The Origami Chip-on-Sensor Concept
for Low-Mass Readout of Double-Sided Silicon Detectors

M. Friedl[a], C. Irmler[a], M. Pernicka[a]

[a] Institute of High Energy Physics, Nikolsdorfergasse 18, A-1050 Vienna, Austria

`friedl@hephy.at`

Abstract

Modern front-end amplifiers for silicon strip detectors offer fast shaping but consequently are susceptible to input capacitance which is the main contribution to the noise figure. Hence, the amplifier must be close to the sensor which is not an issue at LHC, but a major concern at material budget sensitive experiments such as Belle or the ILC detector.

We present a design of a silicon detector module with double-sided readout where thinned front-end chips are aligned on one side of the sensor which allows efficient cooling using just a single, thin aluminum pipe. The connection to the other sensor side is established by thin kapton circuits wrapped around the edge – hence the nickname origami.

I. Introduction

The current Silicon Vertex Detector (SVD2) [1] of the Belle experiment at KEK (Tsukuba, Japan) is composed of four layers with a total of 246 double-sided silicon sensors as shown in fig. 1. Its polar acceptance extends from $17°$ to $150°$, which allows to place the readout electronics outside of that region. This is important as the KEK-B is a low-energy machine ($8\,\mathrm{GeV}$ electrons on $3.5\,\mathrm{GeV}$ protons) which requires careful consideration of material in order to minimize multiple scattering. Up to three ganged (concatenated) silicon sensors are read out by the VA1TA [2] front-end amplifier which has a low noise figure of $ENC = 180\,\mathrm{e} + 7.5\,\mathrm{e/pF}$ due to its slow shaping time of $800\,\mathrm{ns}$.
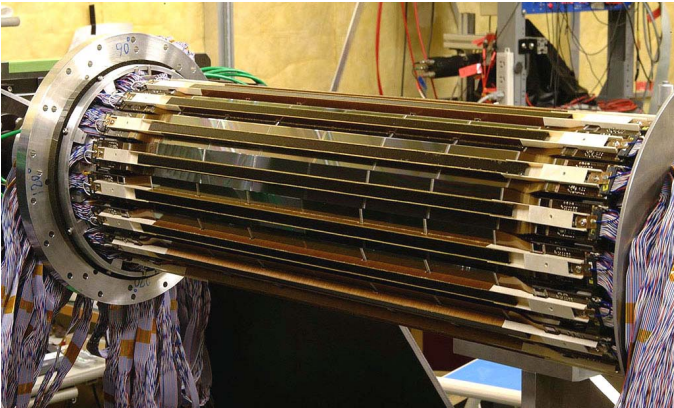

Figure 1: The Belle Silicon Vertex Detector (SVD2) prior to insertion. The active length of the outermost (and thus visible) silicon ladders is about $50\,\mathrm{cm}$. A total area of about $0.5\,\mathrm{m^2}$ is covered by all silicon sensors of the SVD2.

A major upgrade of both accelerator and experiment is planned for the period of 2009-2012. The ultimate goal of the SuperKEK-B is to achieve a luminosity of up to $5 \times 10^{35}\mathrm{cm^{-2}\,s^{-1}}$, which is about 30 times higher than now. Consequently, the SVD readout (as well as many other parts of the experiment) must be replaced by a faster and more powerful one. In particular, the SVD2 is already pushing the limits now in terms of occupancy of the innermost layer (about 10%) and dead time (a few percent). These issues can be addressed by a readout chip with shorter shaping time, faster readout clock and an internal pipeline.

II. APV25

The APV25 [3] readout chip, originally developed for CMS at CERN, was identified to meet all the requirements of the Silicon Vertex Detector of SuperBelle. It has a shaping time of $50\,\mathrm{ns}$ and thus intrinsically reduces the occupancy by a factor of 16 (a factor of 12.5 was found by measurement as one has to consider the actual shaping waveform as well as thresholds). Moreover, it has a 192-cell deep pipeline which allows continuous measurement without dead time. Due to the interpretation of 3-bit symbols on the trigger line, there is in fact a very small dead time of 3 clock cycles, or $75\,\mathrm{ns}$, which is however irrelevant in the experiment.

The APV25 is designed for operation at $40\,\mathrm{MHz}$ and offers the possibility to read out several sampled values along the shaping curve with each trigger. By means of post-processing, this feature allows to determine the actual particle timing with a precision of a few nanoseconds [4] and will also be implemented in SuperBelle using look-up tables in FPGA devices.

Unfortunately, compared to the slow VA1TA chip, a higher noise figure of $ENC = 250\,\mathrm{e} + 36\,\mathrm{e/pF}$ inevitably comes along with the faster shaping. Hence, one has to minimize the capacitive load at the front-end in order to keep the noise as low as possible. While the signal size is given by the canonical $300\,\mu\mathrm{m}$ thickness of the silicon sensors, it is the noise that essentially determines the resulting signal-to-noise ratio (SNR). Thus, if fast shaping is desired, we conclude that the reduction of capacitance is the only practical way to increase the SNR, which influences spatial precision and ultimately the impact parameter resolution. In practice, a minimum cluster SNR (sum of cluster signal divided by the square sum of RMS noise of all strips in the cluster) of ten is considered mandatory for reasonable operation of the silicon detector.

III. LADDER GEOMETRY

Ganging of sensors is possible in the current SVD2 as the noise slope of the VA1TA chip is shallow. In the innermost layer, there are just two sensors per ladder, which are individually read out from either side. As the ladders get longer with increasing radius, readout is still done at the edges in the same way as in the innermost layer while sensors are ganged (concatenated). This results in spatial ambiguities but allows the same number of readout chips for each ladder and avoids any readout electronics inside the active volume. In case of the long strip dimension, ganging is simply accomplished by connecting adjacent sensors with bond wires. For the opposite side, with strips orthogonal to the ladder axis, thin flex circuits are used to contact the same strip of each sensor to a single readout channel. The layer four ladders are composed of six silicon detectors of which three are ganged and read out from either side. Fig 2 shows a photograph of all four ladder types.

Figure 2: The ladders of innermost (bottom) to outermost (top) layer of the current Belle Silicon Vertex Detector (SVD2). In the longest ladders, three sensors are ganged and read out from either side.

In SVD2, the outermost layer is located at a radius of $8.9\,\mathrm{cm}$. Despite of ganging, the minimum cluster SNR is in the order of 13 due to the slow shaping. For SuperBelle, the SVD shall extend to significantly larger radii ($\approx 14\,\mathrm{cm}$) and consist of five layers, where the innermost is a double layer and will probably be equipped with DEPFET pixel detectors [5].

IV. GANGED SENSORS WITH APV25

As indicated in the previous section, the noise figure of the APV25 is significantly influenced by the load capacitance presented at its input which is the sum of the total strip capacitance and the intermediate connections (flex circuits). Hence we can suspect that ganging of sensors would lead to a poor SNR, especially in larger structures than the current SVD2.

Using prototype double-sided silicon sensors made from 4" wafers for the future SVD, we built a module with two ganged sensors read out by four APV25 chips on each side. In order to compare the SNR of a single sensor with that of two ganged sensors, the connections between first and second sensors were only partially bonded. Thus, we indeed found that the SNR of a single sensor is good, but with two ganged sensors it already drops to an unacceptably low value, as shown in tab. IV.. Conse-

quently, there is no chance to operate the APV25 with more than one sensor, especially as those might be made from 6" wafers in the future.

	ganged		single	
	p-side	n-side	p-side	n-side
Cluster SNR	9.4	10.1	13.1	13.9

Table 1: Cluster SNR comparison between a single and two ganged 4" double-sided sensors using the APV25. The readout pitch is $50\,\mu\mathrm{m}$ on the p-side and $152\,\mu\mathrm{m}$ on the n-side; both are equipped with a floating intermediate strip. The values were obtained in a beam test with $3\,\mathrm{GeV}$ electrons.

V. CHIP-ON-SENSOR

Hence, the only way to maintain a good SNR is to put the APV25 as close as possible to the sensor strips, which leads to the chip-on-sensor concept. This means that the readout chip, together with its hybrid circuit, sits on top of the sensor such that there is no need for long flex fanouts anymore. A schematic view of such an assembly is shown in fig. 3, where the "hybrid" is made of a (double-layer) kapton circuit which is separated from the sensor by a layer of rigid foam called Rohacell [6] which provides both electrical isolation and thermal insulation. The APV25 chip can be thinned down to $100\,\mu\mathrm{m}$ (or less) to reduce the overall material budget and make the whole assembly more homogeneous. Mechanically, the structure is held by both the cooling pipe and the plastic rib in parallel which could be made of Zylon [7], which was already used in the SVD2, or another stiff and light-weight material.
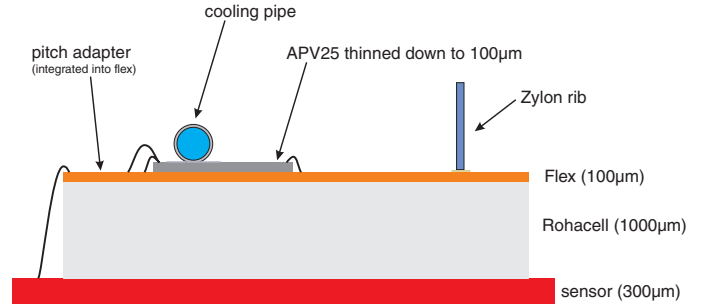

Figure 3: Schematic cross-section of the chip-on-sensor assembly. Please note that this drawing is not to scale. See main text for explanation.

A. Flex_Module

In 2006, we built a demonstrator prototype module called "Flex_Module" where the short strips (n-side) are read out using the chip-on-sensor concept, while the long strips on the p-side are connected in a conventional way even though the associated "hybrid" is also made on kapton. This module uses a carbon fiber pipe for both cooling and as a stiffener, which turned out to be more massive than other solutions to be discussed later. Fig. 4 shows photographs of both sides of this module.

As expected, this concept of short front-end connections on the n-side yields much better signal-to-noise than the traditional

assembly where all strip signals are routed to the side of the sensor using long fanouts. Tab. A. compares the cluster SNR values of this module to the ones of the conventional module with a single sensor already shown earlier in tab. IV..

n-side
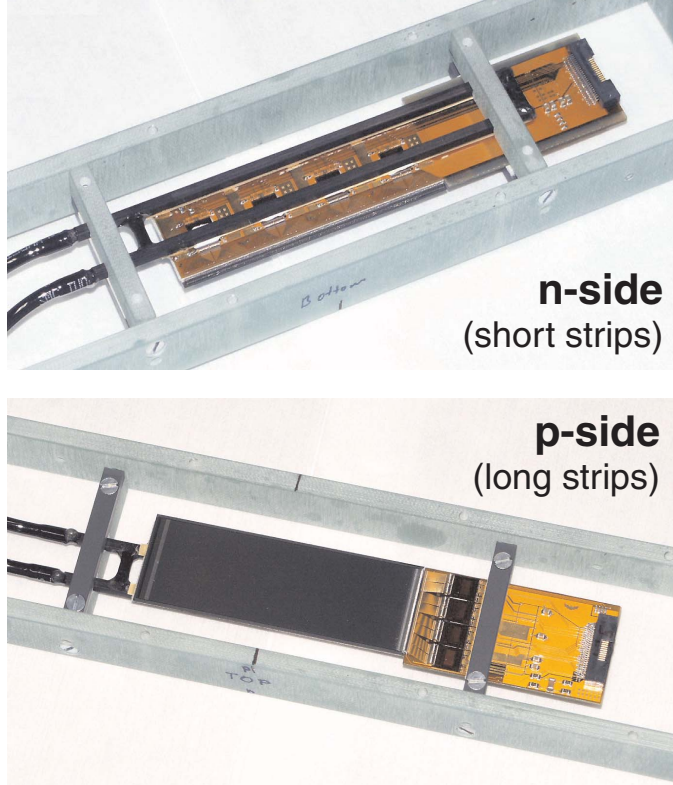(short strips)

p-side
(long strips)

Figure 4: Both sides of the Flex_Module. The n-side (top) is built according to the chip-on-sensor concept, while the p-side (bottom) is conventionally read out from the edge.

	Flex_Module		Conventional	
	p-side	n-side	p-side	n-side
Cluster SNR	13.8	18.4	13.1	13.9

Table 2: Cluster SNR comparison between the Flex_Module utilizing the chip-on-sensor concept for the n-side readout and the conventional readout where all strips are fanned out to the edge of the sensor. These measurements use the same sensor design and were performed in the same beam test setup as described in the caption of the previous table.

Obviously, there is little difference on the p-side as both modules are read out from the side in the same way with just slightly different lengths of the fanouts. On the n-side, however, the difference is substantial and moreover, the chip-on-sensor figure does not depend on the number of detectors lined up in a ladder, as each sensor is read out independently.

B. Cooling

As each APV25 dissipates about $350\,\mathrm{mW}$ and finally there will be several thousand chips built into a small and sealed volume, cooling of the chips is absolutely mandatory. In order to explore various cooling options, we built a thermally insulated cooling channel with dimensions that resemble those of the space around a ladder in the future SVD. 24 SMT resistors are lined up on a carrier board and fed by a current as to dissipate the same amount of power as APV chips. Four different methods of cooling were tested:

- Air

- Water

- Heat pipe

- TPG

Air cooling suffers from the problem that its heat capacitance is very low. Hence, guided air would need extreme pressure and flow which is not feasible in practice. Alternatively, even free air flow without a pipe would need flow rates that could become dangerous to bond wires and moreover, there would be the problem of bottlenecks at inlet and outlet. Thus we conclude that air cooling is unrealistic.

Water has an enormous heat capacitance and thus is well suited for cooling requiring very little flow. Hence, one can use a pipe of small diameter and with thin walls in order to achieve a low material budget. We found aluminum pipes with an outer diameter of $2\,\mathrm{mm}$ and a wall thickness of $0.2\,\mathrm{mm}$ to be sufficient. Using a flow of only $1\,\mathrm{ml/s}$, we could achieve a coolant temperature difference of only $3^{\circ}\mathrm{C}$, which is in good agreement with theoretical calculation. Clearly, water cooling has a significant damage potential in case of leakage. In particular, there are worries about the long-term behavior of aluminum with water and it was suggested to use paraffin oil instead, which has already successfully been used to cool the beam pipe of the Belle experiment.

We also tried heat pipes which contain a liquid in a sealed volume where heat transfer is obtained by means of internal evaporation on the hot side and condensation at the cold side. Unfortunately, thin heat pipes did not meet the requirements of heat transportation, whereas thick heat pipes present too much material, as they are usually made of copper which has a low radiation length. Even though special aluminum heat pipes were made for space applications, their long term reliability remains doubtful.

The last option relies on heat conduction rather than the flow of a coolant and thus would completely avoid any risk of leakage. We used a bar of thermal pyrolytic graphite (TPG) [8], which features extremely high thermal conductivity – about four times higher than that of copper – along two axes. Despite that property, we did not achieve satisfying results with a cross-section of $5 \times 2\,\mathrm{mm}^2$ as the conductance is still too low. This could be improved by using a larger cross-section, but then also the material budget would increase accordingly. In conclusion, the TPG might be marginally suitable for an inner layer with limited length and heat load, but certainly not for the outer layers.

Overall, liquid cooling appears to be the only feasible option, but great care has to be taken to avoid potential risks due to leakage.

C. Origami

The chip-on-sensor technology presented above allows readout of single-sided silicon detectors, but it is not straightforward to apply this concept to double-sided sensors. This would imply a duplication of "hybrid" as well as cooling and moreover present some challenges in routing of fanouts and/or the cooling pipe when dealing with the long strips (parallel to the ladder axis).

The Origami chip-on-sensor concept is a solution which overcomes these limitations by putting all the chips aligned onto the sensor side with the short strips, which allows a single straight cooling pipe to serve all dies. The short strips are routed to the chips with an integrated pitch adapter as done in the Flex_Module (fig. 4), whereas the long strips of the opposite side are connected by flex fanouts wrapped around the edge, as shown in fig. 5. A 3D rendering in more realistic colors can be found in fig. 6.

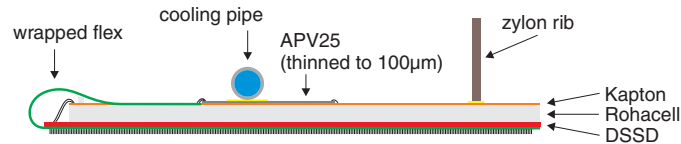## a) Top view
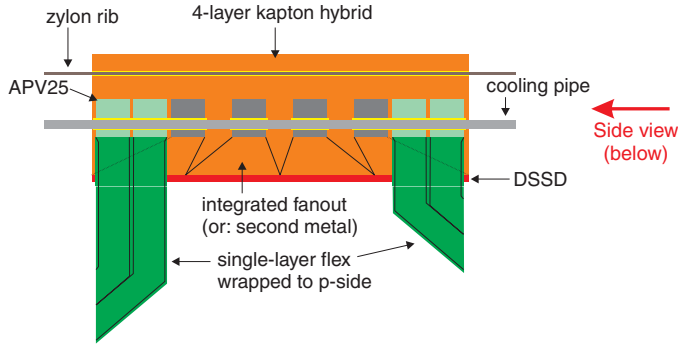
b) Side view (cross-section)

Figure 5: Drawings of the Origami chip-on-sensor assembly. Please note that both top and side views are dimensional, but on a different scale. a) Top view: The four APV25 chips which read out the strips on the opposite side are shown in green for clarity and the flex pieces to be wrapped around the edges are straightened out. b) Side view: This drawing closely resembles the single-sided chip-on-sensor concept (fig. 3) except for the additional wrapped flex donating the name "Origami".

The APV25 chips reside at the voltage potential of the strips that they read out, which implies that readout chips for p and n sides, even though they are all lined up, are separated by the bias voltage of the sensor, typically $80\,\mathrm{V}$ at Belle. Hence, the kapton hybrid will be manufactured with four layers, where two layers are devoted to each side in order to achieve the best isolation possible. Naturally, the cooling pipe must also be isolated from the chips, which can be achieved by a thin heat-conductive foil with high electric strength.

One could argue that the cooling pipe, being round, cannot establish a good thermal contact with the chip. This can easily be solved by slightly flattening the pipe locally at the chip positions, which we confirmed with test in the thermal channel. Moreover, thermal grease will be used to improve the heat transfer. The position of the pipe will be chosen such that the preamplifier/shaper part of the chip is touched which is not only the noise-sensitive element but also dissipates about half the power consumed by the whole chip.
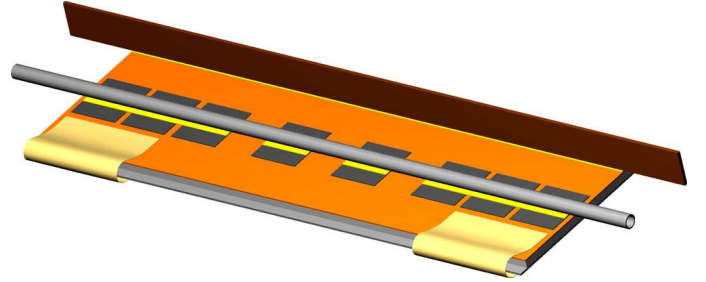

Figure 6: Perspective view of the Origami module.

Clearly, the most challenging aspect of the Origami module is its assembly. Several jigs will be required for proper handling and the order of gluing and wire-bonding steps must be well considered. Probably the most critical item is the wrapping of the p-side flexes which will then get close to the underlying wire bonds of the n-side strips (see fig. 5 b). One option would be to protect those by a glob-top, but there are some worries about its long-term reliability, especially in conjunction with thermal stress and radiation. Hence, we prefer a purely mechanical solution such as a small piece of Rohacell glued onto the fanout behind the bond wires which would simply avoid the bent flex from falling down onto the bond wires. The third and most expensive solution would be to integrate the fanout into the sensor by using a second metal layer routing. The drawback of that option would be an increased capacitive load which is exactly what should be avoided by the chip-on-sensor concept.

D. Material Budget

Obviously, the chip-on-sensor method adds material to the sensitive volume at the gain of significantly improved signal-to-noise ratio. A calculation considering all components involved in both the conventional and the Origami concepts is summarized in tab. D. and yields a relative increase of about 50% for the latter.

	Conventional	Origami
Material Budget [$\%X_0$]	0.48	0.72

Table 3: Comparison of the averaged material budget between conventional and Origami chip-on-sensor concepts.

The Origami chip-on-sensor concept is a trade-off between material budget and signal-to-noise. A simulation of the Super-Belle Silicon Vertex Detector showed that additional material is prohibitive in the innermost two layers due to degradation of the vertex resolution by multiple scattering, but no problem

for layers three to five, where the impact onto the vertex accuracy is limited. Fig. 7 shows a possible layout of the SuperBelle SVD, where DEPFET pixel detectors are foreseen in the innermost (double) layer. All double-sided silicon strip sensors are assumed to be made from 6" wafers and thus will likely need six readout chips on each side. We currently do not anticipate any problems arising from the transition from four to six APVs regarding the Origami concept.
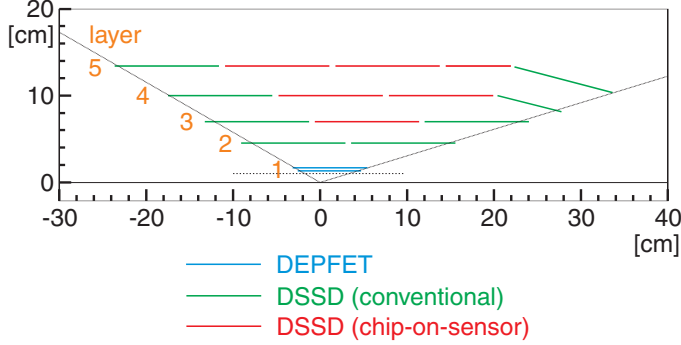

Figure 7: Conceptual design of the Silicon Vertex Detector for Super-Belle, consisting of a double pixel layer surrounded by four silicon strip layers. The edge detectors will be conventionally read out from the sides, whereas the inner sensors will make use of the Origami chip-on-sensor concept. All silicon strip sensors are made from 6" wafers.

As seen in fig. 7, the sensors closest to the edge will be read out in a conventional way since sufficient signal-to-noise can be achieved in that configuration (see tab. A.). The detectors inside the ladders, however, are read out using the Origami chip-on-sensor concept as ganging is not possible with the APV25 chip. Thanks to the choice of 6" wafers, the arrangement is perfectly compliant with the simulation requirements mentioned above, i.e. additional material is only introduced from layer three on.

VI. SUMMARY AND OUTLOOK

Motivated by the upgrade of the Belle Silicon Vertex Detector – but not restricted to – we have presented the Origami chip-on-sensor method which can be used to read out double-sided silicon detectors with cooling of the readout chips. The main focus lies on minimization of the material in the active volume which is achieved by aligning all thinned APV25 chips on one side and wrapping kapton flexes around the edge to connect half of the chips to the strips on the opposite side.

The averaged material budget of such a module was calculated to be $0.72\% \, X_0$ in comparison to 0.48% found for the conventional construction where the readout hybrid sits on the side and thus, in case of Belle, outside of the active region. Such an assembly, in conjunction with sensor ganging, is possible with slow amplifiers, but prohibitive with fast shaping which is inherently more susceptible to noise and thus capacitive load. Yet, fast shaping is a requirement imposed by high luminosity and thus the Origami chip-on-sensor concept appears to be the only viable solution for silicon strip detector readout in SuperBelle.

In the near future, we will design the flex circuits and build a prototype Origami module in order to gain experience with the challenging assembly procedure which may lead to design optimizations. Source and beam tests will be performed to evaluate the performance of the module, even though we have no doubt about its functionality, as we already performed measurements with prototype assemblies such as the "Flex_Module" which employed the chip-on-sensor concept on one side.

Eventually we will move to a 6" sensor design implying the readout of six APV25 chips for each side, which appears to be straightforward from our current point of view.

REFERENCES

[1] Z. Natkaniec *et al.*, **Status of the Belle Silicon Vertex Detector**, Nucl. Instr. and Meth. A 560 (2006), 1-4

[2] **VA1TA Chip**, `http://www.ideas.no/products/ASICs/pdf/Va1Ta.pdf`

[3] M.French *et al.*, **Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker**, Nucl. Instr. and Meth. A 466 (2001), 359-365

[4] M.Friedl, C.Irmler, M.Pernicka, **Time Resolution of a Few Nanoseconds in Silicon Strip Detectors Using the APV25 Chip**, 12th Workshop on Electronics For LHC and Future Experiments (2006), 460-464

[5] P.Fischer *et al.*, **Progress towards a large area, thin DEPFET detector module**, Nucl. Instr. and Meth. A 582 (2007), 843-848

[6] Rohacell (`http://www.rohacell.com`) is a rigid type of styrofoam produced by Degussa.

[7] Zylon (`http://www.toyobo.co.jp/e/seihin/kc/pbo/`) is a stiff, but light-weight material made by Toyobo.

[8] Thermal pyrolytic graphite (TPG) (`http://www.advceramics.com/products/tc1050/index.html`) is manufactured by Momentive.