Bringing Silicon Microsystems to Space

Manufacture, Performance, and Reliability

BY

JOHAN KÖHLER

ACTA UNIVERSITATIS UPSALIENSIS
UPPSALA 2001
Dissertation for the Degree of Doctor of Philosophy in Engineering Science with Specialisation in Microsystems Technology presented at Uppsala University in 2002

ABSTRACT


The incorporation of extremely compact multifunctional microsystems is a highly profitable long-term approach in spacecraft design. These systems bring substantial launch-cost reductions, and enable exciting space exploration and science missions.

Silicon microsystems technology is an adequate choice for the multifunctional microsystem development. However, the development of basic microsystems technology cannot be financed within application-specific space missions. Rather, the microsystems technology should be matured through fundamental research.

Silicon microsystems technology was used to develop a cold gas microthruster system suitable for minute movements of spacecraft (low ΔV). In a hybrid integration, the system unit contains three silicon microsystem parts with four individual thrusters in total, together with external control electronics. The total mass is 0.35 kg.

Further integration will result in a mass of 0.08 kg. Complete system integration means that all package and interconnection levels are integrated into the silicon microsystem units. Several vital issues must be addressed, e.g. the reliable bonding of silicon wafers, the microfabrication process compatibility, and the manufacture process sequence. A graphical tool is introduced for process sequence evaluation.

Wafer bonding is used as fabrication process, assembly tool, and packaging technique. The quality and reliability of the bonded interfaces must be assessed in order to secure the operation of the microsystems in space. Therefore, statistical methods for burst test evaluation have been developed.

Weibull fracture probability functions have been derived in order to interpret the bond quality. In addition, rank-sum tests on spot series and analysis of variance are performed for bond quality diagnostics. The dependence on annealing temperature and surface-activation are presented, together with diagnosed degradation of insufficiently annealed bonds due to different spaceflight environments (thermal cycling, vibration, γ-irradiation).

Johan Köhler, The Angström Space Technology Centre, Uppsala University, Box 534, SE–751 21 Uppsala, Sweden

© Johan Köhler 2001
ISSN 1104-232X
ISBN 91-554-5196-9

Printed in Sweden by Kopieringshuset AB, Uppsala 2001
It was on a dreary night of November, that I beheld the accomplishment of my toils.

Victor Frankenstein
PAPERS INCLUDED IN THIS THESIS

These papers are appended to the thesis, and will be referred to in the text by their Roman numerals:

I. **Multifunctional Design of Microsystems for Space Applications**
   Johan Köhler and Lars Stenmark

II. **A Hybrid Cold Gas Microthruster System for Spacecraft**
    Johan Köhler, Johan Bejhed, Henrik Kratz, Fredrik Bruhn, Ulf Lindberg, Klas Hjort, and Lars Stenmark

III. **The Manufacture of an Integrated Silicon Cold Gas Micronozzle System for Spacecraft**
     Johan Köhler, Johan Bejhed, Ulf Lindberg, Klas Hjort, and Lars Stenmark

IV. **Mechanical Reliability of Silicon Microsystem Fusion Bords in Spaceflight Environment**
    Johan Köhler, Kerstin Jonsson, Fredrik Bruhn, Laurent Marchand, and Lars Stenmark

V. **Weibull Fracture Probability for Silicon Wafer Bond Evaluation**
   Johan Köhler, Kerstin Jonsson, Staffan Grell, and Lars Stenmark

VI. **Oxygen Plasma Wafer Bonding Evaluated by the Weibull Fracture Probability Method**
    Kerstin Jonsson, Johan Köhler, Christer Hedlund, and Lars Stenmark

VII. **Weibull Fracture Probability for Characterisation of the Anodic Bond Process**
     Åse Richard, Johan Köhler, and Kerstin Jonsson

VIII. **Silicon Fusion Bond Interfaces Resilient to Wet Anisotropic Etchants**
      Johan Köhler, Carola Strandman, Örjan Vallin, Christer Hedlund, and Ylva Bäcklund
The author’s contribution to the appended papers

I. Major part of the planning, invention, evaluation, and writing.

II. Significant part of planning and experimental work. Major part of evaluation and writing.

III. Major part of planning, experimental work, and evaluation. All of the writing.

IV. Substantial part of the experimental work. Major part of planning and evaluation.

V. Significant part of experimental work and writing. Substantial part of evaluation. Major part of planning.

VI. Part of the experimental work. Significant part of planning and writing. Substantial part of evaluation.

VII. Significant part of planning and writing. Substantial part of evaluation.

VIII. Substantial part of planning, experimental work, and evaluation. All of the writing.

Paper II is reproduced with permission from Elsevier Science. Paper IV is reproduced with permission of the International Astronautical Federation. Paper V is reproduced with permission from The Electrochemical Society, Inc. Papers VII and VIII are reproduced with permission from the Institute of Physics Publishing (www.iop.org/Journals/JMM).
# Table of Contents

ABSTRACT .............................................. 2

PAPERS INCLUDED IN THIS THESIS ............ 4
  THE AUTHOR’S CONTRIBUTION TO THE APPENDED PAPERS 5

TABLE OF CONTENTS ................................. 6

2. MULTIFUNCTIONAL MICROSYSTEMS .......... 8
  2.1 MULTIFUNCTIONAL MICROSYSTEM STRATEGY 10
  2.2 MULTIFUNCTIONAL MICROSYSTEM DESIGN ... 11
  2.3 CURRENT AND FUTURE COMPETITION ... 12

3. WHY SPACE? ......................................... 14

4. MICROSYSTEM MANUFACTURE ................. 16
  4.1 MERITS AND FLAWS OF SILICON MICROSYSTEMS TECHNOLOGY 16
  4.2 MICROSYSTEM FABRICATION STRATEGY AND PROCESS SEQUENCE EVALUATION 19
  4.3 MICROSYSTEM PERFORMANCE EVALUATION 24

5. EXAMPLE: MICROPUSION SYSTEM ............. 26
  5.1 THE HYBRID COLD GAS MICROTHRUSTER SYSTEM 29
  5.2 EVALUATION OF THE HYBRID COLD GAS MICROTHRUSTER SYSTEM 32
  5.3 OTHER MICROPUSION PRINCIPLES 36

6. DEGRADATION OF BOND QUALITY ............. 37
  6.1 POSSIBLE DEGRADATION MECHANISMS .... 37
  6.2 GROUND TESTS OF SPACEFLIGHT ENVIRONMENT 40
  6.3 RANK-SUM TEST DIAGNOSIS 41

7. WEIBULL FRACTURE PROBABILITY METHOD ........ 43
  7.1 MODEL ........................................ 46
  7.2 PARAMETER INTERPRETATION .......... 48
  7.3 ANALYSIS OF VARIANCE ................ 50

8. CHALLENGES AND OPPORTUNITIES OF THE SPACE INDUSTRY 52
  8.1 TECHNOLOGY AS A SCIENTIFIC ENDEAVOUR; BUILDING MATURITY 54

SUMMARY OF APPENDED PAPERS .................. 57
  SUMMARY OF PAPER I 57
  SUMMARY OF PAPER II 58
  SUMMARY OF PAPER III 58
  SUMMARY OF PAPER IV 58
  SUMMARY OF PAPER V 59
  SUMMARY OF PAPER VI 59
  SUMMARY OF PAPER VII 60
  SUMMARY OF PAPER VIII 60

ACKNOWLEDGEMENTS ................................. 61

REFERENCES .......................................... 62

Johan Köhler
1. INTRODUCTION

Louder and funnier!
Fozzie Bear

The tiniest eyes may best study the vastest entities. Space exploration is dying to use the full potential of microsystems technology to its best advantage.

This thesis addresses the bringing of silicon microsystems to space. What are the motivations for such an endeavour, and what niche can be found for microsystem use in space? The results and conclusions of my own research form my basis for discussing these questions.

I relate the steps necessary for eventually entering this niche. The arguments are illustrated by experiences and conclusions from micropropulsion system manufacture and evaluation, and wafer bonding. Indeed, the obstacles perceived and ambiguity felt by the end user towards incorporating microsystems in their spacecraft emerge in all conceivable instances. Examples include the technical performance, the alleged necessity, the durability, the reliability, the risks (to schedule, scientific return, and budget), the maturity, the remaining research and development [1], and probably also the potential glory involved in pulling it off.

The theme of the title is further developed to probe the full potential of microsystems technology for future spacecraft design. I find that the full use of this potential calls for a change of perspective in basic space mission architecture. The fixed nature of the microsystems design demands consideration in all steps of spacecraft conceptualisation, including the very first [Paper I].

Thus, the pursuit of the desired level of integration and miniaturisation clearly requires a new way of thinking. This new way of thinking is the ultimate message of the thesis. This opens exciting and welcoming fields for microsystems to explore.
2. MULTIFUNCTIONAL MICROSYSTEMS

If poor talent, little experience of present things,
and weak knowledge of ancient things make this
attempt of mine defective and not of much utility,
it will at least show the path
to someone with more virtue
Niccolo Machiavelli

The strategy of incorporating extremely compact multifunctional microsystems is identified as a highly profitable long-term approach for spacecraft design. In other words, a multitude of system functions will be merged without separate interconnects into one single microsystem unit. The conclusion is based on experience on building silicon micropropulsion systems, together with an accumulated knowledge of the current needs in spacecraft capabilities, the status of microsystems technology, and the trends in space exploration.

This is a race towards the most heavily miniaturised space systems. Ultimately, the extreme miniature of an entire fully capable spacecraft is on the line [2] (Figure 2.0.1). Consequently, larger spacecraft will benefit as they gain performance with respect to mass and size.

In order to achieve the desired level of miniaturisation, integration, and corresponding performance, all opportunities that improve the microsystem

Figure 2.0.1: The Ångström Space Technology Centre Nanosatellite 1:1 model. This design aims at a fully capable satellite, accomplished by the massive use of multifunctional microsystems. Eager scientists struggle to grasp the full potential of this kind of spacecraft.

Johan Köhler
must be exercised. In this context, any arbitrary decision represents an opportunity wasted. This means that the identification of truly arbitrary decisions on the spacecraft design must be accomplished, regardless of e.g. traditional system architectures or available flight-proven hardware [Paper I].

The multifunctional microsystem may be illustrated by the water example [Paper I]. Water, an extremely versatile fluid, can be used in many subsystems in a future spacecraft. It can be decomposed into oxygen and hydrogen, which are stored for later use in order to generate electrical power when needed. Hereby, the need for large batteries is eliminated. These gases can also serve as fuel in powerful bipropellant thrusters for primary propulsion or in water vapour thrusters for attitude control [3-5].

Water can be used inside the spacecraft thermal-control system for heat transportation and thermal management. The large latent heat capacity of water also makes it suitable for heat storage. As a mechanical material, water can support several structural functions – radiation shields, mechanical stiffening, and damping during launch. The use of water extends the multifunctionality concept to involve the entire spacecraft design, as water becomes an integral part of the spacecraft function.

The heart of the water system is the water-splitting/power-generating multifunctional microsystem. Here, reversible fuel cells are used for water splitting during mission periods when superfluous electrical power is available. The obtained hydrogen and oxygen are stored in suitable tanks and can be fed back when electrical power is needed. The wastewater is returned for next cycle. The multifunctional microsystem will also integrate all necessary valves, water separation structures, pressure sensors, and control electronics in order to form an autonomous functional unit. Additionally, integrated microturbines may complement the system, delivering very high peak power if required [6]. Specific problem areas to be tackled in the realisation of this system are the simultaneous handling of liquids and gases in zero-g, the energy conversion efficiency, and the system integration in the spacecraft structure, with its distributed water storage.

Historically, research on miniaturisation of space systems has been primarily performed on instrumentation and power generation (solar cell panels). Indeed, the payload miniaturisation has been identified as the prime driver towards miniaturisation of entire spacecraft [7]. However, this miniaturisation will in practice only allow the payload to be utilised more efficiently by stuffing additional or better instruments in the available space [8]. Rather, the miniaturisation efforts on payload should be matched by the corresponding miniaturisation of platform systems, e.g. power generation, transceivers, attitude control, propulsion, and navigation. The requirements on payload and

*Bringing Silicon Microsystems to Space* 9
platform are intricately co-dependant, considering mass, space, and power budget [Paper I].

The conclusion from the above philosophy of avoiding arbitrary decisions implemented on microsystems technology is the concept of the multifunctional microsystem. Here, the decisions on the microsystem design and manufacture should aim at integrating as many functions as possible into the system. Simultaneously, all possible synergetic gains to other parts of the spacecraft should be considered. In this way, extremely compact and truly multifunctional microsystems will emerge.

2.1 Multifunctional microsystem strategy

The strategic progress towards achieving the multifunctional microsystems goes through the steps of demonstration, affordability, and profit. This thesis shows on items necessary for the first and second step, while indicating some issues on the third. In a general sense, the demonstration of total integrated microsystems is currently possible.

The first step of the strategy is to successfully demonstrate that these kinds of extremely complex systems can indeed be realised. This is accomplished by prototyping, disregarding the resources employed in order to prove the concept of the microsystems.

The next step aims at the convincing demonstration that these systems are indeed affordable in actual spacecraft design and manufacture. In order to do this, the reliability of the systems must be assessed, and manufacture must be economically sound. Thereby, spacecraft designers can use these solutions without compromising the mission performance, risk, or budget.

Finally, it remains to be demonstrated that the proper utilisation of the systems will be highly profitable, compared to using existing spacecraft designs. This is possible either through opening a gate to new fields of possible exploration, or by using the extremely compact multifunctional microsystem approach to produce spacecraft that unequivocally outperform conventional efforts.

Now, in my opinion the first step is already accomplished. By using enough resources, prototyping can be done. There is no doubt that complex microsystems are conceivable in that sense. The point of multifunctional, yet affordable, microsystems demands extra attention. This step in the development is essential to understand the potential for the multifunctional designs. Without a clear view on microsystem reliability, space mission risk, and the economics of manufacture, spacecraft will never use these systems. Once they can be relied upon, however, the space market has already
recognised the open road to new mission concepts and high-performing nanospacecraft, requiring a fraction of the mass needed today.

2.2 Multifunctional microsystem design

Six levels of packaging are recognised in the classical integrated circuit (IC) hierarchy, L0–L5. These are IC features (L0), IC chip (L1), chip or multichip package (L2), chip package mounting board, e.g. printed wiring boards (L3), chassis, box, or harness (L4), and the entire system, e.g. a computer (L5) [9].

Each level routinely adds interconnects and packages requiring reliable mounting, extra mass, and extra space. Any integration of these levels brings miniaturisation, but the tremendous 100–1000-fold decrease dreamt of by microengineers is achieved only through the successful integration of at least L0–L4 [10].

In the multifunctional microsystem approach, an explicit goal is to integrate all levels, possibly barring L5 [Paper 1]. The completed microsystem will be ready for direct mounting, or already be integrated in the spacecraft structure. The high level of integration efficiently reduces the mass and size of interconnections and eliminates intermediate packages. Extremely compact microsystems should be self-sealed and self-packaged, requiring a minimum of mounting facilities, preferably using the final package as a functional structural element in the spacecraft architecture.

Using this high level of integration, the robustness of the system functions can be significantly increased. The number of interconnects and package-induced failures naturally decrease. Furthermore, the microsystems technology makes the systems spacious at the feature level, thereby enabling reliability management by introducing multiple redundancy.

Wafer bonding [11, 12] is chosen as the basic method in the proposed all-levels integration of multifunctional microsystems. Thus, bonding of silicon and non-silicon parts are utilised as fabrication, assembly tool, and packaging technique. Commonly, wafer bonding is considered for packaging at levels L0, L1, and L2 [9]. Complex systems realised at MIT demonstrate the possibility of multiwafer bonding, thereby merging levels L0–L3[4, 6]. Suga et al. have developed bonding techniques intended for integrating levels L0–L3 [13-15]. The Aerospace Corporation has outlined an integration scheme for levels L0–L5, ultimately suggesting a nanosatellite as a bonded assembly of application-specific integrated microsystems [16, 17].

One major challenge in using bonding for all-level integration is the tremendous demands on silicon microfabrication process compatibility. The primary consideration is which functions that should be integrated on each
wafer, and which should be connected or formed by subsequent bonding. In this context, a dominant characteristic of the integration method becomes vividly clear: The detailed designs of all integration levels strongly influence each other. In other words, for silicon microsystems technology, the master lithographic mask patterns are intertwined from smallest device to the assembled spacecraft subsystem, or even the entire spacecraft. Thus, the incorporation of multifunctional microsystems in a space mission must by necessity be predominant in the spacecraft design from the very beginning [Paper I].

The multifunctional approach to mass and space saving while retaining or extending the capability of the spacecraft may also benefit from development in flexible and inflatable structures, and macroscale integration of functions in the structural elements of the spacecraft. The design philosophy adopted for microsystems may be extended to spacecraft architecture. This claim is supported by the equally important and interdependent miniaturisation of platform and payload, together with the intertwined design of multifunctional microsystems—the basic building block of future spacecraft. This means that the traditional split between payload and platform should be reconsidered, allowing for truly multifunctional entire spacecraft [Paper I].

2.3 Current and future competition

Competitors to the highly integrated multifunctional microsystems are mainly modular hybrid microsystems [8, Paper II], shape deposition manufacturing [18] and advanced fine mechanics [19]. However, these techniques primarily compete with each other. The extremely compact multifunctional microsystems represent the ultimate end of the current miniaturisation efforts. Thus, this approach has no competition with respect to mass and space limitations. Only when considering the current feasibility and price of the different systems do they compete, and compete seriously. The conclusion is clear—the feasibility of massive integration must increase, and the price of multifunctional systems must fall—and the way forward is to increase research and development in this field.

A quick glance on future competitors reveals the nanomechanics [19] discipline that boldly suggests nanosystems. When nanosystems are demonstrated, they will indeed be competitors to the integrated multifunctional microsystems. However, the transition from nanomechanics to nanosystems in a sense analogous to microsystems is by no means a trivial matter.

I predict that a completely different perspective on the system principles must be adopted in order to fully taking the step to nanosystems. For example, a
nanomechanical propulsion system may overcome the fuel supply and regulation issue by making a surprising turn to the yet-to-be-exploited and decidedly dubious vacuum energy [20] speculated on by Dirac [21] and convincingly derived from first principles by Casimir [22]. This so called Casimir effect was experimentally demonstrated by Lamoreaux [23, 24], and has recently been exploited for actuation in a silicon microsystem device [25].

Until such a new perspective on system principles makes proper nanosystems possible, the use of nanomechanics will be an adequate option for multifunctional microsystems, rather than a competitor in itself. Like the compromise of the modular hybrid microsystem approach suffers competition from the conventional methods of fine and mesoscale mechanics, so will insufficiently developed revolutionary nanosystems compete with multifunctional microsystems.

Returning to microsystems, a definition of hybrid integration of microsystems is in order. In this thesis, I use hybrid microsystems to designate the assembly of microsystems by conventional techniques. Common to the hybrid approach is the adding of modular parts in order to form the complete system. Further integrated microsystems, like the extremely compact multifunctional microsystems, are packaged using microsystems technology processes. This does not limit them to the purely monolithic wafer-scale integrated systems.

In this respect, the conventional and hybrid subsystem and spacecraft assembly both adhere to the same basic philosophy of adding separate, self-contained and sealed parts of the complete system together by ordinary – albeit small – tubing and external wiring. Eventually, the total assembly comprises the entire spacecraft, which to a large degree will be occupied by the packaging, tubing, mountings, and interfacial circuitry [26]. The different parts in turn form the independent subsystems, separately packaged and mounted onto the frame of the spacecraft, and further interconnected to other subsystems.

The hybrid microsystem approach does not come close to fulfilling the potential of extremely compact microsystems. Unfortunately, both strategies still suffer from the same limitations, foremost the intricate complexity of production. Despite these shortcomings, the efforts in making a hybrid microsystem solution may be worthwhile. First, the hybrid is evidently a lower technical risk, although still not a negligible one. Second, a hybrid microsystem may comprise a valid development in order to eventually mastering the fully integrated manufacture. Last, naturally, if a feasible and high-performance subsystem can be demonstrated no one will search for competitors that are not readily available, regardless of the technical approach chosen.

*Bringing Silicon Microsystems to Space* 13
3. WHY SPACE?

The multifunctional microsystems have manifold duties in future space missions. In all respects, launch costs will be saved by the consequent mass and space reduction [27, 28]. However, as these systems gradually become available, completely new mission concepts have emerged sharing the common enabling feature of extreme miniaturisation [29].

The disposable spacecraft concept has suggested robust nanosatellite swarm networks with distributed functions. High-risk nanoprobe for planetary exploration can also be allowed if they do not infringe too much on the mass budget of the mother spacecraft. This concept also includes mass-efficient planetary rovers, balloons, and unmanned aero vehicles. Simultaneous multiple spacecraft investigation of celestial bodies has recently produced excellent results by the combined efforts of the Cassini and Galileo rendezvous with the Jovian system [30]. Using miniature spacecraft, future similar missions may be affordable in the future. Precision constellations of satellites or deep-space platforms also require the performance obtained from microsystems, as do deployment and operation of flexible and inflatable spacecraft architectures [31-33]. In general, any mission requiring a multitude of spacecraft, cheap access to space, or a distributed risk will become feasible.

On a more profound level, you could ask what the human interest in space is all about in the first place. A classic answer would conjure the image of the final frontier – the romantic idea of Mankind’s destiny-bound striving for the unknown. Using this imagery, at least three distinct frontiers may be found in space.

First, the exploration of the unknown, simply by discovering new vistas, ultimately the greatest of them all, attracts interest from inventive spacecraft designers, astronauts, and the public. Examples include the Search for Extraterrestrial Intelligence (SETI), the manned spaceflights, and to a certain extent endeavours like the Hubble Space Telescope and premier spaceprobe flights. The common denominator could be the seeking of answers to the basic question: What is out there?

Second, the grand models of science may eventually be unambiguously tested in space, providing definite answers to the great scientific issues, e.g.

Johan Köhler
relativity, gravitation, unification theories, and theories of everything. The basic question may be expressed: How does it work?

Third, many investigators in space phenomena brush the limit to the eternal questions of philosophy (or theology, depending on inclination). I do not propose to state a basic question here, other than an unspecified and symbolic: Why?

Apart from these romantic notions, space can also be exploited for beneficial ends. Examples include climate research and monitoring, navigation support, telecommunication links, surveillance, mapping, earthquake prediction, space weather research, natural disaster monitoring, and meteorology.
4. MICROSYSTEM MANUFACTURE

Truth emerges more readily from error than from confusion.
Francis Bacon

With the goal of making extremely compact multifunctional microsystems for operation in space, a suitable basic manufacturing technology needs to be selected. No available technology can readily produce these microsystems, which implies that research and development efforts must be put into the technology of choice in order to reach the goal. Thus, the most promising option must be found and advanced.

The choice must consider the harsh environment, both in space and during launch, together with the extreme requirements on reliability [34]. Furthermore, there should be an obvious potential for integration at all system levels. Finally, the technology should supply an inherently compact and feasible possibility for system assembly.

The obvious option is silicon microsystems technology, and that is also the choice made in this work. It is not the only possibility, however, and the basic reason for deciding on silicon microsystems technology was not the overall performance of the technology but rather the outstanding maturity of available individual fabrication processes.

Having made this choice, the wafer bonding assembly of separate parts was recurrently emerging as an issue of concern. Central questions were how to investigate the quality of the bonds in a silicon microsystem for space operation, how to accomplish high bond strength at low process temperatures, and how the bonds would endure the space environment and the trials of launch. These concerns will be addressed in the following, and are further detailed in the appended papers (Paper IV, V, VI, VII and VIII).

4.1 Merits and flaws of silicon microsystems technology

The potential of direct wafer bonding, patterned thin film interconnections, and the formation of fully internal microsystem tubing support the choice of silicon microsystems for the desired massive integration.

Alternative microsystem technologies have been disqualified mainly due to inexperience, missing facilities, structural properties, space compatibility, or unexplored potential for integration. These include flexible printed circuit boards [35], diamond microengineering [36], smarter materials (e.g. PZT) [37], quartz [38-40], glass [41], surface micromachining [42], plastic
replication, LIGA [43], photosensitive glass [19], and high-aspect-ratio photoresists [44, 45].

Microsystem engineering has traditionally profited and trumpeted the excellent mechanical properties of single-crystalline silicon [46] in order to support the claim of rugged silicon microsystems, together with the ability of forming features in the micrometer range. However, even though microfabrication accomplish excellent absolute tolerances (e.g. sub-micron), the relative accuracy is commonly low (e.g. 1%). Here, traditional fine mechanics outperform microsystems technology, successfully obtaining accuracy in the ppm range [19].

Furthermore, complex microsystems are joined from mainly silicon parts, e.g. by direct wafer bonding and soldering. The durability of these kinds of bonds determines the robustness of the system function, analogous to the weakest link of a chain. Fortunately, this potentially fatal feature of the system can be relieved by subsystem redundancy. Silicon microsystems furthermore allows for easily designed redundancy, due to the space available through the extreme miniaturisation. However, it is necessary to carefully isolate the weak links themselves and multiply these, if random effects are responsible for their failure, or make efforts to substantially strengthen them, if the weakness is inherent to the link in question. This is particularly true for failures in the joint of microsystem building blocks e.g. silicon wafers.

Direct wafer bonding provides a precise method of joining silicon wafer parts of a microsystem together [11, 12]. No adhesives are needed; the clean, planar silicon surfaces bond spontaneously to each other at room temperature. The bond strength and long-term stability are commonly improved by annealing at elevated temperatures [47, 48].

Cleanliness and care are of utmost importance for maintaining the yield in the bonding process. Trapped particles, chemical residues, or surface damage (general roughening or individual scratches) cause unbonded areas between the wafers, known as voids or interface bubbles. Such occurrences can partly be overcome by subsequent heat treatments, annealing, which works to strengthen and increase the bonds between the wafers. The annealing also allows residual gas to diffuse in the interface, hopefully reaching an etched cavity or structure edge. Thus, annealing can heal some voids, while other residuals (notably hydrocarbons) are believed to nucleate interface bubbles from diffusing gas in the interface at elevated temperatures [49-51].

The ultimate bonding strength is mainly dependent on annealing temperature, but lower annealing temperatures can still give reasonable bonding strength if long annealing times are allowed. This is due to two features of the annealing process. First, the heat of activation provided at the chosen temperature
determines the eventual bonding type, and second, that the bond formation and reformation process rate is dominantly diffusion-limited, which allows for a partial trade-off of temperature and time [52]. In addition, the direct bond is target of fatigue and consequent crack growth [48, Paper IV, V].

In a microsystem fabrication process, the choice of annealing temperature is also restricted by the materials added to or introduced in the silicon structures. Change of dopant profiles, surface-coating diffusion, material melting, and thermal expansion can severely damage or even destroy the intended microsystem.

Three system characteristics together theoretically determine the fracture behaviour of a bonded microsystem [Paper V, VI]. First, the stress distribution at the bond, i.e. the stress as a function of position shows if the load anywhere becomes too severe. Second, the defect distribution at the bond, i.e. the frequency at which defects of certain sizes occur, describes the existence of the weak spots initiating fracture. Third, the fracture surface energy of the bonded interface as a function of location provides the adhesion responsible for joining the wafers.

Three major physical phenomena have been identified that influence the obtained bond quality [Paper VI]. First, the surface roughness of the joined wafers together with trapped particles affects the defect distribution [53-55]. Second, the occurrence of nucleation sites for void formation further inflicts defects in the bond [47, 49, 56]. Last, the ability of water, hydroxyl groups, and other species to exist, react, and move in the bond interface greatly influence the obtainable fracture surface energy, and also the void formation [52, 55, 57, 58].

Spacecraft often experience large and rapid thermal variations, or exhibit substantial thermal gradients within their structures. This is due to their limited thermal mass, high power-density, and inhomogeneous incident radiation. Here, the problems connected to different coefficients of thermal expansion emerge as primary issues. Silicon microsystems technology can be aimed at reducing these problems by using the silicon for all bulk structures exposed to thermal variation. Furthermore, silicon is a good heat conductor, which allows for straightforward thermal equilibrium design in the separate microsystems, if desired. On the other hand, if heating is contemplated, this heat conductivity can be a considerable obstacle to the microsystem design [59, Paper II, II].
4.2 Microsystem fabrication strategy and process sequence evaluation

The successful delivery of a single working silicon microsystem forms the natural starting point for a general evaluation of the fabrication process. In order to benefit fully from the promises of massive integration, reliable manufacture of microsystems must be achieved. Here, the process time, fabrication yield, and microsystem cost must be considered. The resources necessary for development must be correctly assessed and valued. In this context, the detailed characteristics of the complete line of processing must be evaluated as a whole, in order to determine which alternate line uses the available resources most efficiently.

The process yield of the described fabrication sequence is a major issue for any microsystem manufacture. For IC manufacture, yield models are used with respect to substrate area, line width, routing density, feature size, typical defect size, process cleanliness, and number of process steps [60, 61]. These models are used mainly to estimate the economics of production, relying on large batches of identical devices [62].

In the current state of multifunctional silicon microsystem manufacture, on the other hand, the perspective is primarily technical – what is a feasible line of processing? The complex sequence of a multitude of advanced microfabrication techniques easily results in an overall low yield, suggesting a shaky reliability to the casual observer. However, the development of multifunctional microsystems is still in its infancy. The inherent yield losses are often associated with specific procedures, non-uniformity problems, or random misfortunes. Such occurrences should be identified and evaluated in order to advise on efficient microsystem manufacture and process-risk mitigation.

Here, the efficiency of the manufacturing may be substantially influenced by the chosen design of the individual wafers in a multiwafer stack. By proper integration of structures on the wafer level, the number of necessary processes can be minimised. In this manner, the requirements on process compatibility can be relaxed. The wide variety of silicon microfabrication processes will still be available through the bonding of differently processed wafers. In other words, the peak performance of the fabrication process sequence will not be obtained by wafer-scale integration.

The production cost is also of prime interest when considering the microsystem manufacture. The feasibility of microsystem fabrication by a certain process should always be weighed against the cost of using that process. Naturally, the cheapest feasible processes are preferred. Furthermore,
the total manufacture time will also be an item for optimisation, as a shorter manufacture turnover is desirable from a logistic point-of-view.

If an adequate total cost model is desired as a basis for pricing of the Microsystems, the costs of documentation, goods, design, mask manufacture, and verification testing should also be included. Naturally, this can be done, but falls outside of the present discussion. When a precise price is requested, the manufacture would normally already be mastered. For research-intensive development projects this may not be true. The development of manufacturing processes will make the process choice, process duration, and process yield items of recurrent optimisation.

However, the three main issues – yield loss, process cost, and process duration – are not themselves de-coupled, independent quantities ready for optimisation. In mature IC production, management decides the priority of the separate optimisation of each item from a purely profit-based point of view [61]. Currently, the technology for fabrication of highly complicated silicon Microsystems is still in a development phase, implying that yield is comparably low, large-scale production is rare and require hefty investments, and the market is small or slumbering. In other words, the production of multifunctional Microsystems (for space) is rarely economic [7]. Using yield-modelling terms, the systematic yield can be suspected to be low, although perhaps easily amended if identified [60]. The Microsystems technology has yet to gain acceptance and maturity in order to attract the sponsors necessary to overcome these issues. Until then, manufacturers have to work within these limits, perhaps eventually bringing about the desired change.

Now, the coupling of the three issues can be represented by the process sequence figure of merit. This is a measure of the total amount of resources put in a single produced pack at any time in the line of processing. This value is the amount at stake in the consecutive process step, target of the yield loss gamble.

Three quantities describe each separate process in the process sequence. First, the process duration for a batch is summed over the entire sequence (Equation 4.2.1). Second, the process cost per unit is quite straightforward (Equation 4.2.2). Third, the yield loss is described (essentially the inverted yield of the process). This parameter can be normalised with respect to the batch-size for convenience, forming the batch-size yield-loss factor, which is the inverted yield of possible units from a full batch (Equation 4.2.3).

All these quantities can be established by careful investigation of laboratory log journals and offers of processes from foundry-service suppliers. These parameters quantify the process sequence overview and allow for derivation
of the process sequence figure of merit (Equation 4.2.4 and Figure 4.2.1) under the specific boundary conditions set by available laboratory equipment.

\[
Accumulated\ process\ time = \sum_{processes} \frac{process\ time}{batch} \tag{4.2.1}
\]

\[
Cost\ of\ processing = \frac{Cost}{time} \tag{4.2.2}
\]

\[
Yield\ loss\ factor = \frac{1}{yield \cdot batch\ size} \tag{4.2.3}
\]

\[
Figure\ of\ Merit\ for\ process\ sequence = \sum_{processes} \frac{process\ time}{batch} \cdot \frac{Cost}{time} \cdot \frac{1}{yield \cdot batch\ size} \tag{4.2.4}
\]

The figure of merit is given in monetary units, but cannot be interpreted as the process cost of the manufacturing sequence. Rather, the suggested figure of merit of the process sequence establishes the progressive value of each pack during the manufacturing process. In doing this, the yield of each separate process step is targeted, not the total yield of the process sequence. Furthermore, all process steps are weighted to represent the manufacture of a complete batch.

The choice of parameters is made considering which values that are readily extracted from on-going research and development, and not in order to establish a future price. The figure of merit implies that all flawed samples from each process step are removed immediately from further processing. This is obviously not a true representation of the real case. However, during manufacture development, the inspections are quite frequent and failures are often recognised at these inspactions. In this context, the implied approximation of automatic screening of failed samples may still be valid.

Now, the figure of merit uses a weighted sum, where each process cost is modified by the process-step yield and the number of packs (wafers or bonded wafers) in a complete batch. This is a fair model in the evaluation of the process sequence during development, as it allows an equally important contribution to the process sequence figure of merit from each process step. However fair the model, this also means that the figure of merit can never be used for predictions of microsystem cost, or fabrication cost. In order to fulfil
**Micro-fabrication processes**

<table>
<thead>
<tr>
<th>Process</th>
<th>Duration [h]</th>
<th>Batch capacity [wafers]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxidation</td>
<td>5.5/25</td>
<td></td>
</tr>
<tr>
<td>Plating</td>
<td>4.5/1</td>
<td></td>
</tr>
<tr>
<td>Mask etching</td>
<td>4.5/6</td>
<td></td>
</tr>
<tr>
<td>Oxide removal</td>
<td>0.5/25</td>
<td></td>
</tr>
<tr>
<td>Oxidation</td>
<td>33.5/25</td>
<td>high risk</td>
</tr>
<tr>
<td>Plating</td>
<td>4.5/1</td>
<td>medium risk</td>
</tr>
<tr>
<td>Pre-shape etching</td>
<td>20.5/8</td>
<td>low risk</td>
</tr>
<tr>
<td>Etch depth validation</td>
<td>2/1</td>
<td></td>
</tr>
<tr>
<td>Oxide removal</td>
<td>5.5/6</td>
<td></td>
</tr>
<tr>
<td>Oxidation</td>
<td>5.5/25</td>
<td></td>
</tr>
<tr>
<td>Plating</td>
<td>2.5/1</td>
<td></td>
</tr>
<tr>
<td>DRE</td>
<td>4.2/1</td>
<td></td>
</tr>
<tr>
<td>VL interferometry</td>
<td>2/1</td>
<td></td>
</tr>
<tr>
<td>Oxide removal</td>
<td>0.4/25</td>
<td></td>
</tr>
<tr>
<td>Aligned bonding</td>
<td>2.5/1</td>
<td></td>
</tr>
<tr>
<td>Inspection</td>
<td>1/1</td>
<td></td>
</tr>
<tr>
<td>Annealing and oxidation</td>
<td>6/25</td>
<td></td>
</tr>
<tr>
<td>Inspection</td>
<td>1/1</td>
<td></td>
</tr>
<tr>
<td>Copper evaporation</td>
<td>2/4</td>
<td></td>
</tr>
<tr>
<td>Residual electroplating</td>
<td>2/1</td>
<td></td>
</tr>
<tr>
<td>Inspection</td>
<td>1/1</td>
<td></td>
</tr>
<tr>
<td>Plating</td>
<td>2.5/1</td>
<td></td>
</tr>
<tr>
<td>Inspection</td>
<td>1/1</td>
<td></td>
</tr>
<tr>
<td>Platinum deposition</td>
<td>2/4</td>
<td></td>
</tr>
<tr>
<td>Lift-off</td>
<td>0.5/6</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4.2.1:** The figure of merit of the process sequence in the progress of fabrication. The example is from the manufacture of an integrated silicon micronozzle system for cold gas micropropulsion [Paper II, III].
this, the batch size would have to be progressively reduced by the yield of each process step. Still, the economic soundness of the microfabrication process sequence is considered, even if production break-even is far from achieved, and much less addressed.

In Figure 4.2.1, the figure of merit is presented as a function of the stepwise fabrication progress. The graph further expresses the qualitative process risk, the separate process durations, and the batch capacity. This graph was extracted from lab journals for process times and yield estimates, and requests for quotations on microfabrication processes for the approximate process costs.

The inspection of the figure of merit graph first conveys which steps that cause the most rapid increases of the figure of merit. These processes should naturally prioritised targets of research. Second, the most precarious processes are clearly identified. Finally, the process duration and batch capacity can be read. Thus, the graph allows for straightforward identification of the risky bottlenecks (high yield loss), the expensive processes, and the long manufacture times. The graph helps in allocating resources to increase the lowest yields, and up-scaling low batch-capacity equipment or speeding up the slowest serial processes.

Considering the figure-of-merit concept, some recommendations and observations can be made. First, expensive high-yield processes should be preferred late in the fabrication sequence, while initially the cost of the processing is more important. An initial low yield does not risk a valuable pack, and can thus be accepted. Second, the production capacity gain from using batch processes is more valuable early in the processing rather than late, as a single process failure here will likely destroy the entire batch. A serial process, preferably even a direct-write process, will only destroy a single element or unit at the time, which protects valuable batches of potential microsystems late in the manufacture from total annihilation. Third, the figure of merit is kept as low as possible by using cheap processes early in the line of processing. Later, the added value from using more expensive processes is acceptable if the yield is high enough to protect the valuable packs characterised by considerable manufacturing efforts.

The direct bonding of silicon wafers deserves a special attention here. This process is characterised by being not repeatable and by drastically increasing the figure of merit, literally adding packs together (Figure 4.2.1). Thus, the bonding needs to be a highly reliable process, and efforts should be allocated to optimise bonding yield, and to develop equipment accomplishing this. An alternate approach that is pursued is achieving dismountable wafer bonding [14, 57, 63].
The process sequence overview of the figure of merit illuminates the sceptic view of the commercial manufacturers on microsystems technology for product fabrication. By interpretation, this tool indicates the suitable market niche for the particular microsystem investigated. In the case of a micropropulsion system in particular and a multifunctional microsystem in general, this niche should be characterised by small production series of units that can charge a high cost for unique performance. The space industry springs readily to mind.

In this context, further consideration on the results suggests what areas of the microsystem fabrication that need to be altered, and how, in order to fit the users’ desires. Will the yield be too low, and what processes cause that? Is the production time too long? Are the processes too expensive? What is the optimum batch size considering the available facilities? Which processes must be replaced? In this way, the required direction of the microsystem fabrication sequence can be found in order to spite the sceptics.

4.3 Microsystem performance evaluation

The aim in the microsystem design should always be to gain performance by subsystem integration, either with respect to a critical metric (e.g. mass, volume, or power), or by enabling desired applications. A system function must not be understood as simply the sum of all parts, but rather as a larger whole. The individual performances of e.g. nozzles, valves, and filters do provide necessary experimental data for system evaluation, but these data are useless until subsystem compatibility is demonstrated.

The demonstration of simultaneous integrated fabrication of subsystems is the key achievement, adding intangible items of process compatibility as constraints to the whole. Further, a proper evaluation of the fully assembled microsystem cannot merely quote figures of merit from isolated subsystems, but needs to combine them to form key figures applicable to the entire system. In this context, it is necessary to take the imaginary step back in order to grasp the whole.

Drawing from examples obtained while investigating cold gas micropropulsion systems, possible key figures can be discussed. These may be specific to the type of microsystem in question, in this case propulsion, but the main theme drowns the particulars – these kinds of key figures are desired for any complex microsystem. In the particular application, the relevant key figures grasping the performance of the whole need to be established and determined. So, what are these key figures that a system boasts in order to prove its excellence?

Johan Köhler
In spacecraft design, these are not new concerns. From the system perspective, many parameters of this kind have been established. In the following, propulsion will be used as an example.

First, the capability of changing the velocity of the spacecraft is simply described by the quantity $\Delta v$ [64]. For electric propulsion, where electric fields generate the impulse used for moving the spacecraft, $\Delta v$ is limited by the amount of power available. For chemical propulsion, on the other hand, where the propellant is expelled using the energy stored in the propellant itself, $\Delta v$ is limited by the propellant mass.

$\Delta v$ describes the incremental velocity applied to a spacecraft of specified mass in order to perform a specific mission, and can thus be used to set the main capability requirements of the propulsion system.

Second, the specific impulse of a thruster is a widely used figure of merit for propulsion systems, characterising the efficiency of the thrust achieved. The specific impulse ($I_{sp}$, in seconds) is defined as the total amount of impulse delivered with respect to the propellant weight (in N) [65]. This is equivalent to the exit velocity divided by the acceleration of gravity ($g$). However, this quantity is independent of the system assembly, and thus inappropriate for system comparisons. But the specific impulse can still be valuable as a limiting design parameter, stating a lowest acceptable specific impulse in order to secure the mission requirements.

Now, given an educated estimate or recorded value of the total propulsion system mass, the total impulse delivered by the system per unit mass of propulsion system (including propellant) can be derived [66]. For chemical propulsion (e.g. cold gas propulsion) this system-specific impulse ($I_{esp}$) is a function of total spacecraft mass and mission $\Delta v$-requirement. For electric propulsion systems, the specific power (power to power-generation-system mass ratio) and the power-to-thrust ratio are useful for system evaluation [67].

The space required on-board for the propulsion system can also be estimated by calculating propulsion-system mass-fraction (including propellant) of the total spacecraft mass. Thus, an early design estimate on the size of the spacecraft can be made. The mass fraction of a chemical propulsion system is also a function of spacecraft mass and mission $\Delta v$-requirement.

All these parameters should be used for propulsion microsystems competing for a flight opportunity. However, the proper estimation, or better still calculation, of the system-specific parameters requires a high degree of readiness in the design of the entire system. This includes the electronics, propellant feed system, power supply, housing, etc.
5. **EXAMPLE: MICROPROPULSION SYSTEM**

The development of cold gas micropropulsion systems by silicon microsystems technology forms the origin, motivation, and ultimate goal of my efforts in bringing silicon microsystems to space. Thus, the subject of this chapter is intertwined with the themes developed previously, while indicating the relevance of the following matters.

I have pursued the fabrication of both a hybrid and a compact integrated microthruster system. In order to accomplish the extremely compact system, several areas need further efforts. Identified vital issues are the development of multichip module technology for efficient control electronics miniaturisation, the securing of microfabrication process compatibility and reliability, and the robustness of different wafer bonding and thin-film soldering processes.

In the following, I concentrate on the published hybrid integration solution [Paper II, III], i.e. different microsystem units connected and mounted using conventionally machined housing parts. The microsystem units are served by external electronics rather than integrated or multichip module devices. This hybrid cold gas microthruster system exhibits highly integrated micromechanical parts together with conventional components.

The underlying goal of merging the parts of the hybrid design into a highly compact microsystem unit will accomplish a further substantial reduction of mass, power, and space. Here, the electronics are made using multichip module technology, and the entire microthruster system will be contained within a spherical shell of radius 41 mm.

A suggested design is presented in Figure 5.0.1, where a direct-bonded five-wafer silicon stack with soldered piezoelectric valve actuators contain all microsystem functions of the hybrid design, together with integrated pressure sensors. Apart from the obvious demagnification, performance may increase, e.g. due to shorter rise and fall times and shorter data and control buses.

In this design, the integration of microsystems has allowed extreme close-packing of functions, and the full advantage of redundancy potential can be employed. For example, the multichip module electronics can be made redundant in a 3- or 4-way system without much loss of space or gain of mass.

---

26

Johan Köhler
The estimated mass of each complete extremely compact microthruster system is 80 g, excluding propellant, tank, and propellant tubing. The main achievements remaining to be mastered are the reliable multiwaefer bonding, the electric connector introduction, and the full integration of the multichip modules. The manufacturing process sequence is also a target of further optimisation.

The successful miniaturisation of these micropulsion systems basically leads to the possibility to develop smaller fully capable spacecraft and providing better stability and precision to larger ones as well. Thus, the systems aim at enabling new concepts in space exploration, requiring features like precision formation flying of spacecraft clusters, extreme stabilisation for ultra-sensitive measurements on earth phenomena, or excellent pointing accuracy for long distance observations.

The basic propulsion principle utilised is cold gas propulsion. Here, the expulsion of matter, specifically gas of modest temperatures, at high velocity generates the desired impulse to the spacecraft. The gas is accelerated in a converging nozzle part until reaching the sonic transition in the throat, where the gas enters a diverging nozzle part, allowing for the supersonic speed to increase even further.

This kind of cold gas micropulsion, using proper microsystem solutions, was first demonstrated by Bayt et al. [68, 69]. Here, an essentially two-dimensional converging-diverging nozzle was employed, made by etching through one silicon wafer and sealing the system by anodic bonding on both sides. The specific impulse of the unheated device reaches 60 s at optimum conditions for nitrogen propellant. This figure is improved to 90 s by the later integration of an ingenious heater [59] utilising the intrinsic point of silicon in

---

*Bringing Silicon Microsystems to Space*
Figure 5.1.1: The hybrid cold gas microthruster system, containing four independent thrusters. A) Overview showing the gross dimensions of the system. B) Cross-section showing the different parts, e.g. the microsystems: Nozzle unit, valve assemblies, and particle filters. (Excerpt from Paper II).
order to gain some autonomy. However, this device suffers from power inefficiency due to heat loss.

The hybrid cold gas microthruster system investigated here use a similar nozzle design, although only silicon is used. A single precision silicon wafer bonding process assembles the nozzle. Furthermore, the nozzle height dimension has been made smaller. This probably explains the lower specific impulse achieved, due to the viscous-loss phenomena reported by Bayt et al. [69].

5.1 The Hybrid Cold Gas Microthruster System

The cold gas microthruster system releases minute amounts of gas at supersonic speed through micronozzles. Piezoelectric valves – using feedback from differential pressure sensors straddling each nozzle – proportionally modulate the gas flow to the nozzles. Heat exchangers with thin film heaters and temperature sensors in the nozzle unit increase the efficiency of the thruster. Micromachined filters protect these microsystem units from degrading particle contamination due to incoming gas and ambient atmosphere on ground.

The mechanical design of the microthruster system pod is quite conventionally made by fine mechanics machining. An overview is given in Figure 5.1.1. The pressurised housing consists of four major parts. First, a cylindrical bottom part contains the gas inlet and the interface connector. The bottom surface with four threaded holes and two guide pins provides the mechanical interface to a mounting bracket. Second, the cylindrical lower part contains the filter stack. Third, the cylindrical centre part contains the valve assemblies and the nozzle microsystem. Last, the upper half-sphere houses the pressure sensors.

O-rings keep the silicon microsystems floating between the aluminium sections while sealing. In this manner, mechanical stress due to different thermal expansion coefficients can be avoided.

The outer coating to be used is a functional surface, exhibiting different optical properties for different wavelengths. The present most promising film should maintain an equilibrium temperature in direct solar irradiation below 50°C without significant internal power dissipation [70]. Besides assisting in the thermal control, this coating is mechanically extremely durable, chemically inert, and electrically conductive which prevents electrical charging in space.

The complete microthruster unit weighs approximately 0.15 kg, without the external electronics and propellant storage. The electronics – excluding some
separate microsensors and preamplifiers – are contained outside of the mechanical envelope presented above, with an estimated mass of 0.2 kg.

The electronic system integration of the hybrid microthruster is divided in two conventional systems, a digital and an analogue part with a unique interconnection. The analogue control electronics that carry signals to and from the thrusters in closed loops need an advanced control system with reference point input. These reference points are provided by a microcontroller through a number of digital-analogue and analogue-digital converters.

The primary material for the different microsystem units of the hybrid cold gas microthruster system is single crystalline silicon, shaped by anisotropic wet etching and/or deep reactive ion etching (DRIE). Direct bonding joins structured silicon wafers in order to eventually forming the complete units. In this context, the bonding temperature is of primary importance for process compatibility [Paper VI], and the issue of the endurance of bonded microsystems in space was questioned [Paper IV]. Evaporation and patterning of different metals provide conductors and solder-pads for electronics, sensors, and PZT actuators. Other thin film techniques give wear-resistant coatings on particularly exposed details.

![Figure 5.1.2: The nozzle and heat exchanger structure etched approximately 50 \( \mu \text{m} \) into one silicon wafer by deep reactive ion etching. The throat width is 11 \( \mu \text{m} \), the heat exchanger fins are 3 \( \mu \text{m} \) wide and spaced 7 \( \mu \text{m} \) apart. (From Paper III).](image)

---

Johan Köhler
The nozzle unit comprises four independent nozzles, divided by 90\(^\circ\)-angles on a 39-mm diameter disc. Heat exchangers are integrated directly upstream of each nozzle, and platinum thin film heater elements combined with temperature sensors are included on the outside surfaces of all heat exchanger structures. The nozzle and heat exchanger are contained in a free-hanging tubular structure in order to restrict the heat loss to the supporting bulk silicon.

The nozzle shape is essentially a two-dimensional converging-diverging outlet with rectangular cross-section. Throats as narrow as 11 \(\mu\text{m}\) have been demonstrated. Each nozzle and heat exchanger is made from two mirrored halves bonded together (Figure 5.1.2). The nominal height of two such halves is 100 \(\mu\text{m}\).

The valve unit is a normally-closed proportional piezoelectrically actuated device. Five multilayered piezoelectric lead-zirconium-titanate (PZT) elements perform the valve action, lifting a central silicon cap from the coated silicon valve seat (Figure 5.1.3). The multilayered design of the actuators limits the required drive voltage to 50 V for a maximum stroke of 4 \(\mu\text{m}\).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure5_1_3}
\caption{The valve unit. Five multilayered PZT actuators (4 mm high) provide the lifting action to the valve cap. The electrodes and connector circuitry is clearly visible on the silicon valve part and the sides of the actuators. The assembly is made by thin film soldering.}
\end{figure}

The actuators are assembled using a multilayer thin film soldering system. This technique allows for precision mounting of the actuators between the silicon valve and yoke, which is necessary for the successful operation of the valve. The internal electrodes of the multilayer structure are connected using evaporation of gold through shadow masks onto the sides of the actuators. This efficient use of thin films has substantially facilitated the piezoelectric actuator integration.
Figure 5.1.4: The filter unit. Top: Filter cross-section. The upper bond interface contains the v-groove part and the lower interface contains the slot part. Left: The v-groove filter part (top view close-up). Right: The slot filter part (top view close-up). (From Paper II).

The filter unit is realised from combining two different types of filters, forming a 3-wafer stack. The active filter parts are located in the bond interfaces (Figure 5.1.4). First, a fenced slot that may convey flakes of radius \(-5\) \(\mu\)m and thickness \(-1\) \(\mu\)m filters the gas; then, a vast array of crossed v-grooves screen all debris larger than \(-1\) \(\mu\)m diameter.

Now, the successful manufacture of the microsystems of the hybrid design requires a solid knowledge in complex sequential microfabrication processes in order to secure reliable production. From the efforts of establishing this knowledge, the thoughts on microsystem fabrication strategy and process sequence evaluation evolved (paragraph 4.2) [Paper I, III]. In future efforts these tools will greatly facilitate the demonstration of yet more complex microsystems.

5.2 Evaluation of the Hybrid Cold Gas Microthruster System

The basic functions and performances of the microsystems in the hybrid cold gas microthruster system have been experimentally verified [Paper II, III]. For the complete system evaluation, the specific impulse of the nozzle, the heater function, and the valve leakage characteristics are of primary interest. Furthermore, a tank has to be decided upon, and the mass necessary for this must be estimated in order to address system characteristics.

The specific impulse is calculated from experimental measurements of the thrust and the mass flow at different inlet pressures (Figure 5.2.1). At inlet pressure 4 bar, the specific impulse of 45 s was achieved.
Figure 5.2.1: Thrust force (dashed) and exit velocity (solid) as functions of inlet pressure for nozzle with throat dimension 35x100 µm, without the heat exchanger structure. The data are presented with 95% confidence intervals.

The leakage of the valve has been investigated for different valve seat coatings in extensive wearing. For good coatings, the leakage remains below $10^{-5}$ sec/s of helium.

A common spherical high-grade titanium tank has been assumed in the system analysis. The storage system mass ($m_{ST}$) is approximately proportional to the mass of the propellant neglecting auxiliary tubing.

Now, the system-specific impulse can readily be derived as [66] [Paper II]

$$I_{sp} = \frac{I_{sp} \cdot m_p}{4m_i + 4m_e + m_p + m_{ST}}$$  \hspace{1cm} (5.2.1)

where $I_{sp}$ is the specific impulse, $m_i$ is the thruster system mass (four in all on the spacecraft), $m_e$ is the electronics mass, and $m_p$ is the required mission propellant mass according to [64]

$$m_p = m_{i/e} \left( \frac{\Delta v}{w} \right)$$  \hspace{1cm} (5.2.2)

---

*Bringing Silicon Microsystems to Space* 33
Here, $m_{\text{sc}}$ is the total spacecraft mass including propellant, $\Delta v$ is the velocity increment requirement of the thruster mission, and $v_e$ is the exit velocity of the propellant. This expression for the propellant mass is further adjusted for leakage based on the experimental results above.

Finally, the mass fraction of the entire propulsion system with respect to the spacecraft mass can be calculated [66]

$$\frac{m_{ps}}{m_{sc}} = \frac{I_{sp}}{I_{sp}} \left( 1 - e^{-\frac{\Delta v}{v_e}} \right)$$  \hfill (5.2.3)

The system-specific impulse and the propulsion-system mass-fraction, both functions of mission $\Delta v$ and spacecraft mass, are presented in Figure 5.2.2.

![Figure 5.2.2: The system-specific impulse predicted for the hybrid cold gas microthruster system (solid lines), together with the estimated mass fraction occupied by the propulsion system (dashed lines). The estimates are presented as a function of mission $\Delta v$ requirement and total spacecraft mass (from Paper II).](image)

From Figure 5.2.2 it is evident that the hybrid cold gas microthruster system is best suited for low-$\Delta v$ assignments, like attitude control, fine stabilisation, drag compensation, or short-term free-flying missions. But how can the system parameters be improved?
The most obvious improvement would be a change of tank, for example using a light-weight carbon-fibre reinforced resin.

Another approach would be to change the propellant. Carbon dioxide can be stored in solid phase without a high pressure assembly or heavy storage facilities.

The total propulsion system mass can be reduced by a factor of four by using the compact system suggested in Figure 5.0.1.

Further alternatives include improving the specific impulse, either by heating, using another gaseous propellant, or optimising the nozzle shape by inventive micromachining. For a converging-diverging nozzle, the ideal maximum specific impulse obtainable by complete conversion of the chemical energy stored in the pressurised gas is given by [65]

\[
I_g = \frac{1}{g} \left[ 2 \cdot \frac{k \cdot R_m \cdot T_0}{k - 1} \cdot \left( 1 - \left( \frac{P_{out}}{P_{in}} \right)^{\frac{k-1}{k}} \right) \right]^{1/(k-1)}
\]

(5.2.4)

where \( g \) is the gravitational acceleration, \( k \) is the ratio of specific heat capacities \( (c_p/c_v \text{ for } \text{N}_2 \text{ } k=1.4) \), \( R_m \) is the gas constant with respect to gas mass \( (\text{N}_2: R_m=297 \text{ Nm/kgK}) \), \( T_0 \) is the inlet gas temperature, \( P_{out} \) is the pressure at the nozzle outlet and \( P_{in} \) is the inlet pressure. In vacuum, the surrounding pressure does not determine the actual pressure ratio \( P_{out}/P_{in} \), but rather the area expansion of the nozzle sets this ratio.

For nitrogen at room temperature, the ideal specific impulse for an optimum area expansion is 80 s. However, the maximum experimentally verified specific impulse is 73 s, implying an efficiency of 91% [31]. For the present area expansion, the ideal specific impulse using Equation 5.2.4 is 70 s. Thus, the measured specific impulse of 45 s yields an efficiency of 66% [Paper II]. Obviously, there may be room for improvement of the obtained specific impulse.

Inspecting Equation 5.2.4, the specific impulse can evidently be improved by optimum choice of propellant (considering the \( k \) ratio and the molecular mass), heating of the gas, and by obtaining the optimum nozzle shape. This shape would probably be a bell-shaped or conical perfectly rotational symmetric converging-diverging nozzle.
5.3 Other micropropulsion principles

Unsurprisingly, there are many propulsion alternatives available or under development that are considered or should be considered for total integration by microsystems technology. Here follows a summary description of commonly encountered micropropulsion options, with brief comments on their suitability for microsystem adaptation.

First, the principle of using accelerated ions from various sources to generate momentum defines the field of electric propulsion by ionisation. These options generally exhibit high specific impulses. They all include an ionisation source, an accelerator arrangement, and a neutraliser in order to avoid continuous charging of the spacecraft by the expulsion of ions. A common complication in miniaturising these devices are high-voltage management (e.g. in the accelerator) and erosion of integrated field emitter arrays for ion beam neutralisation [71]. In certain applications, plasma confinement may also limit the miniaturisation efforts [32]. Examples of this micropropulsion principle include the micro ion thruster [32], field emission electric propulsion (FEepy) [67, 72, 73], pulsed plasma thrusters [74-76], and the colloid microthruster [77, 78].

Second, further chemical micropropulsion options imply the use of energy stored in the propellant for expulsion in order to deliver momentum to the spacecraft. Often, these options require heating of the exhausted mass, either by combustion, catalytic decomposition, or resistive heating. Main problems for extremely compact microsystems are heat loss by conduction, thermal endurance, and aggressive fuel handling. The chemical propulsion microsystem options include the bipropellant engine [3, 4], different resistojet [5, 79], solid combustion thruster arrays [80], and various alternate fuel options [31, 81].

Last, momentum transfer propulsion is an exotic propellantless system, where the impulse is delivered by momentum from a source outside the spacecraft. Two basic types are proposed: either the momentum is transferred by the solar wind, or by a collimated photonic or material beam. Existing examples include the solar sail [64] and the magnetospheric momentum transfer [82], while the particle beam suggestions [83, 84] belong to the future.

Additionally, for attitude control, there are specific commonly employed options that do not require the use of propulsion at all. These use manipulation of the moment of momentum, centre of gravity, or magnetic field alignment to rotate or merely stabilise the spacecraft. Examples include magnetorquers, reaction wheels, gravitational stabilisation, and control moment gyroscopes [28, 85, 86]. The rotating devices are generally not suitable for extreme miniaturisation, as their performance improve with their mass and radius.

Johan Köhler
6. Degradation of Bond Quality in Spaceflight Environment

To conquer Death you only have to die.
Jesus Christ Superstar

The harsh environment in space and while going to space put substantial requirements on the endurance of space systems. In silicon microsystems (e.g. the cold gas microthruster system), the weak links of the bonded wafers interfaces are a prime concern. Here, the greatest stresses are induced at the potentially most defect-populated areas.

Spaceflight requires these microsystems to remain reliable during launch and in space. This includes trials like extensive vibration, thermal variations (abrupt and long-term), charged-particle impingements (e.g. solar protons, trapped electrons of the van Allen belts, and cosmic ray heavy nuclei), prevalent γ-irradiation, and micrometeoroid impacts [34, 87].

In the cold gas micropropulsion system, the major loads on the bonded microsystems are induced by internal gas pressure. A burst test would thus suit the quality assessment of the bonds in this application. Here, the burst pressures of test samples containing a specified cavity at the bond between two wafers are recorded, both for a reference series and for series subjected to spaceflight environment ground tests. These data series are then compared using statistical rank sum diagnostics in order to assess the significance of the fracture behaviour variation.

A clear example of the effects of spaceflight environment on the bond quality is the test performed for γ-irradiation [Paper IV]. Here, the bond strength first degrades at dose 6 krad, then actually increases at 22 krad, and has finally degraded again at 100 krad (Figure 6.0.1). All these changes were found to be significant.

6.1 Possible degradation mechanisms

The quality of the bond between silicon wafers is subject to degradation from a selection of different spaceflight simulations. The thermal environment has already been mentioned e.g. [52], and the bonding process introduces point defects in the silicon lattice and at the bond by itself [56]. Vibration, thermal cycling, and different radiation exposures can also cause degradation [Paper IV, V], as well as applied sub-critical cyclic loads, that can eventually separate even high-temperature annealed wafers [48, 63].
Figure 6.0.1: The changing bond strength due to $\gamma$-irradiation of silicon-silicon samples annealed at 700°C. The 95%-confidence intervals are given. All changes due to increasing dose have been diagnosed significant by the rank-sum test (Paragraph 6.3).

In the view of these results, I conclude that the bond interface should be considered a fundamentally dynamic system, subject to environmentally induced change. However, the long-term stability is obviously improved (and the dynamics correspondingly suppressed) for high-temperature annealed bonds [48].

Two typical scanning acoustic microscopy images of the bond test sample are presented in Figure 6.1.1. Here, the defect population is clearly depleted near the stressed edges of the test cavity. This is due to the ability of residual species to diffuse into the cavity during annealing. It should be noted that the observed defects are small, and thus, semi-macroscopic defects, such as void healing or propagation are not probable causes for bond degradation.

Under the small defects assumption, several mechanisms may be responsible for the degradation. Further point defects may be generated (Figure 6.1.2), e.g. by $\gamma$-irradiation [88], reactions of species or point defects in the bond interface system may be activated, and slow fracture may occur if water residuals are still present at the bond (Equation 6.1.1) [52].
Figure 6.1.1: Scanning acoustic microscopy images of bond test samples. The lack of observed defects close to the bond interface edges is apparent.

\[
\text{Si} - \text{OH} + \text{OH} - \text{Si} \xrightarrow{\text{polysiloxane}} \text{Si} - \text{O} - \text{Si} + \text{H}_2\text{O} \quad (6.1.1)
\]

The point defects introduced in silicon by γ-irradiation include the vacancy-oxygen complex, which may be easily formed in the oxygen-rich bond interface. The bonding-process induced defects also contain these complexes, preferably located close to the interface [56].

Figure 6.1.2: Schematic illustration of point defects in silicon, specifically the vacancy (A), the divacancy (B), the oxygen impurity (C), and the vacancy-oxygen complex (D), in conjunction with an amorphous silicon oxide at the wafer bond interface. The oxygen-rich region may accumulate vacancies in vacancy-oxygen complexes.
For different kinds of spaceflight environments, other degradation mechanisms may be relevant. A broad survey of observed mechanical degradation of silicon or silicon-related materials in radiation environments is made in Paper IV. For vibration and thermal cycling, the slow crack propagation at subcritical applied loads may occur [48, 63, 89].

6.2 Ground tests of spaceflight environment

For a preventive understanding of the influence of spaceflight environments, suitable facilities for re-creating the proper conditions must be found on ground. Here follows a quick list of available test-beds.

The launch vibrations can be addressed in programmable vibration equipment, sweeping through vibration frequencies at different magnitudes for extensive numbers of cycles.

The thermal variations may easily become quite severe on a spacecraft, especially a nanospacecraft due to the low thermal mass. The variations are caused by travelling through the earth shadow, by making slow manoeuvres, or by power dissipation fluctuations from onboard systems. The thermal cycling is readily accomplished in a thermal test chamber, or rapid thermal annealing (RTA) furnace for quicker swings, on component level.

\(\gamma\)-irradiation can be obtained from a Co\(^{60}\)-source, yielding electromagnetic radiation of 1.17 and 1.33 MeV.

Proton irradiation can be achieved by using a cyclotron source, together with an accelerator storage ring for convenience, e.g. the Celsius ring in Uppsala. Proton irradiation in space comes mainly with the solar wind, and can reach extreme fluxes during solar proton events.

Energetic electron irradiation can be produced in a linear accelerator. Electrons are commonly encountered in the van Allen belts.

Heavy ions may impact spacecraft from the solar wind or as cosmic rays. Heavy ion irradiation on earth can be accomplished using the cyclotron source (see protons) or a tandem accelerator. The specific species and energies may by varied and even combined into an “ion cocktail”. Normally, the heavy ions themselves do not reach the bond interface. Rather, the secondary effects may cause degradation of the bond.
6.3 Rank-sum test diagnosis

Two rank-sum tests that respectively detect significant variation from the reference fracture pressure level and scattering are suggested [90]. The spot series subjected to different spaceflight environmental tests can diagnose the possible degradation from the reference series. Parametric statistical tests are not suitable due to the unbalanced series sizes.

The fracture pressure results from each wafer pack is considered a spot series, and tested for significant variations from the larger reference series using statistical rank sum tests [Paper IV]. The test uses the test variable $U$, according to [90]

$$U = n_1 n_2 + \frac{n_1 (n_1 + 1)}{2} - r$$  \hspace{1cm} (6.3.1)

where $n_1$ is the number of samples in the spot series, $n_2$ is the size of the reference series, and $r$ is the rank sum of the spot samples in the combined ordered series. $U$ can be compared to tabulated critical values $u$, dependent on the size of the spot series and the reference series, respectively, and the significance level. The critical value is derived from basic combinatorial issues on how the obtained rank sum may be formed. Thus, the test variable $U$ yields the significance of the hypothesis of the spot series belonging to the same distribution as the reference series by the comparison of $U$ and $u$ [90].

Degradation from the fracture pressure level is diagnosed by using the ordinary rank sum of the spot series in the sorted union of the reference series and the spot series samples. Degradation diagnosis of the fracture pressure scatter uses the fracture pressure data corrected for the separate series mean. Rank numbers are attributed by circling origin, beginning with the lowest corrected measurement value, followed by the highest and the second highest, proceeding with the second lowest and third lowest, etc. The last rank number is omitted if odd.

These rank sums of the spot series are subsequently calculated. Furthermore, the different rank sums, both denoted $r$, follow

$$r \in Z \frac{1}{n_1, n_2 \text{increase}} \in N(\mu, \sigma)$$  \hspace{1cm} (6.3.2)

The $Z$ distribution of the rank sums is determined solely by combinatorial results on picking $n_1$ positions from a set of $n_{\text{tot}}$ ($n_{\text{tot}} = n_1 + n_2$) and adding their respective position rank numbers. $Z$ approaches the normal distribution as $n_1$ and $n_2$ increase.

---

*Bringing Silicon Microsystems to Space*  
41
Now, for the rank sum of a spot series of \( n_1 \) samples — belonging to the Z distribution — from a total data series of \( n_{\text{tot}} \) samples, the mean rank sum estimate \( m^* \) and estimated standard deviation \( s \) of the approximate normal distribution of Z can be derived.

\[
m^* = \frac{n_1(n_2 + 1)}{2}
\]  
\[
s = \sqrt{\frac{n_1 \cdot n_2 \cdot (n_{\text{tot}} + 1)}{12}}
\] (6.3.3) (6.3.4)

The significance of the hypothesis \( H_0 \), that the spot series has not degraded with respect to the reference, can now be tested by the approximate test variable \( U^* \), thus

\[
U^* = \frac{r - m^*}{s} \in N(0,1)
\] (6.3.5)

Using \( U^* \) in the standard normal distribution yields the probability \( p \) of the hypothesis, for both the fracture pressure level and the fracture pressure scatter. However, the test must be double-sided, so the significance level will be at double \( p \).

In this way, the hypothesis can be tested by proper comparison of the test variable to tabulated critical values, or else the approximate test variable belonging to the normal distribution can be used. The approximation is better for increasing series sizes.
7. WEIBULL FRACTURE PROBABILITY METHOD FOR WAFER BOND EVALUATION

It’s getting late for scribbling and scratching on the paper.
Something’s gonna give under the pressure
And the cracks are already beginning to show.
Derek William Dick

The fracture strength of bonds needs to be evaluated in order to obtain mechanically reliable silicon microstructures and to correctly assess the effects of changes in the bond brought about by different treatments or environments. The previous chapter described a diagnostic tool for detecting bond degradation. More thorough inspection of the wafer bond quality is required in order to establish its detailed characteristics. Here, I will outline the basics of a robust evaluation method, developed from the basic blister test [91] for use on wafer bonding in particular [92], and adapted to the needs of quality assessment for space applications [Paper V].

The method is based on fracture probability measurements on direct bonded structures. The strength of these components, as for any manufactured from brittle materials, depends on the combination of defect size frequencies, fracture surface energy distribution, and stress distribution. When the stress around a defect reaches a critical level, fracture is initiated. If the stress is well defined and known, the fracture strength can be determined by the joint frequency distribution of defects and fracture surface energy distribution.

The Weibull fracture probability method proposes to indirectly estimate these distributions at the interface of bonded silicon wafers. Thus, when the bond limits the strength of a component, the problem of finding the strength will be the problem of finding the stress distribution at the bonded interface.

Once fracture is initiated at the bond interface, the fracture propagates spontaneously, either in the bond or into the bulk materials bonded together. The propagation tendency into the bulk is governed by the parameter X, which is described by the relation [93]

\[
X = \frac{G}{2 - 4\nu}
\]  

(7.0.1)

where the fracture preferentially propagates into the material of lowest X. G is the shear modulus and \(\nu\) is Poisson’s ratio. For anodically bonded Pyrex™ (X=28 GPa) to silicon (X=92 GPa) samples, the fracture thus propagates into the glass. This holds for spontaneous crack propagation where the bond is strong enough not to permit propagation in the interface. Even so, the fracture
depends on the quality of the bond interface, which is indicated by inspection
of the fractured samples (Figure 7.0.1) and e.g. the fracture data from test
series annealed at increasing temperatures (Fig 7.0.2) [Paper V, VI].
Evidently, for a weak or in other respects flawed bond, the fracture may
indeed propagate in the bond interface, regardless of the X parameter of the
bulk materials. Perhaps the equivalent of the X parameter can be derived for
the interface, making the spontaneous crack propagation tendency truly
predictable.

The experimental fracture probability, or fracture failure ratio of the test
series, are determined using the interval mid-rank position, recommended to
minimise errors in the subsequent Weibull analysis [94], i.e.

$$P_{r}^{exp} = \frac{i - \frac{1}{2}}{n}$$ (7.0.2)

Here, $P_{r}^{exp}$ is the experimental fracture probability, i is the index of ascending
order in the sample series, and n is the number of samples in the series.

Figure 7.0.1: Top views of fractured samples from different tests. A) Silicon-
silicon fusion bonded sample annealed at 700°C. Fracture was initiated at the
bond interface at the middle of the square side of the cavity and propagated
into the bulk silicon. B) Silicon-silicon fusion bonded sample annealed at
300°C. The fracture has propagated along the bond interface before being
diverted into the bulk. C) Anodically bonded Pyrex<sup>TM</sup>-silicon sample. The
fracture propagated into the bulk glass. (From Paper V and VII)

By the Weibull fracture probability analysis, it is hereafter possible to derive
geometry-independent parameters, which makes comparisons and predictions
of the fracture behaviour for differently shaped and sized structures possible.
This is demonstrated by the results in Figure 7.0.3, where the fracture
behaviour of differently sized square cavities has been successfully predicted
from the fracture pressure data on the middle-sized cavity samples [Paper V].
Figure 7.0.2: The fracture behaviour of silicon-silicon fusion bonded samples annealed at increasing temperatures. Top: The fracture probability as a function of applied pressure for test series annealed at 120°C, 300°C, 700°C, 800°C, and 1050°C. Bottom: A comparison of mean fracture pressure, mean fracture tensile stress of unit length obtained from the Weibull analysis, and fracture surface energy as a function of annealing temperature. The fracture surface energy data are reproduced with permission from Q-Y Tong [95]. (From Paper V and VI)
Figure 7.0.3: Fracture probability prediction for samples of different cavity sizes, i.e. 4-mm, 5-mm, and 6-mm squares. The predictions were made on the 4-mm and 6-mm squares, using data obtained from the 5-mm squares. (From Paper V)

7.1 Model

The fracture strength of brittle materials or structures is conveniently described by the Weibull distribution [96, 97]

$$P_r = 1 - \exp\left[\int_{\Omega} \left(\frac{\sigma(t,\phi,\zeta) - \sigma_0}{\sigma_0}\right)^m \cdot dV\right]$$

(7.1.1)

where $P_r$ is the fracture probability, $\sigma$ is the stress distribution in the sample, $\sigma_0$ is the limit under which no fracture occurs, commonly and hereafter assumed to be zero. $\sigma_0$ is the Weibull normalising factor and $m$ is the Weibull modulus. These Weibull parameters are independent on size, geometry, and load case [98, 99] which make them ideal for characterising the influence of treatment on bond strength.

46  Johan Köhler
The integration in Equation 7.1.1 is made over the stressed volume, as fracture might initiate from any stressed defect in the sample. However, if the fracture dominantly initiates from a surface or an edge, Equation 7.1.1 can be rewritten. Analogous to surface fracture [98], the formula for edge fracture is

$$P_e = 1 - \exp \left( - \int_C \left( \frac{\sigma_C(s)}{\sigma_0} \right)^m \cdot ds \right)$$  \hspace{1cm} (7.1.2)

where \(C\) is the stressed edge and \(\sigma_C(s)\) is the stress distribution on \(C\), using \(s\) as a position parameter at the edge.

Based on the assumption of conditional survival probabilities [100], clamped plate approximation [101], and numerical approximation of the integral in Equation 7.1.2, the final expression for the fracture probability become [Paper V]

$$P_e = 1 - \exp \left( - \frac{24 \cdot a^2 \cdot q}{\pi^2 \cdot \sigma_0 \cdot h} \cdot 8 \cdot a \cdot T \cdot \exp(-z \cdot m) \right)$$  \hspace{1cm} (7.1.3)

for equally thick bonded wafers of thickness \(h\), containing square pressurised cavities (pressure \(q\), cavity side \(a\)). The parameters \(T (= 0.271)\) and \(z (= 0.937)\) are obtained in the numerical approximation. Using Equation 7.1.3, the Weibull parameters \(\sigma_0\) and \(m\) can be readily extracted by the least square adaptation to experimental results.

The convenient alternative Weibull parameter \(\bar{\sigma}_{fc}\), the mean fracture stress of unit length, is correlated to the obtained result by [Paper V]

$$\bar{\sigma}_{fc} = \frac{0.4011 \cdot e^t}{(c \cdot T)^{\frac{1}{m}}} \cdot \sigma_0 \cdot \Gamma \left( 1 + \frac{1}{m} \right)$$  \hspace{1cm} (7.1.4)

where \(\Gamma\) is the Gamma function. This strangely named parameter, based on former derivations [94, 100], makes comparison of Weibull parameters straightforward, as its unit (Pa) is independent of the Weibull modulus \(m\), in contrast to \(\sigma_0\).
7.2 Parameter interpretation

In Figure 7.0.2 the obtained values of $\bar{\sigma}_{fc}$ follow the behaviour of the mean fracture pressure and fracture surface energy alike. This means that the fracture behaviour characterised by $\bar{\sigma}_{fc}$ is determined by the fracture surface energy as measured by the crack opening method [102]. The influence from the frequency distribution of defects is negligible or constant over the annealing temperature interval, or possibly somehow transmitted and confounded into the measurement of the fracture surface energy.

In Figure 7.2.1, the fracture behaviour of fatigued samples is shown; here, the lower fracture pressures and consequently the lower $\bar{\sigma}_{fc}$ are assumed to be due to crack propagation [Paper V].

![Graph showing fracture failure ratio vs. applied pressure](image)

Figure 7.2.1: The degradation of the fracture probability due to vibration tests or thermal cycling of silicon-silicon samples annealed at 700°C. (From Paper V).

It should be possible to detect significant contributions from either the fracture surface energy distribution or the frequency distribution of defects to the fracture behaviour of bonded samples. The Weibull modulus $m$, being a scatter parameter correlates to the frequency distribution of defects present at the bond. Together with $\bar{\sigma}_{fc}$, $m$ thus describes the joint fracture surface-energy distribution and the defect population of the bond interface. An example from low-temperature bonding investigations using oxygen plasma activation is presented in Figure 7.2.2 [Paper VI].
Figure 7.2.2: Fracture probability for silicon-silicon samples activated by oxygen plasma and bonded in vacuum. One series is stored at room temperature; the other is annealed at 300°C. The scatter of the series differs substantially. (From Paper VI).

Here, the fracture behaviour of oxygen plasma activated samples stored in room temperature is compared to samples annealed at 300°C. The higher annealing temperature expectedly yields a higher fracture pressure (and $\bar{\sigma}_f$) due to increasing fracture surface energy, but the data simultaneously become more scattered (lower m). This is attributed to either the annihilation of the dominant defect size responsible for fracture present at room temperature [Paper VI], or the workings of the slow fracture mechanism [52, 54], making the fracture surface-energy distribution non-uniform, or even introducing nucleation sites for microcracks and voids. The latter suggestion would be consistent with the reported occurrence of voids at the 300°C annealing of these samples [Paper VI].

The reliable estimate of the Weibull parameters requires a good amount of samples. In my experience, 20 samples are a minimum. Naturally, the parameters can be extracted using less data, only the reliability of the results, and thus the confidence in them, will suffer.

The Weibull modulus appears as a power number in an exponential function (Equation 7.1.1). This implies that m is a highly compressed figure, and thus difficult to resolve at high values. In other words, the significance of absolute
differences in \( m \) depends on the magnitude of the involved moduli. At larger \( m \), the difference becomes increasingly insignificant. This means that as \( m \) increases, the absolute value obtained becomes experimentally more unreliable. For practical reasons, \( m \)-values higher than 20 may be difficult to compare and above 30 they can be considered equal for the series sizes and accuracy of my experiments.

The same reason makes the \( \overline{\sigma}_{\text{fc}} \) (or \( \sigma_0 \)) the dominant Weibull parameter at high Weibull moduli \( m \). As \( m \) increases, the value of \( \overline{\sigma}_{\text{fc}} \) simply become decreasingly sensitive to variations in \( m \). Simultaneously, the value of \( \overline{\sigma}_{\text{fc}} \) approaches \( \sigma_0 \) as \( m \) increase.

In the present study, a two-parametric Weibull distribution has been used. Equation 7.1.1 states a three-parametric function, where the third parameter \( \sigma_u \) is routinely kept out of further analysis. The interpretation of this parameter is the stress under which no fracture occurs, a kind of safe limit, which normally would not be allowed for brittle materials. On the other hand, this is a precaution based on the macroscopic properties of most engineering brittle materials, which do exhibit fatal flaws, thus making the assumption valid.

What if the bond interface of silicon wafers does not adhere to the macroscopic features of brittle materials? Single-crystalline silicon itself is brittle enough, but does not exhibit these flaws over standard wafer dimensions. Thus, it could be justified to introduce this parameter in an extended study, trying to establish if there exists a safe limit on the stress of a bonded silicon microsystem. In this way, perhaps the concluded dynamic bond interface (paragraph 6.1) may be experimentally stabilised.

### 7.3 Analysis of variance

The fracture data can be subjected to other statistical tools as well. Using a generalised linear model, an analysis of variance (ANOVA) is possible. Thus, the total variance of the response (fracture pressure) of all series can be separated into the factors describing the bonding parameters. This approach has been executed on anodically bonded test series (Paper VII), using the factors Temperature (four levels, 300-450°C) and Method (two levels, shock or ramp).

The separated variances are subsequently tested against the supposed pure experimental error. Here, the experimental error comprises both the scatter from the defect distribution in the bond and measurement errors. Thus, this combination overestimates the true random error, yielding a conservative test of significance in the analysis of variance. The generalised model also included factor interaction of temperature and method. The results show that
the two factors significantly interact [Paper VII]. Accordingly, there is a significant contribution to the variance from both main effects too.

The standard deviations of the different test series are plotted in Figure 7.3.1. Bartlett’s test for variance homogeneity indicates that the variances differ significantly. This normally limits the use of ANOVA, as the experimental error seems to vary for different test series. However, the use of fixed effects factors tends to make the analysis less sensitive to inhomogeneity of variances, if the samples in question are of about the same size. Indeed, the spread of data is expected to vary with experimental conditions, as this would correlate to different defect distributions in the bond.

Figure 7.3.1: Standard deviation plot of different test series with 95% confidence intervals. The standard deviations seem lower and narrower for the shock method as compared to the higher and wider variance of the ramp method. Bartlett’s test shows a significant variation of variances (P < 0.05). (From Paper VII).

The homogeneity of the series variances can be tested by Bartlett’s test [103]. The null hypothesis $H_0$ – that all variances are equal – can be tested by a tabulated test variable.

The result from Bartlett’s test is valid even for unbalanced test series sizes, and can consequently be brought to bear on the diagnostics of significant variations in the data scatter. This, by implication, points to a significant difference in the Weibull modulus value $m$. In figure 7.3.1, the result from Bartlett’s test together with visual inspection suggest a significant difference of the Weibull moduli dependent on the choice of bonding method.
8. CHALLENGES AND OPPORTUNITIES OF THE SPACE INDUSTRY

The remarkable constraints on autonomy, data-handling and data-transmission capability, power requirements, operational reliability, available space, and mass allowance involved in enabling space missions have always invited massive research and engineering efforts in order to succeed. However, in any engineering or research project of this magnitude, there is risk present: risk threatening the schedule, the technical performance, the project budget, the life and health of those involved, or the scientific, strategic, or commercial output of the mission [104].

The risks to the mission are posed by the unforeseen snags in developing new systems prior to launch and by the hazards of spaceflight itself. Naturally, there are no commonly available repairmen in space either, barring the International Space Station and the occasional extra-vehicular activities from the Space Shuttle. So, all systems are expected to be delivered punctually at launch date, survive launch, and operate flawlessly, unattended, and preferably autonomous, while delivering state-of-art data for scientists or others.

Originally, at the dawn of space exploration, new science and technology were researched and developed in parallel and enormous results were gained (e.g. Man on the Moon). However, as space travel stagnated into an accepted feature of everyday life, the research in technology lagged behind. This is probably due to a very natural precaution from basic risk analysis schemes.

In conventional risk analysis, you simply list all identified risk items and grade their respective severity and probability of occurrence. Two medium grades, or a combination of one high and one low grade, classically designate the item as unacceptably risky and demands risk mitigating actions. In this way, risk can even be translated into the expected financial loss [105] due to a particular risk item.

Once space was accessible, the sky was suddenly not only the literary limit to the science that could be done supported by the available technology. Then, why run the risk of ruining your scientific outcome by using unproven technology? In the commercial sector this hold even truer. In the following, I will limit myself to the scientific example.

Johan Köhler
So, while the development risk or mitigation investment for scientific payloads was routinely accepted, platforms were delivered using technology refined from the 1960’s. Inventive engineering was still employed, but the research into new basic technologies decreased over the years. Still, any new device had to be painstakingly tested in order to build confidence in its reliability under the extreme space conditions.

Eventually, new ideas surfaced with the changing times. The possibilities of miniature spacecraft and spacecraft constellations were predicted and gained credibility during the 1990’s. Although conventional miniaturisation could press the limits at first, and really small but comparably low-performing minute satellites kept up appearances for a while, it is ultimately evident that new miniaturisation technology is needed for these kinds of spacecraft.

In order to introduce the new miniaturisation technology, the lengthy protocols of quality assessment was targeted as a main obstacle. Simultaneously, the increasing cost of developing and launching huge and extremely competent spacecraft like the Hubble Space Telescope, the Galileo mission to the Jovian system, and the Cassini/Huygens probe to Saturn/Titan raised concerns about the soundness of this trend in space exploration. The proposed solution was the Faster, Better, Cheaper (FBC) program at NASA and similar approaches at other space agencies.

The use of FBC expressively aims at accepting risks to the space project in part or as a whole in order to free resources for other aspects such as required mass, development time, quality assurance, cost, or performance [106]. For example, higher performance could be gained by new devices, while accepting the risk that these may malfunction in space. Decreasing the redundancy of subsystems, obviously at a calculated risk, may likewise reduce the mass. The prudent use of the FBC is based on a correct understanding of the present risks and how to efficiently trade them for faster development, better performance, and cheaper total cost. And, at the end of the day, this program coldly anticipates spacecraft failures.

In retrospect, the FBC approach was culminated due to two failed Mars missions, while other missions within the program were highly successful, e.g. the NEAR Shoemaker rendezvous with Eros. However, there is no fault to the reasoning here; fortune simply takes her toll of these promptly developed, highly capable, and cheap spacecraft due to meddling with risk as a resource. Perhaps the return would have been better, had only mature technology been allowed. But then, we would still be refining old technology.

This FBC strategy obviously encouraged the miniaturisation by microsystems technology. Perhaps this was part of what FBC was all about?
Indeed, the objectives are twofold, and must retain some separation. First, new tempting missions demand the introduction of entirely new miniaturisation technology. Second, the cost of space missions seemingly soared and had to be reduced. Obvious parallel efforts were identified: reduced mass means reduced cost, reduced demands on risk mitigation imply shorter development time, which in turn saves money and facilitates the introduction of new technology, and the enabled space missions promised better performance with the courageous use of the smaller spacecraft. This was indeed a deliciously baited trap.

Now, there is still an existing mission initiative, and the FBC approach has not been a disaster. I believe the miniaturised space missions have advantages that make them indispensable to future space exploration. They can even reduce the space mission cost, be produced more swiftly, and outperform the conventional state-of-art spacecraft that we know today. Why, then, would not FBC work?

In my view, the FBC cannot accomplish the simultaneous development of a new basic technology and the production of these exciting spacecraft within its dedicated framework. I do not claim that the so-called failure of the FBC was due to improper introduction of miniaturisation technology. However, in the light of the reported failures that have occurred, the FBC approach will be unable to handle both of these issues based on an immature but enabling technology. The maturity of the microsystems technology must be improved, at its own cost, separated from the space mission cost, otherwise the risk will be unacceptable.

The choice ultimately boils down to one between trusting in your good fortune or making a substantial contribution to the scientific advance of the microsystems technology.

8.1 Technology as a scientific endeavour: Building maturity.

In contemporary coffee-break discussions, the science of technology is recurrently questioned and defended. Some characteristics of science and technology seem difficult to explain within the same set of rules, and the imaginable sub-status of one with respect to the other works both ways.

Descriptive definitions from the dictionary [107] tell that science is “the intellectual and practical activity encompassing the systematic study of the structure and behaviour of the physical and natural world through observation and experiment” while technology means “the application of scientific knowledge for practical purposes, especially in industry”.

54

Johan Köhler
The efforts in science have been described as fact gathering (implying discovery), puzzle solving, and invention of theories or models, all based on a conceptual pattern of thought: the paradigm [108].

These patterns of thought, once established, form the framework within which the scientific exploration continues, meaning the discovery, gathering, puzzle solving, and modelling in the context of that framework.

In technology, inventions are rampant, but these are not uncovered facts of technology, rather they are creations, comparable to the invention of models in science, based on discovered facts about nature.

Invention (of models) in science eventually imply a change of paradigm [108], while creative invention is the core activity in engineering technology, far from introducing new patterns of thought on the level of paradigms.

Now, the advance of science lies in the invention of models and discovery of facts. The inventions of technology, however, serves only to advance the discipline of technology in itself, while the science of technology is advanced by pushing the limits of the solid foundation from which engineering inventions are readily made. In other words, this pushing of limits is analogous to fact gathering and modelling, and thus (in spite of the dated status of Kuhn’s original view on paradigms) a valid illustration of the science of technology, or any particular technology, like the microsystems technology.

Figure 8.1.1: The scientific endeavour of technology: The shovelled ground represents the maturity of the technology. The advance of technology by engineering invention and the advance of the science of technology by pushing the limits of the discovered foundation. (Illustration: Karl Astrand).

---

*Bringing Silicon Microsystems to Space* 55
Using this evocative descriptive definition for the science of technology and engineering invention (Figure 8.1.1), the extent of this solid foundation in any technology represents the maturity of the technology in question. Consequently, the solid foundation also provides the confidence of the end user in the promised engineering possibilities.

Returning to the issue of bringing silicon Microsystems to space, the maturity of the Microsystems technology certainly needs improvement, which is the undertaking of this thesis. As a consequence, this will build confidence in the technology itself, paving the way for silicon microsystem integration in spacecraft design and providing a reliable basis for engineering invention in the field.

Evidently, the long-term exploitation of any new technology requires research in order to grasp properly the foundations of that technology. In this context, the courageous confidence in good fortune seems an ill-advised strategic delay at best and a fatally short-termed waste of efforts at worst.
SUMMARY OF APPENDED PAPERS

Ingen ände är på det mycken bokskrivandet,
och mycket studerande gör kroppen trött.
Predikaren 12:12

The appended papers form the foundation from which this thesis emerged. In the following brief summaries, emphasis will be given to the specific contribution of the particular papers to the theme and conclusions discussed previously.

Summary of paper I: Multifunctional Design of Microsystems for Space Applications

This paper details the benefits of extremely compact multifunctional microsystems for future space missions. Emphasis is put on the profits possible by enabling new missions, reducing launch cost, increasing the mass efficiency, relaxing the spatial requirements, and reducing the life-time cost of multifunctional microsystems by fundamental research in microsystems technology.

The water concept for multifunctional microsystems in spacecraft platforms is described. Here, water is used in a highly integrated microsystem for diverse purposes, e.g. propulsion, energy storage, heat management, and radiation shielding.

The progressive figure of merit is suggested for the manufacture process sequence development for specific microsystems.

This paper presents the central ideas for developing multifunctional microsystems and wafer bonding is identified as a vital part of the suggested extreme integration of microsystems.

Summary of paper II: A Hybrid Cold Gas Microthruster System for Spacecraft

A hybrid cold gas microthruster system suitable for low Δv applications on spacecraft has been developed. Microsystem subsystems together with fine mechanics form the microthruster units, integrating four independent thrusters.

The verification of the thruster provided input for an analysis of the system performance. The total system performance has been estimated in two parameters, the system-specific impulse and the mass ratio of the propulsion
system to the spacecraft mass. These figures can be used in spacecraft design and manufacture.

The demonstration of the hybrid integration of the cold gas microthruster microsystems helped in illuminating the issues of process compatibility and manufacturing sequence, while striving towards the fully integrated version.

**Summary of paper III: The Manufacture of an Integrated Silicon Cold Gas Micronozzle System for Spacecraft**

This paper gives a detailed description of the manufacture of the nozzle unit used in the hybrid cold gas microthruster system. The reasons for deciding on different processes are related, and process monitoring methods for manufacturing development are presented.

The discussion focuses on the line of processing, and process sequence considerations. The concept of the progressive figure of merit in a microsystem manufacturing sequence is used to suggest general recommendations. Expensive high-yield processes should be preferred late in the fabrication sequence, while initially the cost of the processing is more important. The production capacity gain from using batch processes is more valuable early in the processing rather than late, as a single process failure here will likely destroy the entire batch. Cheap processes should be used early in the line of processing. Later, the use of expensive processes is acceptable if the yield is high enough to protect the approved silicon parts.

The graphical representation of the progressive figure of merit was introduced in order to grasp the coupled influences of yield, process time, process cost, and batch size on the pertinence of the line of processing.

**Summary of paper IV: Mechanical Reliability of Silicon Microsystem Fusion Bonds in Spaceflight Environment**

The impacts of expected launch conditions and space environment on the mechanical reliability of silicon wafer fusion bonds were investigated. Rank-sum tests on burst test data were used as a diagnostic tool to detect degradation of the bond quality.

Degradations have been successfully detected in the fracture pressure level for short-duration vibration or thermal cycling (-196°C-200°C) of insufficiently annealed samples, and for low and high doses of γ-irradiation (6 and 100 krad). Intermediate dose of γ-irradiation (22 krad) improves the fracture pressure level. Degradation of the fracture pressure scatter has been detected for Xe-irradiation, and improvement for thermal cycling (-196°C–200°C) of

---

Johan Köhler
insufficiently annealed samples. Vibration of 1050°C annealed samples, more lenient thermal cycling, and other types of irradiation failed to significantly influence the bond reliability.

This investigation was crucial to the conclusion that the bonded interface of silicon wafers should be considered an inherently dynamic system, which must be meticulously secured.

Summary of paper V: Weibull Fracture Probability for Silicon Wafer Bond Evaluation

This work presents a strategy towards determining the mechanical reliability and quality of bonded silicon microsystems. The fracture strength of a bond has been examined using burst tests and Weibull statistics. The testing method in itself exhibits small errors, is independent on operator, and works for weak and strong bonds alike. By the Weibull fracture probability approach, bond characteristics can be derived, which makes comparisons and predictions of differently sized and shaped structures possible.

This paper started my understanding of the fracture behaviour of bonded wafers. The method developed here was central to the degradation diagnostics and further insights from following bond characterisations. The basic application of Weibull fracture probability to the evaluation of bonded microsystems also inspired the suggestion of future use of the three-parameter Weibull distribution.

Summary of paper VI: Oxygen Plasma Wafer Bonding Evaluated by the Weibull Fracture Probability Method

The Weibull fracture probability evaluation method was employed on low-temperature bonded wafers, where the surface activation was accomplished by oxygen plasma treatments. The bond quality was characterised after different annealing processes. The interpretation of both Weibull parameters was used for describing the bond quality in terms of the fracture surface energy and the frequency distribution of defects at the interface.

Oxygen plasma activation for silicon wafer bonding was confirmed to be a promising bonding method, especially if extensive annealing at low temperature is permitted. Bonding strengths comparable to 700°C annealed samples using wet chemical activation were obtained. However, the tendency to form voids during anneal may limit the applications where large bonded interfaces are required.
This paper highlights the influence of different features of the bond interfaces on the resulting fracture behaviour. The major issue of bonding process compatibility was pursued by the investigations into low-temperature bonding. The paper delved further into the interpretation of the Weibull parameters, concluding the models initiated in Paper V.

**Summary of paper VII: Weibull Fracture Probability for the Characterisation of the Anodic Bond Process**

This paper applied the burst test series approach to anodically bonded samples. Two distinct bonding methods were used: shock (applying voltage at process temperature) and ramp (raising the temperature at applied voltage). Four process temperatures were used for each method. The strength of anodic bonds, processed at temperatures between 300°C and 450°C, are comparable to silicon fusion bonds annealed at temperatures above 800°C. This comparison of anodic and silicon wafer bonding is straightforward using the Weibull parameters.

The choice of method had a significant effect on the bond quality, diagnosed by Analysis of Variance (ANOVA). The shock method yields stronger bonds and lesser spread of data. The significant variation of the Weibull modulus (scatter) was indicated by the interpretation of Bartlett’s test for homogeneity of variances.

**Summary of paper VIII: Silicon Fusion Bond Interfaces Resilient to Wet Anisotropic Etchants**

The attacks on exposed bond interface edges from anisotropic etchants have been studied in order to obtain chemically resilient systems. The bond interface can be successfully etched using a bond configuration of two intermediate thin thermal oxide films. These films must be removed by plasma etching.

The results of the investigations show a vivid example on intricate process compatibility, and illustrate that the bond is indeed a weak link in silicon microsystems. Inappropriately chosen bonding and etching procedures cause severe damages to the precision of the resulting microstructures. These damages may be minimised by perfecting the wafer bond strength, specifically the bonding energy of the bulk silicon atoms to the bond oxide.
ACKNOWLEDGEMENTS

Och Knyttet tog av skorna och suckade och sa:
"Hur kan det känns sorgsamt fast allting är så bra?"
Men vem ska trösta Knyttet med att säga "Lilla vän,
vad gör man med en snäcka om man ej får visa den?"
Tove Jansson

Many people deserve my gratitude, and I therefore express this by dedicating central thoughts of my thesis to these worthy persons or entities. First, Lars Stenmark, "an inventor of grand and unusual things" [109], superb supervisor, and a good friend. Any arbitrary decision represents a wasted opportunity.

ESA and the AME Programme at Uppsala University for major financing. The trio Sören Berg, Bertil Hök, and Jan-Ake Schweitz for employment and inspiration. My supervisors of various kinds: Ylva Bäcklund, Lars Rosengren, Christer Hedlund, Ulf Lindberg, Klas Hjort, and Stefan Johansson. The good officers at ESTEC: Martin Lang and Laurent Marchand. The successful delivery of a single working microsystem forms the natural starting point for a general evaluation of the fabrication process.

My co-authors of papers included and omitted. All members of the Ångström Space Technology Centre and the Micromechanics group at Solid State Electronics. Marianne Asplund, Ann-Sofie Lindberg, and Carin Palm for expertise and patience. The engineers at the FTE and the ASTC for skills and advice. Everybody left at the FTE and friends in other parts of the Academy. You all provide guidance, knowledge, skill, and good cheer! The Weibull parameters describe the joint fracture surface-energy distribution and the frequency distribution of defects at the bond interface.

My meticulous proof-readers and commentators: For the text, you did all but write it! And naturally, the obliging librarians at the Ångström Laboratory promptly found me my references! The pursuit of the desired level of integration and miniaturisation clearly requires a new way of thinking.

Members of the African Domino Guild, for supplying mental rest and vivid imagination, and other friends that helped and/or showed an honest interest in my work. And of course, my parents, my in-laws, my brother, and my sister for too much to mention. The inventions of technology only advance technology itself, while the science of technology advances by extending the solid foundation of engineering inventions.

Finally, Annie, Kasper, and Mia, for making it all worthwhile. Bringing Silicon Microsystems to Space: ...
REFERENCES

However Deceived in Generalities
Men Are Not Deceived in Particulars
Niccolò Machiavelli


62

Johan Köhler


Johan Köhler

66  

Johan Köhler