~12 mm-long chip to get a $C_{AC} = 180\,\mathrm{pF}$ decoupling capacitor. These dimensions are prohibitively large given the limited space available in the hybrid region. There would also be a significant increase in the number of wirebonds. Capacitors which are integrated on the detectors are the most compact solution, minimizing the number of wirebonds and yielding a value for the decoupling capacitance of 30–80 pF/cm [Ton94], depending on the implant width.

Bias resistors. The bias resistors must be between 4 and 20 MΩ. The lower limit is determined by two factors. The noise has a $1/\sqrt{R_B}$ dependence, and if several strips are ganged together, the effective resistance is correspondingly decreased. Another factor is the requirement that, for floating strips, the product $R_B \cdot C_{TOT}$ must be much larger than the amplifier peaking time (100–400 ns) to allow for capacitive charge partition. The upper limit (20 MΩ) is dictated by the allowable potential drop due to the strip leakage current, which

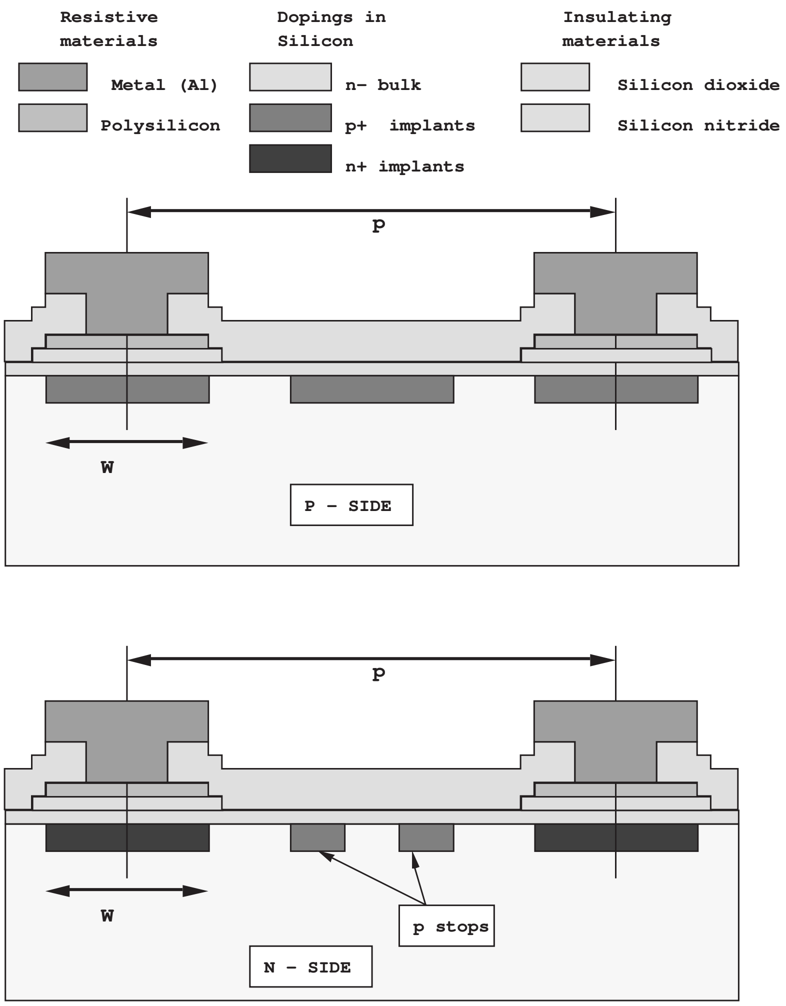

Figure 4-10. *Artist's conception of a silicon detector cross section.*

is taken to be 100 nA at maximum. A good target value is 8 MΩ. A final requirement is that the bias resistor be quite stable for the expected radiation doses.

To meet these requirements, we plan to use polysilicon bias resistors. In prototype CMS detectors [Ton94], values for the sheet resistance of polysilicon of 40 kΩ/□ were achieved, and 50 kΩ/□ is feasible. Thus, it is possible to fabricate an 8 MΩ resistor with a 6 μm-wide, 960 μm-long polysilicon resistor. With a suitable shaping of the polysilicon line, the space required by the resistor will be 480 μm for a 25 μm implant pitch.

For the junction side, punch-through biasing [Hol89] could also be chosen if it is confirmed to be sufficiently radiation hard. Punch-through biasing requires less space, and it can also reduce the parallel noise on the junction side, since the dynamic resistance is much higher at low currents (70 MΩ at 1 nA, decreasing to a few MΩ at 100 nA).

Considering the space needed to accommodate the biasing resistors and to gracefully degrade the electric field close to the edge with a guard ring structure, we specify the active region of the detectors to be 1.4 mm smaller than the physical dimensions (700 μm on each edge).

Optimization of z and ϕ readout strips. A major issue is which side of the detector (junction or ohmic) should read which coordinate (z or ϕ). The capacitance, and consequently, the noise is smaller on the junction side than on the ohmic side, and the strip pitch on the junction side can be 25 μm, while on the ohmic side, it is limited to about 50 μm because of the p-stop implant. For these reasons and because the z vertex measurement is more important from the point of view of physics, we use the junction side for the z strips on the inner layers. The better performance of the junction side also helps compensate for the additional resistance and capacitance imposed by the longer z fanout circuit.

In order to maintain acceptable signal-to-noise ratios for tracks at large dip angles, we employ a 100 μm readout pitch for these z strips with one floating strip interleaved between every two readout strips. We have considered using a wider readout pitch, for example, 200 μm for the very forward and backward regions in order to increase the signal at large dip angles. However, this would involve yet another detector design, and based on our present estimates of achievable electronic noise, it does not appear to be necessary.

Acceptable resolution can be obtained for the ϕ strips on the inner layers using the ohmic side. Two solutions are possible; either a 50 μm readout pitch without floating strips, since there is no room for them on the ohmic side, or a 100 μm readout pitch with one floating strip. Either solution is feasible, and they should give roughly equivalent position resolution for single tracks. Double-track resolution is better for the first solution, and the noise contribution due to detector leakage currents is doubled in the latter solution. Therefore, preference goes to a 50 μm readout pitch without floating strips. Although this choice has twice as many readout channels, the cost implications are not very important because the

Detector Type	z-readout Side			ϕ-readout Side		
	C_{int} (pF/cm)	C_{AC} (pF/cm)	R_{series} (Ω/cm)	C_{int} (pF/cm)	C_{AC} (pF/cm)	R_{series} (Ω/cm)
I	1.3	40	7	2.8	40	7
II	1.3	40	7	2.8	40	7
III	1.3	40	7	2.8	40	7
IV	1.5	80	3.5	1.3	40	7
V	1.5	80	3.5	1.3	40	7
VI	1.5	80	3.5	1.3	30	9.2
VII	1.5	80	3.5	1.3	37	7.5

Table 4-3. *Electrical parameters for the different detector types.*

electronics cost is dominated by the development effort and consequently the per channel incremental cost is not significant.

On the outer layers, the ϕ strips are quite long (up to 24 cm), and their capacitance can become very high on the ohmic side. The position resolution is not as important in these layers, and the maximum track crossing angle is about 45° for the z strips. Therefore, we choose instead to put the ϕ strips on the junction side in order to optimize the signal-to-noise ratio for these very long strips. A 100 μm readout strip pitch with one floating strip is foreseen in the rectangular sections, decreasing to about 65 μm at the ends of the trapezoidal detectors.

For the design of ϕ strips on the trapezoidal detectors, experimental investigations are needed to decide whether to keep the width of the implants constant, the size of the gaps constant, or to maintain a constant ratio of the gap to the implant width. In theory, a constant ratio between width and pitch is favored because most electrical parameters (*i.e.,* depletion voltage and interstrip capacitance) approximately scale with this quantity and therefore would remain constant along the length of the wedge.

For the z strips in the outer layers, the ohmic side is chosen with a readout pitch of 200 μm and one floating strip. The larger readout pitch is employed in order to minimize the ganging of z strips.

The geometrical layout of the seven different types of wafers, including strip pitches and number of floating strips, is summarized in Table 4-1 (Section 4.2). In Table 4-3, the interstrip capacitance, coupling capacitance, and series resistance per unit length are summarized for each detector type.

Det. Type	I	II	III	IV	V	VI	VII
z Length (mm)	47.7×	57.6×	45.5×	46.0×	55.0×	68.1×	63.4×
ϕ Width (mm)	39.8	49.4	65.4	52.5	52.5	35.0–44.0	44.0–52.5
Junction on	z	z	z	ϕ	ϕ	ϕ	ϕ
Layer 1	24	—	—	—	—	—	—
Layer 2	—	24	—	—	—	—	—
Layer 3	—	—	36	—	—	—	—
Layer 4a	—	—	—	24	0	16	16
Layer 4b	—	—	—	8	16	16	16
Layer 5a	—	—	—	9	27	18	18
Layer 5b	—	—	—	0	36	18	18
Total	24	24	36	41	79	68	68
Total w/Spares	28	28	42	46	88	76	76
Batches	3	3	4	4	8	7	7

Table 4-4. *Shape, multiplicity, and number of batches for the different detector types, with one spare module per module type.*

Wafer sizes and quantities. Table 4-4 summarizes the shapes and number of detectors for the baseline design. The quoted multiplicities refer to the installed modules plus one spare per module type. The number of fabrication batches quoted assumes a yield of 50%, *i.e.,* 12 good detectors per batch of 24 wafers.

Using the numbers in Table 4-4, we see that the current design employs seven different types of detectors (*i.e.,* seven sets of masks) and needs 36 fabrication batches, for a total of 340 installed detectors. Having so many types of detectors complicates both the design and production phases, especially for prototype and spare production. A reduction in the number of detector types would be most welcome; however, this represents the minimum which we have been able to achieve in our present baseline design.

4.4.3 Fanout Circuit Design

The SVT front-end electronics will be located outside of the active tracking volume to minimize the amount of material crossed by particles within the acceptance. As a consequence, the signals from both the z and ϕ microstrips must be carried to the front-end chips by flexible fanout circuits. While the ϕ fanout circuits are just one-to-one connections a few centimeters long, the z fanouts are more complicated. They must extend the full length of the detector modules in order to connect all the z strips and to provide interconnections, or ganging, in cases where the number of available readout channels exceeds the number of z

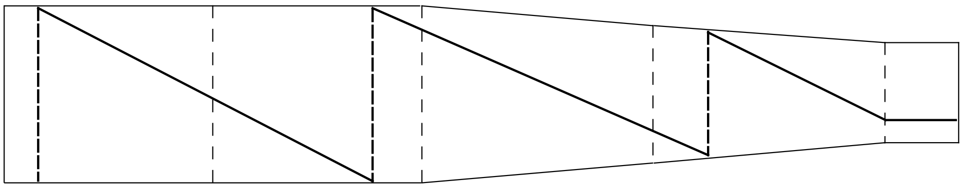

Figure 4-11. *Schematic drawing of fanout connection to z strips showing ganging.*

strips. Figure 4-11 depicts the concept of ganging. Two or three z strips (dashed lines) are connected in series by the fanout traces (solid lines), bringing the signals to the end of the detector module where they can be wire-bonded to the front-end electronics IC.

Table 4-5 gives a detailed list of the required fanout circuits with their geometrical features. There are four possible types of fanout circuits per layer (z/ϕ and forward/backward), though in many cases, the forward (F) and backward (B) circuits are identical. The number of input readout strips is shown together with the number of front-end electronic channels available in the corresponding readout sections. The number of z strips always exceeds the corresponding number of readout channels: therefore ganging is required. The number of ϕ strips is often smaller than the corresponding number of available channels. In such cases, some of the readout channels are not used. Typical pitches at the input and output bonding pads are given; the line pitches can be smaller in critical places. The total number of fanout circuits is 208; the number of different circuit layouts is 28.

Similar circuits have already been realized by industry as very thin, high precision Kapton flex cables [Lev] for the L3 Silicon Microvertex Detector [DiB95] and the Aleph VDET200 detector [Bag94]. Alternative solutions exist but are not suited to our design. Double-metal fanouts integrated on the detectors exhibit much larger parasitic capacitances. A thin glass substrate has interesting properties, but in our case flexible circuits are needed to route the signals to the readout electronics.

The base material for the z fanout of Reference [DiB95] is 50 μm-thick type-E kapton, plated with a 0.25 μm nickel adhesion layer followed by 2.5 μm of electroplated copper. High precision masks were realized by electron beam photolithography with a typical geometrical resolution of 1 μm. The required conducting lines were chemically etched to a width of about 11 μm, starting from line widths of about 20 μm on the masks. The pitch of the lines could be as small as 33 μm. Bonding pads were electroplated with 2 μm nickel followed by 1 μm of high-purity amorphous gold.

This technology is satisfactory both for the mechanical and the electrical properties. The average radiation thickness of the fanout circuit is only $2.3 \times 10^{-4} X_0$, and the degradation of the signal-to-noise ratio due to the fanout contribution to the capacitance and resistance

Layer	Fanout Type	Length (cm)	Number of Readout Strips	Channels	Typical Pitch at Input(μm)	Output (μm)	Number of Circuits
1	z, F+B	12.5	950	768	100	50	12
	ϕ, F+B	3.0	768	768	50	50	12
2	z, F+B	14.5	1150	1024	100	50	12
	ϕ, F+B	3.0	960	1024	50	50	12
3	z, F+B	15.6	1360	1280	100	50	12
	ϕ, F+B	2.0	1280	1280	50	50	12
4a	z, F	19.7	885	512	200	50	8
	z, B	24.3	1115	512	200	50	8
	ϕ, F+B	2.0	512	512	65	50	16
4b	z,F	20.6	930	512	200	50	8
	z,B	24.2	1160	512	200	50	8
	ϕ, F+B	2.0	512	512	65	50	16
5a	z, F	25.2	1160	512	200	50	9
	z, B	25.1	1205	512	200	50	9
	ϕ, F+B	2.0	512	512	65	50	18
5b	z, F+B	26.1	1205	512	200	50	18
	ϕ, F+B	2.0	512	512	65	50	18

Table 4-5. *Summary of fanout circuit characteristics.*

is acceptable. For example, the ALEPH Kapton fanout circuits contribute 0.6 pF/cm to the interstrip capacitance and 0.1 pF/cm to the strip-to-back capacitance. An acceptable production yield (about 50%) could be achieved [DiB95] by allowing a maximum of 1.5% of optically detected defects (shorts or breaks in the lines), most of which could be repaired.

One potential drawback of this approach is that the additional wirebonds required to connect the fanout can result in a loss of channels due to faulty connections if great care is not taken in assembly. These aspects must be further investigated in 1995. It is also necessary to perform more accurate measurements and calculations of interstrip and back capacitance for the fanout circuits. In addition, mechanical and electrical tests will assess possible limits in the sharpness of bending. Other substrates, such as Upilex, may offer superior characteristics and better production yield; these are currently being investigated.

4.4.4 R&D on Detectors and Fanouts

Although the technology is rather stable and reliable, some R&D and prototyping must be done before freezing the design by the end of 1995. We plan to submit a prototype batch of AC-coupled, double-sided detectors early in 1995. The mask design will include several structures, including a detector which is almost full size ($\sim$15 cm^2) for use in a test beam in the summer of 1995, as well as several test structures to measure capacitances and leakage currents for various strip pitches and implant widths. In addition, a small prototype wedge detector and various test devices for process characterization and radiation studies will be included. Prototype fanouts will also be employed in this test beam.

Following the first prototype submission, there will be a second submission consisting of two batches, with one full-size detector per batch. These detectors will be put into a later test beam together with prototype readout electronics.

4.5 Electronic Readout

4.5.1 Introduction

The SVT readout electronics must:

- Process the signals that appear on the readout strips with the aim of digitizing the charge;
- Temporally correlate the detector signals with an externally generated Level-1 trigger;
- Format the vertex data, which consists of the channel ID and the digital values of the charge and time-stamp, and transmit them to the DAQ system upon receipt of a read command; and
- Remain functional up to 10 times nominal background with graceful degradation at higher background levels.

The charge information is retained for several reasons. One important reason is to correct for comparator time-walk; the pulse arrival time must be measured offline as accurately as possible in order to minimize the number of background hits associated with an event. A second reason is to improve, by interpolation, the position measurement of hits for which there is charge-sharing between adjacent readout strips. Finally, charge information may also be used to reject background hits and to correlate ϕ- and z-side hits.

The readout electronics are subdivided into readout sections. A readout section services the ϕ or z strips for one-half of a detector module (either the forward or backward end). The vertex readout electronics consists of a monolithic readout chip; a hybrid circuit, which provides the mechanical support for the chips and the substrate for connections to the chip; a transition card, which interfaces the front-end readout to the DAQ system; and ancillary units, such as power supplies and detector bias supplies. The readout chip, the hybrid, the transition card, and the power supplies will be described in the remainder of this section.

For further information on the choices that underlie the design of the readout electronics and the relevant architectural details, the reader is referred to the relevant *BaBar* notes [Lev94, Roe94b, Joh95a, Joh95b].

4.5.2 Readout Chip

Functional Overview

The readout chip must amplify and digitize the input signals for all channels in parallel. In addition, the signals must be buffered for the duration of the Level-1 trigger latency. For those hits selected by the trigger, the digitized data must be buffered until a read command is received. Upon receipt of the read command, only signals which exceed a predefined threshold should be read out. Since the time between beam crossings will be very small (4.2 ns), none of the operations of the chip can be restricted to occur between beam crossings. Furthermore, the maximum expected Level-1 trigger rate is too high to allow data acquisition to be suspended during readout operation. Therefore, during normal operation of the chip the analog front-end is actively acquiring data while digitization, buffering and readout occur.

Requirements

The requirements for the readout chip are summarized below. For more information, see Reference [Joh95a].

1. Mechanical requirements:
   - Number of channels per chip: 128;
   - Chip size: total width < 6.2 mm, and total length < 9 mm; and
   - Channel-to-channel pitch: $50\,\mu$m.
2. Operational requirements:

- Operating temperature: maximum 40°C;
- Radiation tolerance: up to 2 Mrad (ten years at 200 krad/yr); and
- Power dissipation: not to exceed 2 mW/channel.

3. Dynamic range: Minimum input charge is 0.75 fC (0.20 MIP), and maximum is 19 fC (5 MIP); circuit must accept either positive or negative signals. It is desirable to increase the dynamic range to 38 fC (10 MIP) if feasible.
4. Analog resolution: 5 bits on a linear scale are required to achieve the required resolution of 0.25 MIP for the above dynamic range, while a minimum of 3 bits are necessary for a nearly optimal position measurement with range compression in a nonlinear system.
5. Time-stamp resolution: 100 ns time-stamp resolution is required for the inner layers, increasing to a maximum of 400 ns in the outer layers to match the pulse shaping times.
6. Noise performance: The noise must not exceed 800 e^- rms at a risetime of 100 ns with an external capacitance of 10 pF, an external resistance of 100 Ω, and a leakage current of 100 nA. It must not exceed 1200 e^- rms at a risetime of 400 ns with an external capacitance of 50 pF, an external resistance of 200 Ω, and a leakage current of 200 nA. (These specifications refer to the inner and outer layers of the SVT, respectively.)
7. Maximum data rate: Simulations show that the lost-particle backgrounds dominate the overall rates; at nominal background levels, the maximum hit rate per strip is about 300 kHz.
8. Deadtime limits: The maximum total deadtime of the system must not exceed 10% at a 10 kHz trigger rate and at 10 times the nominal expected background rate.
9. Double-pulse resolution: The maximum inefficiency from pulse overlap must be less than 3%.
10. Trigger specifications: The trigger has a nominal latency of 9.5 μs, a maximum jitter of 0.5 μs, and the minimum time between triggers is 1.5 μs. The maximum Level 1 trigger rate is nominally 2 kHz, but a conservative upper limit of 10 kHz is assumed.
11. Output data format:
    - only hits above threshold are transmitted (sparse readout); and
    - the data packet should include a synchronization code, chip ID, time stamp, data, status word, and a trailer.
12. Calibration requirements: Provision must be made for injecting a known amount of charge at the inputs of an arbitrary, remotely selectable set of channels.

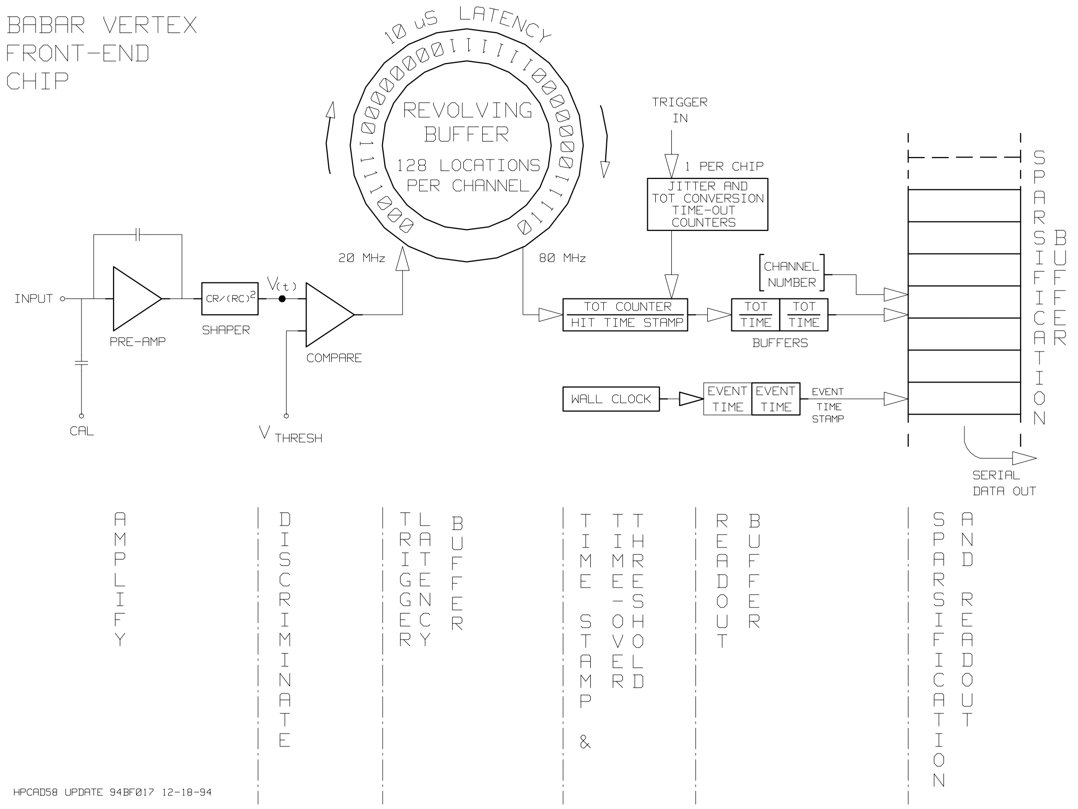

Figure 4-12. *Block diagram of the silicon vertex detector readout chip.*

13. Failure recovery: It is desirable to reduce the risk of failure by employing conservative design methods and building in redundancy where possible.

Readout Chip Implementation

The chip described at a block diagram level in Figure 4-12 is intended for implementation in a $0.8\,\mu$m CMOS radiation hard process and has been designed to comply with the requirements discussed above [Joh95b]. Each chip consists of 128 channels together with additional circuitry that is common to all channels. The chip operation is synchronized by an external clock, which is frequency-divided on the chip to obtain synchronization of different operations at different frequencies.

Each channel consists of an analog section, a comparator with a setable threshold (V_{th}), a trigger latency buffer, a counter, and a back-end buffering section. The analog section consists of a charge-sensitive preamplifier followed by a CR(RC)^2 shaping amplifier, which implements a second-order semi-Gaussian impulse response. The shaping has a twofold

	t_{peak} (ns)	Length (cm)	p/n	Cap. (pf)	ENC(e^-)	Signal Loss
Layers 1–3 z	100	5	p	7.7	400	10%
Layer 3 ϕ	200	15	n	35.5	730	29%
Layer 5 z, x3	400	15	n	29.3	630	10%
Layer 5 ϕ	400	25	p	24.7	590	33%

Table 4-6. *Equivalent noise charge for various strip configurations.*

purpose: to improve the signal-to-noise ratio (S/N) over that obtained at the preamplifier output, and to provide a signal shape suitable for the TOT processing in the next section of the analog circuitry. The equivalent noise charge (ENC) after the shaper depends on the front-end amplifier/shaper characteristics and on the characteristics of the detector to which it is connected. Table 4-6 illustrates the ENC, which is calculated using a detailed SPICE model of both the analog front-end and the detector, for a variety of strip lengths found in the SVT. The simulation assumed an equivalent noise at the input transistor of $1.9\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ and a post-irradiation leakage current of $1\,\mu\mathrm{A/cm}^2$, corresponding to approximately 1.5 Mrad of ionizing radiation. The signal loss is the fraction of the signal which is lost to the back of the detector and therefore will not be collected at the preamplifier. There is additional charge lost to the neighboring strips; however, this is typically smaller and will be seen, provided it is above threshold.

The function of a channel can be understood by following the signal path through it. The fast detector signal with charge Q appears at the output of the shaping amplifier with a time dependence given by:

$$V(x) = k \cdot Q \cdot x^2 e^{-2x}, \tag{4.1}$$

where k (in Volts/Coulomb) is the charge sensitivity of the preamplifier-shaping amplifier combination, and $x = t/t_p$ is the time normalized to the peaking time t_p. The value of t_p is selectable by an external chip control in order to comply with noise, deadtime, and double-pulse-resolution requirements.

The shaped pulse is compared with the preset threshold V_{th}, and if it exceeds V_{th}, it produces a TOT at the comparator output. The TOT is an analog variable which carries information about Q. The relationship between Q and TOT is nonlinear and is not far from a logarithmic dependence. This is desirable because it provides good resolution for small Q and reduced resolution for large Q, effectively compressing the range into fewer bits.

A simulation was performed to evaluate the ratio of the pulse width to the sampling period which was required to achieve good position resolution using TOT information [Roe94b]. The position resolution as a function of the crossing angle between the track and the silicon strip detector was evaluated for several different assumptions: perfect linear analog readout, digital readout (hit/no-hit information only), and TOT with variable t_{peak}/t_{clock}. For the

TOT simulation, a CR-RC^2 shaper was assumed. The results are shown in Figure 4-4; it is observed that TOT with $t_{peak}/t_{clock} = 2.5$ is essentially as good as a perfect linear analog readout.

It should be pointed out that an ENC-based noise specification makes sense so long as it is restricted to the linear section consisting of the preamplifier and shaping amplifier and refers, for example, to the threshold at which the resulting pulse is discriminated. For the accuracy of the TOT-based charge measurement, a different S/N must be defined:

$$(S/N)_{TOT} = TOT/ < \sigma >, \tag{4.2}$$

where $< \sigma >$ is the rms, noise-induced dispersion on the TOT. As a first approximation, $(S/N)_{TOT} = (Q - V_{th}/k)/ENC$, where k is the gain of the preamplifier/shaper combination in volts/Coulomb.

The TOT information lends itself to a straightforward digitization, for instance, by comparison with the period of a reference clock. TOT digitization and trigger delay compensation are performed by the revolving buffer of Figure 4-12, in which the status of the TOT comparator is stored under the supervision of a write pointer operating at the frequency of the reference clock. The number of locations in the revolving buffer is given by the maximum expected trigger delay, taking into account the trigger latency, trigger jitter, and the front-end chip jitter, divided by the period of the writing clock.

The readout operation is initiated when a trigger is received, beginning with the generation of the event time stamp, obtained by latching the contents of a 16-bit counter (one per chip) which continuously counts the write clock (the time stamp counter in Figure 4-12). A timed sequence ensues, controlled by the trigger-jitter counter and the TOT full-scale counter, both of which are common to all channels in the chip. The read pointer for the revolving buffer is set at the position corresponding to the earliest possible data in the buffer, taking into account the fixed trigger latency, the trigger jitter, and the position of the write pointer at the instant of the trigger arrival.

The read pointer is controlled by a clock at a higher frequency than the write pointer in order to minimize deadtime, because no new triggers can be received while the read pointer is active. The data associated with the trigger are retrieved from the revolving buffers by detecting a zero-to-one transition in the revolving buffer. When a zero-to-one transition is found, three actions take place:

- A hit flag is generated;
- The TOT counter is enabled to count the number of ones in the buffer; and
- A hit time stamp is latched.

The number of bits counted corresponds to the digital value of the TOT and hence, of the charge Q. Based on simulation of the TOT technique [Roe94b], it has been determined that an adequate resolution for position measurement is obtained with between 3 and 5 bits, so 4 bits are envisaged to represent the digital value of the TOT. The hit information in the individual channels consists of 10 bits; five of them represent the hit time stamp, four represent the TOT value, and one is for the hit flag.

The back-end section of each channel includes additional buffers into which the hit data can be stored while awaiting readout. If there is no backlog of data to be read out, the hit data are completed by adding the information about the channel number, which requires 7 additional bits, and transferred into a special buffer where sparsification takes place. Removing the hit flag, there are 16 bits per hit channel which are then to be transmitted upon receipt of a data transmit command.

4.5.3 Hybrid Design

Functional Overview

The hybrid is composed of several components which are hybridized into one electrical unit [Col94]. These include the high-density interconnect (HDI), the front-end chips and other components that are mounted on the HDI, the thermal interface between the chips, and the water-cooling system, and the cable or tail which connects the HDI to the transition card.

The hybrid is connected to the silicon detectors by flexible fanout circuits which bring the detector signals to the front-end chips. The hybrid is mechanically mounted onto the fiber composite beams which provide structural rigidity for the detector modules. The mounting of the detector modules onto the cooling/support cones is accomplished with pins and screws which are inserted from inside the cone through the aluminum mounts of the cone. The screws pass through the thermally conductive substrate of the hybrid and into the rigid mechanical structure of the module. This mechanical support thus provides the thermal contact that allows heat from the ICs to flow into the cooling water circulating around the aluminum mounts. From this brief description of the functionality of the hybrid, it is apparent that the hybrid design is complicated by the fact that the requirements imposed by the mechanical support, the electronic readout, and the cooling must all be met in a self-consistent solution. The physical space available for the hybrid is very limited, imposing yet another constraint on the design.

High-Density Interconnect and Thermal Interface

The HDI is a simple fanout circuit on which the readout chips are mounted. It provides connections for each chip to the data and control lines and to the low voltage supplies and grounds. It may also include some passive filtering and any external components required by the front-end chip. It is important that the layout of the HDI minimize crosstalk and noise, especially on the analog power and ground planes. The HDI must also provide connections to the detector for the detector biasing voltages and provide a local, low-impedance connection from the ϕ-side to z-side preamplifiers for the detector current return path. There must also be a low-impedance thermal path from the front-end chips to the aluminum mounts which protrude from the cooling/support cone; this thermal path will either be through the HDI itself or through a substrate to which it is laminated.

There are five different HDI layouts required, one for each of the five layers. The same HDI layout is used for both ϕ- and z-side readout within a layer. There are several options for the implementation of the HDI. It may be a flexible kapton circuit laminated to an aluminum heat sink, or a thick-film ceramic circuit fabricated from a thermally conductive material such as AlN or BeO. Thin-film circuitry on a ceramic substrate is also an option. At present, these design options are being studied, and an optimum choice in terms of performance, ease of fabrication, and cost is being determined.

The thermal path from the front-end chips to the heat sink has been evaluated with Finite Element Analysis (FEA) simulations, assuming a flexible kapton circuit laminated to an aluminum substrate. A simple model was built and tested experimentally, and the FEA calculation was found to agree well with the measured temperatures. Having validated the simulation, a more detailed FEA calculation was performed, taking into account the precise dimensions of the hybrid and all of the layers in the multilayer kapton circuit. The calculations indicate a temperature drop of about 15°C from the IC to the cooling water. Given the maximum chip operating temperature of about 40°C, this is more than adequate. The design is being further refined, and more experimental tests are planned.

Hybrid Tail

The tail is the connection between the HDI and the transition card (described below). It is envisioned as a flexible multilayer circuit which will have power and ground planes and one or more signal planes. It will be fabricated separately from the HDI itself and will be connected to two HDIs by means of bump-bonds, solder, or other permanent connection. The connection to the transition card will utilize a mateable connector. The biggest constraint on the design of the tail is the limited space available for cables from the inner SVT layers to pass between the cooling/support cone and the outer detector modules. This restricts the width of the cable to less than about 0.75 cm, and the thickness to less than about 0.3 mm.

The trace impedance, which is about 30 Ω, must be matched to the drivers and receivers at either end.

4.5.4 Data Transmission

Data Transmission Requirements

The main requirements for the data transmission system are given below.

1. Data rate: between 60 and 80 Mbits/s/readout section;
2. Power consumption (inside the detector volume): below 5 W/readout section;
3. Radiation resistance (inside the detector volume): the system must be functional within the specified performance after being exposed to a 100 krad integrated dose of ionizing radiation; and
4. Error rate: less than 10^{-11}/bit. This would imply roughly 1 bit error in every 10^6 events transmitted.

Functional Overview

Two different designs have been studied for transmitting data from the SVT front-end readout IC to the DAQ system and also to send clock, trigger and control signals from the DAQ system to the readout IC. The two designs differ primarily in where the transition between electrical and optical transmission occurs.

In the first design which was considered, the transition from electrical to optical transmission was made on transition cards which sit inside the support tube at the outboard end of the B1 magnets. A flexible circuit carries electrical signals between the HDI and the transition card, and optical fibers connect the transition card to the DAQ system. This option has the advantage of avoiding noise pick-up in the 5–6 m cable run from the transition card to the outside of the detector, and reducing the volume of cable in the space-constrained area inside of the support tube. However, it requires customized packaging and use of radiation-hardened opto-electrical components on the transition card which are very costly.

In the second design, electrical transmission is used from the HDI all the way outside of the detector to a point where crates can be mounted to support a commercial optical transmission system. The transition cards inside of the support tube are still needed; however they would have little or no active circuitry. Because this option can make use of commercial fiber optic systems, the cost is reduced and it is possible to use optical transmission systems common to other sub-detectors. The opto-electric system is also accessible for repair and does not need to be radiation hard.

The transition cards are required in both designs in order to provide a place where the hybrid tail can terminate. In the second option, if the impedance of the hybrid tail and the transmission cables extending beyond the transition card can be matched sufficiently well, all active components can be removed from the transition card thus improving the reliability of the system inside the support tube where access is limited.

After studying the space budget for cabling beyond the transition card and the feasibility of the readout ICs transmitting signals electrically out of the support tube area, the second option has been chosen as the baseline design.

4.5.5 Baseline Design

The baseline design for transmitting data out of the detector starts with the readout ICs transmitting differential signals on pairs of traces up the hybrid tail to the transition card. The signals are there transferred to conventional paired cables for transmission to an interface card outside of the detector. Likewise, clock and control signals coming from the interface cards will be transmitted as differential signals on conventional cable to the transition card inside the support tube and then transferred onto the hybrid tail to the HDI and the readout ICs.

The hybrid tail is still under development, and its precise electrical characteristics are still to be determined. The design of the hybrid tail is constrained by mechanical requirements and limited by the available technology for producing flexible cables. If the impedance of the hybrid tail can be closely matched to that of available cabling, the signal transmission through the transition card will be only passive. If the impedance mismatches are too great, commercially available bipolar components have been identified to receive and re-transmit signals on the transition card. Both the design of the transmitting and receiving circuits on the readout ICs and the choice of components for the transition card have been made to maintain the consistent differential, low power, low noise transmission signals. The identified bipolar components for the transition card will meet the speed, power and radiation requirements of the system.

The transition cards will not only provide the physical connection between hybrid tail and conventional cable and any necessary impedance matching for signal transmission, it will also provide necessary filtering for all the DC voltages bussed to the detectors and the readout ICs. There will be one transition card for each hybrid tail. They will be located inside the support tube, approximately 20 cm from the outside layer of the SVT. They are mounted

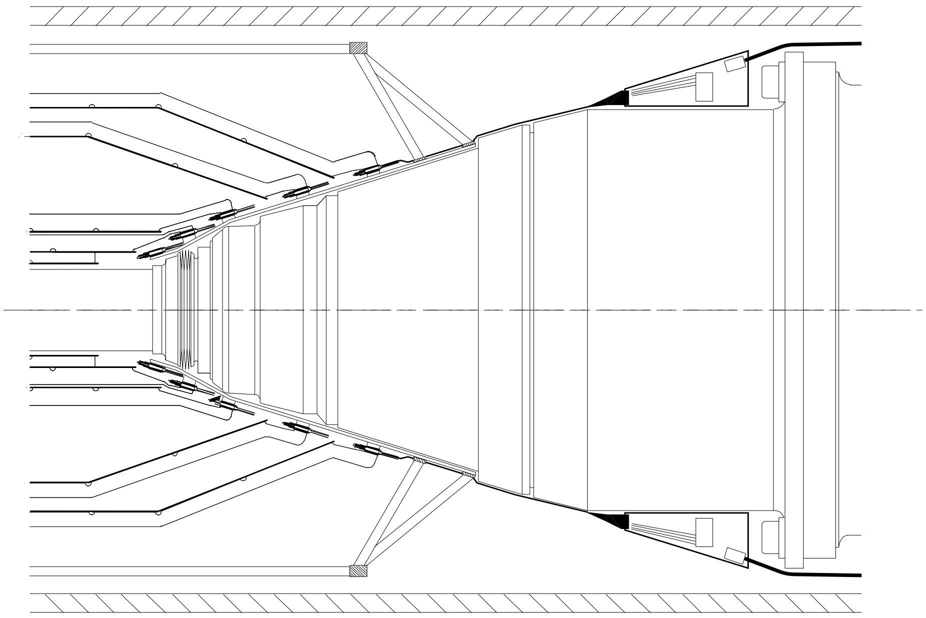

Figure 4-13. *Side view of the connections from the readout electronics to the transition card, which is a triangularly shaped circuit mounted from the B1 magnet.*

from a cooling ring which is attached to the far end of the B1 magnet. See Figure 4-13 for a side view showing the connections from the readout electronics to the triangularly shaped transition card. A multi-conductor cable bundle for each transition card will carry all necessary power, sense and data lines from the transition card through the support tube to a point outside the detector.

The interface cards will provide transition between electrical signals and optical signals. The exact location for these cards has not been determined but they will be located approximately 5–6 m from the interaction point, outside the detector. Eight data lines will be multiplexed onto one gigabit fiber optic link for transmission to the DAQ system. Likewise, clock and control signal coming from the DAQ on a fiber will be fanned out and transferred to electrical cables going into the detector. The exact level of multiplexing between electrical signals and optical is made to match bandwidths and partitioning of DAQ functions.

4.5.6 Power Supplies

The power supply system for the SVT must provide the voltage sources with adequate current levels for two purposes: biasing the silicon detectors, and supplying power to the front-end chip, transition card, and remote opto-electronic circuits. The low-noise electronic readout for the silicon vertex detector imposes stringent requirements on the power supplies, and they must be carefully designed to prevent common-mode noise from entering the system.

Both supply subsystems must be built up from modular, fully floating units without galvanic connection between the utility side and the load side, with the two sides electrostatically shielded. Preference is generically given to continuous-type voltage regulators in view of their lower noise levels. However, tests are foreseen to determine whether smaller switching regulators with better power yield can also provide satisfactory performance.

The power supplies must also feature remotely controlled, digitally preset voltages. Optical connection of the voltage controls is recommended. The power supplies must be capable of sweeping the output voltage throughout the nominal range as specified below.

The specifications for the two different supply subsystems are as follows. For the detector bias supply, the nominal full scale voltage is 80 V per module, with a setting resolution of 7 bits and maximum current of 0.1 A. Ripple must be less than 10 mV peak to peak, and noise must be less than 2 mV rms. For the electronics power supply, the nominal full scale voltage is 6 V per module, with a setting resolution of 8 bits and maximum current of 100 A. Ripple must be less than 2 mV peak to peak, and noise must be less than 500 μV rms.

4.5.7 Electronics R&D

The R&D to support the development of the electronic readout system described above has already begun. Characterization of radiation-hard CMOS processes is well underway. Test structures have been submitted to Honeywell and to UTMC to characterize their radiation-hard processes. This work was begun as part of an R&D program for SDC. The UTMC circuits have already been evaluated, and the work is continuing with irradiation and testing of the Honeywell chips. Honeywell has an 0.8 μm, triple-metal process which is attractive both for its small minimum feature size and for the third metal layer which is useful for minimizing the circuit size and may also prove useful for shielding. Initial results with this process are promising. Figure 4-14 gives comparison of the noise performance of the Honeywell process before and after irradiation, plotted versus frequency. For the 100 ns shaping time which is foreseen for the inner detector layers, the corresponding frequency is about 3 MHz. At this frequency, the noise increases by about 15% after 500 krad and by

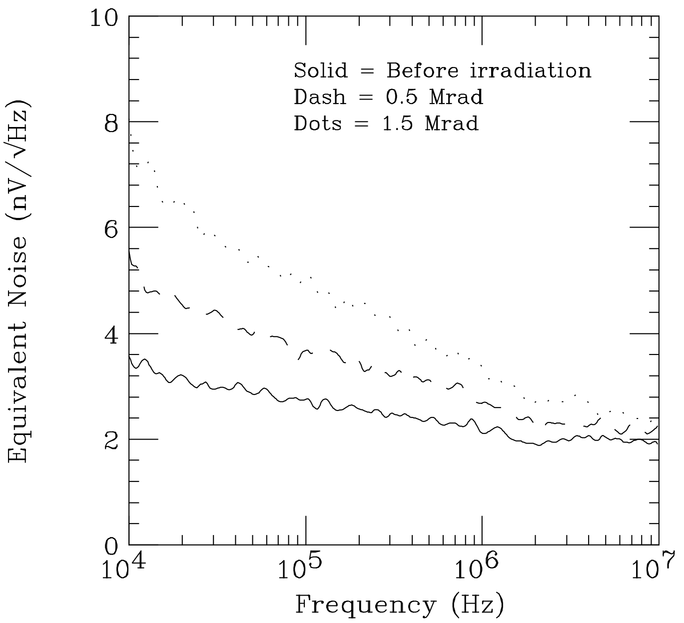

Figure 4-14. *Noise levels for Honeywell transistor before and after irradiation.*

about 30% after 1.5 Mrad of ionizing radiation. Further tests up to a total integrated dose of 5 Mrad are planned.

A first round of test chips on the TOT concept has been submitted and is being tested in early 1995. These test chips will be used to evaluate the TOT technique and to measure the extent to which digital activity on the chip during data acquisition can contribute to front-end noise. Work on a first full-scale prototype of the TOT front-end chip is underway, and the first submission is expected in the spring of 1995. This chip will be operated on a test bench in the fall of 1995 with prototype *BaBar* detectors and will then go into a test beam.

For the hybrid, a mechanical prototype is foreseen early in 1995 as a means of qualifying a vendor. This will be followed by a working electronic prototype in mid-1995, which is designed to work together with the first prototype TOT chip.

4.6 Mechanical Support and Assembly

An overview of the SVT mechanical support is provided in Section 4.2. In this section we provide a more detailed account of the constraints on the mechanical design due to the accelerator components near the IP and describe the details of the detector assembly, installation, survey, and monitoring.

4.6.1 IR Constraints

The support structure design and configuration of the SVT is dictated by the configuration and assembly procedure of the machine components near the interaction point, as well as by the SVT geometry. The close spacing of the bunches in PEP-II (4.2 ns) and the desire to avoid parasitic crossings dictate the need for a pair of permanent dipole magnets located about 20 cm from the IP on either side. These B1 magnets occupy most of the region below 17.2° (300 mr). In order to minimize the mass inside the active tracking volume, it is desirable to mount all of the electronics below the 300 mr line. In the forward direction, this requires that electronics, cooling, cabling, and support be confined in a volume one centimeter thick around the B1 magnet. The use of this small space below the 300 mr line must be carefully coordinated with the needs of the accelerator. In the backward direction, coverage to within 30° of the beam pipe leaves room for some machine components which are moved out of the forward region. The solid angle coverage of the SVT is therefore restricted to the region $17.2 < \theta < 150°$.

A further constraint on the SVT is imposed by the PEP-II support tube, which is a tube of radius 20 cm that aligns the final magnets closest to the IP. The central part of the support tube is a thin carbon-fiber structure, while the outer sections are stainless steel. The SVT must be supported inside the carbon-fiber support tube and therefore cannot be mounted from the drift chamber endplates. In the baseline design, the SVT is supported from the B1 magnets. Since the SVT must be installed with both B1 magnets in place, it must be assembled in two halves and then clam-shelled around the beam pipe. The assembly and alignment of the beam pipe, B1 magnets, Q1 magnets, and SVT inside the support tube will take place in a staging area away from the interaction hall. The entire assembly will then be transported and installed in the interaction hall. This procedure is reversed in order to gain access to the SVT for repair.

4.6.2 Module Assembly

The SVT is constructed from detector modules, each of which is mechanically and electrically independent of the other modules. Each module consists of silicon wafers bonded to fiber composite beams, with a high density interconnect (HDI) electronic hybrid at each end. The HDIs are electrically connected to the silicon strips by means of flexible circuits and are mechanically supported by the fiber composite beams. The entire module assembly is a rigid

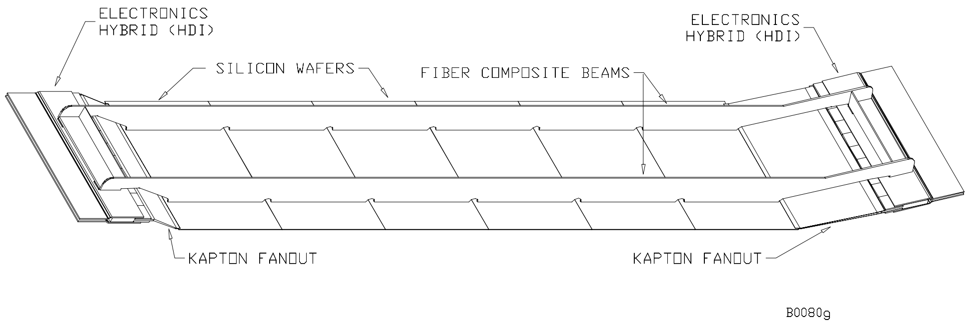

Figure 4-15. *Detector module from Layer 3, consisting of six silicon detectors wirebonded together and read out at each end.*

structure that can be tested and transported with relative ease. A drawing of a detector module from Layer 3 is shown in Figure 4-15.

Assembly of the detector modules begins with the preparation of the necessary parts. The silicon detectors must be fully tested, including a long-term test under full bias voltage. The fanout circuits will be optically inspected and single strip tested for shorts/opens. The readout hybrids must be assembled and tested, starting with the HDI circuit, the front-end chips, any additional passive components, and the hybrid support. Finally, the completed beams, which provide mechanical stiffness, must be inspected to ensure they meet specifications. These individual parts will be fabricated at different institutions from which they can easily be shipped to the institutions at which the module assembly is carried out. The hybrids will be retested after shipment.

The assembly of the inner barrel-shaped modules and the outer arch-shaped modules is necessarily different. However, there are common steps. Generally, the procedure is as follows:

1. The z and ϕ fanouts are glued to the detectors and wire-bonded to the strips. The ganging bonds between ϕ strips are performed.

2. The silicon detectors and readout hybrids are held on a suitable fixture and aligned relative to each other. The fanouts are glued to the hybrids and wire-bonded to the input channels of the readout ICs. Electrical tests, including an infrared laser strip scan, are performed and the detector-fanout assemblies (DFAs) are visually inspected.

3. The final assembly stage is different for different layers. For modules of Layers 1 and 2, the DFA is bonded to the fiber composite beams with appropriate fixtures ensuring alignment between the mounting surfaces on the HDI and the detectors. The module is again tested. A module from Layer 1 and Layer 2 are then joined together by gluing

the beams on the top of Layer 1 to the bottom of Layer 2. The combined structure is called a sextant module. For Layer 3, the DFA is bonded to the fiber composite beams as with Layers 1 and 2.

For the modules of Layers 4 and 5, the DFA is held in a suitable fixture and bent at the corners of the arch and at the connection to the HDI. The module is tested. The fiber composite beams are bonded to the module with fixtures assuring alignment between the silicon detectors and the mounting surfaces on the HDI. This procedure has been performed with real detectors that have been successfully operated in a test beam.

Once completed, these detector modules are extremely rigid devices that can be stored and subjected to long-term testing. They are then shipped to the location for final assembly and installation of the detector.

4.6.3 Detector Assembly and Installation

Half-Detector Assembly

The detector is assembled in halves in order to allow the device to be clam-shelled around the beam pipe. The detector modules are supported at each end by cooling/support cones constructed from double-walled carbon-fiber laminates. Cooling water circulates between the two carbon-fiber shells and around aluminum mounting pieces which protrude through the outer shell. The cones are split along a vertical plane and have alignment pins and latches that allow them to be connected together around the B1 magnets. See Figure 4-16 for a drawing of the cooling cone. The two carbon-fiber support cones are mechanically connected by a low-mass carbon-fiber space frame.

During the half-detector assembly, the two half-cones will be held in a fixture which holds them in precise relative alignment. The detector modules are then mounted to the half-cones at each end. A manipulating fixture holds the detector modules during this operation and allows for well-controlled positioning of the module relative to the half-cones. Pins located in the aluminum mounting pieces provide precise registration of the modules, which are then screwed down. Accurate alignment of the mounting with respect to the silicon wafers is achieved by a pair of mating fixtures. One is a dummy module and the other simulates the mating surfaces on the cone. These fixtures are constructed together and mate perfectly. One is used to verify the machining of the aluminum mounting pieces on the cones. The other is used to position the mounting points on the HDI during the assembly of the modules as mentioned above.

The connection between the module and the cone (called the foot) provides for accurate and reproducible alignment of the module and conduction of heat from the HDI heat sink to the

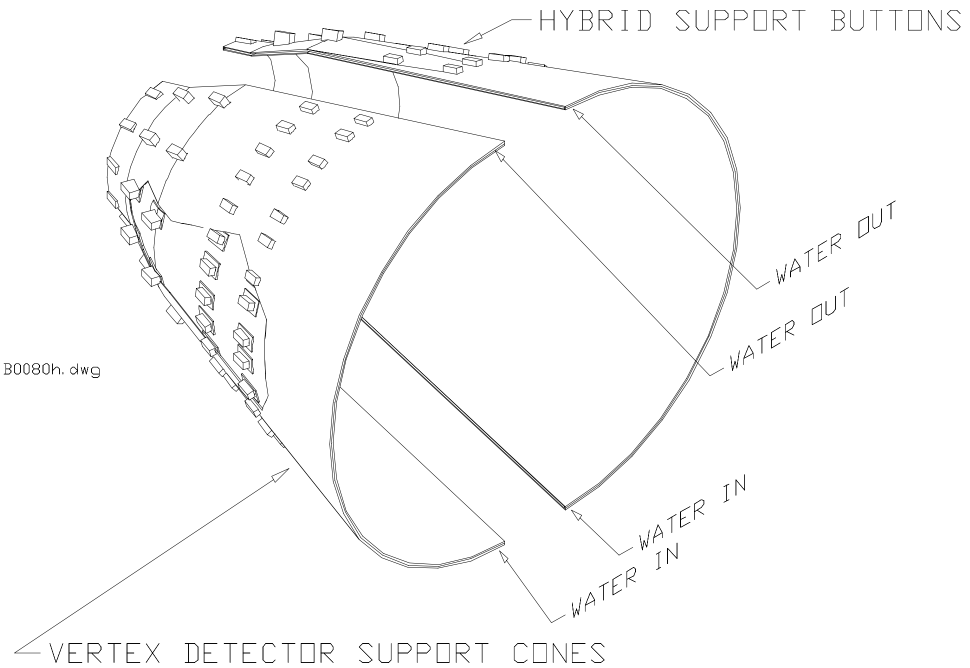

Figure 4-16. *Schematic view of the cooling/support cone. The cone is constructed in two halves and clam-shelled together. Water circulating between the two carbon-fiber shells cools the readout electronics, which are mounted on aluminum pieces protruding through the outer shell.*

cooling water circulating between the cone layers. A detail of the foot region, which contains the readout electronics and the mounting pieces, is shown in Figure 4-17.

After verification of the alignment, the connection between the HDI and the support beams is permanently glued. The glue joint allows for the correction of small errors in the construction of either the cones or the modules. After the beam is glued, the module may be removed and remounted on the cone as necessary. The design of the foot allows this glue joint to be cleaved and remade should major repair of the module be required. After each detector module is mounted, it is electrically tested using a laser scan to verify its functionality. As each layer is completed, it is optically surveyed and the data are entered into a database. Finally, the two half-cones are connected together with the space-frame, resulting in a completed half-detector assembly.

Mount to B1 Magnets and Transport to IR

When the two half-detector assemblies are complete, they are brought to the staging area where both B1 magnets and one of the Q1 magnets have been assembled onto the beam

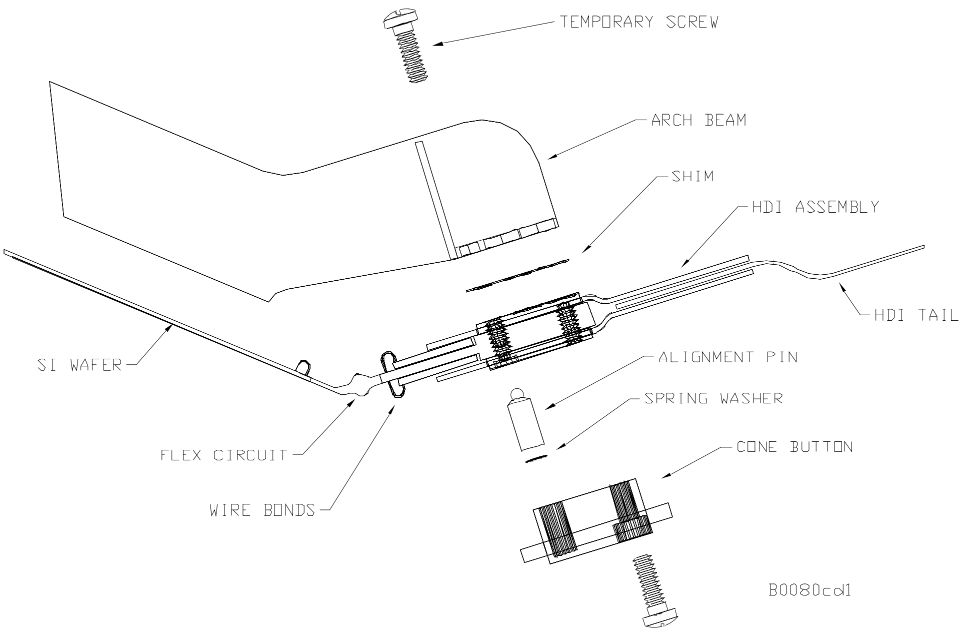

Figure 4-17. *Exploded view of foot showing mounting pins and connection to the detector module.*

pipe. Fixtures are employed to hold the cones as they are brought together and clam-shelled around the beam pipe. The two half-detector assemblies are mated, and the latches between them are closed. The cables from the HDIs are routed to the transition cards, which are mounted in cooling manifolds at the ends of the B1 magnets. The entire detector is then thoroughly tested, and an optical survey is performed. After the survey is complete, the support tube is slid over the entire assembly and mated to the Q1 magnet. The final Q1 magnet is then installed. At this point, the assembly is relatively rigid and can be transported to the interaction hall and installed in the accelerator.

The detector assembly as described above forms a rigid structure as long as the cones and space frame are connected together. This structure is supported on the B1 magnets. During transport of the support tube assembly to the interaction hall, it is possible for the magnets to have as much as a 1 mm relative motion [Bow95]. This motion is reversible, and they will return to their original alignment when installed in the accelerator. In addition, differential thermal expansion may affect the relative alignment of the magnets during periods in which the temperature is not controlled, and relative motion of the magnets and the beam pipe may occur should there be seismic activity.

The support of the detector from the magnets must allow for this motion without placing stress on the silicon wafers. In addition, the position relative to the IR must be reproducible when installed in the accelerator. These constraints are met by mounting the support cones on a pair of gimbal rings. One gimbal ring connects the forward cone to the B1 magnet so as to constrain its center in x, y, and z, while allowing rotation about the x and y axes. A second set of gimbal rings supports the cone in the backward direction in a similar manner, with an additional sleeve that allows both for motion along z and rotation about the z axis, relative to the B1 magnets.

Installation of Complete Assembly into Detector

The clearances between the vertex detector and the beam tube and B1 magnets are on the order of 1 to 3 mm. During transport of the support tube assembly, the critical clearances must be monitored in real time to ensure that no accidental damage to the detector occurs.

In its final position, the support tube assembly will be supported from the ends. Thus it is necessary that it is always supported only from the ends during installation. One possible installation scenario employs beams attached to the ends of the Q1 magnets which support the entire assembly from the ends. The beams are threaded through the drift chamber, resting on wheeled guides or tracks, which have been previously aligned with respect to the *BaBar* detector. Once the support tube is through and supported at both ends, the temporary assembly beams are removed.

4.6.4 Detector Placement and Survey

Placement Accuracy

The SVT must provide spatial resolutions on the order of 10 μm. Final locations of each of the wafers relative to each other and to the IR will be determined by track survey. This requires a certain degree of overlap of the modules within a layer. There must be overlap in z as well as ϕ, so as to accurately locate the z positions of the wafers in a single module with respect to each other. These requirements are discussed in more detail below.

Mechanical tolerances and measurements must be such that the process of track survey converges in a reasonable time. Placement of the wafers within a module should be within 25 μm and optically surveyed to a few μm. Placement of a module relative to other modules should be on the order of 100 μm and be optically surveyed to within 25 μm.

Stability of the positions should be such that resurvey with tracks is rarely required. This leads to the requirement that the relative positions of the various wafers be stable to the 5 μm level over long periods of time (months or more). The position of the entire detector structure with respect to the IR can be followed more easily, so that variations on the order of a day can be tracked. However, longer time constants are desirable.

Stability of the detector components at the 5 μm level requires a stable operating temperature. Preliminary calculations for the thermal expansion of the entire structure predict on the order of 0.5 μm/°C over the length of the active region of the detector. If the temperature inside the support tube is maintained at ±1°C, thermal expansion will not be a problem.

Survey with Tracks

Previous experiments have shown that a silicon vertex detector is best aligned using track data together with self-consistency constraints, because the point resolution in silicon is much more precise than that from a drift chamber. The most powerful alignment constraint comes from the overlap between wafers adjacent in ϕ. Another important constraint comes from muon-pair events, where knowledge of the mass, boost, and common origin of the two-track system provides a means of linking hits on opposite sides of the vertex detector. A more detailed discussion of this subject can be found in Reference [Bro95].

The SVT alignment is parameterized as a pair of nested transformations for global and local coordinates. The global alignment of the SVT describes the position and orientation of each detector layer, considered as a rigid body, relative to the drift chamber. Previous experiments have shown that a good global alignment helps speed the convergence of the local alignment and allows a simple, low-statistics check of time-dependent effects such as detector motion. The procedure depends completely on the drift chamber tracking and simply requires consistency between SVT hits and the extrapolated drift chamber track positions. Systematic effects limit the accuracy of the global alignment to ~50 μm. Assuming an accuracy of ~2 mm for the drift chamber track extrapolation in the z direction, this implies that a global alignment can be performed with ~2000 tracks. Further assuming two useful high momentum tracks per event, we find that 1000 hadronic events will suffice to perform a global alignment.

The local alignment describes the position and orientation of each silicon wafer relative to its nominal position in the layer. The local alignment of the SVT derives most of its statistical precision from the active area overlap between ϕ-adjacent wafers. The overlap directly constrains three of the six geometric degrees of freedom of each wafer, namely those that define its position in the nominal wafer plane. A further strong constraint on the wafer translation in the out-of-plane direction comes from combining the overlap information for all wafers in a given layer. This is essentially a circumference constraint, using the fact that the size of the wafers is precisely known.

The number of events needed to perform the local SVT alignment can be roughly estimated. In the following, we assume that each wafer is aligned separately for six geometric degrees of freedom, and we further assume four useable medium-momentum tracks per event. We require that the alignment should not contribute significantly (less than 5%) to the SVT

Layer	Ω	# of Wafers	# of Tracks/Wafer	# of Hadronic Events
1	2.4%	24	80	20K
2	1.8%	24	80	27K
3	1.8%	36	80	40K
4	4.0% (4.8%)	112	40	28K (24K)
5	2.0% (1.9%)	144	40	74K (77K)

Table 4-7. *Fractional overlaps (Ω) of the SVT layers and the resulting number of hadronic events required for an accurate alignment. The values in parentheses for Layers 4 and 5 correspond to the wedge modules.*

point resolution, implying that each wafer position must be known to better than 30% of its average point resolution. The overlap solid angle and number of wafers are different for each layer and are given in Table 4-7 together with the number of events needed. Layer 5 is seen to set the limit, requiring ~75K hadronic events to be precisely aligned. This is mainly because of the large number of wafers in this layer.

Overlaps do not constrain the relative positions of the different layers, nor the relative positions of the different annuli which comprise the layers. An annulus is here defined as the set of all wafers in a layer having the same z position. The relative positions of these sets of wafers are best obtained using $e^+e^- \rightarrow \mu^+\mu^-$ events; we require an alignment precision of 30% of the intrinsic resolution for each annulus of a given layer. To estimate the approximate number of muon pairs required, we assume that the three parameters defining the muon trajectory have been measured with essentially no error on one side of the SVT, and this trajectory is compared to hits on the opposite side. The tracks are presumed to have 2.5 GeV/c momentum, and the material of each layer is presumed to be 0.5% of a radiation length on average. This gives the resolution of the track extrapolation due to multiple scattering as presented in Table 4-8, which is seen to dominate the total (extrapolation $\oplus$ intrinsic) resolution. Layer 5 is again seen to require the most events to be accurately aligned, namely 7000 muon pairs or the equivalent of 27,000 hadronic events. Thus the overlaps set the overall minimum number of events needed for alignment.

4.6.5 Detector Monitoring

Position Monitoring Systems

Although the final placement of the silicon wafers will be measured and monitored with charged particles which traverse the silicon detector and drift chamber, two displacement monitoring systems will be designed to measure relative changes in the position of the silicon

Layer	Track σ	Hit σ	# of Annuli	# of Muon Pair Events
1	25 μm	10 μm	4	225
2	30 μm	10 μm	4	325
3	50 μm	10 μm	6	1300
4	80 μm	30 μm	7	4000
5	100 μm	30 μm	8	7000

Table 4-8. *Properties of the different SVT layers, and the number of muon-pair events these imply for an accurate alignment.*

detector with respect to the machine elements and the support tube. One displacement monitoring system will be used to monitor relative positions during transportation of the support tube with the silicon detector inside it and during data taking. This system consists of either capacitive displacement monitors or LED-photodiode reflection monitors which are sensitive to relative displacements between the silicon detectors and the machine components such as the beam pipe, magnets, and support tube.

In addition, a laser system will monitor displacements of the outer layer of detectors with respect to the drift chamber during data taking. Given that the SVT layers are not mounted on the same support as the drift chamber, it is possible that motion between the two will occur. To monitor this motion, short infrared laser pulses are brought in with fiber optics (*e.g.,* 50 μm core diameter) which are attached to the drift chamber. The laser light shines through small holes in the support tube and reaches the outer layers of the silicon detector. The resulting signals are read out with the normal silicon detector readout system.

Displacement monitoring systems based on capacitive sensors and laser pulses have been used successfully by many experiments in the past [Acc94, Bin93, Cac92, Bre91]. The two systems are complementary in technology and in their sensitivity to different kinds of displacements. For example, the L3 experiment [Acc94] attains a resolution of a few microns with the laser monitoring system for transverse displacements. With their capacitive monitoring device, they attain a resolution of 1–2 μm in the radial direction and 5–10 μm in the transverse direction.

Radiation Monitoring

To protect the silicon detector system against potentially damaging beam losses, and to monitor the total radiation dose that the detectors and electronics receive, silicon diodes will be installed close to the beam pipe in the vicinity of the SVT. If the radiation dose exceeds a certain threshold, a beam-dump signal will be sent to the PEP-II control room. This sort of radiation protection system is already used in all of the LEP experiments. The silicon diodes will be read out with various gains so that the dynamic range of the full set of sensors covers quiet running and potentially damaging conditions.

4.6.6 R&D Program

The following R&D projects are planned before the design of the SVT mechanical configuration is finalized.

Cable. Prototypes of the cable from the hybrid to the transition card will be constructed. This will allow proof of the details of cable routing and mechanical robustness. It will also allow the electrical properties to be measured to verify simulations.

Hybrid. Realistic mechanical modules of the high-density interconnect (HDI) are required. The HDI is a critical element both in the cooling of the electronics and the mounting of the detector modules. Models will be tested for heat transfer capability and for module mounting schemes.

Inner Layer Sextant. A full-scale mock up of the inner layer sextant will be constructed. It will be used to verify thermal stability calculations and to investigate the effects of nonuniform beam pipe cooling. It will also be used to test and practice assembly techniques.

Arch Modules. Full-scale mock ups of the arch detector modules will also be constructed and used with the prototype cones to verify cooling and mounting techniques.

Cones and Space Frame. A set of prototype cones and a space frame will be built to provide realistic tests of cooling, mechanical rigidity, and thermal stability. In addition, they will be used to design assembly fixtures and test assembly techniques for mounting modules onto cones.

Full-Scale Model of the IR. A model of the B1 magnets and the beam pipe near the IP will be constructed. This will aid in identifying interference problems and verifying mounting schemes. It will also provide a test bed for the design of various installation fixtures.

4.7 Services, Utilities, and ES&H Issues

4.7.1 Services and Utilities

The vertex detector requires the following services, which must be brought inside the support tube to a location near the outboard end of the B1 magnet.

Data and Control Lines. Approximately 624 wire pairs are required on each side of the detector to service the 52 transition cards. These consist of 2 clock, 2 control and 2 data links per readout section providing redundancy for each signal. The interface cards will multiplex these into 26 optical links for data and 26 optical links for control on each side of the detector.

Power. The readout ICs will require three low-voltage power supplies (two analog, one digital), and the transition card will require at least one power supply. This amounts to four power and return cables per readout section, or 416 supplies and returns for each half of the detector. Each line carries only a few watts, except for the transition card supply which could be a few tens of watts. In addition to these low voltage supplies, we require 52 detector bias supplies per detector half. The bias voltage ranges from 40–80 V, and the detector draws very little current. The power supplies will all be specially procured to the vertex detector specifications in order to control electronic noise.

Cooling Water. The readout electronics and transition cards will be water cooled. Two sets of water connections will be required for each cone (since each cone is constructed from two halves), and one water connection for each transition card support. The cooling water will be supplied by a special low volume chiller system dedicated to the vertex detector.

Dry Air or Nitrogen. The vertex detector requires a dry, stable environment, and dry air or nitrogen from each side is planned.

4.7.2 ES&H Issues

There are very few ES&H issues which impact the construction, assembly, and operation of the SVT. The detector bias voltages can exceed 50 V and therefore qualify as high voltage, though they are extremely low current. There is a potential for water leaks from the cooling systems. This water would be confined to the volume inside the support tube and would pose a danger primarily to the vertex detector itself. The laser monitoring system would employ Class 1 lasers which pose essentially no danger.

References

[Acc94] M. Acciarri *et al.,* CERN–PPE/94–122, to be published in *Nucl. Instr. Methods* A.

[Bag94] G. Bagliesi *et al.,* "Test Beam Results from Prototypes of the Upgraded ALEPH Vertex Detector," ALEPH Note 94–069 (1994).
P. Coyle *et al.,* "ALEPH Vertex Detector," presented at the Sixth Pisa Meeting on Advanced Detectors, Elba, Italy (1994).

[Bat92] G. Batignani *et al.,* "Experience with the ALEPH Silicon Vertex Detector," *Nucl. Instr. Methods* **A315**, 121–124 (1992).

[Bat94a] G. Batignani *et al.,* "Double-Sided Silicon Strip Detectors in Pisa," presented at the Sixth Pisa Meeting on Advanced Detectors, Elba, Italy (1994).

[Bat94b] G. Batignani *et al.,* "Specifications of the Silicon Detectors for the Vertex Detector," *BaBar Note* # 196 (1994).

[Bil84] P. Billoir, *Nucl. Instr. Methods* **225**, 352 (1984).

[Bin93] N. Bingefors *et al., Nucl. Instr. Methods* **A328**, 447 (1993).

[Bow95] G. Bowden, "IP Beam Pipe/Vertex Detector Clearance" (1995) (unpublished).

[Bre91] A. Breakstone *et al., Nucl. Instr. Methods* **A305**, 39 (1991).

[Bro95] D. Brown, "*BaBar*Vertex Detector Alignment Requirements," *BaBar Note* # 148 (1995).

[Cac92] M. Caccia *et al., Nucl. Instr. Methods* **A315**, 143 (1992).

[Col94] T. Collins, "*BaBar* Hybrid Design Considerations," *BaBar Note* # 222 (1994).

[DiB95] D. DiBitonto *et al.,* "Ultra-thin, High Precision Flex Cable for the L3 Silicon Microvertex Detector," to be submitted to *Nucl. Instr. Methods* A.

[For94] F. Forti, "TRACKERR Studies for Optimization of Vertex Detector Resolution," *BaBar Note* # 195 (1994).

[Hol89] P. Holl *et al.,* IEEE Trans. Nucl. Sci. **36**, N. 1, 251 (1989).

[Inn93] W. Innes, "TRACKERR, A Program for Calculating Tracking Errors," *BaBar Note* # 121 (1993).

[Joh95a] Front-End Working Group, "Requirements Specifications for Silicon Vertex Detector Readout Chip," edited by R. Johnson, *BaBar Note* # 213 (1995).

[Joh95b] Front-End Working Group, "Target Design Specifications for Silicon Vertex Detector Readout Chip," edited by R. Johnson, *BaBar Note* # 214 (1995).

[Lev] Max Levy Autograph, Inc., Philadelphia, PA (USA).

[Lev94] M. Levi, "The Impact of Backgrounds on the Silicon Architecture and Detector Trigger," *BaBar Note* # 136 (1994).

[Roe94a] N. Roe and L. Kerth, "Vertex Detector Solid Angle Coverage," *BaBar Note* # 160 (1994).

[Roe94b] N. Roe, "Silicon Strip Resolution for Time-Over-Threshold Readout," *BaBar Note* # 161 (1994).

[Sny94] A. Snyder, "The Effect of Vertex Cuts on CP Reach," *BaBar Note* # 177 (1994).

[Ton94] G. Tonelli *et al.,* "Development of Double-Sided μstrip Detectors for High Radiation Environment," presented at the *Sixth Pisa Meeting on Advanced Detectors,* Elba, Italy (1994).