SI-MICROMACHINED KNUDSEN PUMPS
FOR HIGH COMPRESSION RATIO AND HIGH FLOW RATE

by

Seungdo An

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering)
in the University of Michigan
2013

Doctoral Committee:

    Professor Yogesh B. Gianchandani, Chair
    Associate Professor Luis P. Bernal
    Assistant Research Scientist Scott Ryan Green
    Professor Khalil Najafi

To my wife for her love & support,

with all my love & gratitude.

ACKNOWLEDGEMENTS

financial supporters of the LNF expansion, as the working environment at LNF was substantially improved.

I acknowledge partial support by a fellowship from the Electrical Engineering and Computer Science at the University of Michigan.

I would like to thank my group members, lab mates, and office mates at the University of Michigan for sharing life at Michigan together with me. They always came to me when I was in need of help. **Naveen Gupta** was my predecessor for the Knudsen pump. He gave me many bright ideas and guided me through my beginning stages at Michigan, as he trained me on most of the LNF tools. **Yutao Qin**, who continues the Knudsen pump after me, provided many terrific test results from the high flow rate Knudsen pump. **Seow Yuen Yee, Ali Besharatian, Karthik Kumar**, and I participated in the Michigan High-Vacuum Program, held meetings, shared results, discussed, and submitted reports to DARPA together. I will always remember how much you all mean to me.

I thank my previous supervisors at Samsung, **Dr. Sang Kyung Yun** for his inspiration and ambition, **Dr. Jonghyeong Song** for his logical approach, **Dr. Yongsoo Oh** for his hard work, **Prof. Seungseob Lee** for his softness, **Prof. Geunbae Lim** for teaching me bulk micromachining processes in Japan, and **Prof. Ci-moo Song** for his passion and for accepting me into his MEMS team. They have all provided standards of excellence that I hope to follow.

This dissertation is dedicated to my wife **Seongah**. Without her support, I could not have come to Michigan to enjoy its challenges. I thank her and love her. She is the love of my life. She is the best of me. It is my joy to share with her the longest ride.

Finally, I thank **my parents** for their patience and support. Their high expectations of me made this PhD possible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

μGC	Micro gas chromatograph
μMS	Micro mass spectrometer
ALD	Atomic layer deposition
BTE	Boltzmann transport equation
cc	Cubic centimeters
CEA	French Alternative Energies and Atomic Energy Commission
cfm	Cubic feet per minute
CMOS	Complementary metal–oxide–semiconductor
DOE	Design of experiments
DRIE	Deep reactive ion etching
DSMC	Direct simulation Monte Carlo
DTRA	Defense Threat Reduction Agency
EDS	Energy-dispersive X-ray spectrometry
FEA	Finite element analysis
GC	Gas chromatograph
high-Q	High quality factor
KHP	Knudsen high-pressure part
KLP	Knudsen low-pressure part
KP	Knudsen pump

LPCVD	Low pressure chemical vapor deposition
MS	Mass spectrometer
N/MEMS	Nano- and micro-electromechanical systems
NC	Narrow channel
NCH	Narrow channel height
NSE	Navier-Stokes equation
OLED	Organic light emitting diodes
ONO	Oxide-nitride-oxide
PECVD	Plasma enhanced chemical vapor deposition
PIMMS	Planar integrated miniaturized mass spectrometer
polySi	Polycrystalline silicon
RIE	Reactive ion etching
sccm	Standard cubic centimeters per minute
SEM	Scanning electron microscope
Si	Silicon
SIP	Sputter-ion pump
slpm	Standard liter per minute
SOI	Silicon-on-insulator
TEM	Transmission electron microscope
TOF	Time-Of-Flight
WC	Wide channel
WCH	Wide channel height

LIST OF SYMBOLS

$\wp$	Input power
ΔP	Pressure difference between T_{Cold} and T_{Hot}
$\Delta P_{E,}$	Residual error in pressure
ΔP_{Eq}	Equilibrium pressure difference
ΔP_f	Forward pressure difference
ΔP_r	Reverse transpiration pressure difference
ΔP_S	Net pressure difference in a stage
ΔR	Further increment in resistance at P_{Cal} (See Section 4.2.1)
ΔT	Temperature difference between T_{Cold} and T_{Hot}
a	Height of the channel
A_{heater}	Ti/Pt metal heater
a_T	Thermal accommodation coefficient
A_{Total}	Total area of narrow channels (active pumping area)
b	Width of the channel
$c_{th,air}$	Specific heat of air molecules
$C_{th,nc}$	Thermal capacitance of the narrow channels
$c_{th,nc}$	Equivalent specific heat of narrow channels
$c_{th,sw}$	Specific heat of the sidewall
E_1	Standard deviation from the repeatability of responses of each Pirani gauge

E_{1+2}	Total error in dynamic calibration, where the subscripts *1* and *2* indicate two independent variables
E_2	Standard deviation from the mapping of the response of the measured Pirani gauge to that of the reference gauge with linear regression
f	Fill factors for the unblocked gas flow paths
f_{Avg}	Average fill factor of f_{udm} and f_{hs}
f_{hs}	Fill factors for unblocked gas flow paths in the the heat sink
f_{udm}	Fill factors for unblocked gas flow paths in the upper dielectric membrane
g	Pirani sense gap between the supporting membrane and the heat sink
$h_{conv,air}$	Natural convection coefficient of air (10×10^{-6} W/ K-mm^2 [Sai10]).
I	Current
k_B	Boltzmann constant
l	Length of the channel
$\dot{M}$	Mass flow rate
N_{Total}	Total number of narrow channels
P	Pressure
P_{Cal}	Calibration pressure
P_{IN}	Pressure at the inlet
$P_{m,S,i}$	Pressure value converted from the linearly transformed (or scaled) response of the measured Pirani gauge, $(\Delta R/R_0)_{m,S,i,}$ where *i* denotes an index value of the calibration pressure, P_{Cal}; and the subscripts *r* and *m* indicate the reference Pirani gauge and the measured Pirani gauge, respectively
P_{OUT}	Pressure at the outlet
P_{Ref}	Reference ambient pressure

P_{Std}	Standard pressure of 760 Torr
Q_P	Poiseuille flow coefficients
Q_T	Thermal creep flow coefficients
R_0	Unheated resistance
$R_{0,Ref}$	Increment in resistance at P_{Ref} (See Section 4.2.1)
$\mathcal{R}_d$	Thermal resistance of the supporting membrane
$\mathcal{R}_{ext}$	Equivalent thermal resistance to exterior ambient.
$\mathcal{R}_{sg}$	Pressure-dependent thermal resistance of the sense gap
R_T	Total electrical resistance
$\mathcal{R}_{th}$	Thermal resistance
$\mathcal{R}_{th,conv}$	Thermal resistance of the natural convection cooling
$\mathcal{R}_{th,nc}$	Thermal resistance of the narrow channels
$\mathcal{R}_{th,nc}$	Thermal resistance of the narrow channels
$\mathcal{R}_{th,Ref}$	Thermal resistance at the reference pressure
S	First-order coefficient (or slope)
t	time
T_{Avg}	Average temperature of T_{Cold} and T_{Hot}
T_{Cold}	Cold temperature
T_{Hot}	Hot temperature
T_{Std}	Standard temperature of 273.15 K
t_{sw}	Sidewall thickness
V_{Cold}	Volume at the cold chamber

$\dot{V}_{POS}$	Volume flow rate of gas in the viscous flow regime
$\dot{V}_{Std}$	Standard flow rate
$\dot{V}_{Std,Total}$	Total standard flow rate
$\dot{V}_X$	Volume flow rate of gas
$w_{R,hs}$	Width of the repeating unit in the heat sink
$w_{R,udm}$	Width of the repeating unit in the upper dielectric membrane
$w_{ub,hs}$	Unblocked width of the repeating unit in the heat sink
$w_{ub,udm}$	Unblocked width of the repeating unit in the upper dielectric membrane
α	Thermal coefficient of resistance
γ	Flow coefficient ratio, defined as Q_T/Q_P
γ	Proportionality constant, defined as $R_{0,m}/R_{0,r}$
η	Dynamic viscosity of gas
λ	Mean free path of air molecules
λ_{Ref}	Mean free path of air molecules at the reference pressure ($\approx$0.07 μm)
ρ_{air}	Mass density of air molecules
ρ_{Avg}	Mass density at P_{Avg} and T_{Avg}
ρ_{nc}	Equivalent mass density of narrow channels
ρ_{Std}	Mass density at standard conditions
ρ_{sw}	Mass density of the sidewall
σ	Standard deviation
$\sigma_{0,air,th}$	Thermal conductivity of air in continuum flow regime, 0.026 K/m/W
$\sigma_{th,air}$	Thermal conductivity of air molecules

$\sigma_{th,nc}$	Equivalent thermal conductivity of narrow channels
$\sigma_{th,sw}$	Thermal conductivity of the sidewall
τ_{Hyd}	Hydraulic time constant
τ_{th}	Thermal time constant

ABSTRACT

SI-MICROMACHINED KNUDSEN PUMPS
FOR HIGH COMPRESSION RATIO AND HIGH FLOW RATE

by

Seungdo An

Chair: Yogesh B. Gianchandani

This dissertation focuses on Si-micromachined Knudsen pumps. Knudsen pumps exploit thermal transpiration that results from the free-molecular flow in non-isothermal channels. The absence of moving parts, without frictional loss and mechanical failure, provides significantly higher reliability.

For a high compression ratio, 48 stages are cascaded in series in a single chip of 10.35×11.45 mm^2 area. A five-mask, single-wafer process is used for monolithic integration of the designed Knudsen pump. The pressure levels of each stage are measured by integrated Pirani gauges. Using 1.35 W, the fabricated pumps evacuates the encapsulated cavities from 760 to 50 Torr and from 250 to 5 Torr.

Multistage Knudsen pumps are further explored using a two-part architecture. To increase the compression ratio, 162 stages are serially cascaded. The two-part architecture uses 54 stages designed for the pressure range of 760-50 Torr, and 108 stages designed for lower pressures. This approach provides greater compression ratio and

speed than using a uniform design for each stage in the 48-stage Knudsen pump. The design has a footprint of 12×15 mm^2. Using 0.39 W, the evacuated chamber is reduced from 760 to 0.9 Torr, resulting in a compression ratio of 844. The vacuum levels were sustained beyond 37 days of continuous operation.

The dynamic calibration of microfabricated Pirani gauges is explored for increasing pressure measurement accuracy in the 162-stage Knudsen pump. Test results demonstrated that dynamic calibration can be significantly more accurate than conventional static calibration when Pirani gauges are embedded deep within a microfluidic pathway.

Si-micromachined single-stage Knudsen pumps are explored for generating high-flow rates. A high density of thermal transpiration flow channels is arrayed in parallel for combined pumping operation. A design with 0.4×10^6 parallel channels in a footprint of 16×20 mm^2 generates a measured 211 sccm air flow at a pressure difference of 92 Pa, using 37.2 W.

The low-temperature atomic layer deposition (ALD) of Al$_2$O$_3$ is investigated for vacuum seals in wafer-level vacuum packaging applications. The conformal coverage provided by ALD Al$_2$O$_3$ is shown to seal micromachined cavities. Lifetime tests extending out to 19 months are reported.

CHAPTER 1

INTRODUCTION

1.1 Motivation

Knudsen pumps, first proposed and demonstrated in 1909 [Knu09], present an appealing method for obtaining vacuum. Knudsen pumps are motionless, thermomolecular pumps, which do not have any moving parts. These motionless pumps are based on the phenomenon of thermal transpiration (Fig. 1.1a) [Loe34]. If the hydraulic diameter of a flow channel is no larger than the mean free path of gas molecules (i.e., the flow in the channel is confined to the free-molecular or transitional flow regimes), a thermal gradient induces gas streams from the cold end to the hot end of the channel [Loe34, Car11]. As a result, a pressure gradient is established in the same direction as the thermal gradient. (See Section 2.2 for principle of Knudsen pump operation).

Micropumps targeted at a wide range of applications have been reported in literature [Ste93, Sch94, Dar96, Cab01, Sch02, Ngu02, Las04, Kim07a, Tsa07, Kim11]. Most research efforts have been directed at motion-based pumping mechanisms [Cab01, Ngu02, Kim07a, Kim11], which are prevalent in macroscale pumping instruments. The Knudsen pumping method provides advantages over conventional motion-based pumping techniques.
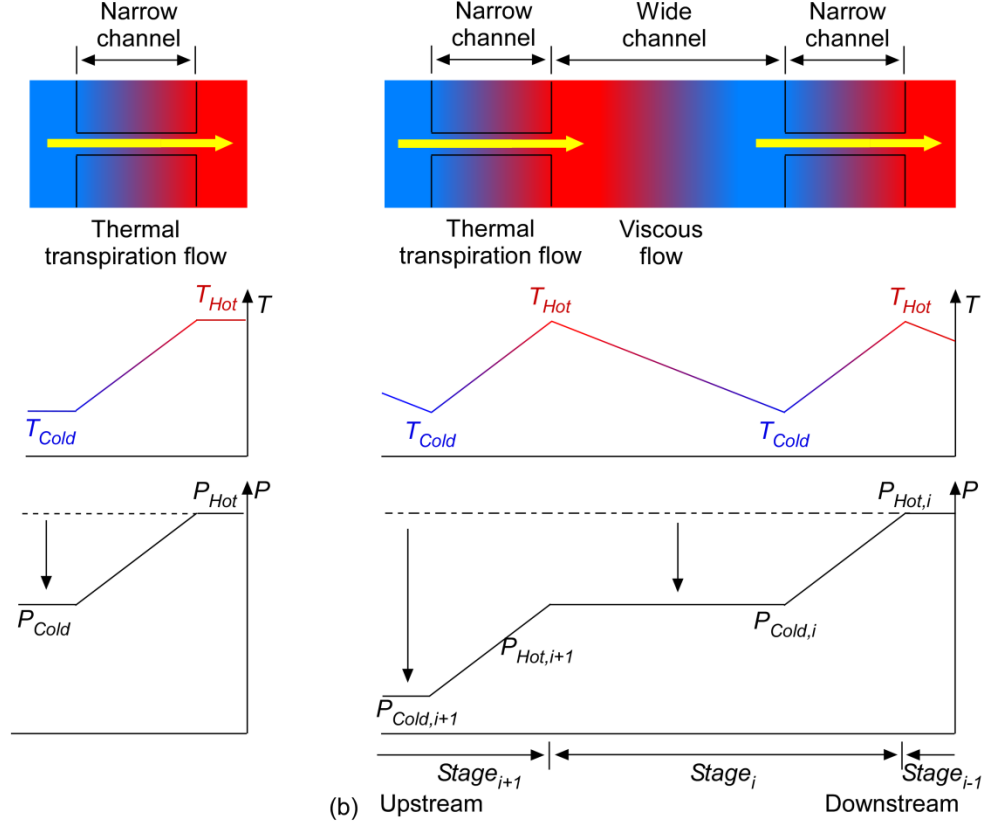
Figure 1.1: Schematic of Knudsen pump operation. (a) Single-stage configuration. (b) Multistage configuration. P and T are, respectively, pressure and temperature, the subscripts *Cold* and *Hot* indicate, respectively, cold and hot chambers, and the subscript i indicates stage number.

First, the absence of moving parts, without frictional loss and mechanical failure, provides significantly higher reliability. Motion-based pumps include, for example, peristaltic and rotary pumps. However, moving parts present challenges at the microscale as the surface area to volume ratio of the components increases. To achieve large compression ratios, it is necessary to build structures that provide relatively large displacement at relatively high frequencies; these are not easily integrated on a microchip. In addition, high frictional losses can compromise operational life. Such challenges from moving parts are absent in Knudsen pump.

Second, Knudsen pumps can potentially have a small form factor [McN05, Mun06]. Miniaturization using lithography-based micromachining processes can enable integration of thermal transpiration flow channels into a chip-scale device.

Third, Knudsen pumps can provide a high compression ratio using serially cascaded configuration [Knud09, You99] (Fig. 1.1b). Viscous flow channels connect thermal transpiration flow channels by restoring the temperature from the hot temperature to the cold temperature with negligible change of pressure. Then, a large pressure difference can be obtained by combining pressure differences developed at each stage.

Finally, Knudsen pumps can generate high flow rates using a parallel configuration of thermal transpiration flow channels. Each thermal transpiration channel has a large flow resistance because its hydraulic diameter is small such that gas flow is confined to the free-molecular or transitional flow regimes. However, by combining a large number of channels in a single stage, the total flow resistance can be substantially reduced, thereby resulting in a high flow rate.

1.2 Vacuum Pumps and Applications

In Fig. 1.2, progress in minimum pressures [Red99] shows that vacuum itself and a vacuum pump were first known in the mid-17th century. Boyle was the first one who made a significant measurement of vacuum, ≈6 Torr, which was generated by a mechanical piston pump. Over time, various vacuum pumps, such as a liquid piston pump, a molecular pump, a diffusion pump, and a getter or cryopump, have been developed. The highest achieved vacuum has been increased continuously. Knudsen, who designed and built the first Knudsen pump, performed comprehensive research on

gas flow related to collisions with surrounding walls. His work is regarded as the beginning of vacuum physics [Jou08]. By the end of the 20th century, the highest vacuum was $\approx 10^{-14}$ Torr using getter and cryopump technology.
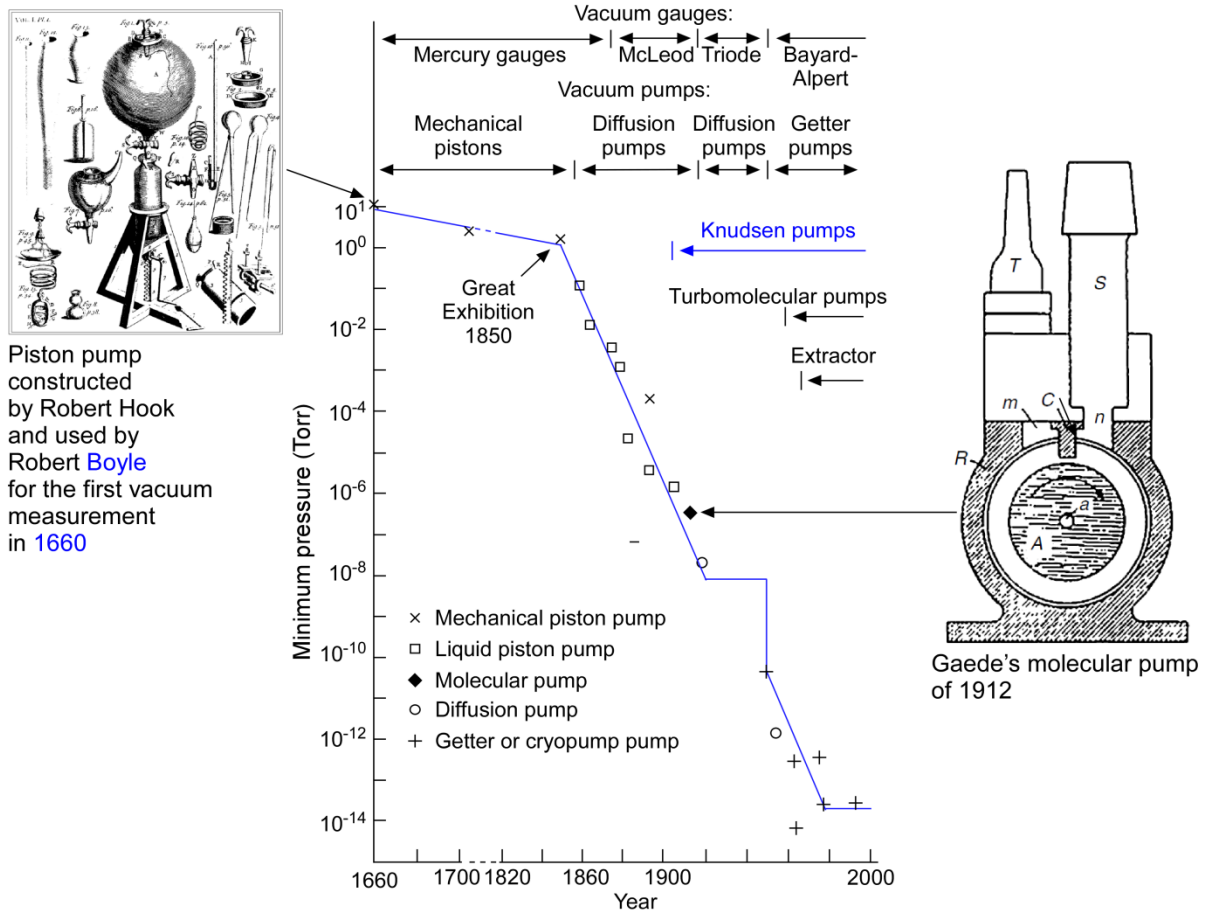
Figure 1.2: Progress in minimum pressures from the mid-17th century to the 20th century [Red99].

Applications range from industry to basic science research (Table 1.1) [Jou08]. Vacuum pumps are important for semiconductor technology, in which pressure ranges are from 10^{-7} Torr to 10^0 Torr. For mass spectrometers in research, pressure ranges are from 10^{-10} Torr to 10^{-5} Torr.

Table 1.1: Applications and pressure ranges of vacuum pumps [Jou08]

Application		Pressure range (Torr)
Industrial area	Semiconductor fabrication	$10^{-7} - 10^{0}$
	Physical vapor deposition	$10^{-5} - 10^{-2}$
	Chemical vapor deposition	$10^{-2} - 10^{1}$
	Vacuum metallurgy	$10^{-5} - 10^{1}$
	Crystal growth	$10^{-6} - 10^{-3}$
Research area	Mass spectrometers	$10^{-10} - 10^{-5}$
	Surface analytics	$10^{-11} - 10^{-3}$
	Plasma research	$10^{-9} - 10^{-5}$
	Elementary particle physics	$10^{-11} - 10^{-8}$

1.3 Vacuum Micropumps and Applications

The term micropump refers to a pump with microscale functional components. It is mostly fabricated by microelectromechanical systems (MEMS) technology. Research on micropumps is directed at meeting challenges for certain applications that require pumps with small size & weight and low power, together with high performance and robustness. Micropumps are potentially useful as part of a tunable vacuum control system in sealed cavities that contain high-Q resonators, such as micro gyroscopes [Liu09] or timing oscillators [Bee12]. Miniaturized analytical instruments, such as micro mass spectrometers (MS) [Hau07, Che10] and micro gas chromatographs (GC) [Pot07, Kim07a, Liu11], also need small vacuum pumps. Miniaturized analytical instruments are being actively researched for portable, commercial or defense applications. One of the challenging requirements for these compact applications is a small vacuum pump of high performance and robustness.

The electrostatically driven comb actuator in a micro gyroscope (Fig. 1.3a) requires a high quality factor for generating a sufficient signal response for a small angular rate input signal of <0.1°/s. A high quality factor much larger than 1,000 is achieved by decreased air damping in the comb actuator at low pressure levels. Typically,

the pressure levels need to be less than 5 mTorr. The present hermetic vacuum sealing method is to bond a device wafer to a cap wafer at high vacuum. Getters are optionally used for absorbing gas molecules, which evolve from interior surfaces. Using a micropump integrated with a microgyroscope, the interior pressure can be controlled and maintained at high vacuum levels, thereby retaining high performance.
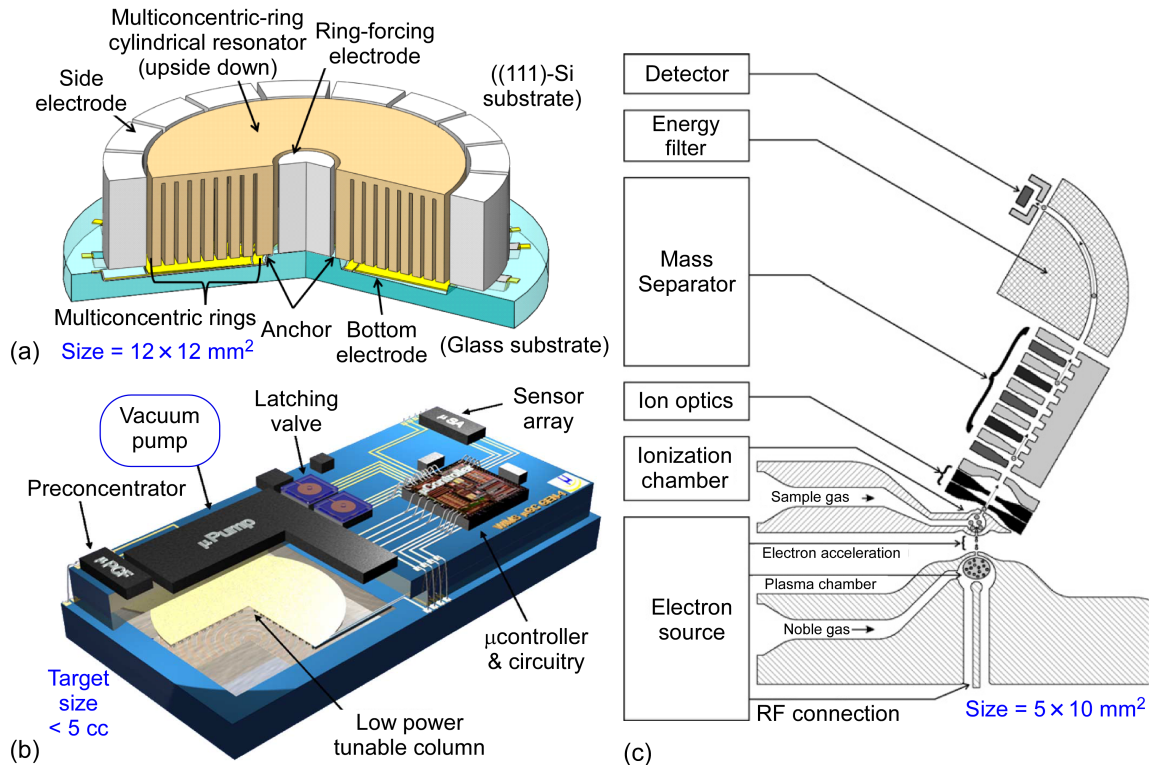
Figure 1.3: Examples of microsensor and miniaturized analytical instruments, in need of a vacuum micropump. (a) Microgyroscope, which demonstrated its rate sensing performance at high vacuum [Cho12]. (b) µGC system showing the relative size of the different components [Pot07]. (c) A single-chip micro mass spectrometer [Hau07].

The micro gas chromatograph (µGC) at the University of Michigan (UM) [Pot07] (Fig. 1.3b) is a miniaturized, portable version of conventional gas chromatographs. Typically, conventional gas chromatographs are bulky and expensive, analytical instruments widely used in laboratory chemical analysis. Gas analyte is carried through a column coated with stationary phase. Gas components are separated using different

6

retention times related to chemical species. The μGC needs a micropump with a high flow rate (> 2 sccm) and a high pressure difference (> 10 kPa) [Kim07a]. In addition, low power (<100 mW) and small volume (< 0.5 cc) are required.

Due to the exceptional analytical performance, diverse types of miniaturized mass spectrometers have been studied [Sym09]. Mass spectrometers electrically charge gas analyte and measure mass-to-charge ratio. The quantitative analysis of molecular fragments results in the determination of the chemical constituents of the analyte. A planar integrated miniaturized mass spectrometer (PIMMS), researched at Hamburg University of Technology in Germany, is based on electron impact ionization and a synchronous, ion-shield mass filter. The PIMMS is microfabricated in a small size of 5×10 mm^2 on silicon-on-glass substrate [Hau07] (Fig. 1.3c). The vacuum chamber should be at ≤ 7.5 mTorr, and the desired pressure of the electron source and ion source is 0.38 Torr. The Time-Of-Flight (TOF) architecture for a micro mass spectrometer, studied by French Alternative Energies and Atomic Energy Commission (CEA), is required to operate at 1 mTorr [Tas11].

Specifications required for micropumps are flow rate, outlet pressure, inlet pressure, compression ratio, size, weight, and power. In vacuum pumps, the flow rate is inversely proportional to the pressure difference, ΔP, between outlet and inlet pressures. For specific applications, a pump performance line and operating point can be determined using a specific design procedure (Fig. 1.4). For example, the gas chromatograph specified by UM, needs a high flow rate (> 2 sccm) and a pressure difference (> 10 kPa) [Kim07a]. Then, the performance line for that application tends to lean toward generating a high flow rate. In the case of a micro mass spectrometer coupled to a gas

chromatograph, (whose specification is released in a Defense Threat Reduction Agency (DTRA) solicitation), both high flow rate and high vacuum are necessary [DTR12]. For this application, the performance line could be designed for generating both a high flow rate and a large pressure difference.
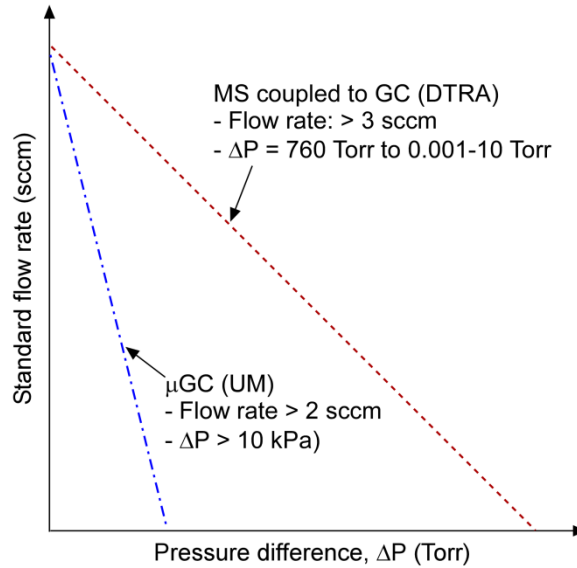
Figure 1.4: Examples of pump performance lines for specific applications

1.4 Previous Efforts in Micropumps

Wide applications ranging from gas sensing and analysis [Ter79, Nag98, Zel04, Wis04] to microfluidic movement [Mar99, Lea03, Wan05, Wan06, Abq07, Tsa07, Wu08, Cho09] have fueled the continuing research efforts in micropumps, which has produced increasing compression ratios and flow rates. The compression ratios achieved in micropumps are summarized with device volume [Ste93, Ger95, Kam98, Wij00, Ngu02, Las04, McN05, Kim07a, Kac08, Web10, Pha10, Gup11a, Gup11b, Mey11, Bes12] (Fig. 1.5). The results of the Si-micromachined Knudsen pumps studied in the work of this dissertation are added in later chapters (See Fig. 2.8, 3.14, 6.15).

A promising result of flow rate and pressure difference for a pump to be used in a micro gas chromatograph was reported by Kim *et al.* [Kim07a] (Fig. 1.6a), using a peristaltic 18-stage micropump driven by electrostatic force. An air flow rate of ≈4.0 sccm and a maximum pressure difference of ≈17.5 kPa were produced, using an input power of only ≈57 mW. The resulting compression ratio was ≈1.1.
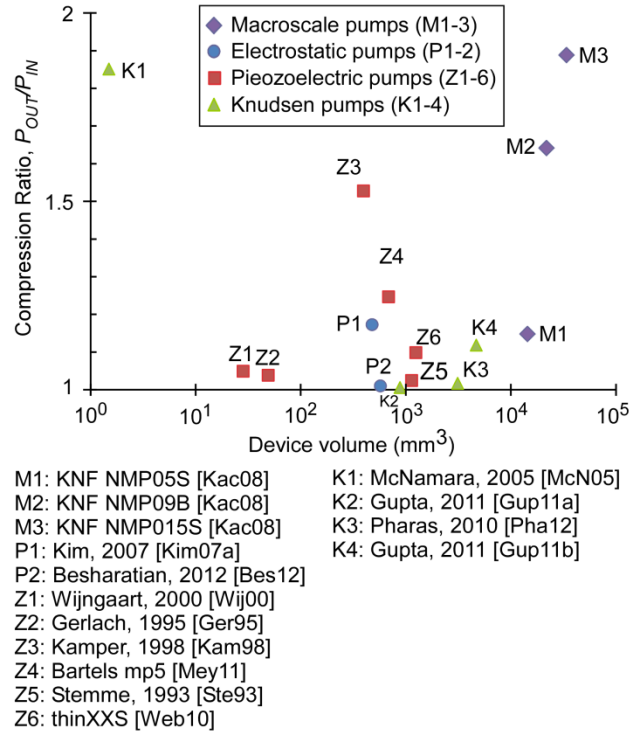
Figure 1.5: Benchmarking of previous micropumps for compression ratio, P_{OUT}/P_{IN}, and device volume.

A compression ratio of ≈4.6 was achieved by the mechanical pump by Zho *et al.* [Zho11a, Zho11b] (Fig. 1.6b). It is not shown in Fig. 1.5 due to the un-published actuator volume located beneath the MEMS structure (Fig. 1.6b). The volume of its MEMS structure only is 1,144 mm^3. The device is designed to be driven by pneumatic actuation using a six-step operation cycle. It achieved 164 Torr at atmospheric ambient pressure. The pressure approaches equilibrium at ≈100 minutes, using the test volume of ≈1 cc.
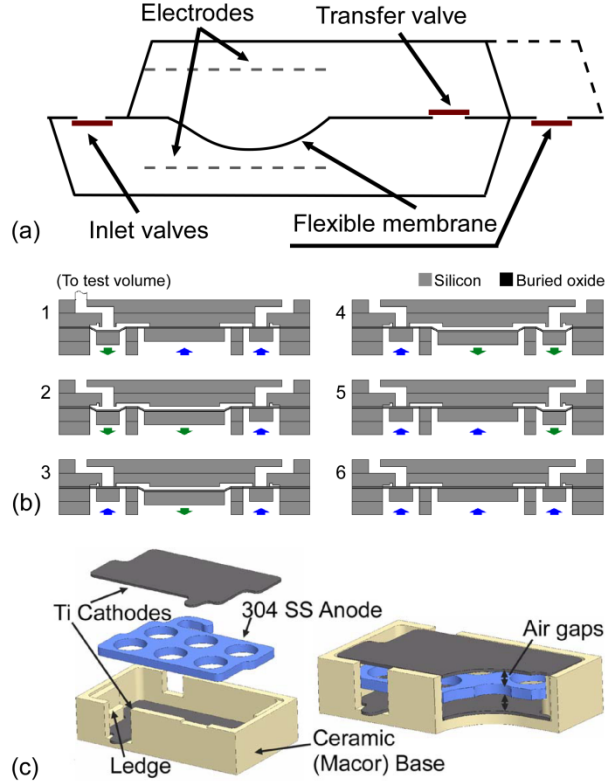
Figure 1.6: Schematics of selected previous micropumps. (a) Peristaltic micro gas pump driven by electrostatic force [Kim07a]. (b) Micromachined mechanical pump using six-step operation cycle [Zho11a, Zho11b]. (c) Chip-scale Penning cell arrays in sputter-ion pump [Gre13].

Another interesting pump that demonstrated pumping at sub-Torr pressure is the sputter-ion pump (SIP). It uses Penning cell arrays [Gre13] (Fig. 1.6c). In a SIP, the continuously sputtered titanium getter layer absorbs ionized gas, resulting in a vacuum in the chamber. This requires a voltage of 450-600 V for plasma ignition. The Penning cell sustains the ionizing efficiency at low pressure levels by trapping the electrons in long orbits, using the Lorentz forces provided by crossed magnetic + electric fields. It demonstrated pressure reductions from 1 Torr to < 200 mTorr and from 115 mTorr to <10 mTorr. This SIP can operate at a starting pressure as low as 1.5 µTorr.

1.5 Previous Efforts in Knudsen Pumps

Martin Knudsen demonstrated the first functional Knudsen pump in 1909 [Knu09] (Fig. 1.7). He used narrow glass tubes (diameter 0.374 mm) for thermal transpiration flow and wide glass tubes (diameter 10.2 mm) for viscous flow channels. Using a platinum wire heater, a thermal gradient was applied to the glass tubes, which were configured as multiple stages operating in series. However, the challenge of fabricating sufficiently narrow channels for operation at atmospheric pressure has limited widespread adoption.
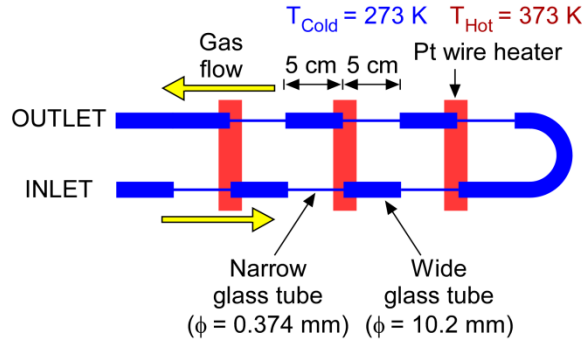
Figure 1.7: The first Knudsen pump studied by Knudsen himself in 1909 [Knu09].

With advancements in microfabrication techniques, this technology has received more attention [Var99, Var01, Mun02, McN05, Col05, Ale06, Kos08]. The past two decades have seen several efforts demonstrating the use of bulk micro- or nano-porous materials for thermal transpiration driven Knudsen pumping [Var01, You05, Han07, Gup08, Guo09, Gup10, Pha10, Gup11a, Gup11b]. However, achieving high thermal isolation and generating high pressure heads using this approach remains a challenge.

The Muntz group at the University of Southern California used porous aerogel material for thermal transpiration [Var99, Var01, Mun02]. Its pore size of 20 nm is smaller than the mean free path of air molecules at atmospheric pressure, ≈ 70 nm.

Therefore, the porous aerogel allows Knudsen pumping (i.e., thermal transpiration) at atmospheric ambient pressure by imposing a thermal gradient across the aerogel membrane, which has a thickness of 520 μm (Fig. 1.8). Using N_2 as the working gas, the resulting pressure difference was $\approx$1 kPa using $\approx$2 W at 760 Torr ambient [Var01].
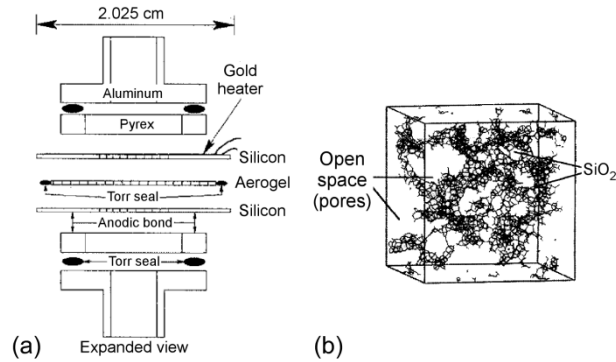
Figure 1.8: Single-stage Knudsen pump using aerogel membrane as thermal transpiration flow channel [Var01].
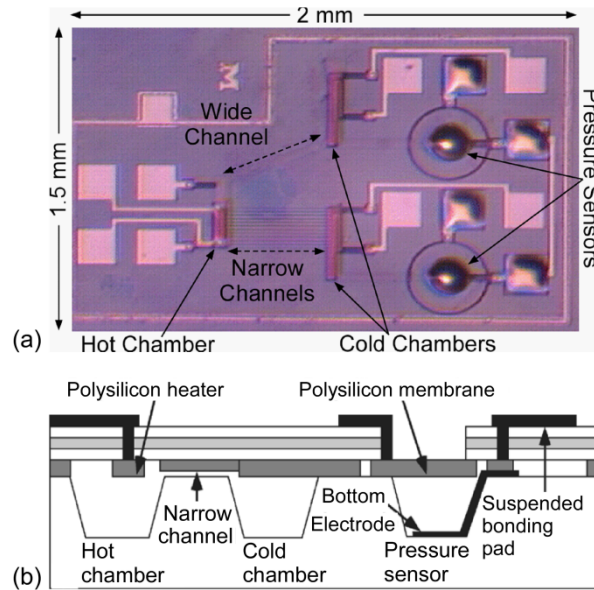
Figure 1.9: The first fully micromachined single-stage Knudsen pump [McN05]. (a) Microscope image of top view. (b) Cross-sectional drawing.

1.6 Previous Efforts in Knudsen Pumps at the University of Michigan

Research of the Gianchandani group at the University of Michigan has resulted in several microscale and mesoscale Knudsen pumps that are appropriate for a wide range

of applications. Using recent advances in microfabrication technologies, efforts were directed toward high flow and high vacuum Knudsen pumps for operating at atmospheric pressure. There has been a decade of research, including the work of this dissertation.

The first fully micromachined Knudsen pump with a compact footprint of 2×1.5 mm^2 was reported in 2003 (Fig. 1.9) [McN05]. This pump was targeted at chip-scale cavity pressure modulation related applications. The calculated temperature of the heater was $\approx$1,100°C, using an input power of 80 mW and a high thermal isolation of 1.4×10^4 K/W. The estimated maximum pump speed was 1 μcc/min. The Knudsen pump could pump down a cavity from 760 Torr down to 345 Torr using an input power as low as 80 mW. The resulting compression ratio was $\approx$1.9; this is the highest value in micropumps, as shown in Fig. 1.5.

A mesoscale Knudsen pump targeted at precision sample movements in microfluidic assays was reported in 2009 [Gup11a] (Fig. 10a, 10b). It was a nanoporous ceramic based 9-stage Knudsen pump. It had a footprint of 25×25 mm^2. While operating at 55 K above room temperature, the pump provided a maximum pressure difference exceeding 12 kPa at its sealed outlet, or a gas flow rate of $\approx$3.8 μL/min. against a pressure head of 160 Pa. This pump was operated continuously for more than 10,000 hours without any deterioration in its performance.

In 2010, a nanoporous polymer based Knudsen pump, foot print 14×14 mm^2, was demonstrated to achieve a gas flow rate of $\approx$1 sccm – the highest ever reported for a Knudsen pump [Gup11b] (Fig. 10c, 10d). Porous materials enable high flow rates due to a large number of channels, operating in parallel. The pore diameter of the nanoporous

polymer was ≈25 nm. This pump operated continuously for more than 6,000 hours without any significant deterioration in its performance.
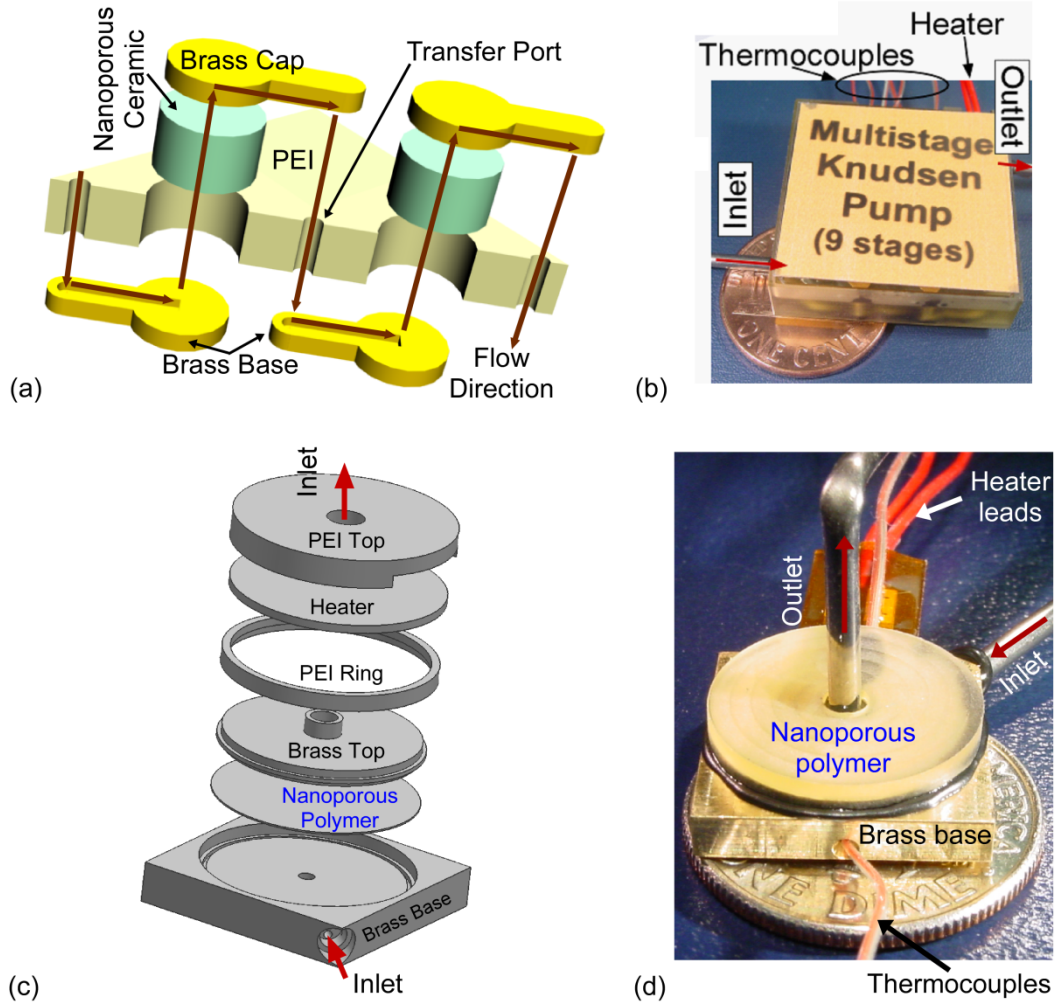
Figure 1.10: Mesoscale Knudsen pumps at the University of Michigan [Gup11a, Gup11b]. (a) Exploded view of two successive stages of nanoporous ceramic based Knudsen pump [Gup11a]. (b) Photo image of a nanoporous ceramic based multistage Knudsen pump with a final packaged volume of $25 \times 25 \times 7.25$ mm^3 [Gup11a]. (c) Exploded view of a nanoporous polymer based single-stage Knudsen pump [Gup11b]. (d) Photo image of a nanoporous polymer based single-stage Knudsen pump with a final packaged volume of $14 \times 14 \times 4.5$ mm^3 [Gup11b]. The arrows show the direction of gas flow through different elements.

By applying Knudsen pumps to μGC, a nanoporous polymer-based Knudsen pump was integrated with a μGC system and demonstrated its capability to separate and detect gas mixtures [Liu11] (Fig. 1.11). Using a six parallel Knudsen pump (KP) array,

flow rates of 1 mL/min and 0.26 mL/min were achieved for helium and dry air carrier gases, respectively. As shown in the test set-up for the μGC system, the Knudsen pump pulls the analyte gas from the injection module to the gas-separating μcolumn and the Fabry-Pérot (FP) optical-detector module. As a result, the μGC prototype system obtained a high resolution detection of 5 alkanes of different volatilities within 5 min.
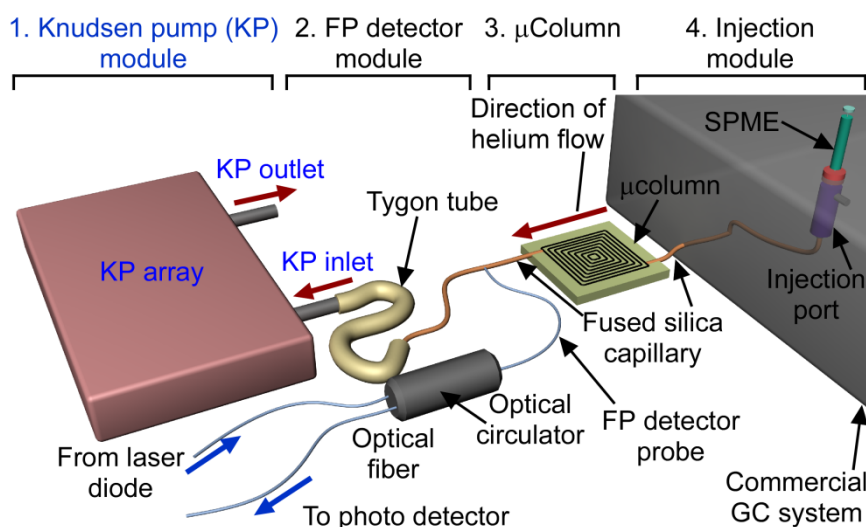
Figure 1.11: An example of high flow rate Knudsen pump applied to micro gas chromatograph [Liu11]. The nanoporous polymer based single-stage Knudsen pump is used.

Although porous materials provided high flow rate, detailed analyses showed that the porous materials have defects that compromise pressure difference and gas flow rate [Gup11a]. In addition, polymer materials withstand limited temperature differences.

1.7 Organization of Dissertation

The following six chapters comprise the balance of this dissertation.

Chapter 2 presents the Si-micromachined multistage Knudsen pump designed for high compression ratio. The 48 stages, operating in series, are monolithically integrated,

using a five-mask fabrication process. The test result shows substantial increases in compression ratio compared to other pumps reported in the literature.

Chapter 3 introduces the two-part architecture into the multistage Knudsen pump design. A two-part architecture is explored for customizing pump stages in order to enhance the compression ratio and the pumping speed over a wide pressure range. The number of stages is increased from 48 stages (as described in Chapter 2) to 162 stages. 54 stages are designed for a pressure range of 760-50 Torr and 108 stages are designed for a lower pressure range. The performance test resulted in a higher compression ratio, using more efficient power. This Knudsen pump is capable of a pressure reduction from atmospheric pressure to sub-Torr pressure.

Chapter 4 introduces a dynamic calibration method for Pirani gauges embedded in fluidic networks. This method is used for enhancing accuracy in pressure measurement in the 162-stage Knudsen pump. This method can be extended to measure vacuum levels in other microsystems, where Pirani gauges are embedded deep within a microfluidic pathway.

Chapter 5 describes low-temperature vacuum sealing of micro cavities using atomic layer deposition (ALD) of Al_2O_3. This technique is used for leak-proof sealing of flow channels in the 48-stage and 162-stage Knudsen pumps, and is very promising for retaining vacuum seals for micro devices.

Chapter 6 presents the Si-micromachined single-stage, Knudsen pumps designed for high flow rates. It explores a large number of parallel vertical flow channels integrated in a silicon-on-insulator (SOI) wafer.

Chapter 7 concludes this dissertation and discusses future work.

CHAPTER 2

A SI-MICROMACHINED 48-STAGE KNUDSEN PUMP
FOR ON-CHIP VACUUM

This chapter[1] describes a monolithic 48-stage Knudsen pump. To achieve a high compression ratio, 48 stages are designed and are cascaded in series in a single chip. A five-mask, single silicon wafer process is used for monolithic integration of the Knudsen pump. It has a footprint of 10.35×11.45 mm^2. The pump has several monolithically integrated Pirani gauges to experimentally measure the vacuum pumping characteristics of the pump. For an input power of ≈ 1.35 W, the fabricated pump evacuates the encapsulated cavities from 760 Torr to ≈ 50 Torr, resulting in a compression ratio of 15. It also pumps down from 250 Torr to ≈ 5 Torr, resulting in a compression ratio of 50.

2.1 Introduction

For Knudsen pumps to achieve high compression ratios, one major challenge is that it is necessary to pack a large number of stages (operating in series) into a small footprint. Micromachining technologies are suitable for integrating these stages in a small area using lithography-based batch fabrication processes. While mesoscale (multistage) Knudsen pumps have demonstrated compression ratios of ≈ 1.2 [Gup11a, You03], the first fully micromachined on-chip single-stage Knudsen pump, reported in

---

[1] This chapter describes joint work with Naveen Gupta

2005, achieved a compression ratio of ≈1.9 [McN05]. This single-stage micromachined Knudsen pump used a six-mask, two-wafer process.

This chapter reports on a 48-stage micromachined Knudsen pump with integrated Pirani gauges, using a five-mask, single-wafer structure. Using a serially cascaded multistage design, this pump achieves substantially higher compression ratios. A pressure drop in excess of 700 Torr is achieved in an air ambient at atmospheric pressure. The pump is operational over a range of ambient pressure starting from atmospheric pressure to less than 10 Torr. The principle of operation of Knudsen pump is described in Section 2.2. The design analysis for the device has been presented in Section 2.3. The details of device fabrication process and results are discussed in Section 2.4-2.5, followed by a discussion of experimental results in Section 2.6.

2.2 Principle of Operation

Knudsen pumps combine thermal creep flow and viscous flow [Loe34]. An imposed thermal gradient results in thermal creep flow along the channel walls from the cold end to the hot end. The pressure gradient induced by the thermal creep flow results in viscous flow in the reverse direction. The Knudsen number, defined as the ratio of the mean free path of gas molecules to the hydraulic diameter of a flow channel, is used to identify the relative contribution of the thermal creep flow and the viscous flow. In a narrow channel, if the Knudsen number is larger than ≈1, a pressure gradient is established in the same direction as the thermal gradient. This is the phenomenon of thermal transpiration. In contrast, a wide channel is dominated by viscous, pressure-driven (Poiseuille) flow.

In a multistage Knudsen pump (Fig. 2.1), a single stage includes a narrow channel and wide channel. The narrow channel generates a forward pressure difference, ΔP_f, under the imposed thermal gradient along the intended forward path. In contrast, the wide channel, by connecting neighboring narrow channels, is intended to restore the temperature from the hot temperature, T_{Hot}, to the cold temperature, T_{Cold}, without a large drop in pressure. At equilibrium, the net pressure difference in a stage, ΔP_S, can be stated as:

$$\Delta P_S = \Delta P_f - \Delta P_r \qquad (2.1)$$

where ΔP_r is the unintended, reverse transpiration pressure difference that could be developed in the wide channel at low pressures, at which the flow is not entirely viscous. A serial cascade increases the compression ratio; $\Delta P_{Total} = P_{OUT} - P_{IN} = \sum_i \Delta P_{S,i}$, where P_{OUT} and P_{IN} are the pressures at the outlet and inlet, respectively, and i denotes the stage number.

Analyses of thermal transpiration flow suggest that the mass flow rate, $\dot{M}$, along a narrow rectangular channel subjected to a longitudinal temperature gradient is [Sha98, Sha99, Jou08]:

$$\dot{M} = \left(\frac{Q_T \Delta T}{T_{Avg}} - \frac{Q_P \Delta P}{P_{Avg}} \right) \frac{a^2 b P_{Avg}}{l} \sqrt{\frac{m}{2 k_B T_{Avg}}} \qquad (2.2)$$

where a, b, and l are, respectively, the height, width, and length of the channel; m is the mass of a single gas molecule; k_B is the Boltzmann constant; ΔT and T_{Avg} are, respectively, the temperature difference and the average temperature of T_{Cold} and T_{Hot}; ΔP and P_{Avg} are, respectively, the pressure difference and the average pressure of the cold and hot chambers; and Q_T and Q_P are, respectively, the thermal creep flow and Poiseuille flow

coefficients. Here, Q_P represents the viscous flow. Note that the terms in ΔT and ΔP have opposite signs to indicate that the viscous flow balances the thermal creep flow.
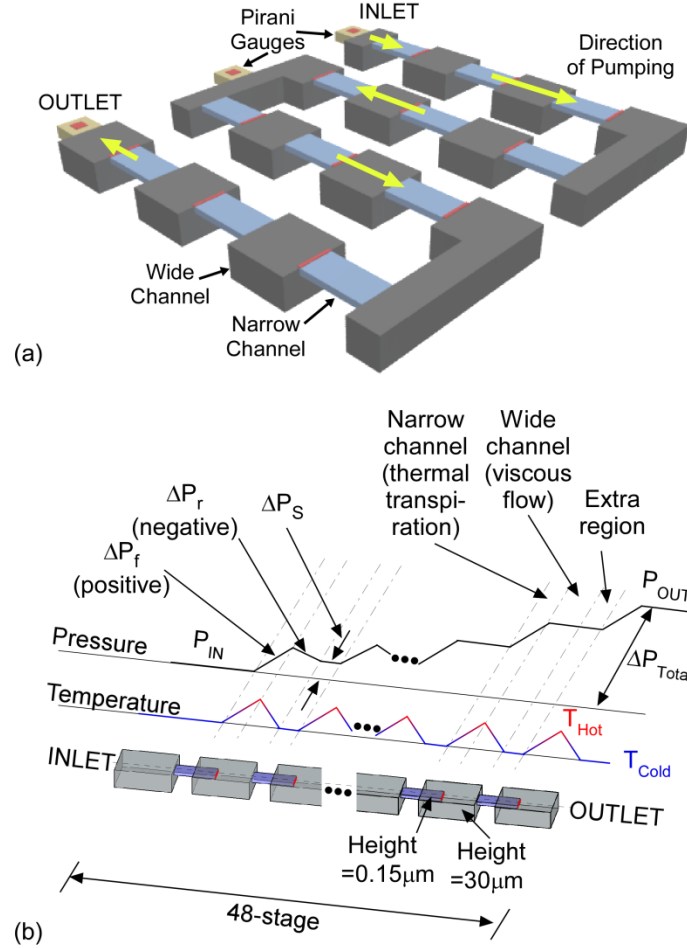
Figure 2.1: Schematic of a multistage Knudsen pump. (a) Perspective view of serpentine arrangement. The hot end of one stage is connected to the cold end of the next stage via a wide channel. Pirani gauges are used to measure the pressure. (b) Forty eight stages operating in series. The narrow channels support thermal transpiration driven Knudsen pumping; whereas the wide channels support viscous flow. Thermal transpiration driven flow causes pressure rise (or drop) in the narrow channels and viscous flow in the wide channels helps retain the pressure rise (or drop). Hence the pressure heads across each of the stages add up, without the need to increase the operating temperature of each of the stages.

The values of Q_T and Q_P are obtained by Sharipov's calculation, which appropriately represents direct simulation Monte Carlo (DSMC) [Mas07]. Sharipov [Sha99] numerically solved the linearized Boltzmann transport equation (BTE), for

extracting Q_T and Q_P from rarefied gas flow. Each flow coefficient was tabulated for the variations in ratio of a to b (= a/b), and the rarefaction parameter. The rarefaction parameter is defined as the inverse of Knudsen number, which is pressure-dependent. For parameters that are unlisted in [Sha99], the values of Q_T and Q_P can be determined from interpolations of the tabulated values.

The pressure difference across a channel, ΔP, depends upon the flow rate, as indicated in Eq. (2.2). When the pressure difference is zero (i.e., there is no pressure head), the flow rate takes the largest value. Over time, when evacuating a blind cavity, gas flow approaches equilibrium (i.e., no net flow), where the thermal creep flow due to ΔT is completely balanced by the viscous return flow due to ΔP. The resulting equilibrium pressure difference, ΔP_{Eq}, is given by:

$$\Delta P_{Eq} = \gamma \frac{\Delta T}{T_{Avg}} P_{Avg} \tag{2.3}$$

where γ is the flow coefficient ratio, defined as Q_T/Q_P, indicating the relative ratio of the thermal creep flow to Poiseuille flow coefficients in a channel [Sha98]. In the free-molecular regime where the Knudsen number is larger than 10, γ takes its largest value of 0.5 for a/b ratio of 1 (square) to 0.44 for a/b of 0 (plate); in the viscous flow regime, where the Knudsen number is smaller than 0.01, γ takes its smallest value of 0 [Sha99]. In the transitional regime, where the Knudsen number is between 10 and 0.1, γ is determined from [Sha99]. (See Fig. 3.2a for determined γ values.)

In the viscous flow regime, where [Sha99] does not fully cover, the Poiseuille equation [Bru08] can be used for describing the flow rate as:

$$\dot{V}_{POS} = \frac{\left\{1 - 0.63\left(\dfrac{a}{b}\right)\right\} a^3 b}{12\eta l} \Delta P \qquad (2.4)$$

where $\dot{V}_{POS}$ is the volume flow rate of gas in the viscous flow regime and η is the dynamic viscosity of gas. The Poiseuille equation is obtained by analytically solving the Navier-Stokes equation (NSE) for incompressible fluids.

2.3 Design Analysis

A 48-stage Knudsen pump with monolithically integrated Pirani gauges is designed for the purpose of this study. Sacrificial polysilicon is used to define narrow channels. Bulk silicon cavities are used to define wide channels. Thin film silicon dioxide/silicon nitride layers are used to encapsulate the wide and narrow channels.
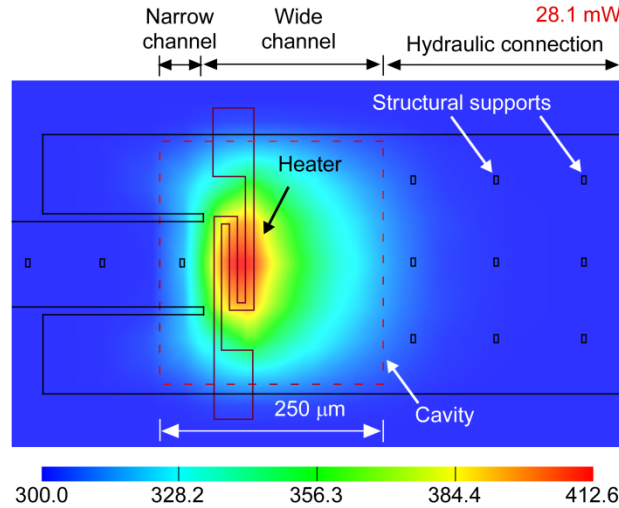
Figure 2.2: Modeled distribution of temperature on the top surface of the Knudsen pump for a given input power of 28.1 mW per stage. The Knudsen pump is designed to limit the maximum operating temperature to <450 K.

Finite element analysis (FEA) is used to investigate the thermal behavior of the device. Figure 2.2 shows the modeled thermal response of each of the stages of the Knudsen pump for ambient temperature of 300 K. For FEA, natural air convection

cooling is ignored and thermal conduction across material layers is considered. A 110 K rise in heater temperature is expected for an input power of 28.1 mW. The thermal map suggests that the temperature gradient is steepest along the narrow channel, indicating that the primary heat loss is along the narrow channel. This temperature gradient is the driving force for gas pumping from the cold end to the hot end. Since the hot end of the narrow channel has about half of the temperature rise of the heater (Fig. 2.2), 355 K is used for T_{Hot} in the following performance estimation. The wide channel from the heater to the cold end of the next narrow channel restores the temperature from T_{Hot} to T_{Cold} with negligible change of pressure. The gas flow in this channel is in the viscous regime.

The Sharipov calculation results of Q_T and Q_P [Sha99] are used to estimate the pressure ratio generated across each of the stages based on the simulated thermal distribution of the device, using Eq. (2.1-2.3). For a Knudsen pump with 30 μm height of wide channel and 0.15 μm height of narrow channel with the hot end at 355 K and the cold end at 300 K, 48 stages are expected to pump down from 760 Torr to 97 Torr. Using the same number of stages, the pressure can also be reduced from an ambient pressure of 250 Torr to 18 Torr. The mean free path of the air molecules at 18 Torr is about 3 μm at 350 K. Hence, for the reduction of pressure to about 18 Torr, the wide channels are required to have hydraulic diameter significantly larger than 3 μm to limit the backflow at low operating pressures.

The operating pressure levels are used to estimate the separation between adjacent structural supports for the diaphragms over the wide and the narrow channels. These structural supports (Fig. 2.2) are required to limit the deflection of diaphragms during the Knudsen pump operation, which would otherwise obstruct the gas flow. A finite element

model is used to estimate the separation of the structural supports for the diaphragms over the polysilicon channels. The device can accommodate a 760 Torr ambient pressure level. Under these operating conditions, the worst-case deflection of the diaphragm is expected to occur at the 48th stage (i.e. the pump inlet). The pressure above the diaphragm is 760 Torr and the pressure under the diaphragm, in the wide and the narrow channels is set to ≈0 Torr for simulation of the maximum deflection. For these operating conditions, the structural supports in the narrow channel are located at 50 μm spacing, limiting the deflection of diaphragm to <7 nm. Similarly, the structural supports in the wide channel are located at every 125 μm, limiting the deflection of diaphragm to <0.1 μm.

For the purpose of this study, the Knudsen pumping stage located furthest downstream, i.e. at the pump outlet, is termed the 1st stage of the 48-stage Knudsen pump. Similarly, the most upstream stage, i.e. at the pump inlet, is termed as the 48th stage. The Pirani gauges are identified by the stage number of the Knudsen pumping stage to which they are connected. For example, if a Pirani gauge is connected to the 48th stage, it is named P48. Similarly, if a Pirani gauge is connected to the 1st stage, it is named P1.

The footprint for each of the stages of the Knudsen pump is 0.3×1 mm^2. The wide channel is ≈300 μm wide and ≈750 μm long; the narrow channel is ≈100 μm wide and ≈250 μm long, whereas the cavity is 280 μm wide and 250 μm long. The footprint of the Pirani gauge is 250×250 μm^2.

The 48 stages are laid-out in a serpentine fashion with eight Knudsen pumping stages in a row. (Note that not all the intermediate stages have integrated Pirani gauges.)

For the ease of testing, the heaters for all the Knudsen pumping stages are connected together so that they could be powered simultaneously using only two contact pads. The serpentine arrangement allows the heaters to be connected in eight parallel branches with six heaters, connected in series, in each branch. The positive and the negative terminals of these eight branches are connected to two contact pads located at the top and the bottom of the chip.
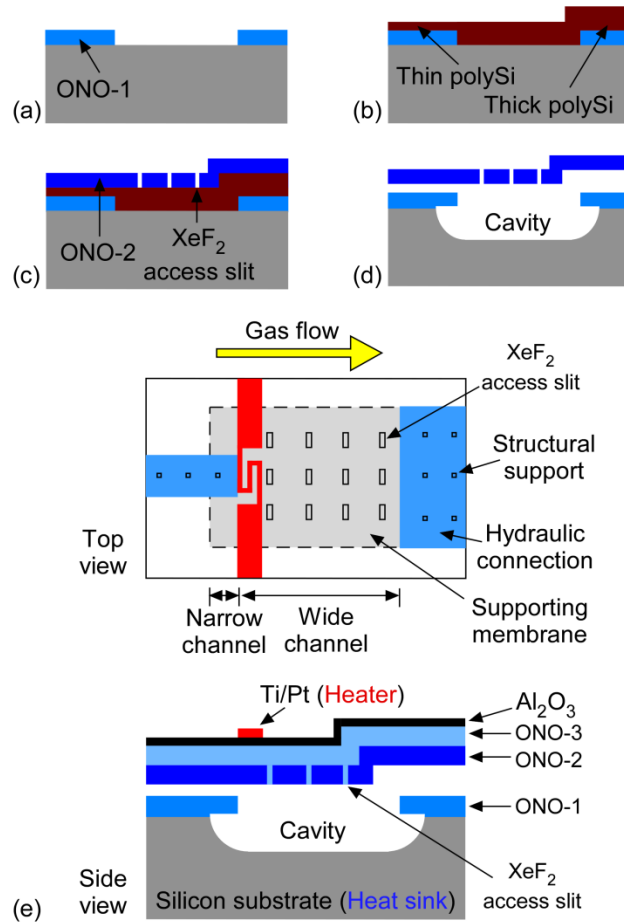
Figure 2.3: (a)-(d) The series of step involved in the fabrication of the 48-stage Knudsen pump with monolithically integrated Pirani gauges. (e) Top and side views of the final device; narrow channels are required for thermal transpiration; wide channels are required to connect successive stages; and cavities are used to increase the effective hydraulic diameter of the wide channel, and for better thermal isolation of the heater and the leading edge of the narrow channel.

2.4 Fabrication Process

The fabrication process selected for the 48-stage Knudsen pump and the monolithically integrated Pirani gauges offers simplicity and high yield. Three important features are: 1) thin (sacrificial) polysilicon is used to define narrow channels; 2) thick (sacrificial) polysilicon is used to define hydraulic interconnects between successive stages; and 3) a cavity is etched in bulk silicon to increase the hydraulic diameter of the viscous flow paths (i.e., wide channels). The walls of the channels are fabricated by stress-compensated low-pressure chemical vapor deposited (LPCVD) oxide-nitride-oxide (ONO) layers.

The microfabrication process begins with deposition and patterning of the first ONO (ONO-1) layer, which is 1.25 μm thick. The pattern defines openings, which are used for etching cavities in the bulk silicon (Fig. 2.3a). The ONO-1 film stress is ≈120 MPa. Sacrificial polysilicon of thickness ≈1.95 μm is used for the hydraulic interconnects, and for the height of the air gap for the Pirani gauges [Dom11] (Fig. 2.3b). A second sacrificial polysilicon layer, thickness ≈0.15 μm, is used to define the narrow channels (Fig. 2.3b). A second ONO (ONO-2) layer, thickness 1.17 μm, encapsulates the sacrificial polysilicon layers (Fig. 2.3c). The ONO-2 film stress is ≈80 MPa. Structural supports (Fig. 2.2) for ONO-2 are defined at regular intervals by patterning rectangular pockets in the polysilicon layers. An array of slits, 1×10 μm^2, are dry etched by reactive ion etching (RIE) of the ONO-2 layer to define access slits for etching sacrificial polysilicon. XeF$_2$ gas is used to etch the thin/thick polysilicon and 30 μm deep cavities in the bulk silicon (Fig. 2.3d). The cavity is ≈30 μm deep and the XeF$_2$ gas etches under the narrow channel by the same extent. Two layers of plasma enhanced chemical vapor

deposited (PECVD) ONO (ONO-3) with an intermediate nitride layer, total thickness of

3.12 μm, are used to block the XeF_2 access slits in ONO-2 (Fig. 2.3e). The ONO-3 film

stress is ≈25 MPa. An atomic layer deposited (ALD) Al_2O_3 layer, thickness 0.58 μm, is

used to hermetically seal the pinholes in ONO-3. The Al_2O_3 film stress is ≈300 MPa.

Finally, lift-off is used to define the Ti/Pt layer, total thickness ≈0.13 μm, for the heaters

and the Pirani gauges. The same Ti/Pt layer is used to define the electrical interconnects,

and contact pads. Figure 2.3e shows the sectional view of the final device, together with

the top view. The diaphragm at the most downstream end is punctured to define the

outlet for this 48-stage Knudsen pump, whereas the inlet is left sealed. The measured

thickness and residual stress of each layer are summarized in Table 2.1.

Table 2.1: Measured thickness and residual stress of each layer.

Layer	Thickness (μm)	Residual stress (MPa)
LPCVD ONO-1	1.25	+120
LPCVD ONO-2	1.17	+80
PECVD ONO-3	3.12	+25
Thin / thick polySi	0.15 / 1.95	-
ALD Al_2O_3	0.58	+300
Ti/Pt	0.030/0.1	-

The use of ONO diaphragm instead of oxide-only diaphragm results in higher

thermal conduction loss because the nitride has an order of magnitude higher thermal

conductivity than oxide. However, stress compensation is critical to avoid unexpected

curling/warping of the suspended structures. The ONO dielectric stacks are stress

compensated to have mild tensile stress, to avoid buckling of the suspended diaphragm

due to compressive stress. The increase in parasitic heat loss due to the additional

dielectric layers is compensated by increasing the thermal isolation of the heater from the

silicon substrate.
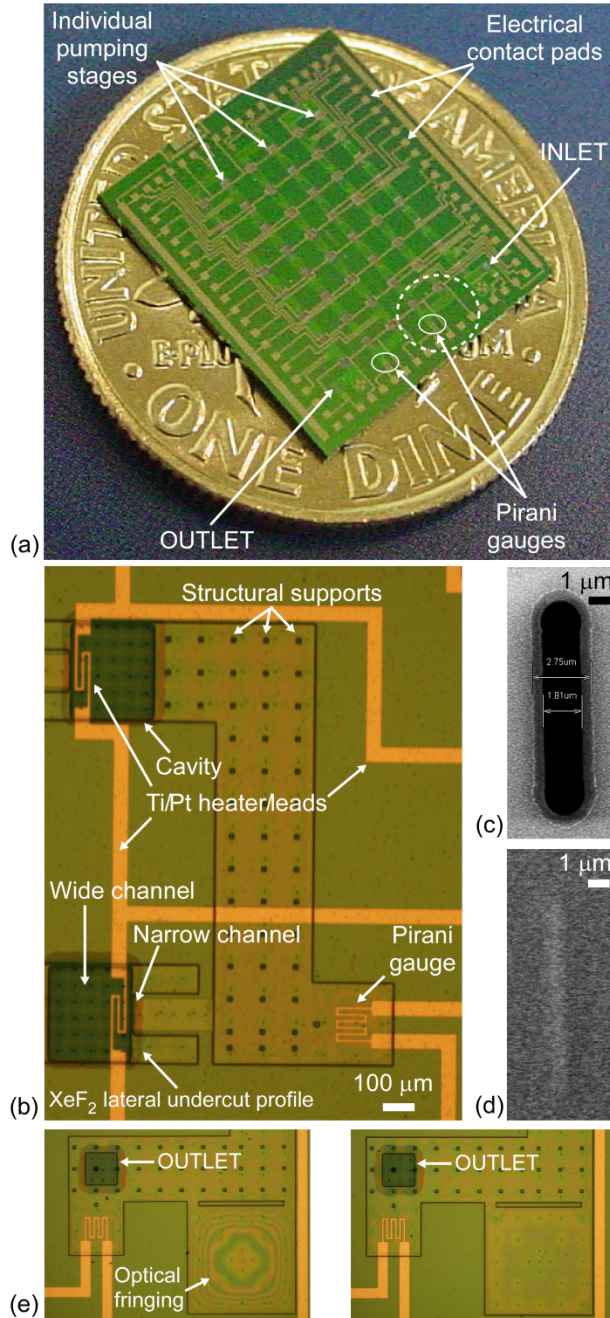
Figure 2.4: (a) Picture of the final fabricated device showing the individual Knudsen pumping stages, the inlet of the 48-stage pump, the outlet of the 48-stage pump, the Pirani gauges and the electrical contact pads. (b) Microscope image of two successive stages, as indicated by the dotted circle in (a). (c) SEM image of the XeF_2 access slit in the ONO-2 – used for the sacrificial etch of the polysilicon layers. (d) SEM image of the XeF_2 access slit in ONO-2 after sealing with PECVD ONO-3. (e) The left microscope image shows the fringes generated after sealing under vacuum, whereas the right image shows cleared fringes after the seal is opened.

Cavities are used to increase the hydraulic diameter of the (wide) channels, interconnecting successive Knudsen pumping stages depending on the desired operating pressure range of the Knudsen pump. The mean free path of the gas molecules is inversely proportional to the pressure. If the hydraulic diameter of the wide channels is not large enough, the wide channels will start supporting (reverse) thermal transpiration flow after attaining a certain vacuum level. The cavities also serve to thermally isolate the heating element for each of the Knudsen pumping stages. The heating elements for each of the stages are located on the diaphragm suspended over the cavity to minimize the parasitic heat loss into the substrate through the wide channel. The cavities are designed to undercut the narrow channel partially, such that the heated end of the narrow channel is suspended. This partially suspended architecture of the narrow channel is necessary to sustain high temperature gradient along the narrow channel. High temperature gradient along the narrow channel translates into high pressure gradient along the narrow channel.

Figure 2.4a shows the picture of the fabricated 48-stage Knudsen pump. It has a footprint of 10.35×11.45 mm^2. Figure 2.4b shows the close-up view of two successive stages, which are located on the side of the Knudsen pump. The etched cavity under narrow channel is indicated by the XeF$_2$ lateral undercut profile. Figure 2.4c shows the SEM image of one of the 1×10 μm^2 slits in the ONO-2 diaphragm for XeF$_2$ gas etching of the sacrificial polysilicon layers. Figure 2.4d shows the SEM image of a slit, similar to that shown in Fig. 2.4c, after PECVD ONO-3 sealing step. Figure 2.4e shows one of the diaphragms after ALD Al$_2$O$_3$ sealing. The optical fringing in the diaphragm in the left microscope image in Fig. 2.4e symbolizes the hermetic sealing of the diaphragm under

vacuum in the ALD tool. The fringing is because of the out-of-plane deflection of the diaphragm due to the vacuum created during ALD deposition of Al_2O_3. The structure retains the fringe pattern for several days, demonstrating that the diaphragm is perfectly sealed and allows no leakage of air. The fringe pattern disappears only after the seal is opened intentionally, as shown in the right microscope image in Fig. 2.4e. Although the SEM image in Fig. 2.4d shows (visually) that the XeF_2 access slits are sealed, acetone liquid seeps into the diaphragm during the lift-off process without the deposition of an ALD Al_2O_3 layer. This suggests the presence of pinholes in ONO-3 even after the deposition of a >3.8 µm thick, seven layered, ONO-3 dielectric stack. ALD Al_2O_3 is effective in sealing the pinholes in ONO-3 layer (See Fig. 5.6 for cross-sectional images of a sealed access slit).

2.5 Experimental Procedures and Results

The fabricated chip is evaluated in a test chamber to permit the control of ambient pressure. The chip is wire-bonded for electrical connections from measurement tools to Ti/Pt metal pads. The Knudsen pump is mounted in an electronic package (Part#: SQ1000-190-12GC by Sinclair Manufacturing Company, MA) for the purpose of testing. Silver grease (AIT Electro-grease ELGR8501) is used to enhance the thermal coupling between the chip and the electronic package for thermally grounding the chip to the room temperature in laboratory environment. The thermal conductivity of the grease is >8.6 W/m/K. The operating medium is laboratory air. The inlet of the Knudsen pump is sealed, while the outlet is open and vented to the test chamber. The pump is operated by

providing a constant voltage to the Ti/Pt heater. The input power levels, as indicated in test results (Fig. 2.6 and 2.7), are for the heated resistances at equilibrium pressures.

Ambient pressure is controlled by a rotary pump (Edwards[TM] E2M30) and monitored, using a commercial calibrated pressure sensor (Adixen[TM] AHC-2010) located within the test chamber, volume ≈185 L. The test chamber has several electrical feed-throughs, which are used to make electrical connections to the electronic package and hence to the device. These electrical feed-throughs are used to power the Knudsen pump and the Pirani gauges, and to record the signal from the Pirani gauges. Three Agilent[TM] 34401A multimeters are used to read the output signal from the three Pirani gauges. These multi-meters are further connected to a computer operating a LabVIEW[TM] program, through a National Instruments GPIB-USB-HS data transfer cable.

2.5.1 Pirani Gauge Characterization
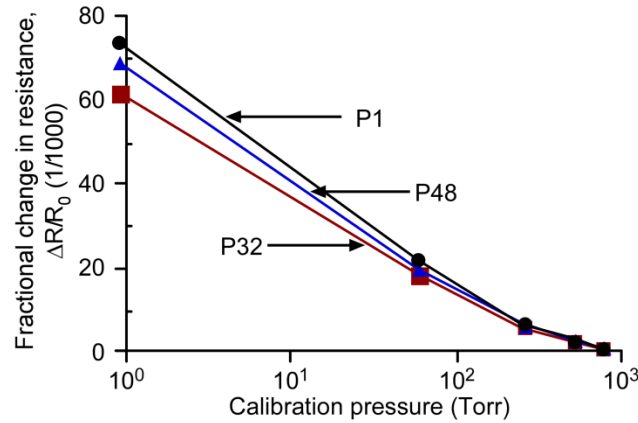
Figure 2.5: Response of the monolithically integrated Pirani gauges with the variation in pressure. Pirani P1 is connected to the 1st stage, i.e. the outlet of the multistage Knudsen pump. Pirani P32 is connected to the 32nd stage. And Pirani P48 is connected to the 48th stage, i.e. the inlet of the multistage Knudsen pump.

The Knudsen pumping stages are maintained in off state during the initial calibration of the Pirani gauges. The Pirani gauges of interest are connected in series,

and a constant current of 5 mA is provided.   The voltage drop across each of them is

recorded as a function of pressure in the test chamber.   The resistance of each Pirani

gauge is ≈155 Ω.   The power consumption for each of these Pirani gauges is ≈3.9 mW,

and varies in less than 10 percent with changes in pressure levels.   Figure 2.5 shows the

fractional change in the resistances of the Pirani gauges with pressure (See Section 4.2.1

for Pirani gauge theory).   The Pirani gauges are sufficiently sensitive to allow the

detection of the pressure changes from 760 Torr down to sub-Torr levels.   Using these

calibration points, the measured pressures are obtained using interpolations between the

neighboring points.   Hence, the maximum error range in each measurement from

interpolation is between the neighboring points in Pirani calibration.

2.5.2 Knudsen Pump Operation
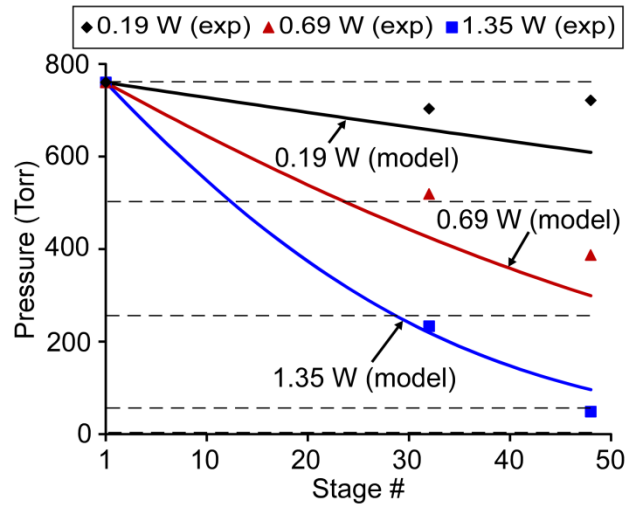
Figure 2.6:   Variation in the experimentally measured and theoretically modeled pressures with stage number for different input power. The pressures shown here are the pressures at the inlet of the respective stages. The pressure at the outlet of the Knudsen pump is the pressure at the 1th stage in this plot. The horizontal dotted lines are the Pirani calibration points which are used for interpolations for the pressure measurements.

32

The inlet of the Knudsen micropump is sealed and the outlet is opened to the 760 Torr ambient pressure. Figure 2.6 shows the variation in the pressure in different stages of the Knudsen pump for different input powers. For an input power of 1.35 W, the pressure is reduced from 760 Torr to ≈50 in experiment and to ≈98 Torr in theory at its sealed inlet with air as the operating medium. In Fig. 2.6, the isolated points represent the experimentally measured data, whereas the continuous line represents the theoretical estimate of the pressure along each of the stages.
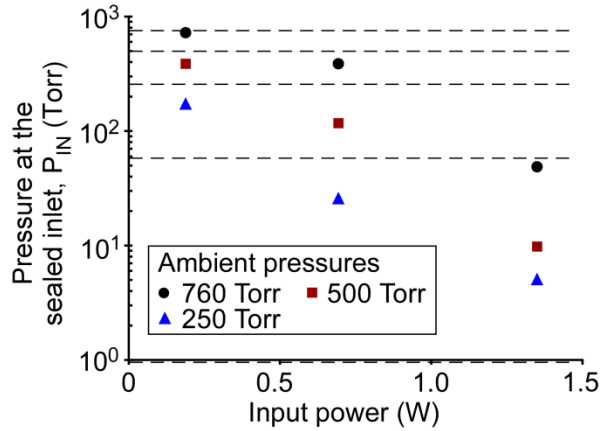
Figure 2.7: Variation in the pressure at the sealed inlet of the Knudsen pump with the input power for different ambient operating pressures. For an input power of 1.35 W the pump pumps down from 760 Torr to ≈50 Torr. For the same input power, while operating at an ambient pressure of 250 Torr, the pump evacuates the cavity at the upstream of the Knudsen pump down to ≈5 Torr. The horizontal dotted lines are the Pirani calibration points which are used for interpolations for the pressure measurements.

The fabricated devices are also tested at the ambient pressures of 500 Torr and 250 Torr to evaluate the performance that may be possible with the use of separate roughing pumps. Figure 2.7 shows the variation in the (steady state) vacuum generated by the device, at its sealed inlet, with input power, for different ambient operating pressures. For an input power of 1.35 W, the pump typically evacuates from 760 Torr to ≈50 Torr, resulting in a compression ratio of 15, and from 250 Torr to ≈5 Torr, resulting in a compression ratio of 50. Hence, two such chips can be potentially integrated in

33

series to design a Knudsen pump that can pump down from 760 Torr to <5 Torr pressure levels.

2.5.3 Discussion

The plots comparing the models and measurements (Fig. 2.6) suggest that, for the stages that are more upstream, at 0.19 W and 0.69 W, the experimentally measured pressure is higher than the idealized theoretical estimates. At 1.35 W, the measured pressures show smaller deviations from the idealized theoretical estimates. One of the possible reasons of larger deviations at 0.19 W and 0.69 W could be smaller fractional changes in the resistance of the Pirani gauge at the high pressure range from 760 Torr to 250 Torr (Fig. 2.5). This could result in larger errors in measurements at the high pressure range than those in measurements at the low pressure range, where the fractional changes in resistance are considerably larger. Another source of mismatch could be chip-to-chip and stage-to-stage variations of the suspended length of the narrow channel formed by XeF$_2$ gas etching of the bulk silicon. These variations could result in errors in the temperature estimations of the hot end of the narrow channel at an applied power, because the hot end of the narrow channel is located at a halfway point from the heater to the thermal ground. The variation in the suspended length of the narrow channel in a wafer is approximately ±20% of the average length of 30 μm, which results in ±10% error in the temperature estimation of the hot end of the narrow channel.

The primary thermal conductance path is from the Ti/Pt heater and the suspended narrow channel, where the silicon substrate is the heat-sinking thermal ground (Fig. 2.2). The relatively high power budget of the device is mainly due to the high thicknesses of

the PECVD ONO-3 and ALD Al_2O_3 layers in the thermal conductance path for hermetic

sealing of the XeF_2 access slits. In this fabrication attempt, the thicknesses are not tuned

to find the minimally required thicknesses for sealing. In future generations of Knudsen

pump, higher thermal isolation could be obtained with lower thicknesses of the layers for

lower thermal budget. The suspended length of the narrow channel can also be increased

for additional thermal isolation.
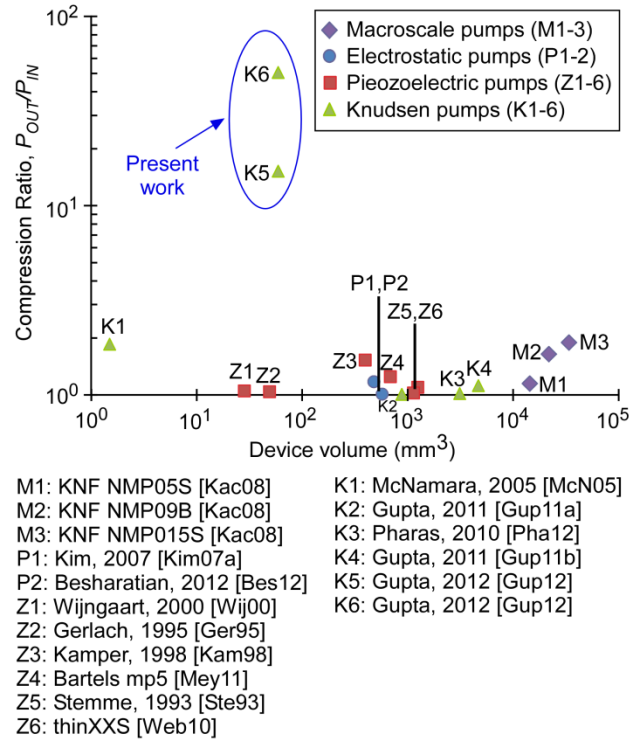
Figure 2.8: This work has resulted in a compression ratio of 15 (if the pump is operated in 760 Torr air ambient) and a compression ratio of 50 (if it is operated in 250 Torr air ambient). A comparison to the published literature suggests that Knudsen micropumps can achieve superior compression ratios.

Benchmarking illustrates a 24 times improvement in the gas compression ratio

compared to the other micropumps [Ste93, Ger95, Kam98, Wij00, Ngu02, Las04,

McN05, Kim07a, Kac08, Web10, Pha10, Gup11a, Gup11b, Mey11, Bes12] previously

reported (Fig. 2.8). The integration of 48 stages into one silicon chip, using the simple

silicon-based micromachining and effective channel design, enabled substantially higher performance.

The type of pump described in this paper is intended for applications such as cavity pressure modulation and tunable vacuum control. These applications typically require the maintenance of a certain level of vacuum, or a limited amount of pressure modulation; repeated pump-down from atmospheric pressure is not anticipated. Hence, the pumping rate is not a dominant consideration. For the design implemented in this work, the total dead volume of gas that must be evacuated is 1.7×10^{-4} cm^3, distributed over 48 stages. The theoretical mass flow rates, calculated using Eq. (2.2), at the high-pressure end (760 Torr) and the low-pressure end (50 Torr) are 8.4×10^{-13} kg/sec and 1.4×10^{-13} kg/sec, respectively. The resulting theoretical pump down time is ≈20 min. The increase in the mass flow rate at the low-pressure end compared to that at the high-pressure end improves the overall pumping rate. Broadening the narrow channel for the low-pressure stages can increase the mass flow rate without compromising the compression ratio per stage, because the mean free path of gas molecules increases as the pressure decreases.

2.6 Conclusion

A five-mask process is described to fabricate a single-wafer Knudsen pump structure with 24 times better gas compression capability than other micropumps reported in the literature. This single-wafer structure is instrumental in mitigating yield problems. This effort demonstrates the feasibility of Knudsen pump with monolithically integrated Pirani gauges for high-vacuum pumping applications. The calibration of Pirani gauges is

used to measure pressures from 760 Torr to 1 Torr pressure level. The 48-stage Knudsen pump with monolithically integrated Pirani gauges has a footprint of 10.35×11.45 mm^2. For an input power of 1.35 W, it pumps down 760 Torr to ≈50 Torr and from 250 Torr to ≈5 Torr with ambient air as the operating medium. The study suggests that the thickness of the ONO-3 and the ALD Al_2O_3 layers could be reduced to limit the power budget of the device without affecting its compression performance in the compression ratio.

CHAPTER 3

A SI-MICROMACHINED 162-STAGE TWO-PART
KNUDSEN PUMP FOR ON-CHIP VACUUM

This chapter investigates a two-part architecture for a Knudsen vacuum pump. For further increasing the compression ratio, the previous 48 stages in Chapter 2 are increased to 162 stages. The two-part architecture utilizes 54 stages designed for the pressure range from 760 Torr to ≈50 Torr, and 108 stages designed for lower pressures. This approach provides greater compression ratio and speed than using a uniform design for each stage as in the 48-stage Knudsen pump. Finite element simulations and design analyses are presented. A five-mask single-wafer fabrication process, studied in Chapter 2, is used for monolithic integration of the two-part Knudsen pump. The Knudsen pump has a footprint of 12×15 mm^2. The pressure levels of pumping stages are measured by integrated Pirani gauges.

3.1 Introduction

A monolithic Si implementation that used one stage was reported in 2005 and demonstrated a compression ratio of ≈1.9 [McN05]. The compression ratio can be increased to a limited extent by increasing the operating temperature. However, a more effective and scalable way to increase compression ratio is to cascade stages. Miniaturized multistage Knudsen pumps have been reported in the past [Gup11a, Gup11b, You05], but the early efforts were limited in scope because they did not utilize

lithographic fabrication. As described in Chapter 2 in this dissertation, a monolithically fabricated 48-stage Knudsen vacuum pump was studied, resulting in a high compression ratio of 15 at atmospheric ambient pressure.

Together with the compression ratio, the pumping rate in a multistage Knudsen pump can be increased by adjusting the channel hydraulic diameter for the steady state operating pressure of each stage. An increased hydraulic diameter provides lower hydraulic resistance (i.e., higher pumping rate) as long as gas flow is confined to the free-molecular or transitional flow regimes for Knudsen pumping. The Knudsen pump described in Chapter 2 used a uniform design for 48 stages, demonstrating the evacuation of on-chip cavities from 760 Torr to <50 Torr. The hydraulic diameter of the channels in this pump was designed to ensure the free-molecular or transitional flow at atmospheric pressure. For pumping to lower pressures, the hydraulic diameter of the upstream stages must be increased in order to increase the pumping rate. The compression ratio is not sacrificed because in steady state these upstream stages maintain relatively low cavity pressures at which the mean free path is much longer than at atmospheric pressure. However, the larger hydraulic diameter cannot sustain pumping in downstream stages. For this reason, a multistage Knudsen pump intended for a wide pressure range should be partitioned, so that both pumping rate and compression ratio can be enhanced for different parts of the pressure range.

In this context, this chapter explores two-part customization of stages for a micromachined Knudsen pump. Section 3.2 describes the two-part multistage architecture of a Knudsen pump, the heater design, and the Pirani pressure gauge.

Section 3.3 details microfabrication processes and results. Test methods and results are

presented in Section 3.4, followed by conclusions in Section 3.5.

3.2 Design

3.2.1 Two-Part Multistage Architecture
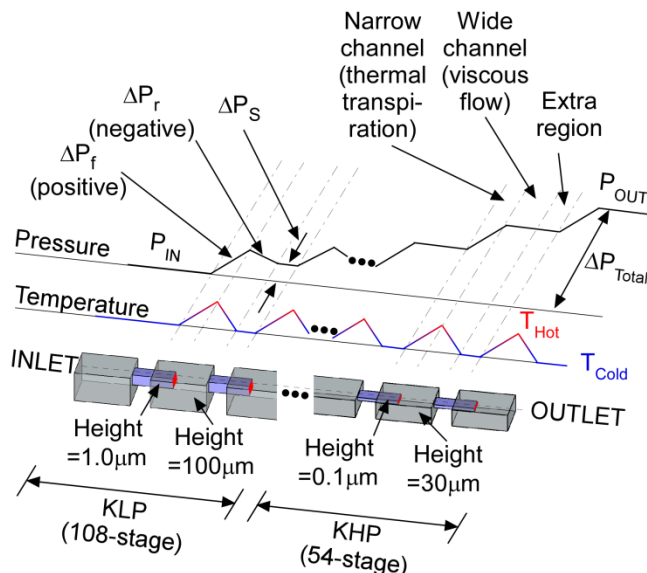
Figure 3.1: Schematic of a two-part multistage Knudsen pump. KHP and KLP stages are serially cascaded. Temperature and pressure profiles are shown along the dashed line in stages. The dot-dash lines indicate correspondence along the flow path.

The proposed two-part architecture (Fig. 3.1) is intended to enhance both

pumping rate and compression ratio within an operating pressure range for each part. 1)

The Knudsen high-pressure part (KHP), located downstream, is customized for the range

from atmospheric pressure to 50 Torr; the KHP is similarly adapted from the 48-stage

Knudsen pump described in Chapter 2. 2) The Knudsen low-pressure part (KLP), located

upstream, is intended for the range from 50 Torr to sub-Torr pressure. The KHP and the

KLP are comprised of serially cascaded single-stage Knudsen pumps, using 54 and 108

stages, respectively. The number of stages in the KLP is double that in the KHP,

reflecting the larger compression ratio in the KLP than that in the KHP. In principle, the narrow and wide channels for each stage can be designed to provide the highest compression for the intended steady state operating conditions of that stage. However, this amount of dimensional diversity can significantly increase the manufacturing complexity. In this effort, two types of stage designs are pursued, as a compromise.

The following flow equation, which is obtained by dividing Eq. (2.2) by mass density, describes the volume flow rate of gas, $\dot{V}_X$, from the cold end to the hot end in a long rectangular channel:

$$\dot{V}_X = \left(\frac{Q_T \Delta T}{T_{Avg}} - \frac{Q_P \Delta P}{P_{Avg}} \right) \frac{a^2 b P_{Avg}}{l} \sqrt{\frac{m}{2k_B T_{Avg}}} \frac{1}{\rho_{Avg}} \tag{3.1}$$

where ρ_{Avg} is the mass density at P_{Avg} and T_{Avg}. Instead of the volume flow rate, the standard flow rate, $\dot{V}_{Std}$, can be used to comprehend molecular flow through a multistage Knudsen pump where each stage is at a different pressure. To obtain $\dot{V}_{Std}$, the mass density, ρ_{Avg}, on the right side of Eq. (3.1), is substituted with the mass density at standard conditions, ρ_{Std}, provided by:

$$\rho_{Std} = \frac{m P_{Std}}{k_B T_{Std}} \tag{3.2}$$

where P_{Std} is the standard pressure of 760 Torr and T_{Std} is the standard temperature of 273.15 K. The standard flow rate allows easy comparison of molecular counts, whereas the volume flow rate is indicative of the swept volume. In Eq. (3.1), the subscript X in $\dot{V}_X$ denotes non-standard conditions defined by P_{Avg} and T_{Avg}. In a given narrow channel, as pressure decreases, the standard flow rate naturally decreases as well. However, the volume flow rate increases because Q_T increases relative to Q_P. This is reflected in Eq. (3.1) and (3.2).
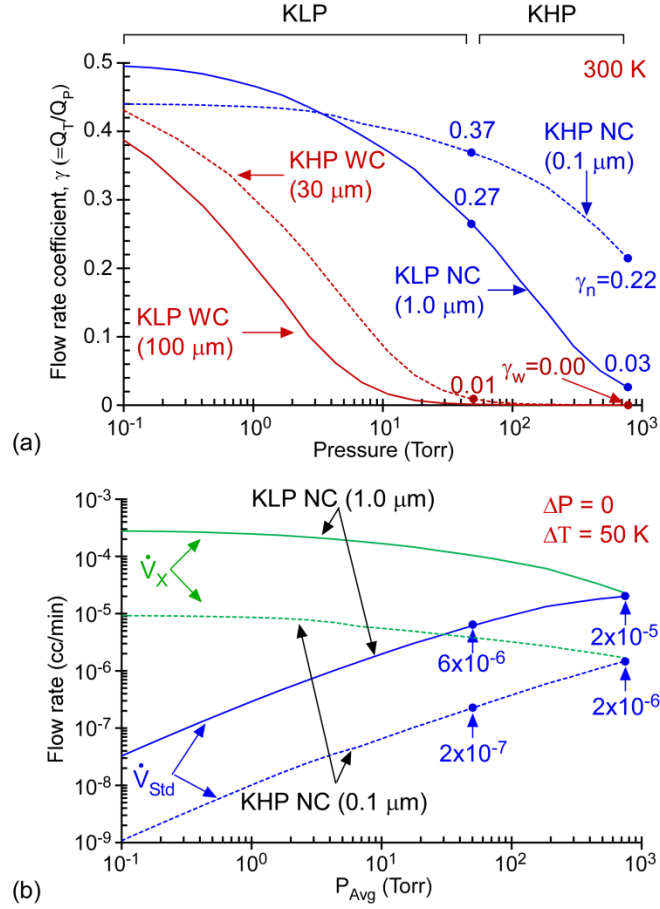
Figure 3.2: Calculated flow parameters using the designed channel heights in Table 3.1. (a) Flow coefficient ratios, γ_n and γ_w, as a function of pressure. (b) $\dot{V}_X$ in cc/min and $\dot{V}_{Std}$ in sccm as a function of P_{Avg}, using Eq. (3.1) and (3.2). Narrow and wide channels, abbreviated as NC and WC, are indicated by the dotted and solid lines, respectively.

Table 3.1: The designed channel heights, number of stages and the mean free path, λ, of N_2 gas for KHP and KLP in a pressure range of 760-0.1 Torr.

Part	Narrow channel height (μm)	Wide channel height (μm)	Number of stages	Pressure (Torr)	λ (μm) at 300 K
KHP (760Torr -50Torr)	0.1	30	54	760	0.07
				200	0.28
				50	1.12
KLP (≤50Torr)	1.0	100	108	1	49
				0.1	511

42

3.2.2 Selection of Channel Dimensions

Channel dimensions selected for the narrow and wide channels are summarized in Table 3.1. At 760 Torr, a narrow channel height of 0.1 μm in the KHP is similar to the N_2 mean free path of 0.07 μm (Table 3.1); the resulting γ_n value of 0.22 provides a reasonable forward pressure difference in the narrow channel (Fig. 3.2a). At 50 Torr, the γ_n value in the KHP increases to 0.37. However, at 50 Torr, in the KLP, where the narrow channel height of 1 μm is similar to the mean free path of 1.12 μm, the resulting γ_n value is 0.27, which is smaller than the γ_n value of 0.37 for the KHP at the same pressure. While the loss in the ΔP_f due to the diminished γ_n value from 0.37 to 0.27 is small, the standard flow rate of 6×10^{-6} sccm at 50 Torr in the KLP is 30 times larger than 2×10^{-7} sccm at 50 Torr in the KHP and even 3 times larger than 2×10^{-6} sccm at 760 Torr in the KHP (Fig. 3.2b). The significantly larger value of the standard flow rate in the KLP is due to the quadratic dependence of the flow rate on the channel height, *a*, as shown in Eq. (3.1).

In the KHP, the wide channel height is 30 μm, considerably larger than the mean free path of 0.07 μm at 760 Torr; the γ_w value of 0.00 sufficiently suppresses the ΔP_r (Fig. 3.2a). In the KLP, the wide channel height is 100 μm, considerably larger than the mean free path of 1.12 μm at 50 Torr. This reduces the magnitude of ΔP_r, as shown in the plots of γ_w (Fig. 3.2a). The wide channel height in the KHP is sized smaller than that in the KLP to reduce the dead volume represented by the wide channels, thereby resulting in faster pumping time with a negligible loss in suppressing the ΔP_r.

In this manner, the two-part design is accomplished by customizing the narrow channels for higher standard flow rate and the wide channels for smaller dead volume, while achieving high compression ratios in both KHP and KLP stages.

3.2.3 Calculated Equilibrium Pressures

The equilibrium pressures at each stage for the KHP and the KLP are theoretically calculated using Eq. (2.3) (Fig. 3.3). The intended design for the Knudsen pump (Table 3.1) results in the upstream pressures of 65 Torr for the KHP and 0.39 Torr for the KLP, using ΔT of 50 K and T_{Cold} of 300 K (represented by blue solid lines in Fig. 3.3). All calculations assume that the KHP vents to a downstream pressure of 760 Torr, whereas the KLP has a downstream pressure of 50 Torr. The possible variations in wide channel height for the KLP and ΔT for both KHP and KLP are separately calculated as explained below.

First, using the fixed ΔT of 50 K, the effects of the increasing wide channel heights from 30 μm to 400 μm for the KLP on pressures are calculated (Fig. 3.3a). The upstream pressure for the KLP improves from 0.96 Torr for a wide channel height of 30 μm to 0.39 Torr for 100 μm, and further to 0.22 Torr for 400 μm. The hypothetical case in which ΔP_r is neglected is also plotted in Fig. 3.3a. This shows that the ΔP_r is increasingly noticeable at sub-Torr pressure as the wide channel height is decreased.

Second, using the fixed channel dimensions listed in Table 3.1, the effects of the increasing values of ΔT from 50 K to 150 K on pressures for both KHP and KLP are calculated (Fig. 3.3b). For the KHP, by increasing values of ΔT from 50 K to 100 K and to 150 K, the upstream pressure decreases from 65 Torr to 6 Torr and to 2 Torr,

respectively. For the KLP, the upstream pressure decreases from 0.39 Torr to 0.08 Torr

and to 0.04 Torr, respectively, but by smaller ratios than those for the KHP. The

response for larger ΔT, especially at sub-Torr pressure, illustrates that the ability to

reduce pressure by increasing ΔT is neutralized by ΔP_r.
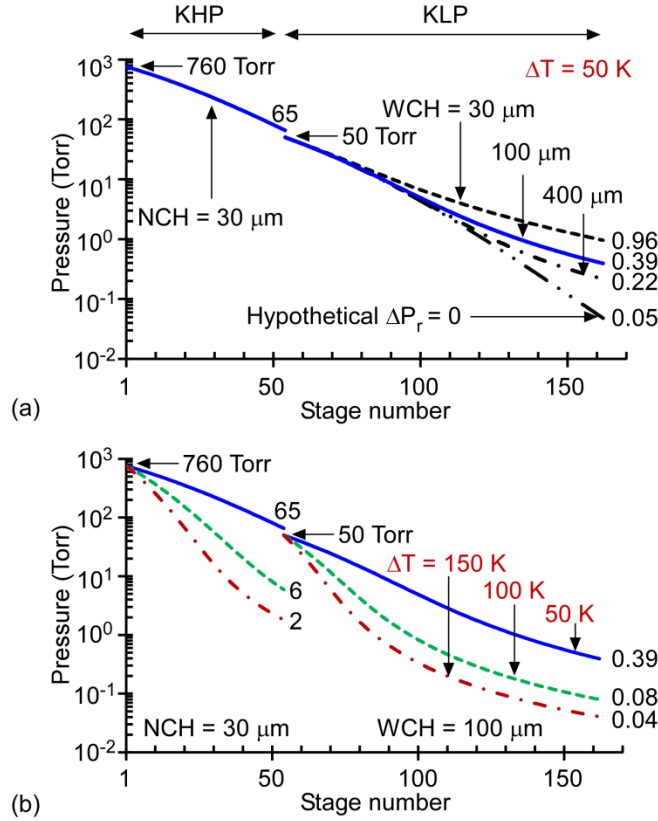
Figure 3.3: Calculated equilibrium pressures at each stage. (a) Variation in wide channel height in the KLP, using the fixed ΔT of 50 K. (b) Variation in ΔT, using the fixed dimensions in Table 3.1. Narrow and wide channel heights are abbreviated as NCH and WCH, respectively.

3.2.4 Heater Design and Pirani Gauge

The layout of a single stage in the Knudsen pump is illustrated in Fig. 3.4. The

narrow channel is formed by dielectric thin films located on the surface of the device,

whereas the wide channel extends into a cavity in the wafer. A thin film metal trace

serves as the heater integrated into each stage. The heater is suspended on a dielectric

membrane at the boundary between the narrow channel and the wide channel.   This

location provides the maximum useful temperature gradient.
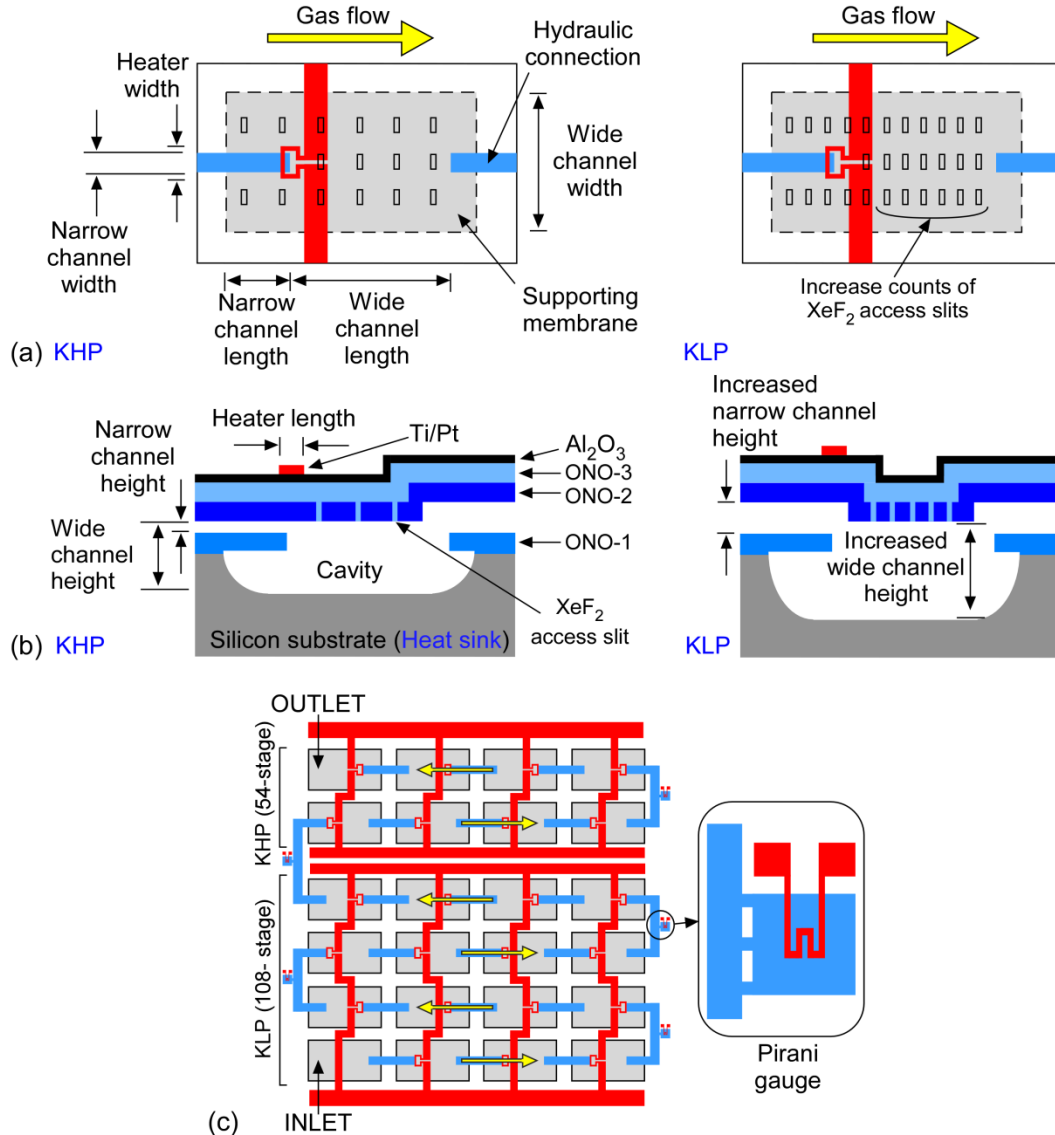
Figure 3.4: (a) Representative layouts of single stages for KHP (left) and KLP (right) in a Knudsen pump. (b) Cross sections along longitudinal axes for KHP (left) and KLP (right). (c) Layout of overall KHP/KLP. In (c), the right side inset indicates a magnified view of a Pirani gauge. For KLP in (a) and (b), only differences in structure from KHP are indicated.

46

Table 3.2: DOE method for low power heater design. Twelve design factors of arbitrary low and high values are used to find ΔT for an input power of 2.4 mW. The four main factors, which are boxed, are fine-tuned for final dimensions. In the 48-stage Knudsen pump, as described in Chapter 2, the narrow channel length, ONO-3 thickness, Al_2O_3 thickness, and heater width are 30 μm, 3.8 μm, 0.5 μm, and 100 μm, respectively. These values are changed in the 162-stage Knudsen pump for saving input power.

Factors		Input 2-level		Output	Final
		Low (μm)	High (μm)	ΔT (K)	(KHP/KLP) (μm)
Wide channel	Width	300	400	-0.02	300
	Length	250	350	0.06	250
	Height	30	200	0.38	30/100
Narrow channel	Width	100	120	-0.05	22/12
	Length	30	200	9.23	100
	Height	0.1	0.2	0.04	0.1/1.0
Insulator thickness	ONO-1	1.2	1.5	-0.22	1.2
	ONO-3	1.3	3.8	-10.75	1.9
	Al_2O_3	0.0	0.5	-6.14	0.2
Heater	Length	25	50	-0.10	25
	Width	50	100	-3.72	25
	Thickness	0.1	0.2	-0.59	0.1

The input power for the Ti/Pt heater is minimized by increasing thermal isolation, using a design of experiments (DOE) method [Ant03]. Using 12-factor, unpaired, 2-level DOE, finite element analysis (FEA) is performed, allowing each structural factor (parameter) to vary arbitrarily. Four primary dependencies are established by evaluating the impact on ΔT. The dimensions are then fine-tuned: 1) the narrow channel length; 2) the thickness of the oxide-nitride-oxide (ONO) layer (ONO-3); 3) the thickness of the Al_2O_3 layer; and 4) the heater width. For an imposed input power, a temperature increase for KHP can be regarded as similar to that for KLP because impacts of differences in height and width for the fabricated wide channel (described in Section 3.3) are small, as can be seen from Table 3.2. Using the final values in Table 3.2, an input power of 2.4

mW can achieve 331.4 K if the lower surface of the device is held at 300 K (Fig. 3.5). For the 162-stage pump that is experimentally evaluated in Section 3.4, the resulting power is 0.39 W.
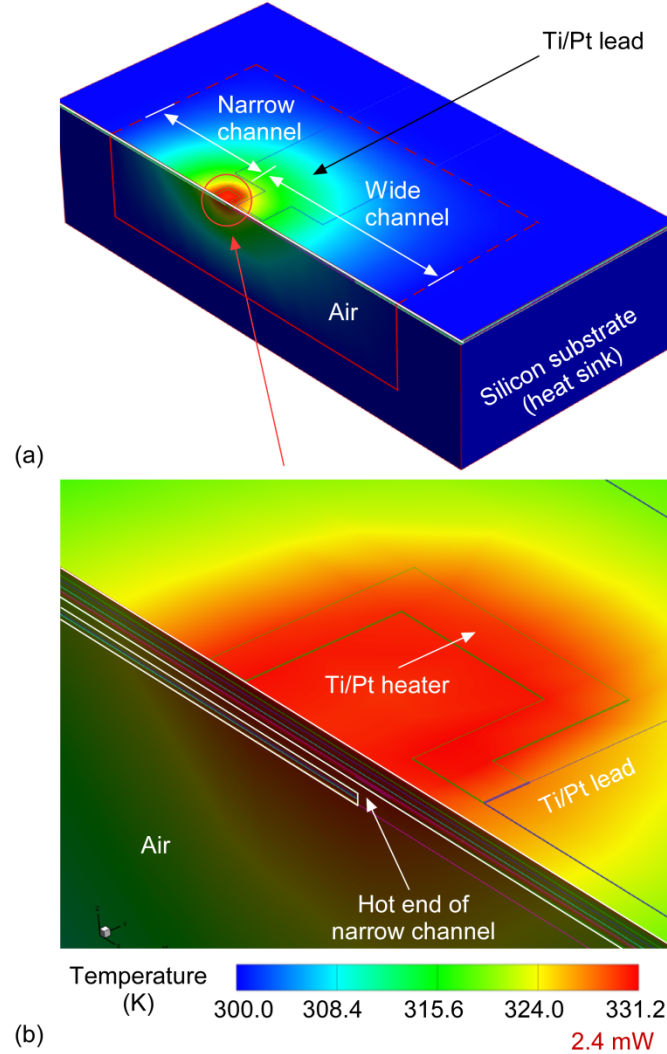
Figure 3.5: Temperature distribution obtained by heater simulation of the designed KLP stage, using FEA. (a) Perspective view of a half-sliced single stage. (b) Magnified view of the T_{hot} position.

Integrated Pirani gauges, located adjacent to pump stages at the perimeter of the array, are used to measure the vacuum levels (Fig. 3.4c). The Pirani gauge converts pressure within a sense gap into a fractional change in electrical resistance of a localized

Joule heater; this is caused by a change in the thermal conductance change of the sense gap with pressure [Mas91]. Four Pirani gauges, connected to stages 1, 54, 99, and 162, are selected for representing the pressure distribution in the overall KHP/KLP layout. These gauges are named P1, P54, P99, and P162, respectively. To enhance the sensitivity for a range from 760 Torr to 1 Torr, the sense gap is designed to be ≈1 μm [Dom05]. This sense gap is identical to the narrow channel height in the KLP, so their fabrication steps are identical.

3.3 Fabrication

To implement the Knudsen pump, a single-wafer five-mask fabrication process is used (Fig. 3.6). The important aspects in the channel fabrication are: 1) the narrow channels for the KHP are defined by a thin (sacrificial) polySi layer (Fig. 3.6b), whereas the narrow channels for the KLP are defined by both thin and thick (sacrificial) polySi layers (Fig. 3.6b, 3.6c); 2) cavities for wide channels are formed by partially etching the bulk silicon (Fig. 3.6e); and 3) the walls of the narrow channels are composed of stress-relieved oxide-nitride-oxide (ONO) layers. The (sacrificial) polySi layers are also used for hydraulic connections between KHP and KLP and between Pirani gauges and nearby stages.
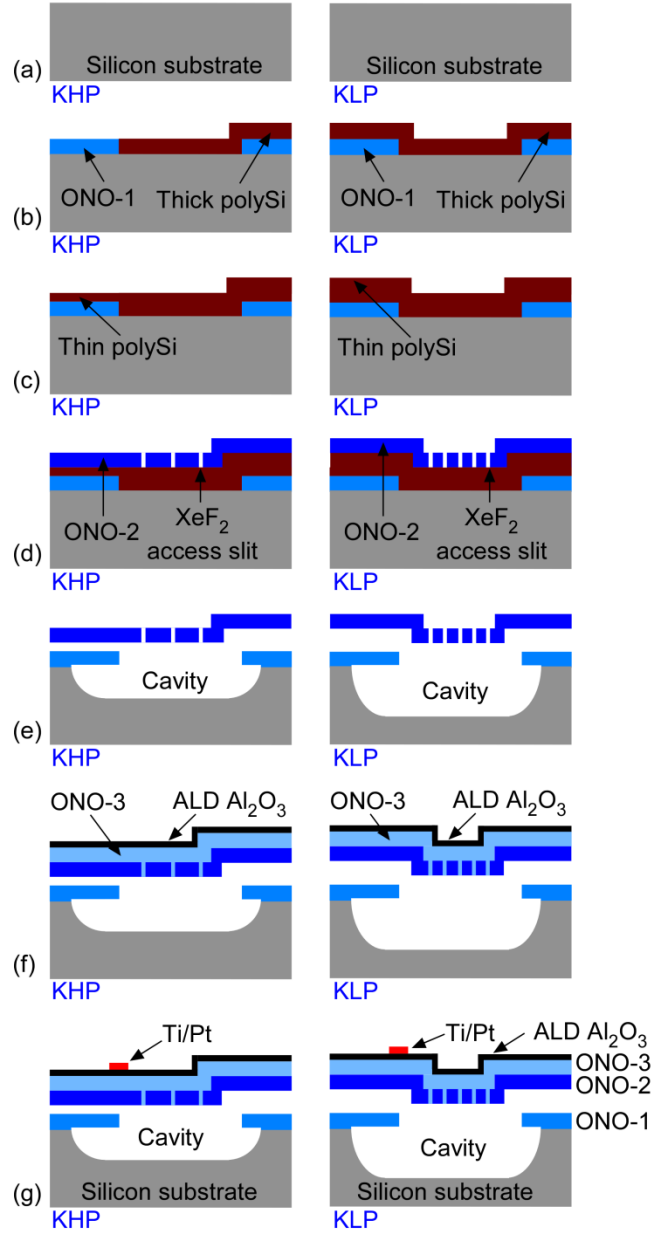
Figure 3.6: The five-mask single-wafer process for KHP (left) and KLP (right). (a) Silicon wafer. (b) Deposition and patterning of (Mask 1) LPCVD ONO-1 and (Mask 2) thick LPCVD polySi. (c) (Mask 3) Deposition & patterning of thin LPCVD polySi. (d) (Mask 4) Deposition and patterning of LPCVD ONO-2. (e) Sacrificial etching of polySi and bulk silicon using XeF_2 gas. (f) Deposition of PECVD ONO-3 and ALD Al_2O_3. (g) (Mask 5) Ti/Pt metallization.

The microfabrication process is initiated by deposition and patterning of the low-pressure chemical vapor deposited (LPCVD) first ONO layer (ONO-1) which is later used as a mask for the cavity etch of the single-crystal silicon substrate (Fig. 3.6a). The

process continues with the deposition and patterning of a thick LPCVD polySi layer of ≈0.7 μm (Fig. 3.6b), followed by the deposition and patterning of a thin LPCVD polySi layer of ≈0.1 μm (Fig. 3.6c). After deposition of the LPCVD second ONO layer (ONO-2), an array of slits, 2×10 μm^2, are patterned using reactive ion etching (RIE); these are intended to provide XeF$_2$ access in the next step (Fig. 3.6d). Then, the sacrificial polySi layers are etched away by XeF$_2$ dry gas; the bulk silicon is partially removed in this step (Fig. 3.6e). The next step is to seal the access slits with the third ONO layer (ONO-3) that is ≈1.9 μm-thick. It is deposited by plasma-enhanced chemical vapor deposition (PECVD) because, with a deposition temperature of 380°C, it provides lower thermal stress to the released membrane than LPCVD, which is typically performed at >800°C. This is followed by a ≈0.2 μm-thick atomic layer deposited (ALD) Al$_2$O$_3$ layer that provides a hermetic seal (Fig. 3.6f). Finally, a metal layer of Ti/Pt (25/100 nm) is patterned, using lift-off process (Fig. 3.6g). This is used for heaters, Pirani gauges, and wire bonding pads.

Table 3.3: Measured thickness and residual stress of each layer.

Layer	Thickness (μm)	Residual stress (MPa)
LPCVD ONO-1	0.5/0.2/0.5	+42.5
LPCVD ONO-2	0.5/0.2/0.5	+42.5
PECVD ONO-3	0.7/0.5/0.7	+42.1
Thin / thick polySi	0.1 / 0.7	-
ALD Al$_2$O$_3$	0.2	+304
Ti/Pt	0.025/0.1	-

At each step in the process, the thickness and residual stress of each layer are monitored (Table 3.3). The oxide and nitride layers in each ONO layer are thickness-controlled to have a mild tensile residual stress of ≈42 MPa to avoid a buckling of the suspended membrane due to compressive residual stress. The total thickness of

sacrificial polySi layers for the narrow channel in the KLP is $\approx$0.8 µm, lower than the design value of 1 µm in Section 3.2. The fabricated chip has a footprint of 12×15 mm^2. The KHP, KLP, individual stages, inlet, outlet, and Pirani gauges are shown in Fig. 3.7. The fabrication yield is typically 60-80% with the available tool set.
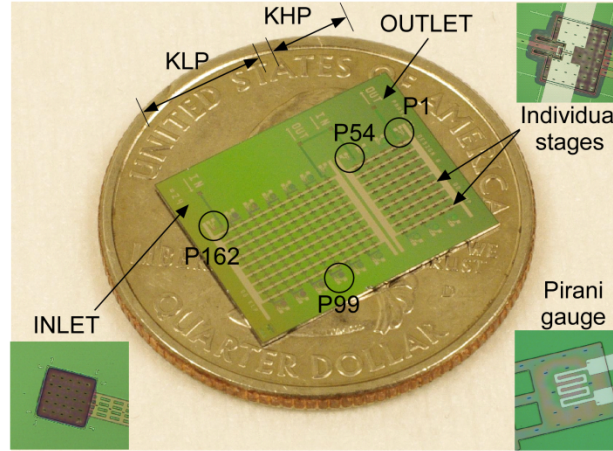
Figure 3.7: Photograph of an as-fabricated chip. The upper right inset shows a stage, the lower right inset a Pirani gauge and the lower left inset the inlet.

During fabrication, the main differences between the KHP and the KLP regions of the chip (Fig. 3.8) are: 1) the areal density of XeF$_2$ access slits is eight times larger in the KLP than that in the KHP; and 2) the narrow channels are formed by sacrificially etching the thin polySi layer for the KHP and the thin and thick polySi layers for the KLP. The height and the lateral undercut of the wide channel are controlled by the areal density of the XeF$_2$ access slits and etching time. Hence, the wide channel in the KLP is etched deeper than that in the KHP, and the XeF$_2$ lateral undercut profile in the KLP is wider than that in the KHP.
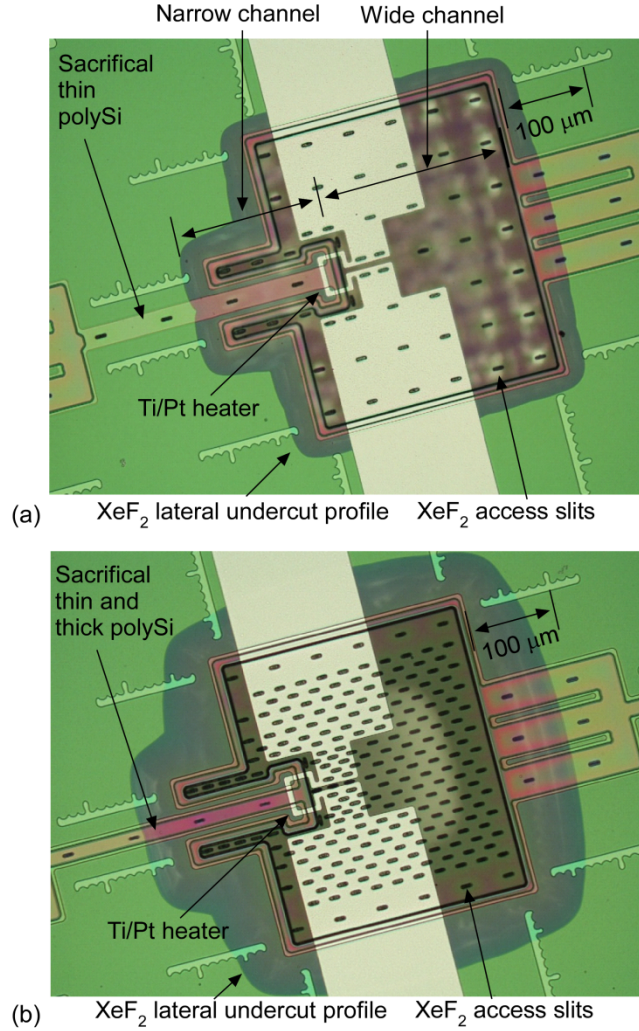
Figure 3.8: Microscopic images of fabricated stages in an identical scale. (a) A KHP stage. (b) A KLP stage.

The cross sections of the narrow and wide channels in the KHP and the KLP are examined with scanning electron microscope (SEM) images (Fig. 3.9). The difference between the cross section of narrow channel in the KHP and KLP is evident from Fig. 3.9. The KHP uses a narrow channel that is 0.1 μm in height and 22 μm wide, whereas the KLP uses this, further enhanced by an opening that is nominally 0.8 μm in height and 12 μm wide. In the wide channel (Fig. 3.9c, 3.9d), the eight times larger areal density of XeF_2 access slits in the KLP produces a height of 108 μm compared to that of 36 μm in

the KHP. The larger density also produces a wider channel width of 422 μm in the KLP

than that of 358 μm in the KHP, evident as the silhouette of the undercut in the top view

(Fig. 3.8); the defined width of the cavity opening in the photomask (Fig. 3.6a) is only

300 μm in both KHP and KLP.
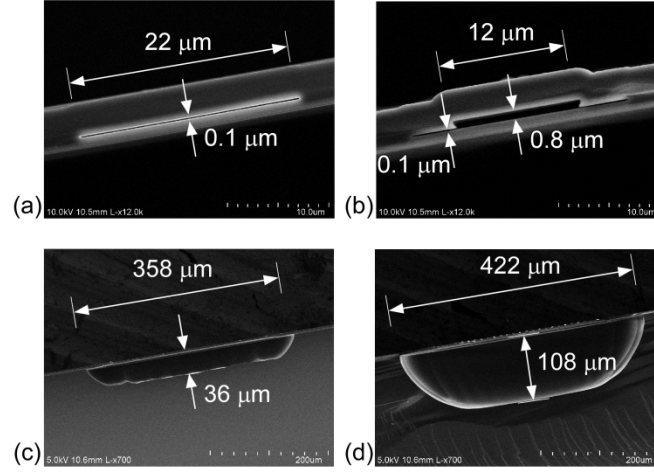
Figure 3.9: SEM images of narrow and wide channels in KHP and KLP. (a) A narrow channel in KHP. (b) An increased narrow channel in KLP. (c) A wide channel in KHP. (d) An increased wide channel in KLP.

3.4 Test Results

3.4.1 Methods

The fabricated chip is evaluated in a test chamber, as described in Section 2.5.1.

The pump is operated by providing a constant voltage to the Ti/Pt heater; the total input

power to 162 stages is 390 mW and the input power in a stage is ≈2.4 mW. The

temperature increase of the heater is calculated by dividing the fractional change in

resistance by the thermal coefficient of resistance, α. The input power levels, as

indicated in test results (Fig. 3.10-3.13), are for the heated resistances at equilibrium

pressures.

For measuring the vacuum levels at evacuated stages, a constant current of 4 mA is provided to each Pirani gauge and the measured fractional change in resistance is correlated to pressure. The input power level for the heated resistance is ≈2.5 mW, which is modest relative to 390 mW for the Knudsen pump heater.

The Pirani gauges are calibrated by a special method, named dynamic calibration, which accommodates the reduced interior pressure and the ambient exterior pressure of the Pirani gauge during pump operation; this is described in Chapter 4 in detail [An13b]. Here, the method is briefly illustrated for P162 and P1; P54, and P99 follow the case for P162. The method uses the response of P162 and P1 under rapid modulation of ambient pressure. 1) P162 provides a response that represents the combined effect of modulated exterior ambient pressure and nearly unchanged interior pressure; this is due to the relative slowness in pressure variation in the interior, as compared to the rapid modulation of the ambient pressure. 2) P1 provides a response that always represents the rapidly modulated ambient pressure because the interior, together with exterior, of P1 is directly exposed to ambient pressure. The multistep-modulation of the ambient pressure yields an equilibrium pressure for which the calibrated responses of P162 and P1 are equal. This equilibrium ambient pressure is exactly equal to the interior pressure of P162. Process-induced variations in responses of P162 and P1 are fitted by a linear regression model. In evaluating the pump performance, the measured pressure values are indicated with error bars for $\pm 2\sigma$ (where σ is the standard deviation), obtained by error analyses of the residual non-linearity in the regression model, together with repeatability of Pirani gauge sensors. The error bars for $\pm 2\sigma$ represent a 95.4% confidence interval, assuming normal (Gaussian) distribution [Kre06].

3.4.2 Pumping Results

Two sets of tests are performed to evaluate the fabricated devices. The first test evaluates the combined operation of the KHP and KLP. For this test the inlet at stage 162 remains sealed, while stage 1 serves as the outlet. The second test evaluates the KLP by itself. For this test, stage 162 remains as the sealed inlet, while stage 54 serves as the outlet. In order to facilitate this, the supporting membrane at P54 is broken to allow it to vent to the ambient. Therefore, this is a destructive test.
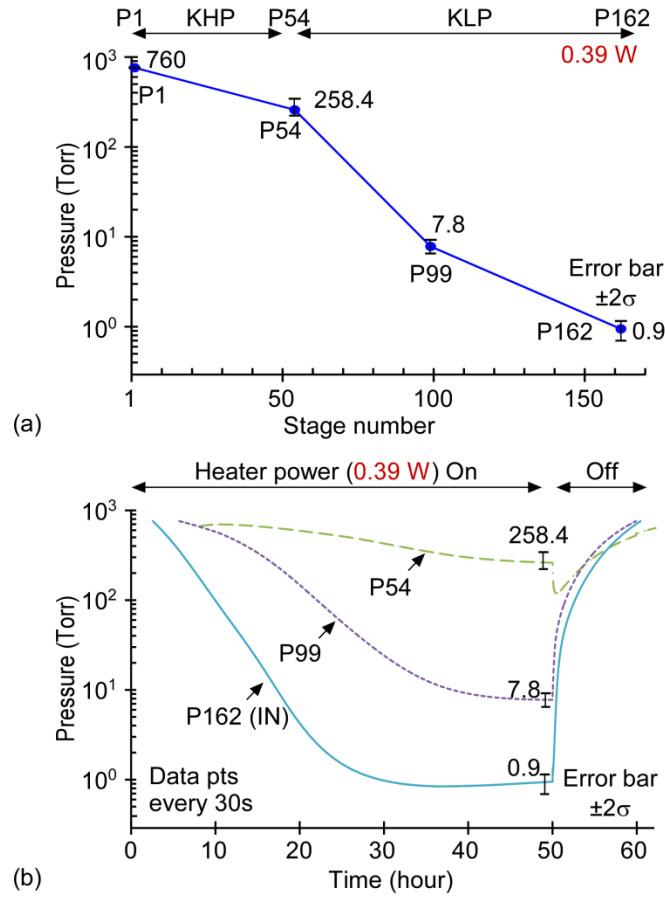
Figure 3.10: Typical test results for the combined operation of KHP and KLP. (a) Equilibrium pressures at 760 Torr ambient. (b) Transient responses at 760 Torr ambient.

The unheated resistances are ≈126 Ω for the Knudsen pump heater and ≈140 Ω for the Pirani gauge. The experimental value of α is 2,314 ppm/K for the Ti/Pt material

56

used in both heaters and gauges. The temperature increase of the Pirani gauge from the typical fractional increase in resistance of 0.12 at 0.1 Torr is ≈52 K.

For the first set of tests (Fig. 3.10a), 0.39 W was applied to the heater at atmospheric ambient pressure. The temperature increase of the heater was typically ≈56 K; this increase is twice that of the theoretical estimate using FEA (Fig. 3.5). The pressure at P162 was 0.9 Torr, which corresponds to a total compression ratio of ≈844. The equilibrium pressures at intermediate stages were 258 Torr for P54 and 7.8 Torr for P99.
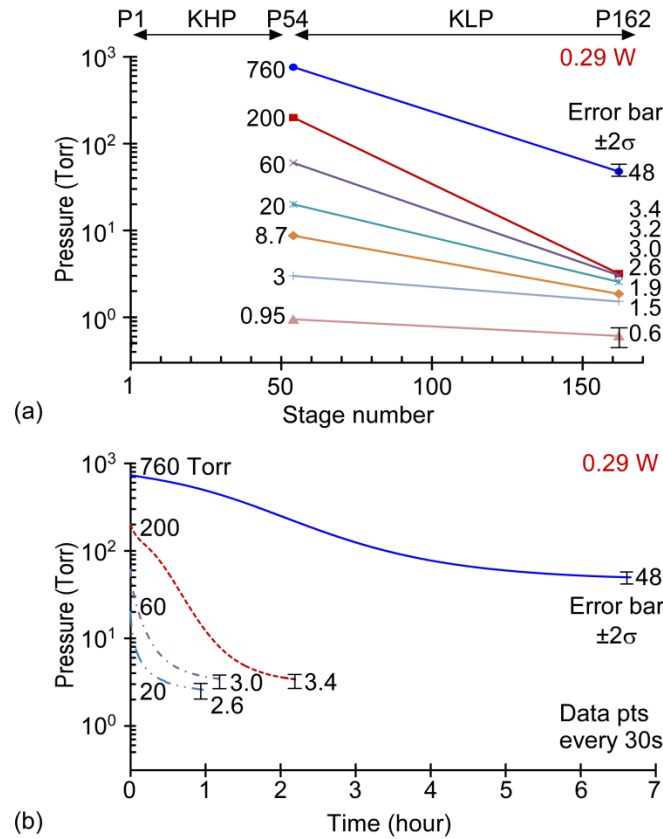
Figure 3.11: Typical test results for the operation of KLP by itself. (a) Equilibrium pressures at various ambient pressures. (b) Transient responses of only P162 at each ambient pressure.

For the second set of tests (Fig. 3.11a), an input power of 0.29 W was applied. The temperature increase of the heater was typically ≈68 K. The equilibrium pressures of

P162 were measured at various ambient pressures to mimic various downstream conditions for the KLP. At atmospheric pressure, the equilibrium pressure at P162 was 48 Torr. At ambient pressures from 200 Torr to 3 Torr, the equilibrium pressures ranged from 3.4 Torr to 1.5 Torr, respectively. At an ambient pressure of 0.95 Torr, the pressure reduced to 0.6 Torr.

The pre-equilibrium transient behavior was also evaluated during the two sets of tests. Using a LabVIEW^{TM} program, the transient response was automatically recorded every 30 seconds after applying the input power. For the first set of tests at atmospheric ambient pressure, P162 took ≈25 hours to achieve equilibrium (Fig. 3.10b). P54 typically took ≈50 hours, i.e., twice as long. Once the heater was turned off, the pressure at P162 returned to the ambient value in ≈10 hours, 5 times shorter than the evacuation time of ≈50 hours for P54. The transient response for the second set of tests was measured at ambient pressures of 760 Torr, 200 Torr, 60 Torr, and 20 Torr (Fig. 3.11b). To achieve equilibrium, P162 took ≈5 hours, ≈2 hours, ≈1 hour, and ≈0.5 hour, respectively.

The impact of varying input power was investigated using the first set of tests at atmospheric ambient pressure (Fig. 3.12). The input power was increased to 0.41 W, 0.69 W, and 1.08 W. The temperature increases of the heater went up to ≈68 K, ≈120 K, and ≈174 K, respectively. The equilibrium pressures for P54 were 96.8 Torr, 9.3 Torr, and 4.7 Torr, respectively. The pressures for P162 were 0.90 Torr, 0.80 Torr, and 0.84 Torr, respectively. The highest compression ratio was approximately 760/0.80, i.e., 950, for 0.69 W. For the higher power of 1.08 W, the compression ratio was lower.
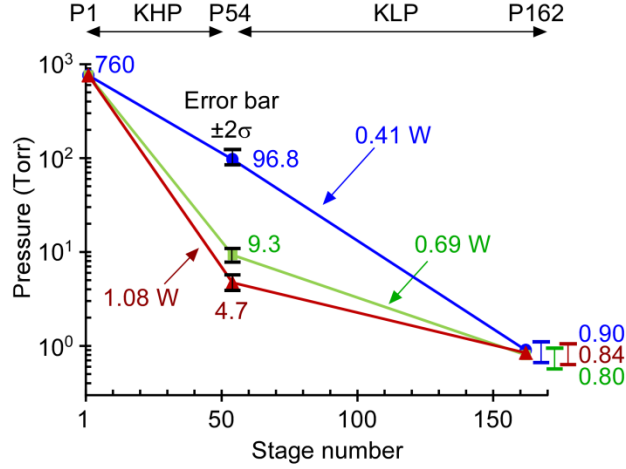
Figure 3.12: Variation in input power for the combined operation of KHP and KLP at 760 Torr ambient.

3.4.3 Reliability Tests
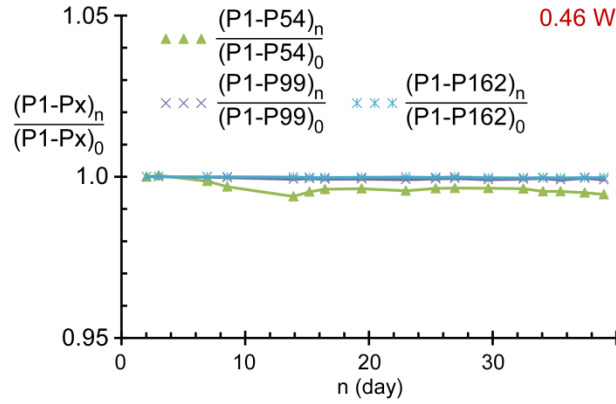
Figure 3.13: Reliability test for 37 days of continuous operation. The subscript n denotes the elapsed days. Each pressure difference from Pirani gauge P1 to Pirani gauges P54, P99, and P162 is normalized to that at the beginning, 0.

To investigate the reliability of a micromachined Knudsen pump, one sample of the full 162 stage Knudsen pump was continuously operated for 37 days, at an input power of 0.46 W. Figure 3.13 shows that the evacuation levels for each of Pirani gauges P54, P99, and P162 remained within 1% of the value on the first day. Since mechanical failure is unlikely because of the absence of moving parts, the primary failure mechanism could stem from the heating of the Knudsen pump. However, the thermal degradation of

the various layers is expected to be very modest because the temperature increases by only 71 K, i.e., to about 96°C, at 0.46 W. In contrast, the deposition temperature of the various thin films used in the fabrication were 910°C for LPCVD oxide, 802°C for LPCVD nitride, 380°C for PECVD oxide, 380°C for PECVD nitride, and 250°C for ALD Al_2O_3. In case of Ti/Pt metal, annealing at 600°C for 5 hours stabilizes electrical properties [Qu00]. Hence, the Knudsen pump, without any substantial failure mechanism, can provide high reliability and long term operation.

M1: KNF NMP05S [Kac08]
M2: KNF NMP09B [Kac08]
M3: KNF NMP015S [Kac08]
P1: Kim, 2007 [Kim07a]
P2: Besharatian, 2012 [Bes12]
Z1: Wijngaart, 2000 [Wij00]
Z2: Gerlach, 1995 [Ger95]
Z3: Kamper, 1998 [Kam98]
Z4: Bartels mp5 [Mey11]
Z5: Stemme, 1993 [Ste93]
Z6: thinXXS [Web10]

K1: McNamara, 2005 [McN05]
K2: Gupta, 2011 [Gup11a]
K3: Pharas, 2010 [Pha12]
K4: Gupta, 2011 [Gup11b]
K5: Gupta, 2012 [Gup12]
K6: Gupta, 2012 [Gup12]
K7: An, 2013 [An13a]

Figure 3.14: Benchmarking with other pumps for compression ratio, P_{OUT}/P_{IN}, and device volume. The present work, for the combined operation of KHP and KLP, achieves a compression ratio of ≈844 (= 760/0.9), using an input power of 0.39 W. Note that both x and y axes are log-scaled.

3.5 Discussion and Conclusion

The compression ratio and device volume are benchmarked with other previously reported pumps [Ste93, Ger95, Kam98, Wij00, Ngu02, Las04, McN05, Kim07a, Kac08,

Web10, Pha10, Gup11a, Gup11b, Mey11, Bes12, Gup12]. Using an input power of 0.39 W, the compression ratio of $\approx$844 (= 760/0.9) exceeds the highest previous compression ratio reported in Chapter 2, by a factor of $\approx$17, with 3.5 times better power efficiency (Fig. 3.14). This performance over the other pumps is enabled by: 1) the monolithic integration of 162 stages into one silicon chip, using silicon-based micromachining processes; 2) the two-part design, separately customizing stages into the KHP and the KLP; and 3) the heater design obtained by DOE and FEA.

The effect of ΔP_r at high vacuum was evident when using varying levels in input power (Fig. 3.12). For P162, the measured pressure remained between 0.9 Torr and 0.8 Torr without noticeable reduction, when the ΔT increased from 68 K to 174 K. In contrast, for a similar ΔT of 50 K to 150 K, the theoretically calculated pressure reduced from 0.39 Torr to 0.04 Torr, which is a larger ratio than that of the measurement. The measured results suggest that the unintended ΔP_r in the KLP wide channel fully neutralizes the ΔP_f in the narrow channel at pressures less than $\approx$1 Torr. For the theoretical calculation of Q_T and Q_P, it was assumed that T_{Hot} gradually decreases to T_{Cold} along the wide channel (Fig. 3.1). However, in the pressure range of 1 Torr to 0.1 Torr, the corresponding mean free path from 49 μm to 511 μm (Table 3.1) encompasses the wide channel length of 250 μm. Therefore, air molecules could be transported from the hot to cold end without collision, resulting in hot molecules entering the cold end of the next narrow channel, and vice versa. This collisionless transport, which could break from the assumed gradual temperature change in the wide channel, could explain why the performance was less than predicted using theoretical calculations. For this reason, a

more advanced level of design and experimentation at these low pressures is necessary for further improvement of compression.

The ΔP_r at high vacuum was also evident from the pressure response of P54. Using similar input powers of 0.39 W and 0.41 W (Fig. 3.10a, 3.12), the pressures of 258.4 Torr and 96.8 Torr at P54, respectively were reduced to ≈0.9 Torr at P162. The different pressures at P54 might be caused by sample-to-sample variations in channel dimensions. Figure 3.11 also shows similar P162 pressures of 3.4 Torr and 3.2 Torr when the initial pressures are 200 Torr and 60 Torr, respectively. It appears that the pressure at P54 bottoms near this level because of ΔP_r in the wide channel.

The pre-equilibrium transient behaviors at different levels of ambient pressure were consistent with the trends anticipated in Section 3.2.1. For example, when the KLP was tested by itself, lowering the ambient pressure reduced the time needed for achieving equilibrium. Additionally, when the pump was tested as a whole, upstream stages achieved equilibrium before downstream stages.

In summary, a two-part (KHP and KLP) architecture has been investigated for the 162-stage Knudsen pump. A five-mask single-wafer process was used to fabricate the pump in a small footprint of 12×15 mm^2. Most notably, the high compression ratio of ≈844 was achieved at atmospheric ambient pressure; the input power was limited to ≈0.39 W; and the pump was reliable, with sustained operation over 37 days. Subsequent analyses suggest that at sub-Torr pressure ΔP_r in the wide channel neutralizes further evacuation. At the cost of additional masking steps in the fabrication process, the architecture may be extended to more than two partitions to improve pumping rate and compression ratio.

CHAPTER 4

A DYNAMIC CALIBRATION METHOD FOR PIRANI GAUGES EMBEDDED IN FLUIDIC NETWORKS

This chapter describes *in situ* dynamic calibration of microfabricated Pirani gauges, which are commonly used to measure vacuum levels in microsystems that may include pumps, valves, and other pressure modulation elements. Calibration is accomplished with a configuration where two sides of the supporting membrane of a Pirani gauge are at different pressures: an interior pressure within the microsystems and an exterior ambient pressure that can be directly controlled. This particular configuration is increasingly common as Pirani gauges are embedded in complex fluidic pathways such as those found in micro gas chromatographs and multistage gas pumps. In the dynamic calibration the ambient pressure is rapidly modulated, while the interior pressure in the sense gap of the Pirani gauge remains relatively unchanged. The exterior pressure that is equal to the interior pressure is determined by a regression model. The dynamic calibration procedure and subsequent error analysis are illustrated by application to a 162-stage monolithic Knudsen vacuum pump presented in Chapter 3. For this device, dynamic calibration improves the estimate of the upstream pressure from 30 Torr (as suggested by uncorrected static calibration) to 0.9 Torr, with a 95.4% confidence interval from 0.7 Torr to 1.1 Torr, assuming normal (Gaussian) distribution. Dynamic calibration can be significantly more accurate than conventional static calibration for certain types of devices.

4.1 Introduction

Pirani gauges, which transduce gas pressure into changes in electrical resistance [Ubi51a, Ubi51b, Wil04], have been widely used in microsystems [Her87, Mas91, Wen94, Cho97, Ell00, Sta03, Cha05, Dom05, Mit08, Xia11, San11]. Typically, the Pirani gauge has a resistive heating element that is suspended on a thermally isolated membrane. The heat generated in the operation of the Pirani gauge flows across a sense gap beneath the membrane, to a substrate, which serves as the heat sink. As the vacuum level or gas pressure changes in the sense gap, the resulting change in thermal conductance causes a change in the temperature of the membrane (or alternatively, a change in the power level necessary to maintain it at a fixed temperature). The change in temperature can be detected by the resistive heating element, or alternatively by another resistor. The structure and operation of a typical Pirani gauge are illustrated in Fig. 4.1. Pirani gauges are favored as embedded sensing elements within microfluidic systems and micromachined pumps in [Dom07] and Chapter 2-3, because of simple fabrication requirements that are relatively easy to integrate into other manufacturing processes.
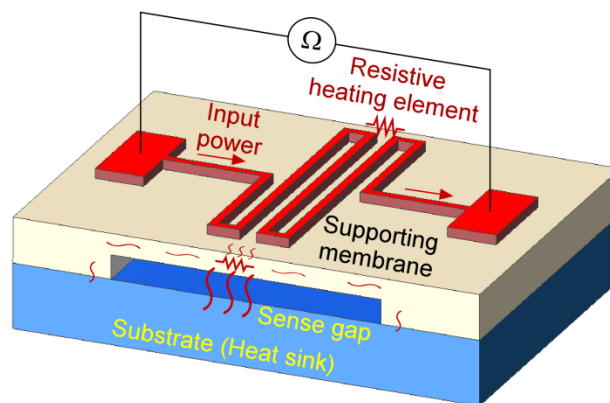
Figure 4.1: Schematic of a Pirani gauge.

For the accurate measurement of pressure with a Pirani gauge, its resistance variation must be calibrated over the operating pressure range. This is necessary because process-induced sample-to-sample variations in dimensions and material properties can cause variations in the response. A Pirani gauge is typically calibrated by varying the pressure in the sense gap. In most cases, the structure is designed so that the heat loss from the suspended heating element is dominated by thermal conduction through the sense gap; i.e., the heat losses to the exterior of the structure, and the conductive heat losses along membrane on which the Pirani gauge is suspended, are relatively small. During calibration, the interior of the membrane and the exterior of the membrane are both at a common pressure controlled by a calibration pump. In this method, the dominance of the heat loss through the sense gap is retained, and the thermal losses to the exterior [Kad85, Sai10] are accommodated in the calibration (Fig. 4.2a). For the purpose of this paper, this calibration procedure is named static calibration. In contrast, when the Pirani gauge resides on a thermally-insulating membrane which completely separates the underlying sense gap from the exterior environment as in Chapter 2-3, static calibration is likely to be inaccurate because of the lack of direct access to – and control of the pressure in – the sense gap. Further, if the interior pressure is low, the heat losses to the exterior could represent a significant un-calibrated component.
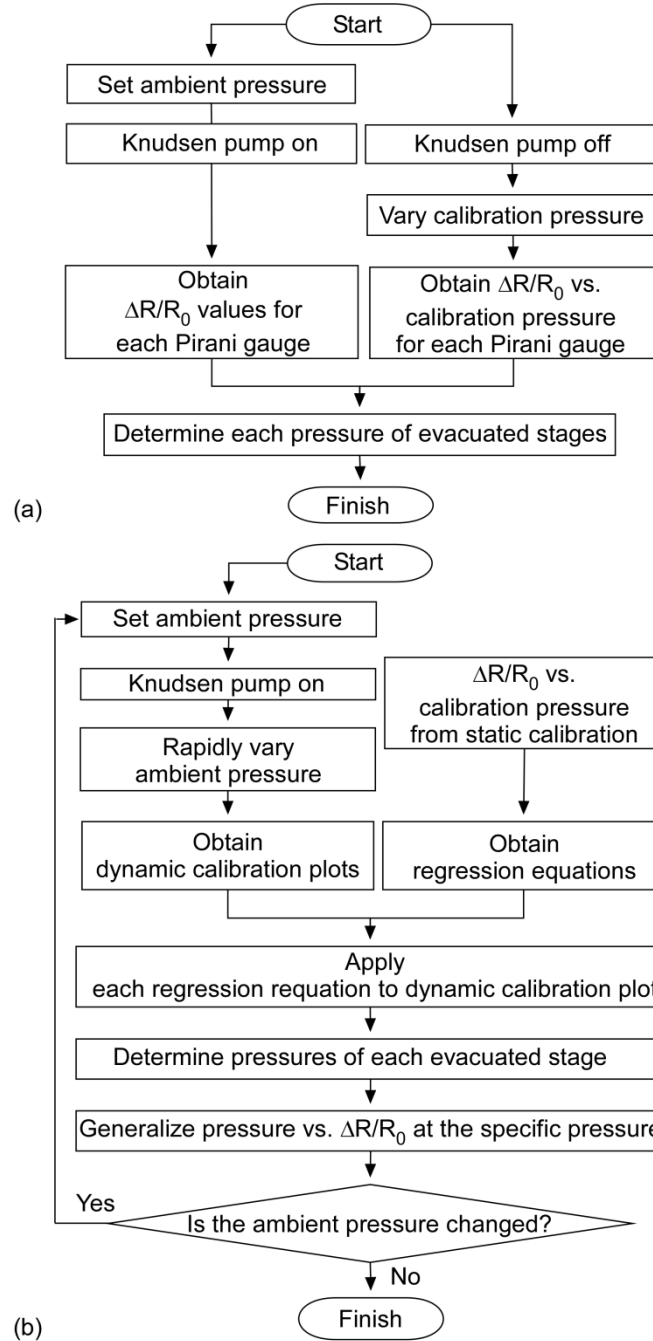
Figure 4.2: Flow charts for pressure measurements with Pirani gauges. (a) Using static calibration. (b) Using dynamic calibration.

Structures in which the Pirani sense gap is inaccessible are increasingly common in microsystems with complex or tortuous fluidic pathways, such as micro-gas chromatographs and multistage gas pumps [Kim07a, Kim07b, Liu11]. This chapter

describes a dynamic calibration method used to calibrate Pirani gauges embedded in such structures (Fig. 4.2b). This method exploits the relative slowness of the pressure variation in a sense gap that is embedded deep within a microfluidic pathway, as compared to the speed with which the exterior pressure can be modulated. The value of the (directly controlled) exterior pressure that is equal to the interior pressure (which is not directly controlled) is determined by a regression model that is fitted to a reference Pirani gauge.

The dynamic calibration method is demonstrated using a Si-micromachined, monolithic 162-stage Knudsen pump described in Chapter 3. Pirani gauges are embedded within the pump to provide *in situ* pressure measurement at various intermediate points within the flow channel. Each Pirani gauge is positioned next to a pump stage to be measured. The pressure in the sense gap of each Pirani gauge is determined by the Knudsen pump operation, whereas the exterior is at ambient pressure, which is controlled by a calibration pump.

The theoretical operation of the Pirani gauge, the dynamic calibration concept, and the linear regression model are described in Section 4.2. Section 4.3 describes the test setup and methodology, followed by the experimental results from the calibration of the sample Knudsen pump by both static and dynamic methods. The possible errors are addressed in estimating the pressure. Conclusions are presented in Section 4.4.

4.2 Theoretical Background

This section describes the theoretical considerations for a Pirani gauge in the context of a differential pressure across the supporting membrane. The dynamic

calibration is described, followed by the linear regression model of Pirani gauge responses.

4.2.1 The Pirani Gauge

As pressure is reduced in the sense gap, the decrease in the density of gas molecules causes an increase in its thermal resistance [Wil04, Mas91, Wen94]. The temperature of the supporting membrane and electrical resistance of the temperature-sensing resistor increase as a fixed input current is supplied to the resistive heating element. This development follows the particular case that the heating resistor is also the temperature-sensing resistor, which is the case for the Knudsen pump demonstration of the dynamic calibration method. In this situation, there is an increase in resistance associated with the heating caused by the current used to operate the Pirani gauge. Calibration is specific to the ambient pressure at which it is performed. Therefore, the total electrical resistance, R_T, can be expressed as:

$$R_T = R_0 + R_{0,Ref}\left(I, P_{Ref}\right) + \Delta R(I, P_{Cal})$$ (4.1)

where I is the current; P_{Ref} is the reference ambient pressure; P_{Cal} is the calibration pressure, the pressure of interest lower than P_{Ref}; R_0 is the unheated resistance; $R_{0,Ref}$ is the increment in resistance at P_{Ref}; and ΔR is the further increment in resistance at P_{Cal}. For the purpose of this paper, I and P_{Ref} are fixed at 4 mA and 760 Torr, respectively, for reasons explained in Section 4.3. In constant current mode, the total electrical resistance, R_T, is provided by [Sen01]:

$$\frac{R_T - R_0}{R_0} = \frac{\alpha I^2 R_0 \Re_{th}}{\left(1 - \alpha I^2 R_0 \Re_{th}\right)}$$ (4.2)

68

where $\Re_{th}$ is the equivalent thermal resistance of the Pirani gauge, which is a function of pressure. The temperature coefficient of the heating/sensing resistor is denoted by α; for a 25/100 nm Ti/Pt thin film, such as used in the sample Knudsen pump, its value is 2,314 ppm/K. The thermal resistance, $\Re_{th}$, is the parallel combination of the constituent thermal resistances:

$$\Re_{th}^{-1} = \Re_d^{-1} + \Re_{sg}^{-1} + \Re_{ext}^{-1} \tag{4.3}$$

where $\Re_d$ is the thermal resistance of the supporting membrane, $\Re_{sg}$ is the pressure-dependent thermal resistance of the sense gap and $\Re_{ext}$ is the equivalent thermal resistance to exterior ambient. However, $\Re_{ext}^{-1}$ can be neglected, assuming $\Re_{ext}$ is large relative to $\Re_d$ and $\Re_{sg}$. Using the thermal conductivity equation in [Dom07], the $\Re_{sg}$ is written as:

$$\Re_{sg} = \frac{1}{\sigma_{th,air,0}} \frac{g + 2\lambda(2/a_T - 1)}{A_{heater}} \tag{4.4}$$

where $\sigma_{th,air,0}$ is the thermal conductivity of air in viscous flow regime, 0.026 W/m-K [Ste85], a_T is the thermal accommodation coefficient (0.75 for air on bright platinum at 280 K [Gri46]), λ is the mean free path of air molecules, which is proportional to temperature and inversely proportional to pressure [Vin75], g is the Pirani sense gap between the supporting membrane and the heat sink, and A_{heater} is the area of the Ti/Pt metal heater. If the pressure within the sense gap is high, $\lambda \ll g$, $\Re_{th}$ becomes nearly equal to $\Re_{sg}$. In contrast, if $\lambda \gg g$, $\Re_{th}$ becomes nearly equal to $\Re_d$. Between these two limits, the fractional change in resistance, caused by the variation in $\Re_{th}$ with pressure, is experimentally characterized in calibration and is used for pressure measurement. At the

reference pressure, assuming that the mean free path of the gas molecules is small, i.e. λ $\ll g$, the $\mathfrak{R}_{th}$ is approximately equal to the thermal resistance of the sense gap:

$$\mathfrak{R}_{th,Ref} \approx \frac{1}{\sigma_{th,air,0}} \frac{g + 2\lambda_{Ref}\left(2/a_T - 1\right)}{A_{heater}} \tag{4.5}$$

where $\mathfrak{R}_{th,Ref}$ is the thermal resistance at the reference pressure and λ_{Ref} is the mean free path of air molecules at the reference pressure (≈ 0.07 µm).

The response of greatest interest is $\Delta R/R_0$, which can be correlated to the calibration pressure. Based on Eq. (4.1-4.2), it can be written as:

$$\Delta R/R_0 = \frac{\alpha I^2 R_0 \mathfrak{R}_{th}}{\left(1 - \alpha I^2 R_0 \mathfrak{R}_{th}\right)} - \frac{\alpha I^2 R_0 \mathfrak{R}_{th,Ref}}{\left(1 - \alpha I^2 R_0 \mathfrak{R}_{th,Ref}\right)} \tag{4.6}$$

where the second term on the right side is $R_{0,Ref}/R_0$. When the applied current, I, is small (i.e, the denominators on the right side of Eq. (4.6) are close to 1), the $\Delta R/R_0$ is linearly proportional to the increment in thermal resistance from $\mathfrak{R}_{th,Ref}$. However, as the applied current is increased, the drop in the values of denominators introduces non-linearity in $\Delta R/R_0$. This non-linearity affects the linear relation of $\Delta R/R_0$ to the increase in thermal resistance with pressure, and varies between different Pirani gauges; this is addressed in Section 4.2.4.

4.2.2 Mechanical Deformation Caused by Unequal Pressures

If the pressure in the sense gap is smaller than the ambient pressure, it may lead to a deflection of the supporting membrane that decreases the sense gap, thereby decreasing $\mathfrak{R}_{th}$, as indicated by Eq. (4.4). The ensuing smaller fractional change in electrical resistance indicates higher pressure than the true pressure in the sense gap. The error

would be increased by any mechanical contact of the supporting membrane to the substrate as a result of the deformation. The mechanical deformation of the supporting membrane in the sample Knudsen pump is addressed in Section 4.3.

4.2.3 Dynamic Calibration

Figure 4.2b describes the steps in the dynamic calibration of a Pirani gauge. The Pirani gauge to be measured and calibrated is assumed to be embedded deep within a microfluidic pathway, well-separated from the outlet (see Fig. 4.6a). Its supporting membrane separates the Pirani sense gap from ambient pressure. In the steady state condition, an internal pump − a Knudsen pump in this example − maintains the pressure in the sense gap, and an external calibration pump controls the ambient pressure. A second Pirani gauge, the reference gauge, is designed such that both sides of its supporting membrane are at ambient pressure. The rapid modulation of the ambient pressure across a series of values is used to identify an ambient pressure (the equalization pressure) at which the responses of both Pirani gauges are equivalent. In the 162-stage Knudsen pump, a Pirani gauge located at the outlet of the pump is used as the reference gauge.

During calibration, the ambient pressure is modulated in multiple steps. Each step provides a response from the reference gauge that is either greater or less than the responses from the gauge to be calibrated. By interpolating between responses from both gauges through successive modulation steps, the ambient pressure is found at which the Pirani gauges have equivalent responses. At this equalization pressure, the thermal

resistances of the two Pirani gauges, $\mathfrak{R}_{sg}$, are also identical, and the membranes are undeflected.

In practice, the Pirani gauge responses may vary even under identical pressure conditions because of minor manufacturing variations. This can be corrected using a linear regression method that is explained next.

4.2.4 Mapping of the Response of Pirani Gauges with Linear Regression in Dynamic Calibration

Since the responses of the measured and reference Pirani gauges are compared to each other in dynamic calibration, the response of each Pirani gauge to be calibrated is first mapped to the response of the reference Pirani gauge. This mapping, performed using linear regression, is used to account for the process-induced sample-to-sample variations in the responses. Process-induced variations in the metal width, film thickness or material properties can cause the sample-to-sample variations in R_0, $\mathfrak{R}_{th}$, or α. The use of linear regression is justified by showing that the responses with pressure change, $\Delta R/R_0$, are linearly related between gauges. Two cases for the process-induced variations are analyzed: variation in the baseline electrical resistance, R_0; and variation in the thermal resistances.
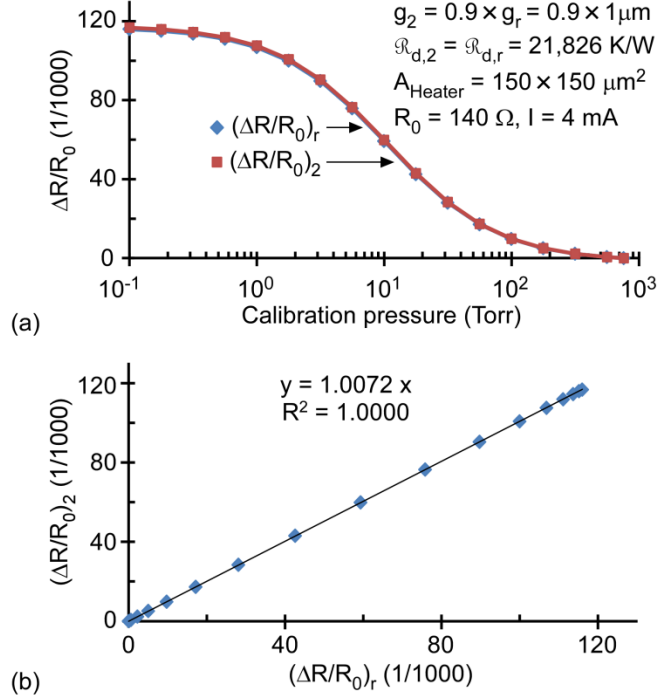
Figure 4.3: The relation between the calculated responses of two hypothetical Pirani gauges, with varying calibration pressures from 760 Torr to 0.1 Torr. (a) The responses of two hypothetical Pirani gauges as a function of calibration pressure. (b) The linear regression fit of the responses of two hypothetical Pirani gauges. Subscripts *r* and *2* indicate two Pirani gauges. The sense gap of the Pirani gauge 2, g_2, is assumed to be smaller by 10% than that of the reference Pirani gauge, g_r, due to the process-induced variation in sacrificial material for the sense gap.

4.2.4.1 Variation in R_0

The unheated resistance, R_0, varies typically from 130 Ω to 150 Ω, even within one Knudsen pump. Using Eq. (4.6), if the 2nd term on the right side is ignored for simplicity by assuming that $\mathscr{R}_{th}$ at low pressure values is much larger than $\mathscr{R}_{th,Ref}$, the ratio of $\Delta R/R_0$ between two Pirani gauges can be simply stated as the following Taylor series expansion:

$$\frac{(\Delta R/R_0)_m}{(\Delta R/R_0)_r} = \gamma\left\{1-(1-\gamma)(\Delta R/R_0)_r+(1-\gamma)^2(\Delta R/R_0)_r^2\cdots\right\} \tag{4.7}$$

where the subscript *m* indicates the Pirani gauge of interest, the subscript *r* indicates the reference Pirani gauge and γ is the proportionality constant, defined as $R_{0,m}/R_{0,r}$. By

73

substituting $(\Delta R/R)_r$ with 0.12, which is a typical maximum value in the test of Pirani gauges that are further described in Section 4.3, and γ with 0.87, which is the ratio of 130 Ω to 150 Ω, the 2nd order and 3rd order terms become 1.6% and 0.02% of the 1st order term, respectively. Because the non-linear terms are relatively small, the relation between the responses of two Pirani gauges with slightly different unheated resistances can be regarded as approximately linear.

4.2.4.2 Variation in $\mathfrak{R}_{sg}$ and $\mathfrak{R}_d$

The process-induced variations from $\mathfrak{R}_{sg}$ and $\mathfrak{R}_d$ are evaluated separately. Consider a hypothetical situation in which the Pirani sense gap is 1 μm, the area of the heater, A_{heater}, in Eq. (4.3-4.4) is 150×150 μm^2, and the value of $\mathfrak{R}_d$ is 21,826 K/W. These values are representative of the Knudsen pump example that is described in Section 4.3. This hypothetical Pirani gauge, denoted as the reference gauge, is compared with another, denoted as gauge 2, in which the sense gap is 10% smaller. As shown in Fig. 4.3, the pressure responses of both devices are almost identical, and the relationship between the fractional changes in resistance $(\Delta R/R_0)_2$ and $(\Delta R/R_0)_r$ can be represented by a straight line with a slope of 1.0072. This fit (Fig. 4.3b) has an R^2 value of 1.0000 over the pressure range of 0.1 Torr to 760 Torr, indicating that a linear regression model can be used for mapping of the response of a Pirani gauge to that of the reference Pirani gauge in the same chip. Another hypothetical case, in which the thickness of the supporting membrane is increased by 10%, thereby reducing $\mathfrak{R}_d$, is shown in Fig. 4.4. Once again, the relation between $(\Delta R/R_0)_3$ and $(\Delta R/R_0)_r$ is approximately linear, as evidenced by the R^2 value of 0.9992 and slope of 0.8918, over the pressure range of 0.1

Torr to 760 Torr. In practice, the process induced variations are even lower because the variation in sense gap and supporting membrane thickness is generally <3% rather than the variation of 10% assumed in the hypothetical examples. The modest nature of the variation is supported by the small size of the device; for the Knudsen pump, the die size is 12×15 mm^2. Therefore, the relation between nearby Pirani gauges can be assumed to be linear.
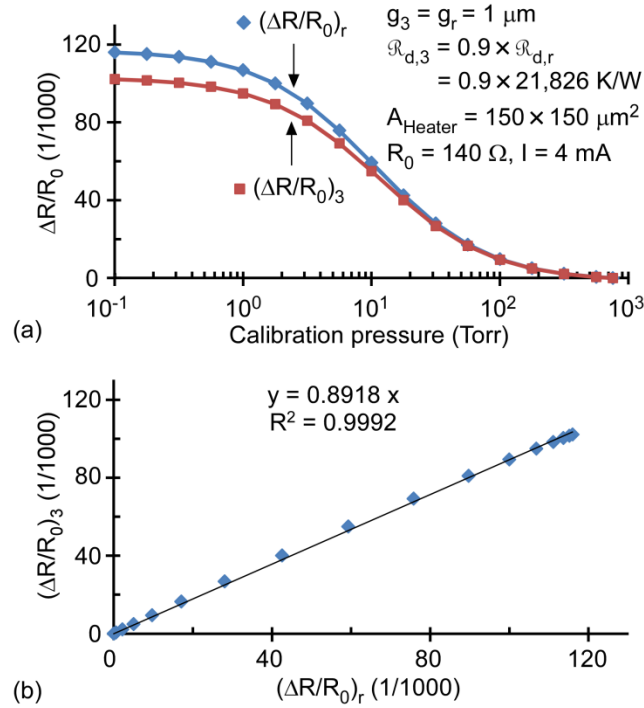
Figure 4.4: The relation between the calculated responses of two hypothetical Pirani gauges, with varying calibration pressures from 760 Torr to 0.1 Torr. (a) The responses of two hypothetical Pirani gauges as a function of calibration pressure. (b) The linear regression fit of the responses of two hypothetical Pirani gauges. Subscripts *r* and *3* indicate two Pirani gauges. The $\mathcal{R}_{d,3}$ value are assumed to be smaller by 10% than the $\mathcal{R}_{d,r}$ value, due to the process-induced variation in thickness for supporting membrane.

The linear transformation of the response of each Pirani gauge to the equivalent response of the reference pressure sensor results in the following relationship:

$$\log_{10}(P_{m,S,i}) = \log_{10}(P_{Cal,i}) + \frac{\log_{10}(P_{Cal,i}) - \log_{10}(P_{Cal,i-1})}{(\Delta R / R_0)_{r,i} - (\Delta R / R_0)_{r,i-1}} ((\Delta R / R_0)_{m,S,i} - (\Delta R / R_0)_{r,i}) \quad (4.8)$$

where $P_{m,S,i}$ is the value converted from the linearly transformed (or scaled) response of the measured Pirani gauge, $(\Delta R/R_0)_{m,S,i}$; i denotes an index value of the calibration pressure, P_{Cal}; or the fractional change in resistance, $\Delta R/R_0$; and the subscripts r and m indicate the reference Pirani gauge and the measured Pirani gauge, respectively. The residual error in pressure, ΔP_E, is defined as the difference between the scaled response and the calibration pressure:

$$\Delta P_{E,i} = P_{m,S,i} - P_{Cal,i} \quad (4.9)$$

The linear transformation of each Pirani gauge response to the equivalent response of the reference Pirani gauge is provided by:

$$(\Delta R / R_0)_{m,S,i} = S(\Delta R / R_0)_{m,i} \quad (4.10)$$

where S is the first-order coefficient (or slope).
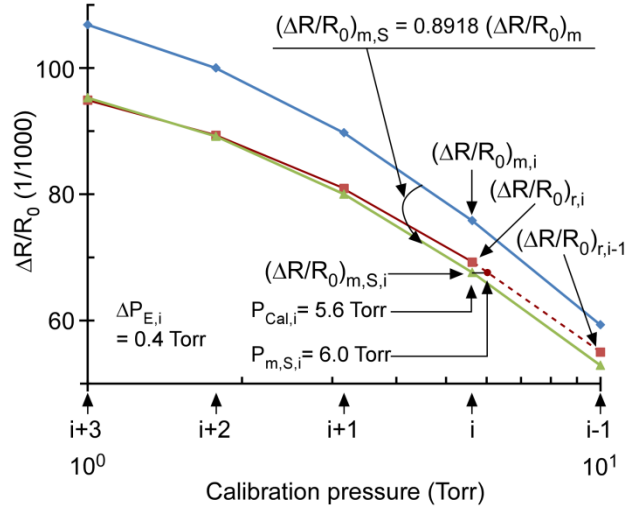
Figure 4.5: An illustrative example of the mapping of the measured Pirani gauge to the reference Pirani gauge and the resulting residual error in pressure, using the hypothetical data in Fig. 4.4. The values of $(\Delta R/R_0)_m$ are linearly transformed to those of $(\Delta R/R_0)_{m,S}$ with linear regression model, as indicated in this graph, and the discrepancies with $(\Delta R/R_0)_r$ are translated into the residual errors in pressure, ΔP_E.

A practical example of the linear transformation of the response of a measured Pirani gauge to the equivalent response of the reference pressure gauge is illustrated as follows. Using the hypothetical data in Fig. 4.4, the un-scaled value of 5.6 Torr (i.e., $P_{Cal,i}$ of 5.6 Torr) is transformed to the scaled value of 6.0 Torr. As illustrated in Fig. 4.5, first, the $(\Delta R/R_0)_{m,i}$ value is linearly transformed to the $(\Delta R/R_0)_{m,S,i}$ value, using Eq. (4.11). Second, the $(\Delta R/R_0)_{m,S,i}$ value is converted to the $P_{m,S,i}$ value of 6.0 Torr by finding the horizontal intercept from $(\Delta R/R_0)_{m,S,i}$ to the interpolated curve for $(\Delta R/R_0)_r$, which is represented by the red dotted line in Fig. 4.5. This process, which is the graphical equivalent of solving Eq. (4.8), results in a residual error in pressure of 0.4 Torr.

4.3 Experimental Validation

The experimental validation of the calibration procedure is performed using a 162-stage Knudsen pump, where each stage generates gas flow in the direction of the thermal gradient (Fig. 4.6a) (See also Fig. 3.4). Pirani gauges are located at stages 1, 54, 99, and 162, and are named P1, P54, P99, and P162, respectively. Stage 1 is the outlet to the ambient, whereas stage 162 is a blind cavity that is to be evacuated.
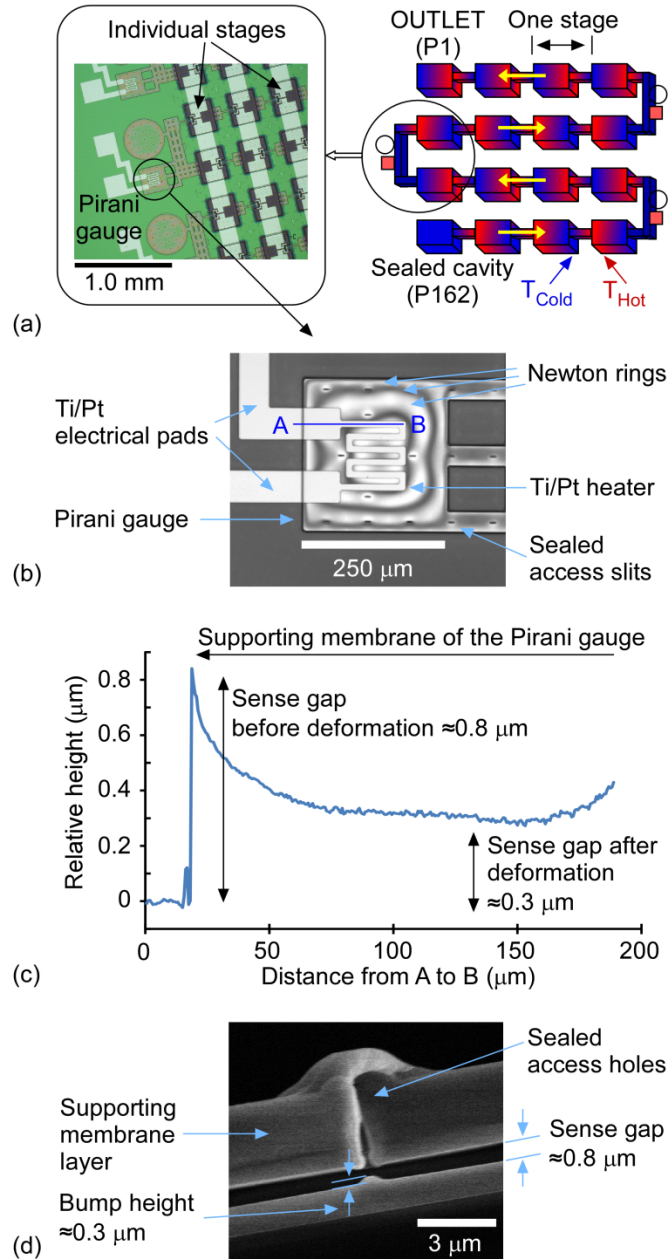
Figure 4.6: Sample Knudsen pump and mechanical deformation of the Pirani gauge when the interior pressure is at high vacuum relative to ambient atmospheric pressure. (a) Schematics of multistage Knudsen pump; the left inset shows microscopic photograph of one side of the serpentine multistage structure. (b) Interferogram of a Pirani gauge obtained with a 407 nm laser. (c) Relative height of the exterior surface of the Pirani gauge by interferogram program. (d) Cross-sectional SEM photograph of a sealed access hole, when the interior is now at an identical pressure with the exterior.

The deflection of the supporting membrane of the Pirani gauge is analyzed by interferograms and scanning electron micrographs (Fig. 4.6). The deflection due to the

pressure difference of more than 700 Torr between the interior and the exterior is measured by a laser interferogram (Olympus^{TM} LEXT OLS3100) (Fig. 4.6b, 4.6c). Each Newton ring (bright or dark) represents about a half integer multiple of the laser wavelength (407 nm) in the sense gap (Fig. 4.6b) [Al06, Smi07]. From the relative height of the supporting membrane surface (Fig. 4.6c), the sense gap in the B region along the A-B line above Ti/Pt metal, after deflection caused by the pressure difference, is ≈0.3 μm. This is smaller than the sense gap of ≈0.8 μm, as shown in the SEM image of the supporting membrane, in the absence of a pressure difference (Fig. 4.6d). However, this sense gap of ≈0.3 μm is nearly identical to the bump that is created by the sealing layers as they are deposited into the access slits for the sacrificial dry etchant (Fig. 4.6d) (See also Fig. 2.3 and 3.6). This bump height suggests that the deflected membrane may contact the bump surface so the equivalent thermal resistance of the sense gap could further decrease. The fractional change in resistance caused by the mechanical strain of Ti/Pt metal is negligible in the deflected membrane.

4.3.1 Set-Up and Methodology

To select the operating conditions for the Pirani gauges, a preliminary evaluation is performed by varying calibration pressure and input current, using the static calibration method (Fig. 4.7). A constant current is provided to the Pirani gauges that are connected in series. From the change in $\Delta R/R_0$ with varying calibration pressure, at each current level, it is evident that the responses from 1 Torr to 100 Torr show greater sensitivity than those from 100 Torr to 760 Torr and from 0.1 Torr to 1 Torr. As the current varies from 2.83 mA to 4.90 mA, the $\Delta R/R_0$ and the ΔT are increased, as anticipated by Eq. (4.6).

The fractional change in resistance is approximately proportional to the square of the current. The maximum values of ΔT are 23.4 K for 2.8 mA, 49.9 K for 4.0 mA, and 68.2 K for 4.9 mA. Of these, the current of 4.9 mA provides the highest sensitivity in $\Delta R/R_0$, as can be seen in the plots (Fig. 4.7). However, in order to reduce the likelihood for thermal aging of the resistor, drift in resistance and non-linearity in fractional change in resistance, an operating current of 4 mA is selected. It is applied to the nominal R_0 of 140 Ω in this test and also to all of the Pirani gauges in the Knudsen pump sample, where the value of R_0 varies from 130 Ω to 150 Ω on a wafer-to-wafer basis.
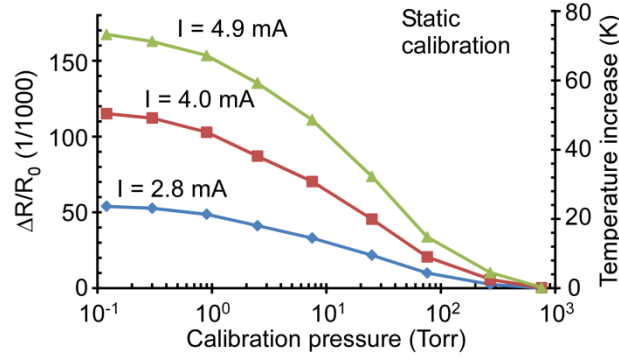
Figure 4.7: A preliminary evaluation of a Pirani gauge. The fractional change in resistance and temperature change of a Pirani gauge are investigated as a function of calibration pressure, with increasing current values. The squared value of the current is approximately linear with $\Delta R/R_0$, as shown in Eq. (4.6). These values are obtained by static calibration.

Using the measured α value of 2,314 ppm/K and the variation in $\Delta R/R$ with the current of 4 mA, the $\mathfrak{R}_d$ value can be estimated for the low pressure limit, with Eq. (4.6). The second term on the right side of Eq. (4.6) is characterized as 10.3×10^{-3}, providing the $\mathfrak{R}_{th,Ref}$ value of 2,002 K/W. At 0.12 Torr, the $\mathfrak{R}_d$ value is nearly identical to the $\mathfrak{R}_{th}$ value of 21,826 K/W.

4.3.2 Pressure Estimation Using Static Calibration

In static calibration (Fig. 4.8a), the responses of each Pirani gauge with variations in calibration pressure are obtained, while the Knudsen pump is off (Fig. 4.8a). For exposing both sides of the supporting membrane in each Pirani gauge to equal ambient pressures, perforations are created close to the calibrated Pirani gauges, thereby enabling the interior pressure to be controlled by the calibration pump. It is that this calibration process is destructive because of the need of these perforations, and is consequently performed after the pumps are tested. As a practical matter, pumps within close proximity are expected to perform similarly, so only a limited number must be sacrificed for calibration.

Each pressure level (Fig. 4.8b) is calculated by interpolation between the neighboring points in static calibration data (Fig. 4.8a). Interpolation is performed using straight lines with calibration pressure in log-scale and $\Delta R/R_0$ in linear-scale. To reduce the error from the straight line between the neighboring points, the calibration pressures are closely spaced at half-decade intervals. The $\Delta R/R_0$ value of 4.3×10^{-3} of P54 is mapped to 430 Torr, the $\Delta R/R_0$ value of 28.2×10^{-3} of P99 to 53 Torr, and the $\Delta R/R_0$ value of 37.7×10^{-3} of P162 to 30 Torr. The pressure at P1 is 760 Torr because it is open to atmospheric ambient pressure. The method for evaluating the error bars is described in Section 4.3.4. As explained in Section 4.1, the true pressures are lower than the inaccurately estimated pressures obtained by this method.
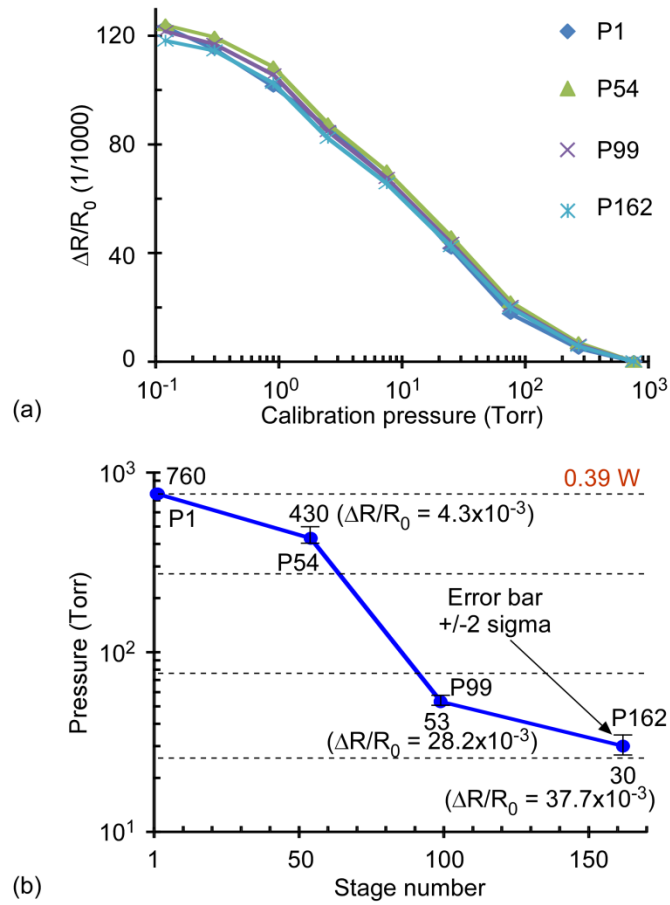
Figure 4.8: The pressure levels of the Pirani gauges, using the static calibration of Pirani gauges of P1, P54, P99, and P162. (a) Each response of the Pirani gauges with varying calibration pressure. (b) Each estimated pressure of the Pirani gauges with the Knudsen pump operation. The horizontal dotted lines in (b) are the static calibration pressures which are used for interpolations.

4.3.3 Pressure Estimation Using Dynamic Calibration

The following four sequential steps describe how to process the measured data and estimate the attained pressures (Fig. 4.2b).
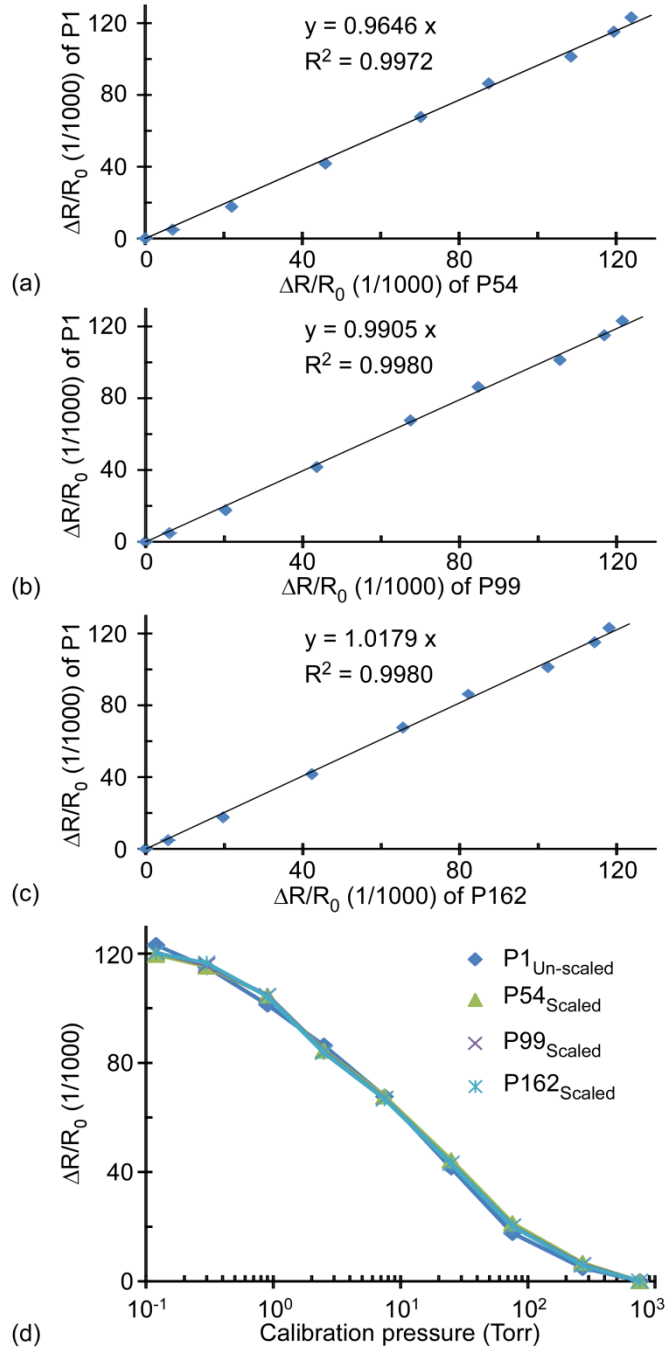
Figure 4.9: Mapping of the Pirani gauges, P54, P99, and P162 to the reference Pirani gauge, P1, using linear regression. (a) The linear regression fit between P1 and P54. (b) The linear regression fit between P1 and P99. (c) The linear regression fit between P1 and P162. (d) Responses of each Pirani gauge after mapping, using regression equations, as indicated in (a-c).

The first step is to obtain the regression equations for mapping the responses of

each measured Pirani gauge to that of the reference Pirani gauge (Fig. 4.9a-4.9c), using

the static calibration data (Fig. 4.8a). Since P1 is the reference Pirani gauge, each of the responses of P54, P99, and P162 is scaled to that of P1. For each regression equation with the offset of zero, the first-order coefficients are 0.9646 for P54, 0.9905 for P99, and 1.0179 for P162. The values of R^2 for mapping P54, P99, and P162 to P1 are 0.9972, 0.9980, and 0.9980, respectively. As can be seen in the plots of the scaled responses of the measured Pirani gauges and the reference Pirani gauges (Fig. 4.9d), the deviations of the fractional changes in resistance with pressure are smaller than those in Fig. 4.8a.
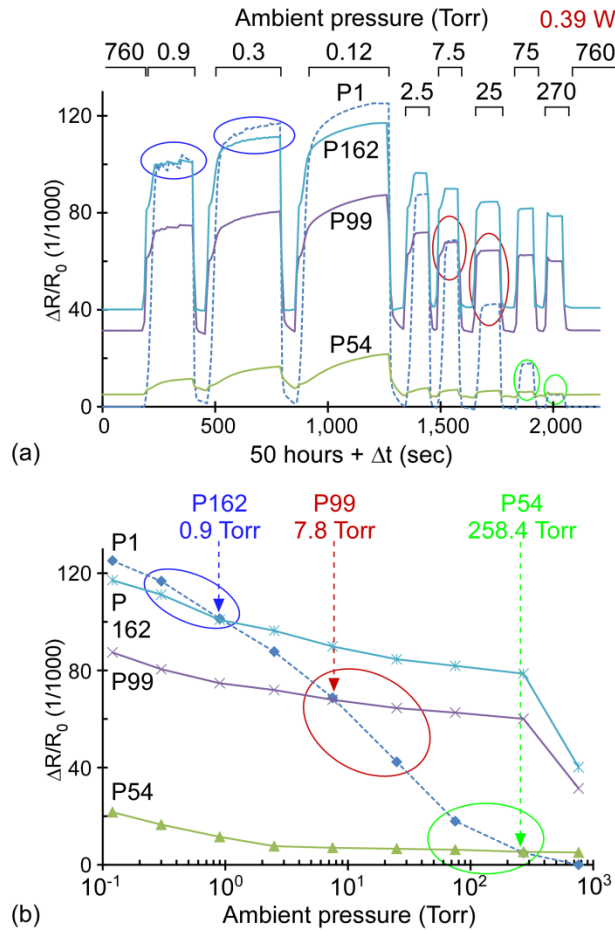
Figure 4.10: Determining the equalization pressures of each Pirani gauge. (a) Scaled dynamic calibration plots; the ambient pressure is rapidly varied from 760 Torr to 0.1 Torr at about half-decade pressure intervals. (b) The equalization pressures are determined, as indicated by the dotted arrow lines between two bounding values of ambient pressure. Blue circles are for P162, red circles for P99, and green circles for P54.

The second step is to apply the linear regression equations obtained by the first step to the measured dynamic calibration plots. The dynamic calibration plots provide equalization pressures at half-decade intervals from 0.12 Torr to 760 Torr. To ensure that the pressure levels in each stage are in a steady state (Fig. 4.10a), the dynamic calibration is performed after 50 hours of Knudsen pump operation. The steady state response of P1 before modulation is zero in the dynamic calibration plots. During modulation, the temporal response of P1 represents the varying ambient pressure, whereas the temporal responses of P54, P99, and P162 represent both interior pressure and exterior varying ambient pressure. The response of P1 reflects the ambient pressure, and matches the static calibration. However, the plots of P54, P99, and P162 in Fig. 4.10a have been corrected, using the regression equations.

The third step is to determine the pressures within the Pirani gauges. As the ambient is rapidly modulated from the original value (which is 760 Torr in this case), the response of P1 is compared to those of P54, P99, and P162. Measurements are taken at the end of each modulation in order to provide the maximum settling time for each reading. The modulated ambient value at which P1 indicates pressure that is higher than the mapped response of any Pirani gauge represents the lower bound of the pressure at that gauge. Similarly, the modulated ambient value at which P1 is lower than the mapped value of any gauge represents the upper bound of the pressure at that gauge (Fig. 4.10a). The equalization pressure of each gauge is calculated using interpolation between these two bounding pressures (Fig. 4.10b). For P162, the upper bound pressure, at which the response of P1 is lower than that of P162, is 0.9 Torr. The lower bound pressure, at which the response of P1 is higher than that of P162, is 0.3 Torr. For P99, two bounding

pressures are 25 Torr and 7.5 Torr. For P54, two bounding pressures are 270 Torr and 75 Torr. The equalization pressures are $\approx$0.9 Torr for P162, $\approx$7.8 Torr for P99, and $\approx$258.4 Torr for P54.
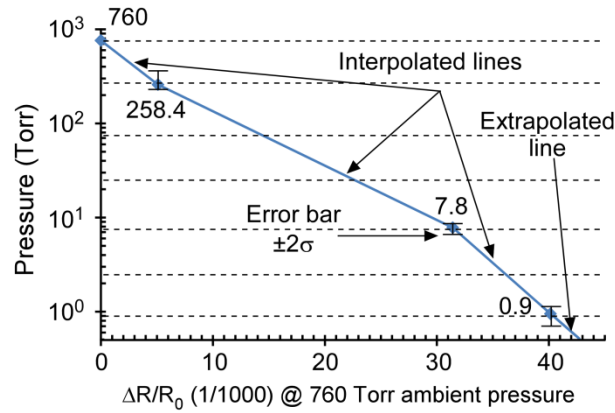
Figure 4.11: The generalized relation of pressure as a function of $\Delta R/R_0$ when the ambient pressure is 760 Torr, obtained by dynamic calibration. The isolated points indicate the equalization pressures (Fig. 4.10b). The solid line indicates interpolations and extrapolations between the neighboring, isolated points, for approximately determining the pressures from any measured $\Delta R/R_0$. The horizontal dotted lines indicate the modulated ambient pressures for dynamic calibration.

The final, fourth step is to generalize the correlation of pressure and fractional change in resistance at the specific ambient pressure. Using the equalization pressures determined in the third step, any value of $\Delta R/R_0$ provided by the Pirani gauges can be mapped to a calibrated value of interior pressure. This value is determined using the closest points from the set 760 Torr, 258.4 Torr, 7.8 Torr, and 0.9 Torr and the corresponding $\Delta R/R_0$ of 0, 4.3×10^{-3}, 28.2×10^{-3}, and 37.7×10^{-3}, respectively (Fig. 4.11). However, this calibration is valid only for an exterior ambient pressure of 760 Torr. For other values of ambient pressure, the rapid pressure modulation and the second, third, and fourth steps of the procedure must be repeated (Fig. 2b). The method for evaluating the error bars is described in Section 4.3.4.

Figure 4.12 compares the fractional change in resistance as a function of pressure obtained by dynamic calibration (Fig. 4.11) to that obtained by static calibration (Fig. 4.8a). As the pressure is reduced, dynamic calibration shows smaller fractional changes in resistance than static calibration.
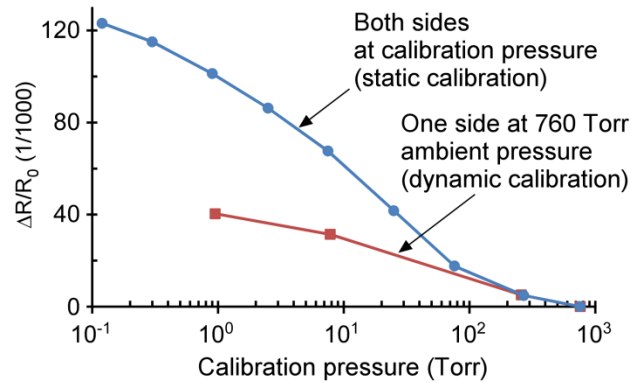
Figure 4.12: The difference in fractional changes in the resistance of a Pirani gauge from static and dynamic calibration methods. The solid circles are obtained by static calibration. The solid squares are obtained by dynamic calibration for 760 Torr ambient.

4.3.4 Typical Errors in Calibration

The two dominant sources of error are analyzed: the repeatability of Pirani gauge sensors and the mapping of the Pirani responses using linear regression. These are used to calculate the standard deviation, σ, of the calibrated response. The uncertainty is represented by an error bar that indicates a range of $\pm 2\sigma$ (indicating a total span of 4σ), which correspond to a 95.4% confidence interval for normal (Gaussian) distribution [Kre06].

The repeatability of the sensor is a source of error for the static and dynamic calibration. The output may vary because of the measurement tools or human factors. Error bars that identify the possible readout fluctuation can be defined for each measurement of the Pirani gauge. To evaluate this uncertainty, one Pirani gauge is repeatedly measured, using identical measurement tools, test chamber, and operator

method, over a period of time. Using the repeated measurements of the $\Delta R/R_0$, the mean values and the error bars for $\pm 2\sigma$ are plotted for a range of calibration pressures (Fig. 4.13a). In a preliminary sample set of 6 measurements, the highest value of the standard deviation, σ, was 1.6×10^{-3}, at 0.12 Torr. The $\pm 2\sigma$ error bars in $\Delta R/R_0$ are converted into the error bars in pressure by using the two closest neighboring calibration points, using Eq. (4.8). Here, $(\Delta R/R_0)_{m,S}$ is replaced by $[(\Delta R/R_0)_r \pm 2\sigma]$ and $P_{m,S}$ by the resulting pressure. The resulting residual errors in pressure, ΔP_E, for the calibration pressure of ≈30 Torr are -3 Torr and +3 Torr, which are indicated as the error bar for the pressure at P162, obtained by static calibration (Fig. 4.8b).

Another source of error is the linear regression performed for dynamic calibration. To estimate this, the responses of multiple Pirani gauges are collected, using identical calibration pressures, measurement tools, test chamber, and operator method. Then, each response is mapped by linear regression to the mean response at every pressure – i.e., the mean is assumed to be the reference response. For a preliminary data set with 12 Pirani gauges, the first order coefficients, as shown in Eq. (4.10), were 0.9482, 0.9152, 0.9397, 0.9657, 0.9482, 0.9152, 0.9657, 1.1412, 1.0766, 1.0795, 1.09121, and 1.0818. The mean and the error bars for $\pm 2\sigma$ are plotted, with varying calibration pressure (Fig. 4.13b). The highest σ value in $\Delta R/R_0$ was 1.4×10^{-3} at the pressure of 0.9 Torr. This error is assumed to be representative of the errors due to linear regression in dynamic calibration.
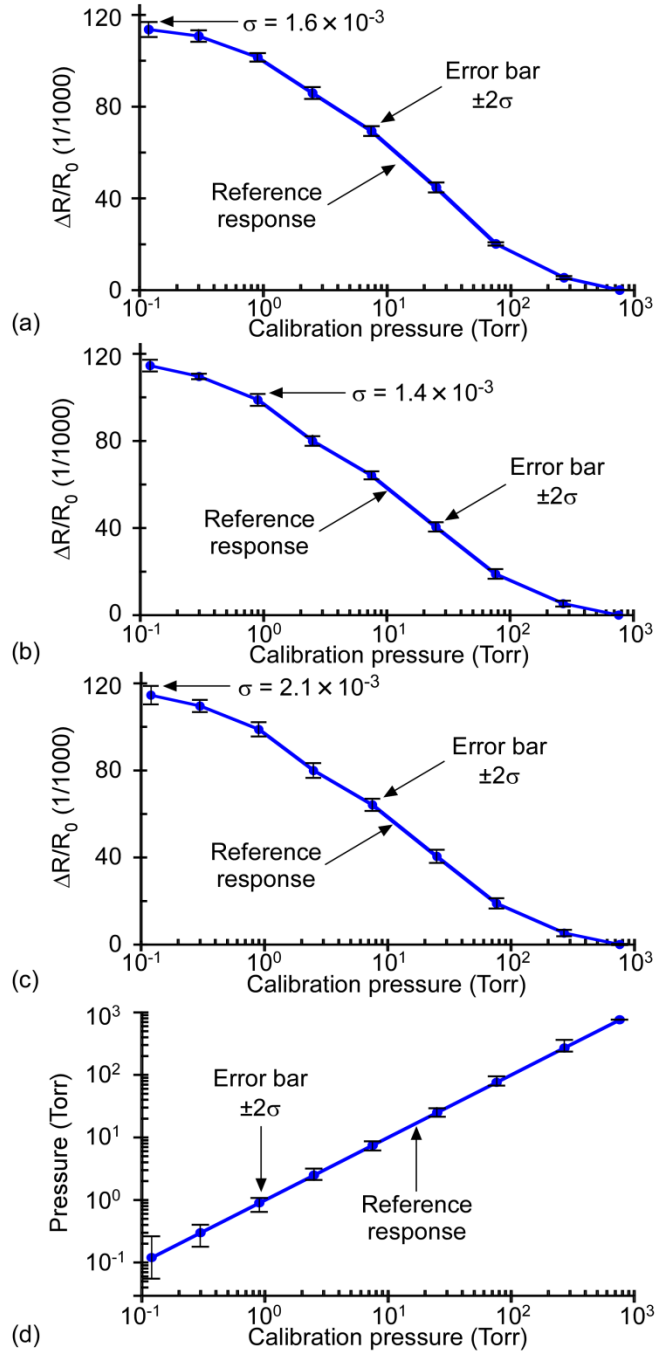
Figure 4.13: Error source analysis for static and dynamic calibration methods. (a) From the repeatability of a Pirani gauge, in $\Delta R/R_0$ as a function of calibration pressure, measured with static calibration method. (b) From the mapping of the measured Pirani gauge to the reference Pirani gauge with linear regression. (c) The combined errors from (a) and (b). (d) The combined errors from (a) and (b) are converted to residual errors in pressure at each calibration pressure. In (a-c), only the highest values in σ are indicated.

The two sources of error are uncorrelated, so the total error, E_{1+2}, is [Ell00, Bel99]:

$$E_{1+2} = \sqrt{E_1^2 + E_2^2} \qquad (4.11)$$

where the subscripts 1 and 2 indicate two independent variables, E_1 is the standard deviation from the repeatability of responses of each Pirani gauge and E_2 is the standard deviation from the mapping of the response of the measured Pirani gauge to that of the reference gauge with linear regression. The total error is plotted in Fig. 4.13c. The highest value of the total error in $\Delta R/R_0$ is 2.1×10^{-3} at the calibration pressure of 0.12 Torr. The resulting residual errors in pressure, ΔP_E, for the calibration pressure of 0.9 Torr are -0.2 Torr and +0.2 Torr (Fig. 4.13d); these errors are indicated as the error bar for the pressure at P162, obtained by dynamic calibration (Fig. 4.11).

4.4 Discussion and Conclusions

The Knudsen pump example demonstrates that dynamic calibration can be significantly more accurate than conventional static calibration for certain type of devices. The pressure of P162 was 30 Torr by static calibration (Fig. 4.8b) and 0.9 Torr by dynamic calibration (Fig. 4.10b). As the interior pressure decreases, the response of a Pirani gauge that has one side exposed to ambient pressure becomes smaller than that of a Pirani gauge for which both interior and exterior are both at the same pressure (Fig. 4.12). A significant contribution to this effect is due to the mechanical deflection of the supporting membrane that contacts the bump surface at the sense gap (Fig. 4.6d). As a result, the smaller fractional change in resistance is mapped to higher pressure than the true value. In dynamic calibration, this error is removed by modulating the ambient pressure around the interior pressure of the Pirani gauge until the responses of both

90

gauges are equal indicating that the supporting membrane is detached from the bump surface.

It is evident from Fig. 4.10a that, when the ambient pressure is returned to 760 Torr after modulation, each $\Delta R/R_0$ returns back to its initial value. The pressure in the sense gaps in stages 55, 99, and 162, located upstream and far from the outlet do not respond to rapid variations in ambient pressure. The effects of the rapid variations in ambient pressure on those interior pressures can be ignored. Therefore, the equalization pressures in dynamic calibration are valid.

One transient response shown in the dynamic calibration plots (Fig. 4.10a) is a large increase in the responses of P99 and P162 when the ambient pressure changes from 760 Torr to 0.9 Torr. This is caused by the supporting membrane detaching from the bump surface resulting in a large change in $\mathfrak{R}_{sg}$. After detaching, different ambient pressures, such as 0.9 Torr, 0.3 Torr, and 0.12 Torr, are expected to cause negligible changes in the responses. This is because the pressures in the sense gaps are unchanged and $\mathfrak{R}_{sg}$ is insensitive to changes in membrane deflection at low pressure in the range 1-10 Torr, where $\lambda \gg g$, as anticipated by Eq. (4.4).

After detaching, the responses of P99 and P162 increase in two ways: 1) from each successive step, 760 Torr to 0.9 Torr, 760 Torr to 0.3 Torr, and 760 Torr to 0.12 Torr, the responses show an initial noticeable increase; and 2) within each modulation step the responses show a gradual increase (Fig. 4.10a). These responses after detaching, which are unexpected from changes in membrane deflection, result from the combined effects of the Knudsen pump heater, which slowly increases pump chip temperature as ambient pressure decreases, and the time it takes for the calibration pump, $\approx$23 cfm, to

achieve full vacuum (>17 s). However, these two effects affect all Pirani gauges equally without changing the found equalization pressures. Therefore, the equalization pressures in dynamic calibration are valid.

The error contributed by measurement uncertainty and the regression model is relatively small near atmospheric pressure, but significant at low pressures, as shown in Fig. 4.13d. The small sensitivity of $\Delta R/R_0$ at low pressure relative to that at higher pressure amplifies the impact of errors in $\Delta R/R_0$, resulting in larger equivalent errors in pressure, especially at 0.12 Torr. These errors may be reduced by increasing the input current to the Pirani gauges at low pressures.

In summary, dynamic calibration has been investigated for the pressure estimation of the pressure modulating devices, with the integrated Pirani gauges. It provides superior accuracy to static calibration by accommodating the impact of the exterior ambient pressure, which is different from the interior pressures of the evacuated stages. The dynamic method is used to determine the interior pressure by rapid modulation in ambient pressure. Using four sequential steps for processing the calibration data, more accurate estimates for the interior pressures are obtained. For the sample Knudsen pump that is experimentally evaluated, the upstream pressure is 0.9 Torr, with a 95.4% confidence interval from 0.7 Torr to 1.1 Torr, assuming normal (Gaussian) distribution. In comparison, static calibration suggests the pressure of 30 Torr, with a 95.4% confidence interval from 27 Torr to 33 Torr.

CHAPTER 5

VACUUM SEALING
USING ATOMIC LAYER DEPOSITION OF AL$_2$O$_3$ AT 250°C

This chapter describes the use of low-temperature atomic layer deposition (ALD) of Al$_2$O$_3$, for vacuum seals in wafer-level vacuum packaging and other applications. This technique is used for vacuum sealing of flow channels in the 48-stage and 162-stage Knudsen pumps. The pin-hole free and highly conformal coverage provided by ALD Al$_2$O$_3$ is shown to seal circular micromachined cavities in this chapter. The cavities are 0.8 μm in height, 400 μm in diameter, and are capped by porous plasma-enhanced chemical vapor deposited (PECVD) dielectrics that form a membrane. The ALD Al$_2$O$_3$ film, of thickness ≈0.2 μm, is deposited at a temperature of 250°C on this membrane. The retention of vacuum is indicated by the deflection of the membrane. Lifetime tests extending out to 19 months are reported.

5.1 Introduction

Vacuum sealing is indispensable for many microsystems applications and is commonly used for performance enhancement and device protection. For example, vacuum encapsulation is required for the high quality (high-Q) factor operation for micro resonant sensors and actuators, such as gyroscopes [Liu09], pressure sensors [Ike90], and timers [Bee12]. Miniaturized analytical instruments, such as gas chromatographs [Liu11] and mass spectrometers [Hau07], together with micro pumps [Kim08] for the instruments,

also need pressure-controlled gas flow in microfluidic channels or cavities. In addition, the techniques for achieving and maintaining vacuum seals can be extended to humidity-sensitive electrostatic devices [Kes98, Wu06].

For vacuum sealing, wafer-scale encapsulation [Esa08] is favored over die-level packaging because it offers greater process compatibility, higher throughput, and smaller form factor. In particular, thin-film sealing [Ike90, Guc84, Guc89] is a natural extension of surface micromachining processes and is preferred to wafer-bonding in many contexts. Typically, an encapsulating membrane is deposited over a patterned sacrificial layer. Holes in this membrane provide access for a gas phase or liquid phase etchant that removes the sacrificial layer. Another thin film is then used to seal the access holes in the membrane. An alternative method is to bond another wafer above the device separated by a spacer layer. Although this circumvents the deposition of a sacrificial layer, it can compromise size and cost, and complicate lead transfer.

Thin-film sealing has been studied for wafer-level vacuum packaging for micromachined resonators [Liu09, Ike90, Bee12, Lin98, Tsu01, Can06, Chi06, Li08, Yon10]. It has also been explored for sealing micro-fluidic channels in neural probes [Wis08]. These efforts routinely used common dielectrics, such as silicon oxide or silicon nitride. The films were, in many cases, deposited at temperatures >800°C, using low-pressure chemical vapor deposition (LPCVD). For lower temperature deposition, plasma-enhanced chemical vapor deposited (PECVD) nitride layers at 450°C [Li08] and 350°C [Liu99] were used. However, with the decrease in deposition temperature, the deposited films were as thick as 2.7 μm and 4.6 μm, respectively. Sputtered AlCu of 1.5 μm thickness, deposited at 350°C, was also reported for sealing porous poly-SiGe

capping membranes [Guo12]. Achieving a vacuum seal with a thin film that is deposited at modest temperatures remains a challenging prospect for conventional deposition methods because of the existence of pinholes in deposited films, as well as limitations to the conformality of the deposition.

Atomic layer deposition (ALD) of films is appealing for low temperature sealing. Sequential, self-limiting surface reactions enable the coverage of topographic variations, with film thickness precisely controlled by the number of deposition cycles [Puu05, Geo10]. The last decade has witnessed increasing efforts to exploit ALD for various sealing purposes that prepare the stage for vacuum sealing. Plasma-enhanced ALD SiO_2 was used to seal mesoporous silica for interlevel dielectrics [Jia06], and ALD TiO_2 was used to seal nanochannel trenches [Nam10]. The deposition temperature of ALD ranges from 25°C to 300°C [Puu05, Geo10, Jia06, Nam10, Wil05, Car09, Lin10, Lee11, Zha11, Keu12, Dic12, Car12]. ALD Al_2O_3 was deposited on Kapton, other polymers, and organic light emitting diodes (OLED) at temperatures as low as 25-125°C to provide a gas diffusion barrier [Wil05, Car09, Lin10, Keu12, Dic12, Car12]. By 2011, it was recognized that ALD layers may be useful for wafer-level vacuum packaging of nano- and micro-electromechanical systems (N/MEMS) [Lee11]. Although the deposition of ALD Al_2O_3 on polymer presented major leakage concerns [Zha11], in 2012, the pressure-induced deflection of an ALD Al_2O_3 membrane with a thickness of only 2.8 nm was reported, indicating that the integrity of the film is sufficient to withstand a pressure differential [Wan12]. In addition, it was reported that a thin film barrier of 20-40 nm ALD Al_2O_3 suppresses the water vapor transmission rate to be on the order of 10^{-6} gm^{-2}day^{-1} [Keu12], which is four orders of magnitude smaller than 100 nm PECVD SiO_x

[Coc12]. These results suggest that ALD Al_2O_3 could be used potentially for vacuum sealing [Keu12, Wan12]. In this context, this chapter describes an ALD Al_2O_3 low-temperature thin-film vacuum seal for wafer-level vacuum packaging. Section 5.2 describes the test structure. Section 5.3 details the fabrication process. The test methods, results and discussion are presented in Section 5.4. Concluding remarks are presented in Section 5.5.

5.2 Test Structure
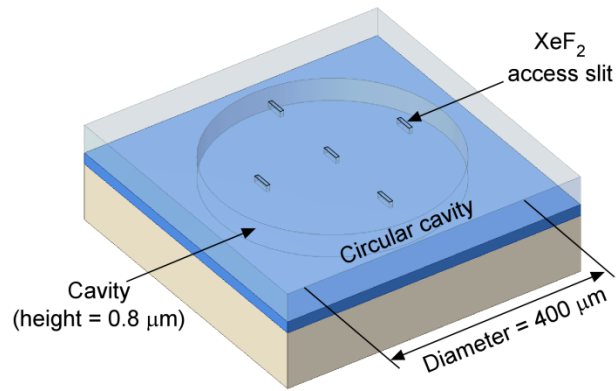
Figure 5.1: Perspective view of a microfabricated test structure with a circular cavity for testing the ALD Al_2O_3 vacuum seal.

To determine the vacuum sealing capability of the ALD Al_2O_3 film, a microfabricated test structure is used in this study (Fig. 5.1). It is composed of a cavity of height ≈ 0.8 µm and an encapsulating membrane. The membrane is deflected by the pressure difference between the interior cavity and the ambient atmosphere. Interference patterns, which are caused by the light rays reflected from the upper and lower surfaces of the cavity, present alternating dark and bright fringes as a function of membrane deflection [Al06]. For the deflection to be easily observable, the diameter of the membrane is as large as 400 µm. The membrane deflections for 160 Torr, 460 Torr, and

760 Torr are estimated using analytical equations in [Gia06] (Fig. 5.2). These estimates take into account the residual stresses of a membrane (which are described in Section 5.3). Although the cavity height is 0.8 µm, deflection is limited by bumps on the lower surface of the cavity (described in Section 5.4). Most fringes are expected to be near the edge; the deflection-limited center region is almost flat. The deflection limit is encountered when the cavity pressure is lower than atmospheric ambient pressure by hundreds of Torr; this provides an indication of the retention of vacuum.
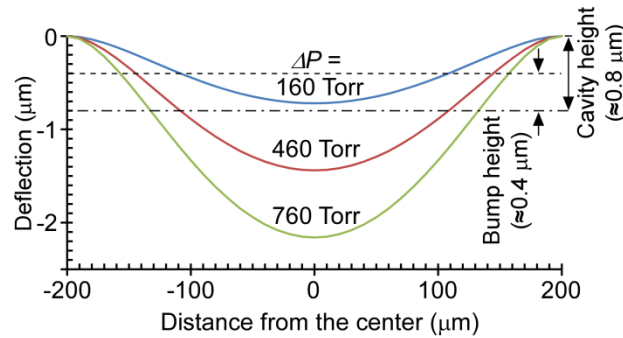
Figure 5.2: Analytical calculation of the deflection of a circular membrane as a function of pressure difference, ΔP. The pressure difference increases from 160 Torr to 760 Torr by 300 Torr. The membrane deflection is limited by bumps on the cavity floor, the heights of which are indicated by the horizontal dotted line.

5.3 Fabrication Process

The test structure is fabricated by a surface micromachining process, using two masking steps (Fig. 5.3). The microfabrication process begins with the LPCVD of the first stress-relieved, dielectric oxide-nitride-oxide (ONO-1) stack on a silicon wafer (Fig. 5.3a). The ONO-1 stack is composed of one nitride layer, thickness ≈0.2 µm and residual stress ≈1.2 GPa, sandwiched by two oxide layers, thickness ≈0.5 µm and residual stress ≈-185 MPa. The oxide and nitride layers are deposited at approximately 910°C and 800°C, respectively. This is followed by the deposition of the sacrificial LPCVD polycrystalline silicon (polySi) layer, of thickness ≈0.8 µm, at 585°C, which is

subsequently patterned (Fig. 5.3a). The cavity height (Fig. 5.1a) is determined by the thickness of the polySi layer. Another LPCVD ONO-2 stack, which is identical to ONO-1, is deposited over the sacrificial polySi layer (Fig. 5.3b). Slits, 2×10 μm^2, are patterned in ONO-2 to provide access for the sacrificial etch. The sacrificial polySi layer is etched away by XeF_2 dry gas (Fig. 5.3b). Following this, a third stack, ONO-3, is deposited to seal the access slits. The ONO-3 stack, which is deposited by PECVD at 380°C (Fig. 5.3c), is composed of one nitride layer, thickness ≈0.5 μm and residual stress ≈300 MPa, sandwiched by two oxide layers, thickness ≈0.7 μm and residual stress $\approx$-50 MPa. Even though the PECVD step is performed at ≈1 Torr pressure, the cavities return to atmospheric pressure because the deposited ONO-3 stack is porous. Finally, an ALD Al_2O_3 layer, with thickness ≈0.2 μm and residual stress ≈304 MPa, is deposited at ≈300 mTorr to seal the pores inherent in ONO-3 (as shown in the inset of Fig.3c).
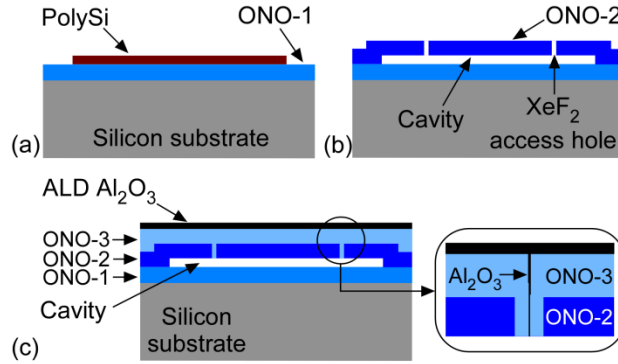
Figure 5.3: Surface micromachining uses two masking steps. (a) Deposition of LPCVD ONO-1, and deposition and patterning (Mask 1) of sacrificial LPCVD polySi. (b) Deposition and patterning (Mask 2) of LPCVD ONO-2, and XeF_2 gas dry etching of polySi. (c) Deposition of PECVD ONO-3 and ALD Al_2O_3. The inset in (c) is a magnified view of a sealed access slit.

The ALD Al_2O_3 film is deposited using Oxford[TM] Instruments OpAL [Oxf13]. Two process options are available: a thermal ALD process performed at 250°C, which uses water vapor; and another ALD process that uses O_2 plasma. The former is selected

because it provides better conformal coverage of topological variations [Puu05]. For this, the base pressure of the ALD process chamber is ≤ 25 mTorr and the chamber pressure during growth cycles is maintained at ≈ 300 mTorr using Ar gas flow. In each cycle for the growth of one atomic layer, the ALD recipe alternates the pulsing of a precursor gas, $Al(CH_3)_3$, and water vapor, H_2O, to create metal oxides [Puu05] (Fig. 5.4). The overall reaction chemistry in a cycle is described by two half reactions [Puu05, Geo10]:

$$AlOH^* + Al(CH_3)_3 \rightarrow AlOAl(CH_3)_2{}^* + CH_4 \qquad (5.1)$$

$$AlCH_3{}^* + H_2O \rightarrow AlOH^* + CH_4 \qquad (5.2)$$
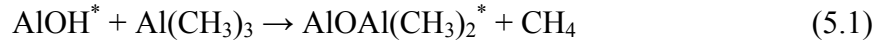
where the asterisk denotes the surface species. The first step in a cycle is to provide a dose of the precursor to the processing wafer for 20 ms for one layer of $Al(CH_3)_2$, as shown in Eq. (5.1) (Fig. 5.4a). Then, the gaseous by-products or unreacted precursors are purged for 4 s (Fig. 5.4a). The second step is to provide a dose of the H_2O vapor for 20 ms for the creation of one OH layer on the $Al(CH_3)_2$ layer, as shown in Eq. (5.2) (Fig. 5.4b). Again, the gaseous by-products or unreacted H_2O are purged for 4 s (Fig. 5.4b). In the first step of the next cycle (Fig. 5.4a), the Al atom within the precursor binds to the O atom by displacing the H atom in the surface adsorbed OH group, as indicated in Eq. (5.1). Locations where the Al-O attachment did not occur in the first cycle are covered in the second cycle [Puu05].

By repeating multiple cycles, the pores in PECVD ONO-3 are sealed with multilayered Al_2O_3 (Fig. 5.4c). However, it is notable that during this process the precursor or H_2O gas molecules might be trapped in the cavity. The total number of cycles for ≈ 0.2 μm-thickness Al_2O_3 is 1727, in which the growth per cycle (GPC) is ≈ 0.12 nm; each cycle takes ≈ 15 s.
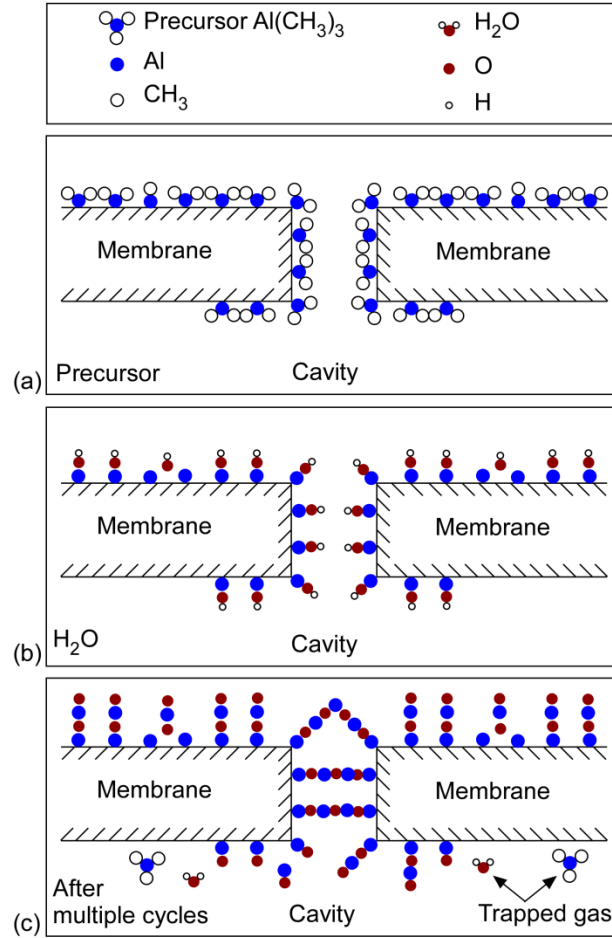
Figure 5.4: Schematic illustration of a sealing process using ALD Al_2O_3 [Puu05]. (a)-(b) One cycle of ALD Al_2O_3: (a) dose of precursor $Al(CH_3)_3$ for 20 ms and purge for 4 s; and (b) dose of H_2O for 20 ms and purge for 4 s. (c) A sealed pore after multiple cycles of ALD Al_2O_3. The figures are not drawn to scale.

5.4 Test Results and Discussion

As noted previously, optical interference leads to Newton's rings [Al06]; each ring represents a half integer multiple of a laser wavelength, which is 407 nm for the interferogram. When the 1st bright fringe is assumed to represent no deflection, the 1st dark fringe corresponds to ½ wavelength ($\approx$203.5 nm); the 2nd bright fringe corresponds to 1 wavelength; etc. Although the cavity height is $\approx$0.8 μm, the membrane ceases further deflection when it touches the ONO-3 bump formed on top of the ONO-1 layer inside the cavity (Fig. 5.6a); this bump is deposited to a height of $\approx$0.4 μm through an

access slit during the ONO-3 filling process (Fig. 5.3c). The membrane deflections were analyzed by laser interferogram (Olympus[TM] LEXT OLS3100) (Fig. 5.5a). The membrane deflection was also clearly identifiable by visual inspection in an optical microscope with a light source (Fig. 5.5b). In some cases, the deflected state was further confirmed by intentionally releasing the vacuum (through a nearby perforation); this flattens the membrane and completely removes the fringes.
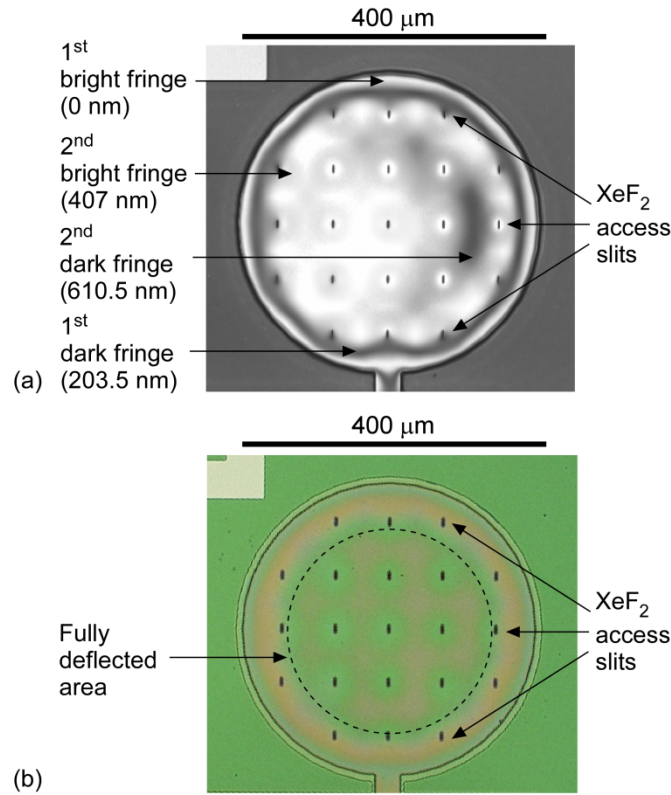
Figure 5.5: Typical images of deflected circular membranes. (a) A laser interferogram, in which the laser wavelength is 407 nm. (b) An optical microscope image. In (b), the dashed circle indicates the deflected area for which the deflection is limited by the ONO-3 bump inside the cavity. These images are typical of structures examined immediately after fabrication as well as 12 months later.

Structures that were not covered with ALD Al_2O_3 showed no deflection, indicating the cavities were at atmospheric pressure after PECVD dielectric deposition. At the end of the fabrication process, after cavity etch, ALD sealing, and dicing, typically

≈69% of the 26 samples in a 100-mm diameter wafer had membranes that were well-formed and appropriately deflected (Fig. 5.5b).

The cross section of a seal was examined by scanning electron microscope (SEM) and transmission electron microscope (TEM), and by energy-dispersive X-ray spectrometry (EDS) (Fig. 5.6). The XeF_2 access slit is clearly closed by the ONO-3 and ALD Al_2O_3 layers, as shown in the SEM image (Fig. 5.6a). The boundary of Al_2O_3 and ONO-3 is not evident in the SEM image. However, the boundary is clearly visible under the magnification provided by the TEM image (Fig. 5.6b). The chemical components in the sealed slit were examined by EDS spectra. The ALD Al_2O_3 spectrum is shown in the first inset of Fig. 5.6b, indicating Al- and O-related peaks, without a Si-related peak. The spectrum of the silicon dioxide in ONO-3 is shown in the third inset of Fig. 5.6b, indicating Si- and O-related peaks, without an Al-related peak. The spectrum of the central zone, as shown in the second inset of Fig. 5b, indicates an Al-related peak, together with Si- and O-related peaks. Such Al peaks are also seen in spectra taken on interior surface of the membrane.
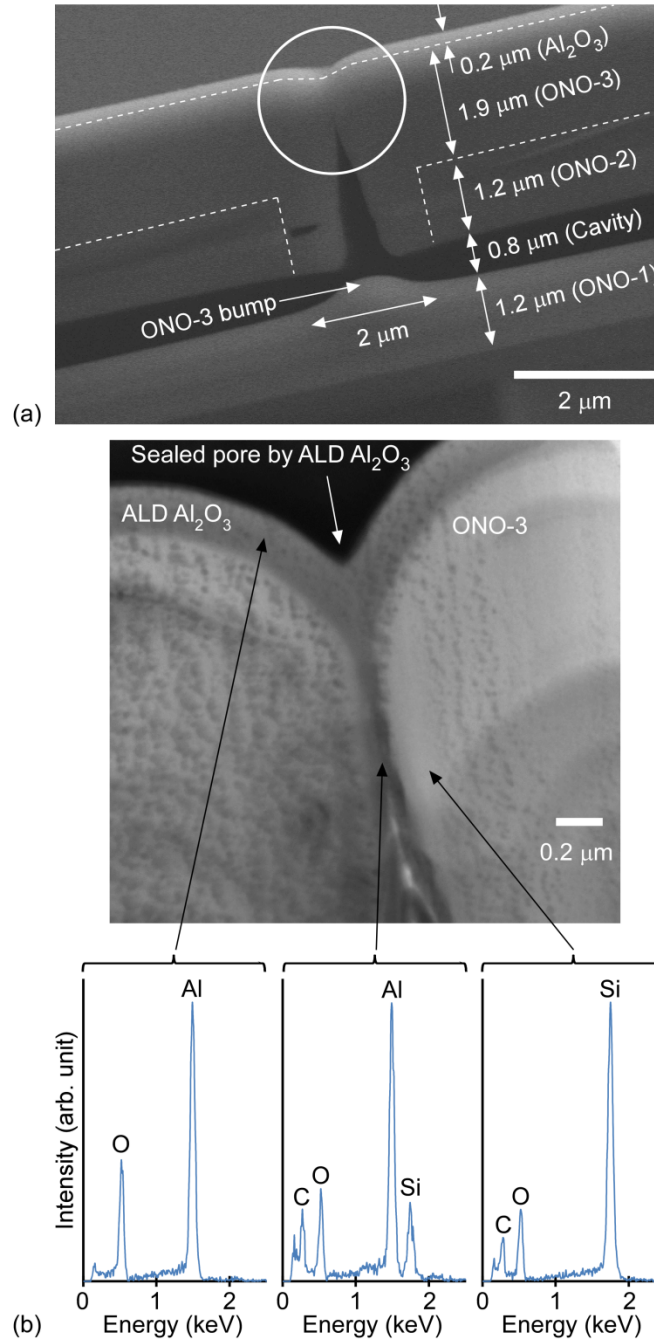
Figure 5.6: Cross-sectional images of a sealed access slit. (a) SEM image with the expected boundaries between layers. (b) TEM image of the circled area in (a), together with EDS spectra for Al_2O_3, sealed access slit center, and ONO-3. C-related peaks come from TEM sample preparation.

Membrane deflections were examined over 19 months while the devices were stored in a clean room. Of the 9 membranes that were examined, each one maintained its deflection without any noticeable change (Fig. 5.7).

The membrane deflection described here is similar to pressure-induced deflection used for bulge tests of deposited films in the past [Wan12]. In contrast to past work, for this paper the cavity was sealed in vacuum and the resulting deflection in the ALD Al_2O_3 membrane was maintained over 19 months.

The growth temperature for ALD Al_2O_3 could be potentially decreased below 250°C to further reduce the thermal budget if necessary. Deposition temperatures less than 100°C for ALD Al_2O_3 have been reported in [Wil05, Lin10, Zha11, Keu12, Dic12]. However, it is possible that the quality of the vacuum seal might be compromised. In addition, the image in Fig. 5.6b suggests that the thickness used for this film might be more than necessary.
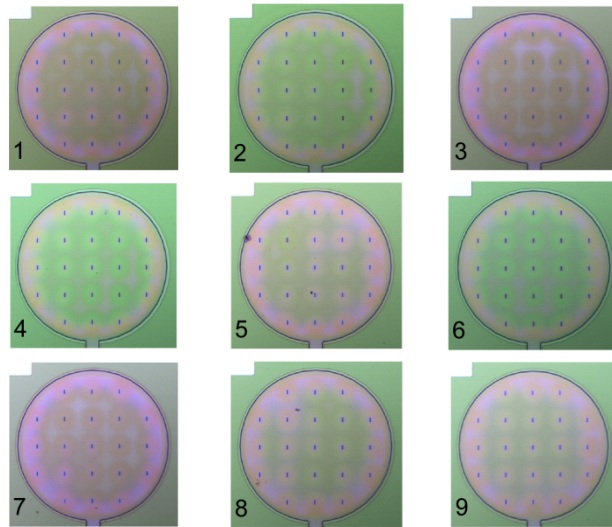
Figure 5.7: Reliability test of the cavity pressures. Microscope images of 9 sealed cavities with circular membranes at 19 months post fabrication are shown.

5.5 Conclusion

Wafer-level vacuum packaging is essential for the performance and protection of many important and widely-used micro-scale devices such as micro-gyroscopes and pressure sensors. However, achieving and sustaining vacuum with films deposited at modest temperatures has been a persistent challenge. Lower temperatures provide compatibility with the fabrication processes used for the devices, which may generally include CMOS electronics. Thin films of ALD Al_2O_3, which are pin-hole free and highly conformal, are very promising for retaining vacuum seals. The results described in this paper demonstrated that ≈ 0.2 µm-thick ALD Al_2O_3 at 250°C is sufficient to provide a seal to vacuum cavities that are ≈ 0.8 µm in height, 400 µm in diameter, and are capped with PECVD dielectrics that are otherwise porous. The sealed cavities retained vacuum for a period exceeding 19 months.

Future efforts can be directed at further reducing the thermal budget of the process. Additional efforts may be directed at using embedded sensors for quantitative analysis of cavity pressures.

CHAPTER 6

A SI-MICROMACHINED SINGLE-STAGE KNUDSEN PUMP DESIGNED FOR HIGH FLOW RATE

This chapter[2] explores single-stage Knudsen pumps that generate high flow rates. These pumps are designed for meeting high flow rate requirements of micro gas chromatographs and micro mass spectrometers. Using Si-micromachining, a high density of thermal transpiration flow channels is formed in a single chip. It has a footprint 16×20 mm^2. A combined operation of all flow channels produce more than 200 sccm air flow rate.

6.1 Introduction

Prior work on high flow rate Knudsen pumps was primarily based on thermal transpiration through polymer membranes [Pha10, Gup11b]. High-density nano-sized pores in the polymer membranes act as narrow channels (i.e., thermal transpiration flow channels) at atmospheric ambient pressure. Imposed thermal gradients across the membranes resulted in 0.4-0.8 sccm air flow [Pha10, Gup11b]. Pharas *et al.* [Pha10] used an input power of 8.6 W in a pump area of 49 mm^2 and Gupta *et al.* [Gup11b] used an input power of 1.4 W in a pump area of 104 mm^2. However, performance was limited by defect-induced leakage in the polymer membranes [Gup11b].

---

[2] This chapter describes joint work with Yutao Qin

In additional prior work, instead of polymer membranes, the narrow channels were integrated in a silicon substrate, using lithography-based batch microfabrication [McN05, Gup12]. It provided deflect-free narrow channels, which were placed in parallel with the substrate surface. The height of the narrow channels was defined by gap between polySi and glass (Fig. 1.9) [McN05] or sacrificial etching of deposited polySi (Fig. 2.3) [Gup12]. Although the fabricated narrow channels in the multistage configuration provided a compression ratio up to ≈ 844 at atmospheric ambient pressure, air flow rate was limited to ≈ 1 µsccm due to a limited density of narrow channels (Chapter 3) [An13a]. For substantially increasing the density of narrow channels that can result in a higher flow rate of at least 1 sccm, a new challenge to array narrow channels perpendicular to the silicon substrate must be met.

In this context, this chapter explores the monolithic integration of high density of narrow channels in a single-stage Knudsen pump designed for high flow rates. Section 6.2 presents the design of narrow channels, a heater, and a heat sink. Section 6.3 details the microfabrication processes and results. Test methods and results are described in Section 6.4, followed by discussion and conclusion in Section 6.5.

6.2 Design

This section describes the concept and the structure of the single-stage Knudsen pumps that are designed for high flow rates. Calculation results based on ideal structure is described, followed by the designed structure.

6.2.1 Concept of Knudsen Pumps Designed for High Flow Rates

6.2.1.1 Single-Stage Knudsen Pump with High-Density Narrow Channels

The required components of a single-stage Knudsen pump are narrow channels, a heater, and a heat sink. Narrow channels are arrayed in parallel in a rectangular matrix (Fig. 6.1a). In each narrow channel, gas molecules are bounded by dielectric sidewalls. The shorter period and longer period of the sidewalls are named the narrow channel height, a, and the narrow channel width, b, respectively. The distance between the heater and the heat sink is named the narrow channel length, l. The names for a, b, and l follow the naming conventions in Section 2.2. By applying an input power to the heater, the temperature difference, ΔT, is generated between the heater and heat sink (Fig. 6.1b). The heater and heat sink are at the hot temperature, T_{Hot}, and the cold temperature, T_{Cold}, respectively. Due to thermal transpiration flow in narrow channels, the pressure difference (pressure head), ΔP, is established (Fig. 6.1c).
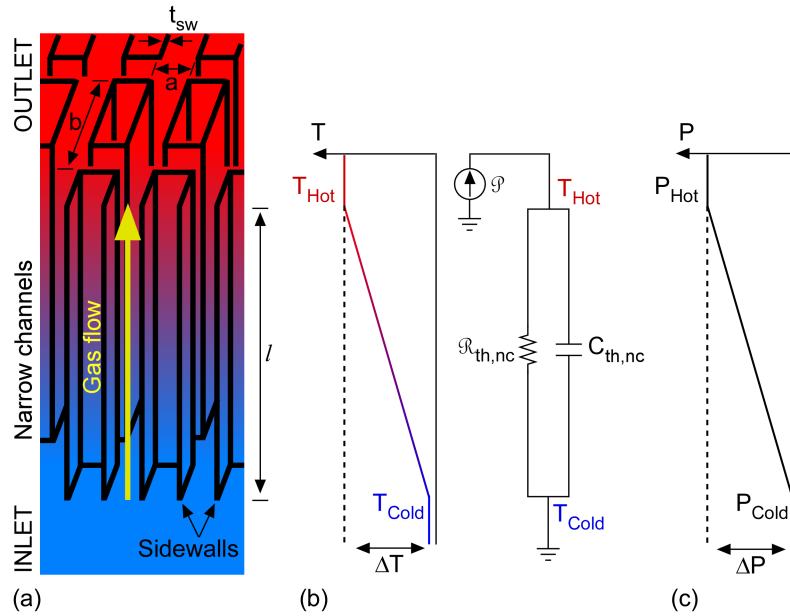
Figure 6.1:   Schematic of a single-stage Knudsen pump with high-density narrow channels. (a) Vertically oriented narrow channels. (b) Steady state temperature profile (left) and thermal circuit (right). (c) Steady state pressure profile.

The standard flow rate in a channel, $\dot{V}_{Std}$, is largest at a pressure difference of 0 Pa (i.e., no pressure head); whereas the standard flow rate is 0 sccm at an equilibrium pressure difference, ΔP_{Eq} (See Section 2.2 and 3.2 for equations). Since a large number of narrow channels are designed in this single-stage Knudsen pump, the total standard flow rate, $\dot{V}_{Std,Total}$, is written as:

$$\dot{V}_{Std,Total} = \dot{V}_{Std} N_{Total} \tag{6.1}$$

where $\dot{V}_{Std}$ is the standard flow rate in a narrow channel and N_{Total} is the total number of narrow channels. Here, N_{Total} is given by:

$$N_{Total} = \frac{A_{Total}}{ab} \tag{6.2}$$

where A_{Total} is the total area of narrow channels (active pumping area). The pressure difference is unrelated to the number of narrow channels, as shown in Eq. (2.2).

6.2.1.2 Steady State Pumping Performance
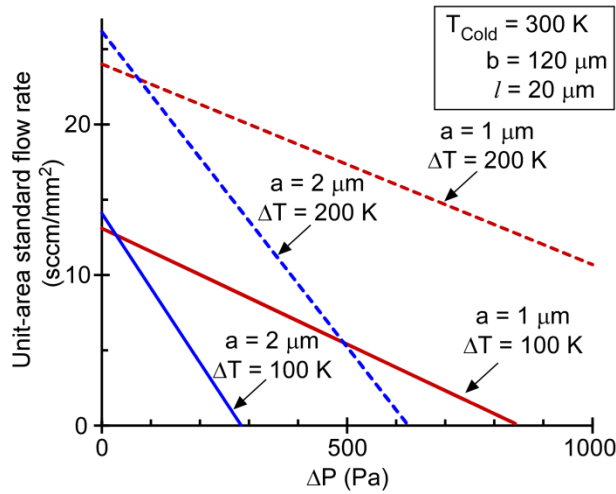
Figure 6.2: Calculated unit-area standard flow rate as a function of pressure difference (pump performance line), with varying a and ΔT. In this calculation, $T_{Cold} = 300$ K, $b = 120$ μm, and $l = 20$ μm.

The relation of ΔP and $\dot{V}_{Std}$ (pump performance line) due to variation in a and ΔT are theoretically calculated using Eq. (2.3, 3.1, 3.2) (Fig. 6.2). Due to microfabrication constraints, b and l are selected as 120 μm and 20 μm, respectively. The narrow channel height, a, is varied in two values, 2 μm and 1 μm. For each a, ΔT is varied in two values, 100 K and 200 K. The T_{Cold} value is 300 K. The unit-area standard flow rate is obtained by division of $\dot{V}_{Std}$ by the product of a and b. By decreasing a or increasing ΔT, ΔP at a standard flow rate increases. At an identical ΔT, the differences in the largest standard flow rates ($\Delta P = 0$) between different a values are small. Due to lithographic limitations in available tool sets, the design value of a is selected as 2 μm.

Joule heating is applied to the heater. The resulting heat is dissipated mainly through narrow channels to the heat sink (Fig. 1.c). To minimize power consumption in a Knudsen pump, the thermal resistance needs to be maximized. By Fourier's law, the steady state temperature difference, ΔT, for an input power, $\wp$, is written as:

$$\Delta T = \Re_{th}\wp \tag{6.3}$$

where $\Re_{th}$ is the thermal resistance. Here, $\Re_{th}$ is the parallel combination of the constituent thermal resistances:

$$\Re_{th}^{-1} = \Re_{th,nc}^{-1} + \Re_{th,conv}^{-1} \tag{6.4}$$

where $\Re_{th,nc}$ is the thermal resistance of the narrow channels, which are composed of air molecules and dielectric sidewalls, and $\Re_{th,conv}$ is the thermal resistance of the natural convection cooling to exterior ambient. The $\Re_{th,conv}$ is given by:

$$\Re_{th,conv} = \frac{1}{h_{conv,air}} \frac{1}{A_{Total}} \tag{6.5}$$

where $h_{conv,air}$ is the natural convection coefficient of air (10×10^{-6} W/K-mm^2 [Sai10]).

The unit-area $\mathfrak{R}_{th,conv}$ is 1×10^5 K/W-mm^2.

The value of $\mathfrak{R}_{th,nc}$ is given by:

$$\mathfrak{R}_{th,nc} = \frac{1}{\sigma_{th,nc}} \frac{l}{A_{Total}} \qquad (6.6)$$

where $\sigma_{th,nc}$ is the equivalent thermal conductivity of narrow channels. Here, it is written as the weighted sum:

$$\sigma_{th,nc} = \frac{t_{sw}\sigma_{th,sw} + (a - t_{sw})\sigma_{th,air}}{a} \qquad (6.7)$$

where t_{sw} is the sidewall thickness; $\sigma_{th,sw}$ is the thermal conductivity of the sidewall; and $\sigma_{th,air}$ is the thermal conductivity of air molecules in the narrow channel, 0.026 K/m-W [Ste85], assuming viscous flow regime (See Section 4.2.1). The sidewall material in this chapter is an amorphous Al$_2$O$_3$ layer ($\sigma_{th,sw} \approx 1.4$ W/m-K at ≈ 270 K [Sta93]). When the sidewall thickness is 0, the unit-area thermal resistance of the narrow channels is 770 K/W-mm^2. The sidewall thickness is selected as 10 nm, as a tradeoff between the thermal resistance and microfabrication constraints. The calculated unit-area thermal resistance, $\mathfrak{R}_{th,nc}$, for the selected sidewall thickness of 10 nm is 608 K/W-mm^2. The $\mathfrak{R}_{th,conv}$ in Eq. (6.4) is neglected because $\mathfrak{R}_{th,conv}$ is two orders of magnitude larger than $\mathfrak{R}_{th,nc}$.

From the performance line (Fig. 6.2), where a is 2 µm, l is 20 µm, and ΔT is 100 K, the calculated results can be summarized as follows: first, the equilibrium pressure difference is 284 Pa and the largest unit-area standard flow rate ($\Delta P = 0$ Pa) is 14.1 sccm/mm^2; and second, the unit-area input power values for the Al$_2$O$_3$ sidewall

thicknesses of 0 nm and 10 nm are 0.13 W/mm^2 and 0.16 W/mm^2, respectively. The Al$_2$O$_3$ sidewall thickness of 10 nm in narrow channels contributes only 0.03 W/mm^2 to the unit-area input power. Using these results based on unit area, the largest standard flow rate and the input power can be calculated. However, these calculated values represent only ideal values. Practically designed structures present performance-reducing factors, which are described in the next section.

6.2.1.3 Transient Pumping Performance

For estimating the transient response of the Knudsen pump, both the thermal time constant and hydraulic time constant must be considered. Physically, the time constant represents the time for a system to reach 63.2% of its steady state response. As shown by calculations in this section, both the thermal time constant and the hydraulic time constant are much less than 0.1 s.

The thermal circuit (Fig. 1c) includes thermal capacitance that is connected in parallel with thermal resistance. The resulting thermal response is written as:

$$\Delta T(t) = \Delta T_f \left(1 - e^{-t/\tau_{th}} \right) \tag{6.8}$$

where ΔT is expressed as a function of time, t; ΔT_f is the final steady state value of ΔT; and τ_{th} is the thermal time constant.

The thermal time constant for the Knudsen pump heater, which considers only heat dissipation through narrow channels, is given by:

$$\tau_{th} = \mathfrak{R}_{th,nc} C_{th,nc} \tag{6.9}$$

where $C_{th,nc}$ is the thermal capacitance of the narrow channels, which are composed of air molecules and dielectric sidewalls. The value of $C_{th,nc}$ is given by:

$$C_{th,nc} = c_{th,nc}\rho_{nc}A_{Total}l \tag{6.10}$$

where $c_{th,nc}$ and ρ_{nc} are, respectively, the equivalent specific heat and mass density of narrow channels. Here, using weighted sum, $c_{th,nc}$ and ρ_{nc} are, respectively, written as:

$$c_{th,nc} = \frac{t_{sw}c_{th,sw} + (a - t_{sw})c_{th,air}}{a} \tag{6.11}$$

$$\rho_{nc} = \frac{t_{sw}\rho_{sw} + (a - t_{sw})\rho_{air}}{a} \tag{6.12}$$

where $c_{th,sw}$ is the specific heat of the sidewall (755 J/kg-K for alumina sidewall [Mun97]); $c_{th,air}$ is the specific heat of air molecules (1.01 J/kg-K [Jou08]); ρ_{sw} is the mass density of the sidewall (3984 kg/m^3 for alumina sidewall [Mun97]); and ρ_{air} is the mass density of air molecules (1.2929 kg/m^3 [Jou08]). The calculated unit-area thermal capacitance for the selected sidewall thickness of 10 nm is 2.0×10^{-9} J/K-mm^2. Therefore, using the calculated unit-area thermal resistance of 608 K/W-mm^2, calculated examples of thermal time constant are 1.2×10^{-6} s for an area of 1 mm^2 and 0.012 s for an area of 100 mm^2.

The hydraulic time constant for gas flow in narrow channels [Roj13] can be directly derived from Eq. (2.2). Using the ideal gas law and including A_{Total}, Eq. (2.2) can be written as the following first-order differential equation:

$$\frac{d\Delta P(t)}{dt} = -\frac{k_B T_{Cold}}{mV_{Cold}}\left(\frac{Q_T}{T_{Avg}}\Delta T - \frac{Q_P}{P_{Avg}}\Delta P(t)\right)\frac{a^2 b P_{Avg}}{l}\sqrt{\frac{m}{2k_B T_{Avg}}}\frac{A_{Total}}{ab} \tag{6.13}$$

where ΔP is expressed as a function of time, t, and V_{Cold} is the volume at the cold chamber. For simplification, Q_T, Q_P, and P_{Avg} are regarded as time-independent variables because they change by negligible amounts. Then, the solution of Eq. (6.13) takes the following form:

$$\Delta P(t) = \Delta P_{Eq}(1 - e^{-t/\tau_{Hyd}}) \tag{6.14}$$

where τ_{Hyd} is the hydraulic time constant. Here, τ_{Hyd} is given by:

$$\tau_{Hyd} = \frac{mV_{Cold}}{k_B T_{Cold}} Q_P \frac{l}{aA_{Total}} \sqrt{\frac{2k_B T_{Avg}}{m}} \tag{6.15}$$

Calculated examples of hydraulic time constant are 0.2 ms when V_{Cold} is 1 mm^3 and A_{Total} is 1 mm^2, and 2 ms when V_{Cold} is 1,000 mm^3 and A_{Total} is 100 mm^2,

6.2.2 Designed Structure
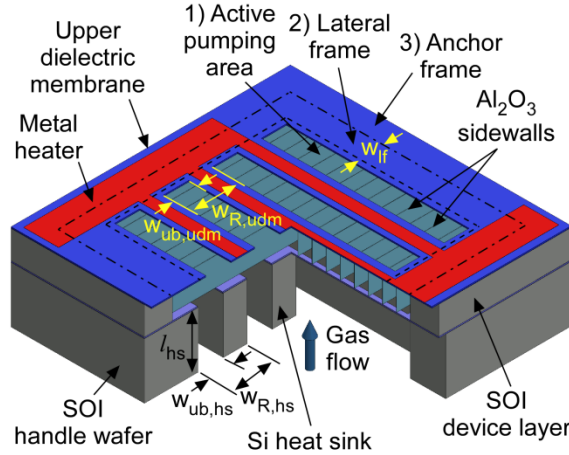
Figure 6.3: Perspective view of Knudsen pump structure. The structure is divided into three regions: 1) active pumping area; 2) lateral frame; and 3) anchor frame (thermal ground).

The Knudsen pump structure in this chapter is designed using a silicon-on-insulator (SOI) wafer (Fig. 6.3). The SOI device layer, thickness of 20 μm, is used for the vertically oriented narrow channels. The SOI handle wafer, thickness of 380 μm, is used for the heat sink. The upper dielectric membrane and the metal resistive heater are placed on the hot end of the narrow channels. The Joule heat is dissipated through

narrow channels and then through the heat sink to the anchor frame that is thermally grounded to ambient temperature.

Although sidewalls are arrayed in the SOI device layer continuously, the upper dielectric membrane and the heat sink partially block gas flow through narrow channels. To account for the unblocked gas flow paths, fill factors, f, are defined as:

$$f_{udm} = \frac{w_{ub,udm}}{w_{R,udm}} \tag{6.6}$$

$$f_{hs} = \frac{w_{ub,hs}}{w_{R,hs}} \tag{6.7}$$

$$f_{Avg} = \frac{f_{udm} + f_{hs}}{2} \tag{6.8}$$

where f_{udm} and f_{hs} are the fill factors for unblocked gas flow paths in the upper dielectric membrane and the heat sink, respectively; f_{Avg} is the average fill factor of f_{udm} and f_{hs}; $w_{R,udm}$ is the width of the repeating unit in the upper dielectric membrane; $w_{ub,udm}$ is the unblocked width of the repeating unit in the upper dielectric membrane; $w_{R,hs}$ is the width of the repeating unit in the heat sink; and $w_{ub,hs}$ is the unblocked width of the repeating unit in the heat sink (Fig. 6.3).

The sub-unity fill factors for unblocked gas flow paths represent the compromise in pump performance. First, the standard flow rate is reduced by the average fill factor, f_{Avg}, as an approximation. Second, the actual temperature difference in the unblocked width is reduced. Finite element analysis (FEA) simulation of a minimum repeating block of narrow channels shows temperature distribution in the unblocked gas flow path (Fig. 6.4a). The heater temperature is assumed to be 400 K and the heat sink temperature is assumed to be 300 K. The average values of T_{Hot} and T_{Cold} in the unblocked gas flow

path are 392 K and 318 K, respectively (Fig. 6.4b). The resulting temperature difference is 74 K out of the input ΔT of 100 K. This translates into $\approx 26\%$ of input power loss.
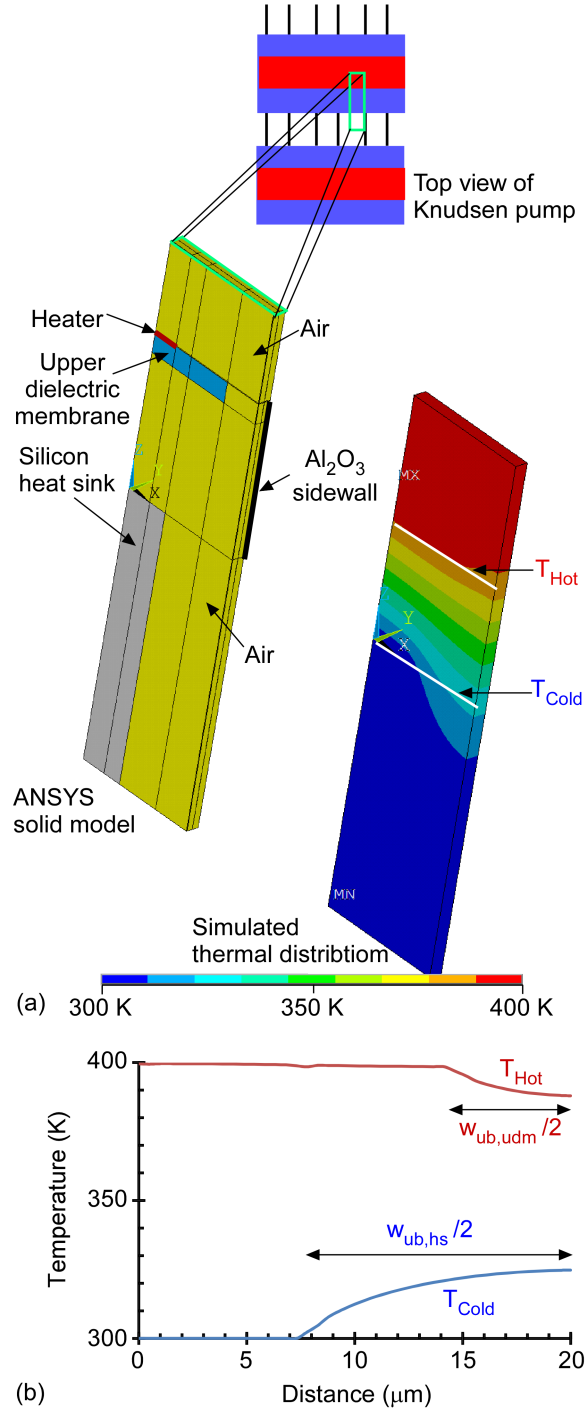
Figure 6.4: ANSYS[TM] FEA thermal simulation of narrow channels. (a) Solid model of a repeating unit of a narrow channel and temperature distribution, assuming that temperatures of the heat sink and heater are 300 K and 400 K, respectively. (b) Temperature profiles along the hot side, T_{Hot}, and cold side, T_{Cold}, in (a).
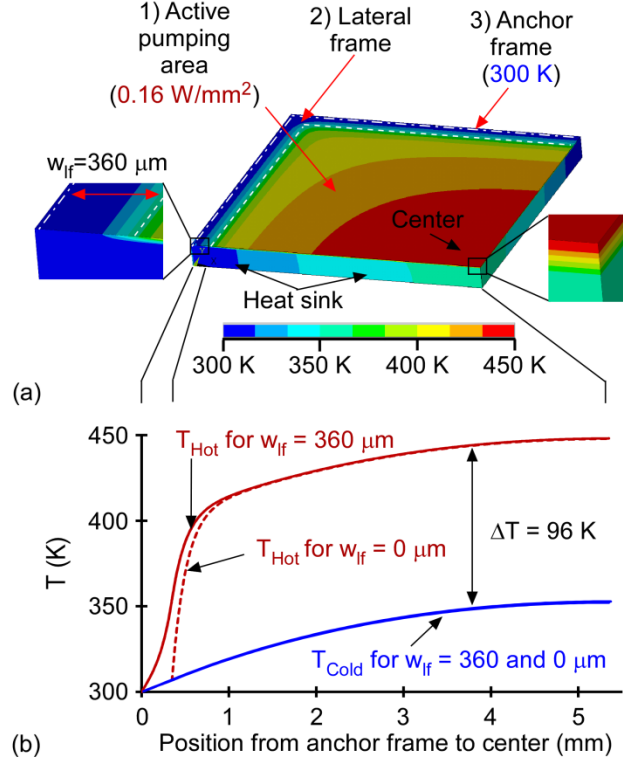
Figure 6.5: ANSYS^{TM} FEA thermal simulation showing the effect of the thermal resistance of the heat sink and the lateral frame on T_{Cold} and T_{Hot}. (a) Thermal distribution of one quadrant of an active pumping area of 1 cm^2 (square) and the lateral frame. (b) Temperature profile of T_{Hot} and T_{Cold} from the anchor frame to center. In (b), the w_{lf} value of 0 µm indicates that the lateral frame with the designed w_{lf} value of 360 µm is replaced with Si from the SOI device layer.

The heat sink is intended to cool the cold end of the narrow channel to ambient temperature using a high thermal conductivity of silicon. However, the distance from the active pumping area to the anchor frame contributes thermal resistance that increases T_{Cold} to a value that is larger than ambient temperature. Moreover, the fill factor of the heat sink affects the equivalent thermal conductivity of the heat sink, which is given by:

$$\sigma_{th,hs} = \sigma_{th,Si}\left(1 - f_{hs}\right) \tag{6.9}$$

where $\sigma_{th,Si}$ is the thermal conductivity of crystalline silicon ($\sigma_{th,Si}$ = 149 W/m-K). As shown by FEA thermal simulation (Fig. 6.5), T_{Cold} increases from the anchor frame to the center of the active pumping area, whereas ΔT remains almost the same. The resulting

pressure difference is nearly unaffected, as anticipated by Eq. (2.3), whereas the resulting flow rate decreases due to increase in T_{Avg} in Eq. (3.1).

The lateral frame is formed in the SOI device layer between the active pumping area and the anchor frame. The width of the lateral frame is denoted by w_{lf} in Fig. 6.3 and 6.5. Although narrow channels are formed in the lateral frame, gas flow is completely blocked by the upper dielectric membrane. The lateral frame prevents a drop in T_{Hot} at the edge of the active pumping area: the resistive heat dissipates through the lateral frame and partially through the heat sink to the anchor frame (the case of $w_{lf} = 360$ μm in Fig. 6.5b). If the lateral frame is replaced with Si from the SOI device layer, a rapid drop in T_{Hot} would appear at the edge of the active pumping area, causing the heat to dissipate through the device silicon layer directly to the heat sink (the case of $w_{lf} = 0$ μm in Fig. 6.5b). This rapid drop in T_{Hot} compromises thermal transpiration, thereby resulting in a smaller flow rate and a smaller pressure difference, as anticipated by Eqs. (2.2, 2.3). The T_{Cold} values are unaffected by the width of the lateral frame (T_{Cold} for $w_{lf} = 360$ and 0 μm in Fig. 6.5b).

For estimating pump performance, average temperatures of T_{Cold} and T_{Hot} in the active pumping area (Fig. 6.5) are used. Then, to reflect the actual temperature difference in the unblocked widths of the upper dielectric membrane and the heat sink (Fig. 6.4b), T_{Cold} is increased by $0.18 \times \Delta T$ and T_{Hot} is decreased by $0.08 \times \Delta T$, where ΔT is the difference of the average temperatures of T_{Cold} and T_{Hot}.

The final Knudsen pump design is summarized in Table 6.1. It includes designed values and fabricated values together. The Knudsen pump is composed of four sub-blocks. The active pumping area in each sub-block is 6×4 mm^2. These sub-blocks are

sized rectangularly (6 × 4) and are separated by an anchor frame for lowering the increase in T_{Cold} due to the thermal resistance of the heat sink (Fig. 6.5). The combined total active pumping area in the Knudsen pump is 96 (= 4 × 6 × 4) mm². Calculated performance based on FEA thermal simulations (Fig. 6.4 and 6.5) using fabricated dimensions are summarized in Table 6.2. For a unit-area input power of 0.16 W/mm², the largest unit-area standard flow rate (ΔP = 0 Pa) decreases from 14.1 sccm/mm² to 4.4 sccm/mm² due to design compromises. The next section describes the microfabrication of the Knudsen pump.

Table 6.1: Summary of dimensions of designed and fabricated structures.

Items		Designed value (µm)	Fabricated value (µm)	Reference figures
Narrow channel	Height, a	2	2	Fig. 6.2, Fig. 6.12b
	Width, b	120	120	
	Length, l	20	20	
	Sidewall thickness, t_{sw}	0.01	0.0107	
Upper dielectric membrane	Unblocked width, $w_{ub,udm}$	12	12	Fig. 6.3 Fig. 6.12a
	Width of the repeating unit, $w_{R,udm}$	40	40	
	Fill factor, f_{udm}	12/40	12/40	
Heat sink	Length, l_{hc}	380	340	Fig. 6.3, Fig. 6.11
	Unblocked width, $w_{ub,hs}$	22	26	
	Width of the repeating unit, $w_{R,hs}$	40	40	
	Fill factor, f_{hs}	22/40	26/40	
Average fill factor, f_{Avg}		17/40	19/40	-
Lateral frame width, w_{lf}		360	>140	Fig. 6.3, Fig. 6.5

Table 6.2: Calculated pump performance using the fabricated dimensions in Table 6.1. The ambient room temperature is assumed to be 300 K. A unit-area input power of 0.39 W/mm² corresponds to the input power of 37.2 W in Section 6.4 (See Fig. 6.13).

Items	Calculation results	
Unit-area input power	0.16 W/mm²	0.39 W/mm²
Active pumping area, A_{Total}	4 × 6 × 4 (=96) mm²	
Number of narrow channels, N_{Total}	0.4 × 10⁶	
Average T_{Cold} in the active pumping area	314.3 K	334.6 K
Average T_{Hot} in the active pumping area	405.6 K	555.6 K
Average ΔT in the active pumping area	91.3 K	221.0 K
$\dot{V}_{Std,Total}$ at ΔP = 0 Pa	424.5 sccm	903.6 sccm
Unit-area $\dot{V}_{Std,Total}$ at ΔP = 0 Pa	4.4 sccm/mm²	9.4 sccm/mm²
ΔP_{Eq} ($\dot{V}_{Std,Total}$ = 0 sccm)	197.5 Pa	555.8 Pa

6.3 Microfabrication

A SOI wafer and four-masking step process is used for microfabrication of the Knudsen pump. The important features include: 1) a SOI device layer mold, formed by channel etching using deep reactive ion etching (DRIE) and conformal refilling using Al_2O_3 sidewall material and a polySi sacrificial layer; 2) a Cr/Pt resistive heater on upper oxide-nitride-oxide (ONO) membrane; and 3) a heat sink, formed by backside DRIE of the SOI handle wafer. In the SOI wafer, the device layer thickness is $\approx$20 μm, the buried oxide thickness is $\approx$0.5 μm, and the handle wafer thickness is $\approx$380 μm.

The process begins with DRIE of trenches for vertical channels, which is in a two-dimensional array of 2×118 μm^2 with spaces of 2 μm (Fig. 6.6a). The buried oxide of the SOI wafer acts as a DRIE etch stop layer. The trenches are covered by Al_2O_3 sidewalls, thickness $\approx$10 nm, using a highly conformal atomic layer deposition (ALD) method (Fig. 6.6b). The process is followed by trench refilling using a low-pressure chemical vapor deposited (LPCVD) polySi layer, thickness $\approx$1.5 μm. Voids (or gaps) could exist in the polySi refilled trenches. The overfilled polySi on the SOI device layer is dry etched until the ALD Al_2O_3 is fully exposed, using an isotropic silicon etching recipe (Fig. 6.6c). The etching recipe does not etch Al_2O_3. The voids may become widened at this step. The upper dielectric membrane is formed by two stacks of PECVD ONO, total thickness $\approx$3.0 μm. Height differences, which are generated during the previous dry etching step (Fig. 6.6c), are planarized by the relatively thick nature of these ONO stacks. Each ONO stack, thickness $\approx$1.5 μm, is composed of one nitride layer, thickness $\approx$0.7 μm, sandwiched by two oxide layers, thickness $\approx$0.4 μm. The RIE of ONO follows (Fig. 6.6d). A lift-off process, using image reversal photoresist (AZ5214),

patterns Cr/Pt resistive heaters and traces, evaporated thickness $\approx$25/100 nm (Fig. 6.6e). Both Cr and Pt are resistant to XeF_2 sacrificial etchant.

Until the last process step on the front side of the SOI wafer, the backside of the SOI handle wafer is covered with Al_2O_3, thickness $\approx$10 nm, and polySi, thickness $\approx$1.5 μm. They are etched in reverse order from the backside of the SOI wafer using RIE and an etching mask of SPR220-7.0 photoresist, thickness $\approx$10 μm. The SOI handle wafer is etched through using DRIE; the SOI buried oxide layer acts as an etch stop layer. The process is followed by RIE of the SOI buried oxide layer, thickness $\approx$0.5 μm. This RIE recipe continuously etches the ALD Al_2O_3 on the bottom of SOI device layer mold (Fig. 6.6f). Al_2O_3, thickness $\approx$10 nm, is conformally deposited using ALD, and etched by RIE. The anisotropic characteristic of RIE enables a higher etching rate of Al_2O_3 on the trench bottom than the sidewall; this process is adapted from [Sun04]. The Al_2O_3, using its chemical inertness to XeF_2 sacrificial etchant, protects the handle wafer during the next sacrificial etching step (Fig. 6.6g). The device silicon and refilled polySi in the SOI device layer mold are sacrificially etched away using XeF_2 dry gas, thereby providing the thermal transpiration channels. The top and bottom views are shown for clarifying the etched patterns. The side view is along the dashed lines from A to A' in the top view and from B to B' in the bottom view.
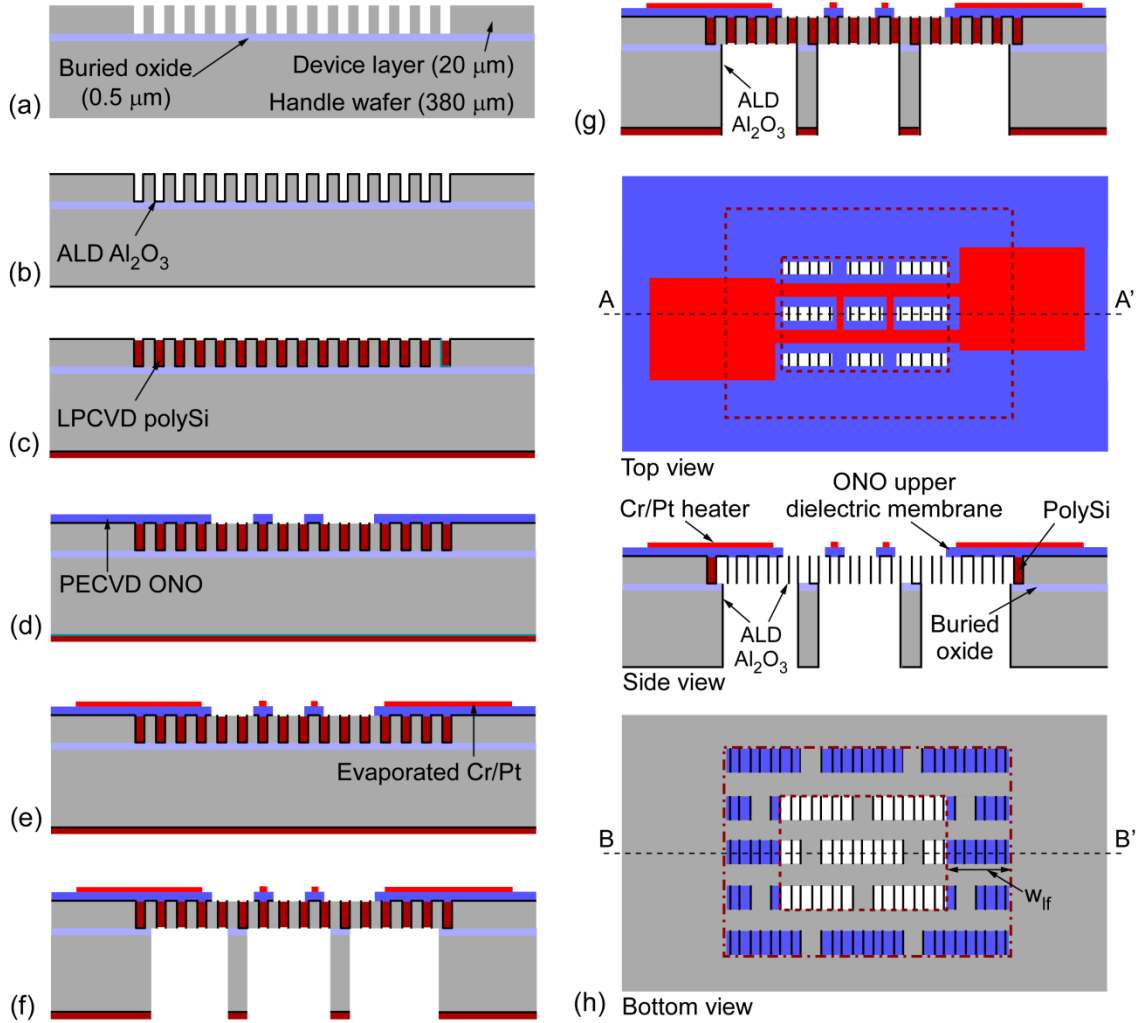
Figure 6.6: Fabrication flow. (a) (Mask 1) DRIE of the SOI device layer. (b) ALD of Al_2O_3 sidewall. (c) PolySi refilling and dry etching. (d) (Mask 2) PECVD and RIE of ONO upper dielectric membrane. (e) (Mask 3) Cr/Pt metallization. (f) (Mask 4) DRIE of the handle wafer. (g) ALD and RIE of Al_2O_3. (h) XeF_2 sacrificial etching.

Knudsen pumps are fabricated using a 4" SOI wafer. The foot print is 16×20 mm^2 (Fig. 6.7). The thickness and residual stress of each layer are monitored and summarized in Table 6.3.
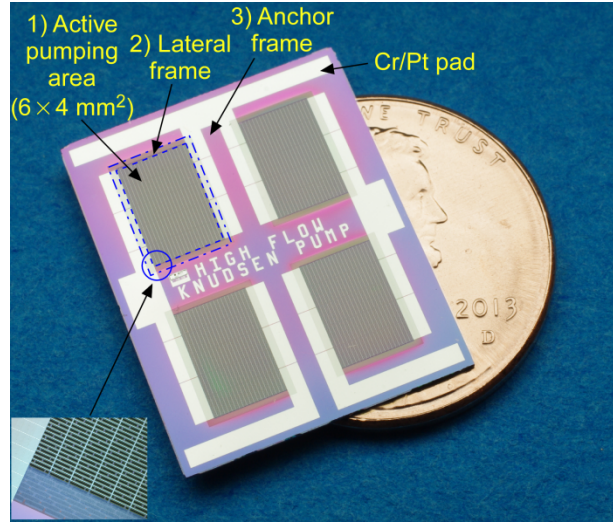
Figure 6.7: A photograph image of an as-fabricated chip. The left-bottom inset shows close-up microscope image on the edge of an active pumping area.

Table 6.3: Measured thickness and residual stress of each layer.

Layer	Thickness (μm)	Residual stress (MPa)
SOI device layer	20	-
SOI buried oxide	0.5	-
SOI handle wafer	380	-
LPCVD PolySi	1.5	-
PECVD two ONO stacks	0.4/0.7/0.4 +0.4/0.7/0.4	+56
ALD Al$_2$O$_3$	0.0107	+304
Cr/Pt	0.025/0.1	+967

To minimize voids in polySi-refilled trenches, DRIE of the SOI device layer is tuned to provide a tapering angle using the Bosch process in STS^{TM} PEGASUS tool. The Bosch process achieves vertical trench etching using repeated cycles of passivation layer deposition and isotropic silicon etching. For tapering, substrate chuck temperature is cooled down to 10°C from 35°C (in the standard recipe). This lowered temperature increases the thickness of passivation layers [Jo05] and results in tapering angle of ≈1° in the vertical trenches (Fig. 6.8).
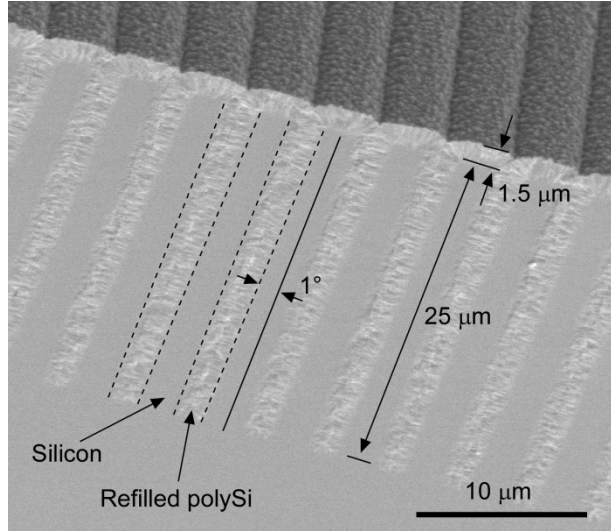
Figure 6.8: SEM image of polySi-refilled trenches in a test silicon wafer. The DRIE recipe is tuned to result in tapering of $\approx 1°$. The dotted lines indicate the boundaries between silicon and refilled polySi.

After dry etching of overfilled polySi, the height differences are planarized using two stacks of PECVD ONO (Fig. 6.9). This planarization prevents metal disconnections in the next metallization process, which uses an evaporated Cr/Pt metal for lift-off. As compared between Fig. 6.9c and Fig. 6.9d, two ONO stacks are necessary rather than one ONO stack for planarization.

In the two ONO stacks of the upper dielectric membrane, thicknesses for nitride and oxide layers are adjusted to provide a mild tensile stress level of ≈ 56 MPa. In each ONO stack, a tensile stressed nitride layer, residual stress ≈ 236 MPa, is sandwiched by two oxide layers, residual stress ≈ -101 MPa. This stack is preferred to a tensile nitride-only layer because an oxide layer is considerably resistant to XeF_2 sacrificial etchant, relative to a nitride layer.
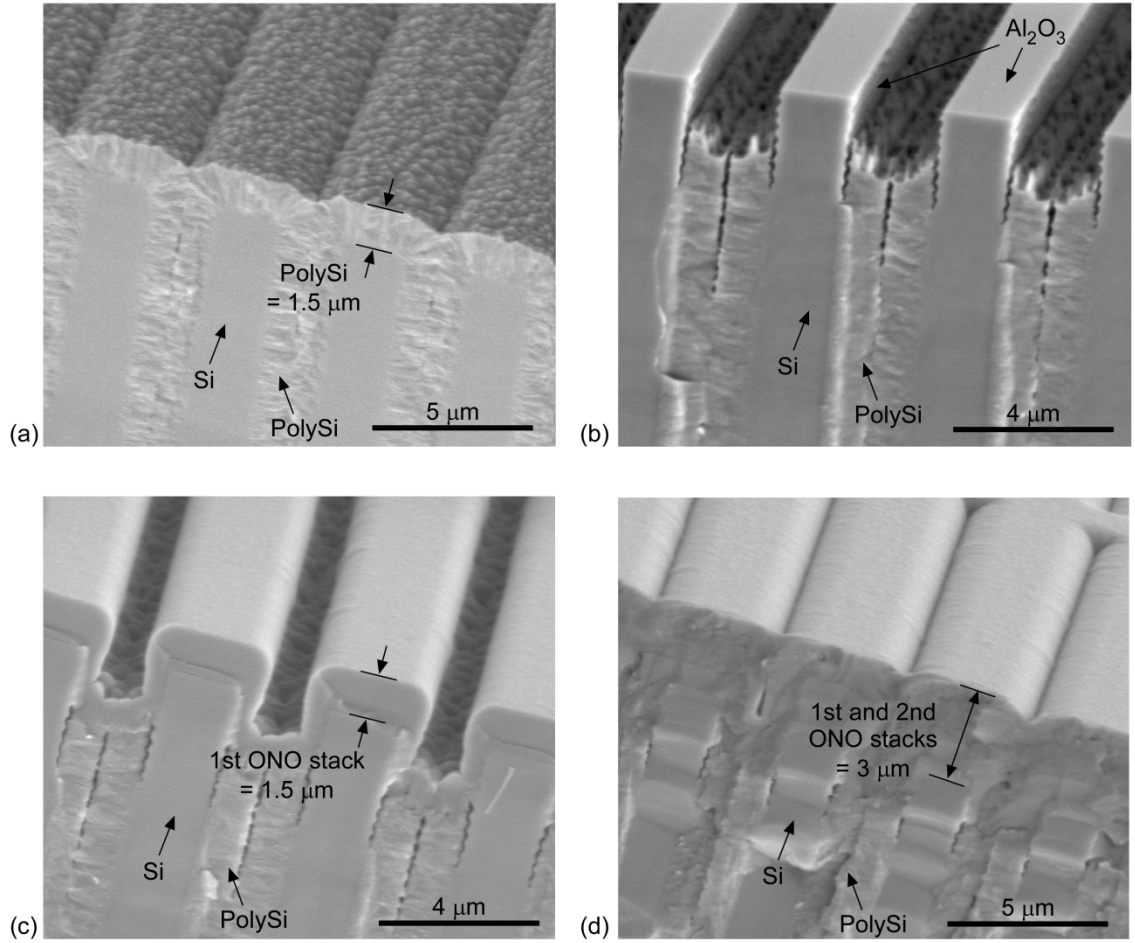
Figure 6.9: SEM images of SOI device layer mold. (a) DRIE of the SOI device layer, ALD Al_2O_3 sidewall, and polySi refilling. (b) Dry etching of overfilled polySi. (c) PECVD of one ONO stack. (d) PECVD of two ONO stacks for planarization.

The sidewall of the backside DRIE trench is protected by ALD Al_2O_3. To investigate the validity of Al_2O_3 as a protective layer during XeF_2 etching, a test silicon wafer is etched, depth ≈ 350 μm, using DRIE and photoresist as an etching mask. Then, the photoresist is ashed, followed by 10 growth cycles of ALD Al_2O_3 at 150°C, resulting in ≈ 1 nm in thickness. Al_2O_3 is etched using a fluoride-based oxide RIE recipe; the etch rate of Al_2O_3 is ≈ 10 times lower than that of oxide. By exposing the test wafer under XeF_2 dry gas, the protection of the sidewall of the DRIE trench and the unprotected etch to the bottom of the DRIE trench are clearly differentiated (Fig. 6.10). Although a

125

thickness of ≈1 nm works, the actual process used a thickness of ≈10 nm, for sufficient

protection of the sidewall.
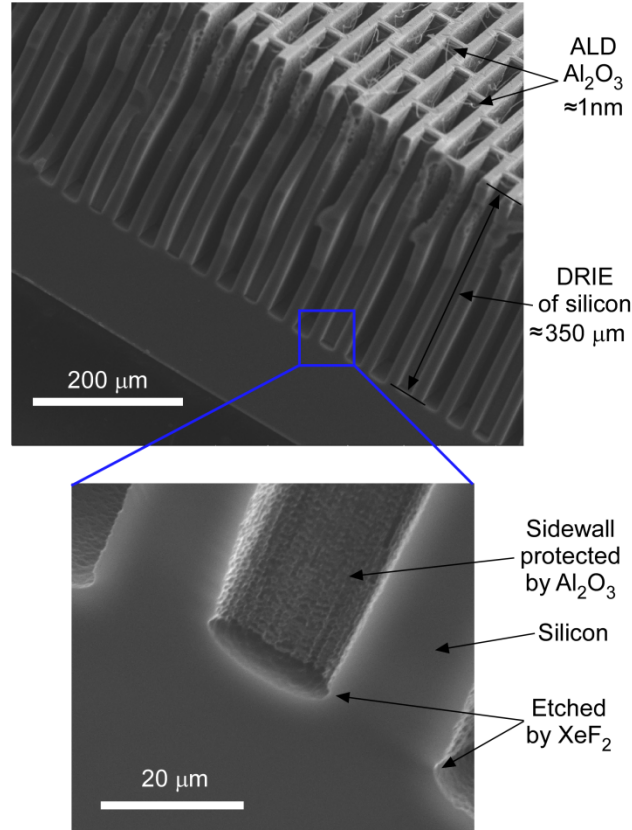
Figure 6.10: SEM images of the XeF$_2$ etched bottom of DRIE trench in a test silicon wafer. ALD Al$_2$O$_3$, thickness ≈1 nm, was used for protecting the sidewall of the DRIE trench.

Figure 6.11 compares the heat sink structures before and after XeF$_2$ sacrificial

etching. Before XeF$_2$ sacrificial etching, sidewalls near the bottom of the SOI handle

wafer are partially damaged by DRIE. It appears that the passivation layer is partially

penetrable by SF$_6$ etching gas, which eventually erodes silicon sidewalls. During the

subsequent anisotropic RIE of Al$_2$O$_3$, the Al$_2$O$_3$ in the eroded region is etched away

because it is similar to Al$_2$O$_3$ in the trench bottom. The silicon at the eroded region is

exposed to XeF$_2$ sacrificial etchant. This reduces the length of heat sink structure from

380 μm to ≈340 μm. The other regions of heat sink structures are effectively protected by the Al_2O_3 layer, and therefore are unaffected by XeF_2 sacrificial etchant.
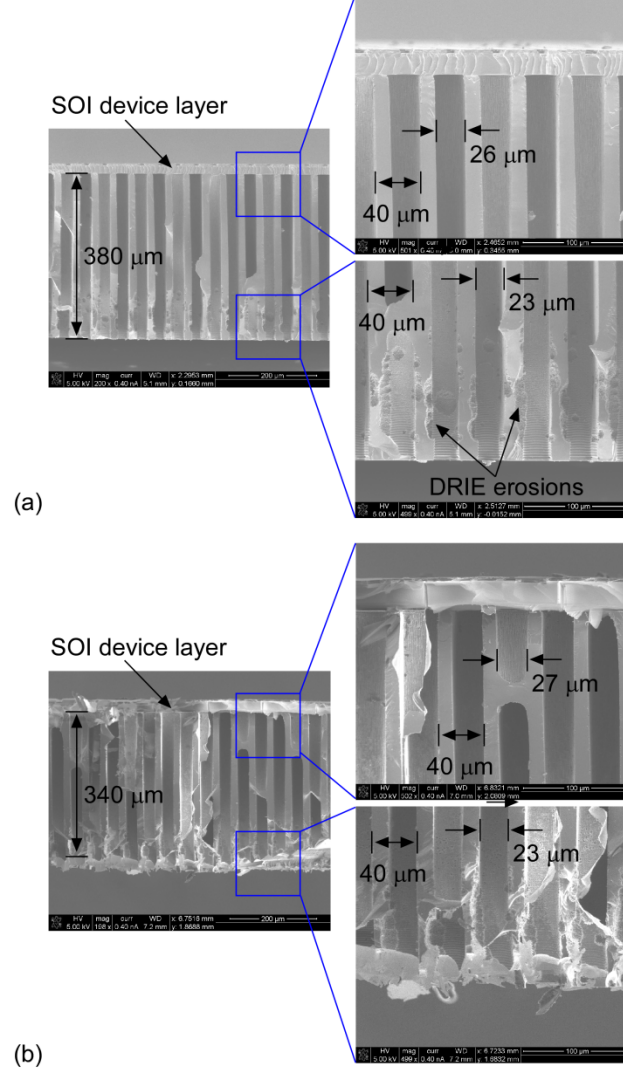
Figure 6.11: SEM images of the heat sink structure. (a) Before XeF_2 sacrificial etching. (b) After XeF_2 sacrificial etching. In each SEM image, $w_{R,hs}$ (= 40 μm) and $w_{ub,hs}$ are indicated.

Fabricated narrow channels were examined using scanning electron microscope (SEM) images (Fig. 6.12). The thickness of the ALD Al_2O_3 layer is measured as 10.7 nm, using a NANOSPEC^{TM} tool. ALD Al_2O_3 sidewalls are released and vertically oriented without noticeable deformation or damage.
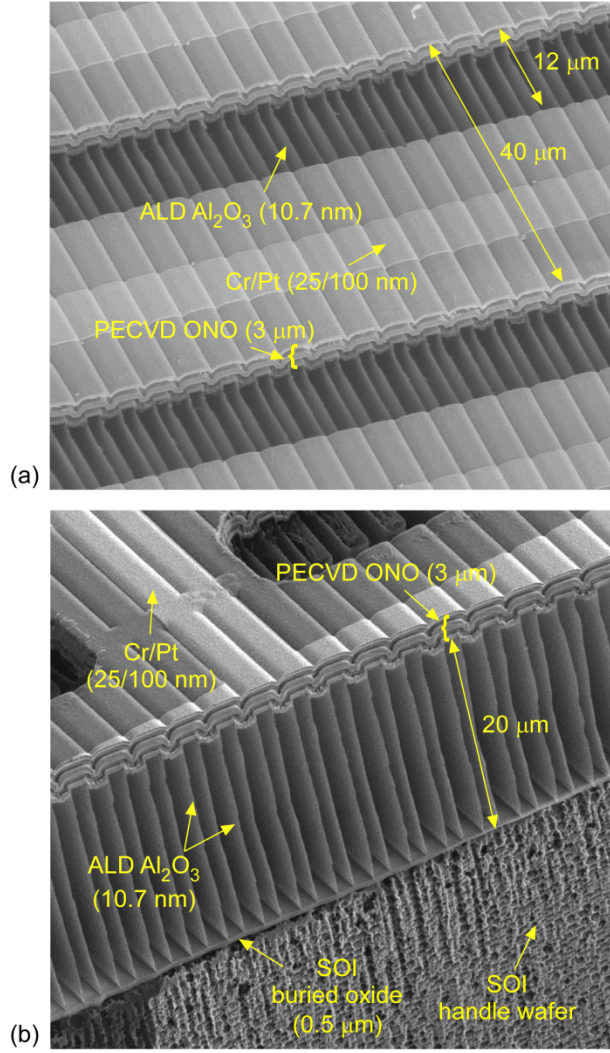
Figure 6.12: SEM images of ALD Al$_2$O$_3$ sidewalls. (a) Top view. (b) Cross-sectional view. To minimize a charging effect in SEM, a gold layer, thickness $\approx$10 nm, is sputtered.

6.4 Test Results

The bottom of the anchor frame of the fabricated Knudsen pump is attached to an external Al block that serves as a thermal ground to ambient room temperature. The Al block is a thin plate, volume $\approx$3 $\times$ 3 $\times$ 1 cm^3. For gas flow to the inlet of the Knudsen pump, the Al block has a machined hole, diameter $\approx$4 mm and length $\approx$15 mm. A commercial flow meter (Alicat Whisper^{TM}: 1 slpm and 400-500 Pa at full-scale flow, 10

ms response time) and pressure sensor (Freescale™ MPX 5010: 1 ms response time) were used to monitor the inlet of the Al block. The outlet of the Knudsen pump was open to the ambient.

The Knudsen pump generated a measured 211 sccm air flow at a pressure difference of 92 Pa, using an input power of 37.2 W; it also generated a maximum ΔP of 351 Pa (Fig. 6.13). The projected largest flow rate ($\Delta P = 0$ Pa) is 287 sccm.
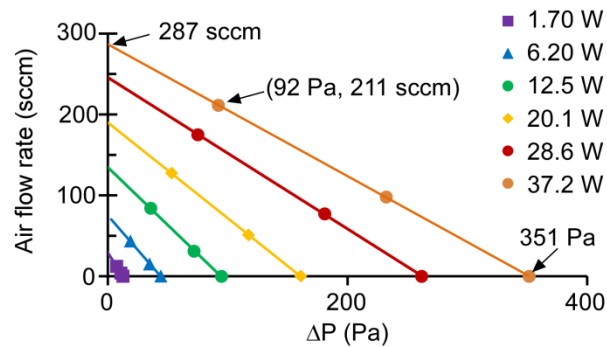
Figure 6.13: Steady state pump characteristics. Highest measured flow is 211 sccm using 37.2 W.
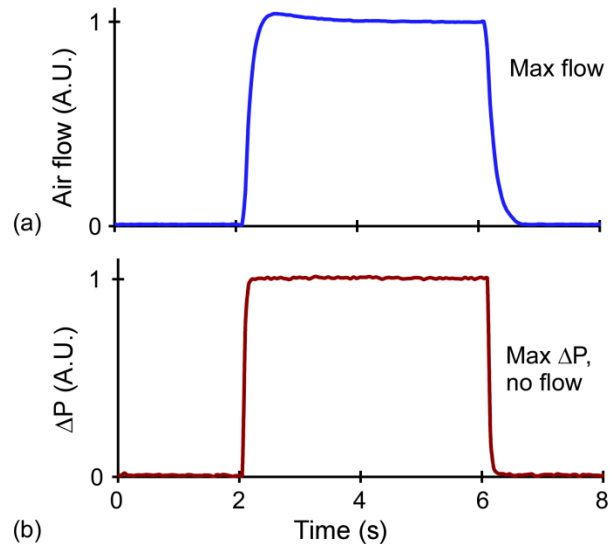
Figure 6.14: Transient response of a Knudsen pump. (a) Air flow rate. (b) Pressure difference. Typical response time for reaching 90% of the final state is 0.1-0.4 s.

The transient responses were tested by using a commercial switch (Panasonic[TM] AQW 215: 0.6 ms response time) to regulate the input power to the Knudsen pump. A customized LabVIEW[TM] program was used to control the switch and to read the flow meter and pressure sensor. Typical transient flow and ΔP responses are plotted (Fig. 6.14). Typical response time for achieving 90% of the final state is 0.1-0.4 s, which is ≈100 times faster than prior polymer-based Knudsen pumps [Pha10, Gup12b].

6.5 Discussion and Conclusion
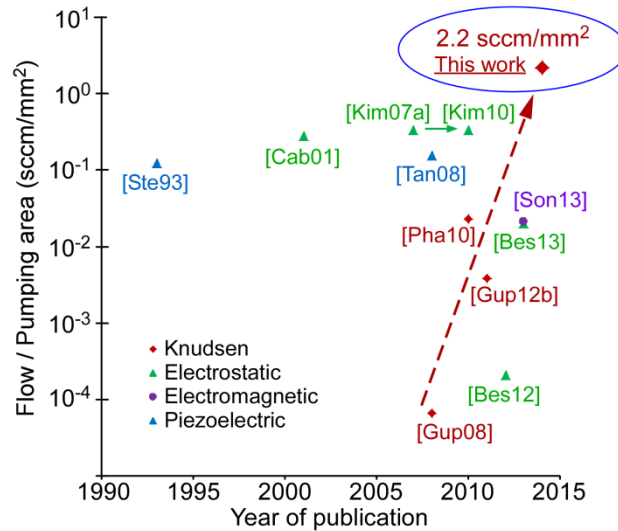
Figure 6.15: Gas flow rates in pumping areas of reported micropumps.

To benchmark various technologies, the flow rates in pumping areas of reported micropumps [Ste93, Cab01, Kim07a, Tan08, Kim10, Pha10, Gup08, Gup11b, Bes12, Son13, Bes13] are compared to the Knudsen pump in this chapter (Fig. 6.15). The unit-area gas flow rate in this chapter is 2.2 sccm/mm^2. Knudsen pumps have shown >10^4× increase in unit-area flow rate since 2008. The flow rate of 211 sccm, together with the

unit-area flow rate of 2.2 sccm/mm^2, are the highest values among micropumps in the literature.

The present design selected 2 μm as the narrow channel height due to fabrication constraints. However, this compromised the pressure difference to a value that is less than 1 kPa (Fig. 6.2). To design a performance line that covers more than 10 kPa that is required for a micro gas chromatograph (Fig. 1.4), the narrow channel height needs to be decreased to sub-micrometer. This can be explored using advanced lithography tools and scaling efforts of the present design in future work.

In summary, monolithic single-stage Knudsen pumps have been investigated for generating high flow rates. A high density of vertically oriented narrow channels was integrated in a single chip of 16×20 mm^2 area, using an SOI wafer and a four-masking step process. A notable result was the measured 211 sccm air flow rate at a pressure difference of 92 Pa, using an input power of 37.2 W. Using higher lithographic resolutions and scaled design, the pressure difference can be increased to a higher value at atmospheric ambient pressure.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

This PhD research explores higher pump performance using advanced design, fabrication, and test methods. The findings advanced the practical application of Knudsen pumps for use in miniaturized analytical instruments such as micro gas chromatographs and micro mass spectrometers. The essential outcomes are summarized as follows.

The previous, first, fully micromachined Knudsen pump which was reported by McNamara *et al.* [McN05] used a single-stage structure and a six-mask two-wafer process. In this work, a five-mask single-wafer process is utilized for fabricating a 48-stage Knudsen pump (Chapter 2), and a 162-stage Knudsen pump (Chapter 3). Pirani gauges are monolithically integrated with Knudsen pump for *in situ* pressure measurement.

For the 48-stage Knudsen pump, the footprint is 10.35×11.45 mm^2. The performance test results in compression ratios of ≈ 15 at 760 Torr, and ≈ 50, at 250 Torr ambient pressures using an input power of ≈ 1.35 W. The compression ratio of ≈ 50 is $24\times$ better than other micropumps reported in the literature.

For the 162-stage Knudsen pump, the footprint is 12×15 mm^2. The pressure at the sealed inlet is reduced from atmospheric ambient pressure to ≈ 0.9 Torr using an input

power of ≈ 0.39 W. The resulting compression ratio is ≈ 844, which is $\approx 17 \times$ larger with $3.5 \times$ better power efficiency than the previous 48-stage Knudsen pump.

The two-part architecture is first introduced to the multistage design of the 162-stage Knudsen pump. This architecture enhances both the compression ratio and the pumping speed in a wide pressure range, from atmospheric pressure to sub-Torr pressure. KHP and KLP are customized for their operating pressure ranges. The design, fabrication, and test for KHP and KLP are demonstrated.

In studying the 162-stage Knudsen pump, a dynamic calibration method for microfabricated Pirani gauges is first explored (Chapter 4). It is used for measuring accurate pressure levels of encapsulated stages in the 162-stage Knudsen pump. For this device, dynamic calibration improves the estimated upstream pressure from 30 Torr (as suggested by uncorrected static calibration) to 0.9 Torr, with a 95.4% confidence interval from 0.7 Torr to 1.1 Torr, assuming normal (Gaussian) distribution. The dynamic calibration can be extended to other certain type of devices, whose Pirani gauges are embedded deep in microfluidic networks.

In studying both the 48-stage and 162-stage Knudsen pumps, a low temperature vacuum sealing method using atomic layer deposition (ALD) of Al_2O_3 is investigated (Chapter 5). Lower temperatures provide compatibility with the fabrication processes used for the micro devices. Thin films of ALD Al_2O_3, which are pin-hole free and highly conformal, are found to be very promising for retaining vacuum seals. Lifetime tests of 19 months are reported.

The previous high flow rate Knudsen pump used nano porous materials [Gup11a, Gup11b]. However, the previous study results show that these materials have defect-

induced leakage. In this dissertation, a large number of parallel thermal transpiration flow channels are integrated directly in the silicon-on-insulator (SOI) device layer (Chapter 6). Deep reactive ion etching (DRIE), channel sidewall deposition, trench refilling by polySi, and subsequent sacrificial etching of silicon and polySi, form a high density of vertical narrow channels in a monolithic manner. The footprint is 16×20 mm^2, and the active pumping area is 96 mm^2. As a result, the flow rate of 211 sccm at a pressure difference of 92 Pa is achieved, using 37 W. The flow rate of 211 sccm, together with the unit-area flow rate of 2.2 sccm/mm^2, are the highest values among micropumps in the literature.

7.2 Future Work

Monolithic, multistage, high compression ratio Knudsen pumps (Chapter 2-3) and monolithic, single-stage, high flow rate Knudsen pumps (Chapter 6) can be combined for future work. Then, a single monolithic, chip-scale Knudsen pump with both high compression ratios and high flow rates can be fabricated.

The narrow channel height in Knudsen pumps needs to be similar to the mean free path of gas molecules for efficient generation of compression in a stage. The mean free paths of air molecules are 0.07 μm, 0.7 μm, and 7 μm at 760 Torr, 76 Torr, and 7.6 Torr, respectively. The lithography technology used in this work is capable of providing pattern resolutions of ≥1 μm. Therefore, when this technology is used, future Knudsen pumps can provide both a high compression ratio and a high flow rate near 76 – 7.6 Torr.

For a Knudsen pump to generate a high compression ratio and a high flow rate at atmospheric pressure, two possible approaches for meeting sub-micrometer narrow

channel height can be considered. First, a nanolithography technique such as e-beam and an RIE technology could provide sub-micrometer narrow channels (Fig. 7.1a). Then, the present design of a high flow rate Knudsen pump (Chapter 6) could be scaled down. Second, using a porous silicon technique, a sub-micrometer narrow channel height could be also possible (Fig. 7.1b). In near future, one of these two techniques or other new advanced techniques could enable a high compression ratio and a high flow rate Knudsen pump at atmospheric ambient pressure.
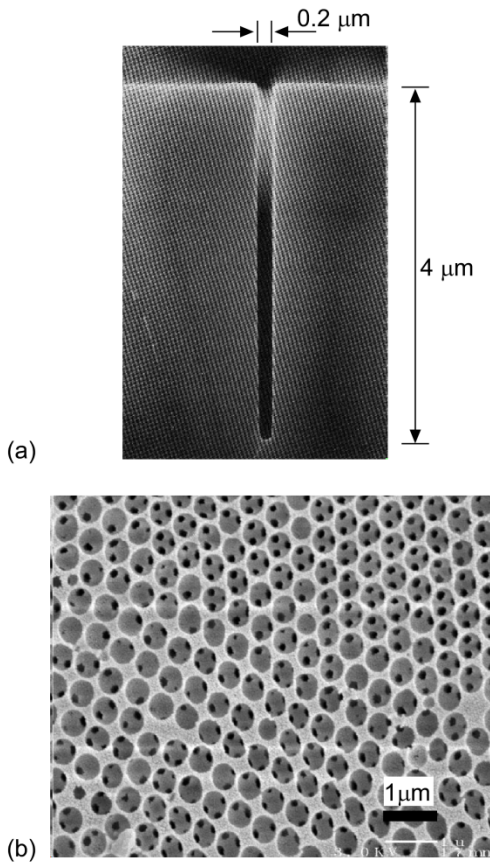
Figure 7.1: Potential candidates for sub-micrometer narrow channels for high compression and high flow Knudsen pump operating at atmospheric pressure. (a) Cross-sectional SEM image of deep trench etch in silicon [Lie05]. (b) SEM image of a porous silicon photonic crystal [Man13].

APPENDIX

THERMOELECTRIC EFFECT
USING KNUDSEN PUMP THEORY

This appendix describes the thermoelectric effect using Knudsen pump theory. Calculation results are presented, followed by conclusion.

A.1 Knudsen Pump Theory

This section briefly reviews background of gas flow regimes and transport equations. Viscous flow and thermal transpiration flow, respectively, for wide channel and narrow channel in a Knudsen pump are described. For simplicity and practicality, it is assumed that the ideal gas law governs in all flow channels.

A.1.1 Gas Flow Regimes and Transport Equation

Gas flow patterns are classified into three flow regimes [Jou08, Laf98]. They are identified using the Knudsen number, defined as the ratio of the mean free path of gas molecules to the hydraulic diameter of a flow channel:

$$Kn = \frac{\lambda}{d_{Hyd}} \tag{A.1}$$

where d_{Hyd} is the hydraulic diameter (or characteristic length) of the channel and λ is the mean free path of gas molecules. Here, λ is written as:

$$\lambda = \frac{k_B T}{\sqrt{2}\pi d_{Col}^2 P} \tag{A.2}$$

where d_{Col} is the collisional diameter of gas molecules. First, when Kn is larger than 10, the flow is in a free-molecular flow (or single-particle motion) regime. In this regime, molecule-to-wall collisions are dominant over molecule-to-molecule collisions. Second, when Kn is smaller than 0.01, the flow is in a viscous flow (or continuum flow) regime, where molecule-to-molecule collisions are dominant over molecule-to-wall collisions. Third, when Kn is between 0.1 and 10, the flow is in a transitional flow (or Knudsen flow) regime. It is between the free-molecular and viscous flow regimes.

The Boltzmann transport equation (BTE) is regarded as the fundamental transport equation that describes the behavior of gas molecules in all flow regimes. It is written as [Kav08]:

$$\frac{\partial f(\boldsymbol{x},\boldsymbol{p},t)}{\partial t} + \boldsymbol{v}_g \cdot \nabla_x f(\boldsymbol{x},\boldsymbol{p},t) + \boldsymbol{F} \cdot \nabla_p f(\boldsymbol{x},\boldsymbol{p},t) = \frac{df(\boldsymbol{x},\boldsymbol{p},t)}{dt}\Big|_{coll} \tag{A.3}$$

where $f(\boldsymbol{x},\boldsymbol{p},t)$ is the particle probability distribution function, whose position and the momentum are $\boldsymbol{x}$ and $\boldsymbol{p}$ at the time t; $\boldsymbol{v}_g$ is the velocity; ∇_x is the spatial gradient; $\boldsymbol{F}$ is the applied force; ∇_p is the momentum gradient; and $df(\boldsymbol{x},\boldsymbol{p},t)/dt|_{coll}$ on the right side is the collisional term, which is the time rate of change of $f(\boldsymbol{x},\boldsymbol{p},t)$ due to collisions. Due to dimensional curse – seven-dimensional space for one particle – in the BTE, the BTE is usually simplified using boundary conditions.

The Navier-Stokes equation (NSE) describes the behavior of gas molecules in the viscous flow regime. The NSE can be derived by applying Newton's second law to fluid motion or can be derived from the BTE in the viscous flow regime limit. For compressible fluids, it is written as:

$$\rho\left(\frac{\partial \boldsymbol{v}_g}{\partial t}+\left(\boldsymbol{v}_g \cdot \nabla\right)\boldsymbol{v}_g\right)=-\nabla P+\eta \nabla^2 \boldsymbol{v}_g+\left(\frac{1}{3}\eta+\varsigma\right)\nabla\left(\nabla \cdot \boldsymbol{v}_g\right)+F_B \tag{A.4}$$

where ρ is the mass density; η and ζ are the viscosity and second viscosity, respectively; P is the pressure; ∇ is the del operator; and F_B is the body force. For incompressible fluids, Eq. (A.4) is simplified as:

$$\rho\left(\frac{\partial \boldsymbol{v}_g}{\partial t}+\left(\boldsymbol{v}_g \cdot \nabla\right)\boldsymbol{v}_g\right)=-\nabla P+\eta \nabla^2 \boldsymbol{v}_g+F_B \tag{A.5}$$

A.1.2 Viscous Flow

In the viscous flow regime, the Navier-Stokes equation (NSE) is used for describing gas behaviors. A channel, where gas flow is confined to the viscous flow regime, is named a wide channel in a Knudsen pump. Due to dominant molecule-molecule collisions over molecule-wall collisions in the viscous flow regime, gas flow is directed toward uniform (kinetic) energies of gas molecules. Since the energy is only a function of temperature, uniform energies of gas molecules are equivalent to uniform pressure distribution, as anticipated from the ideal gas law. Here, since the energy refers to the kinetic energy, the energy density can be expressed as:

$$E_K = n_g \frac{1}{2} m v_g^{\,2} \tag{A.6}$$

where n_g *and m* are, respectively, the number density and the mass of gas molecules.

As a result of energy balance, different pressures drive gas flow such that the pressures become uniform; this flow is called Poiseuille flow or pressure-driven flow. Re-distribution of gas molecules occurs such that the product of n_g and $v_g^{\,2}$ becomes uniform. A non-uniform temperature distribution in a wide channel also causes re-

distribution of gas molecules. This could be in a form of pressure-driven flow, until uniform pressure distribution is attained.

A.1.3 Thermal Transpiration

The Knudsen pump exploits the phenomenon of thermal transpiration in a narrow channel, where gas flow is confined to the free-molecular or transitional flow regimes. Due to dominant molecule-wall collisions over molecule-molecule collisions in those regimes, gas flow in the narrow channel between the cold chamber and hot chamber is determined by gas flux. It is given by:

$$\Gamma = \frac{n_g \upsilon_g}{4}$$

(A.7)

where Γ is the gas flux defined as the number of collisions with a surface per unit area and time. Note that, when the gas is composed of single molecules, gas flux balance is equivalent to momentum balance.

As a result of gas flux balance, a temperature difference between the cold chamber and hot chamber causes the re-distribution of gas molecules that are directed toward equal fluxes – the product of n_g and υ_g – from both chambers. Note that, the re-distribution in the viscous flow regime is related to the product of n_g and υ_g^2. Therefore, the equilibrium, where opposing fluxes from the cold chamber and the hot chamber are equal, must result in a pressure difference or energy difference. The equilibrium pressure ratio, $(P_{Hot}/P_{Cold})_{Eq}$ can be written as:

$$\left(\frac{P_{Hot}}{P_{Cold}} \right)_{Eq} = \left(\frac{T_{Hot}}{T_{Cold}} \right)^{\gamma}$$

(A.8)

where γ is the flow coefficient ratio, defined in Section 2.2 [Sha96]; and T_{Cold} and T_{Hot} are, respectively, the temperatures at the cold and hot chambers.

A.2 Gas Transport to Electron Transport

Gas flow described in the previous section can be similarly applied to electron flow as long as the free electron gas model [Kav08] is used for describing electron transport in solid metal and/or semiconductor. In this case, pressure and flow rate in gas transport are replaced by voltage and current, respectively, in electron transport.

A.3 Electron Transport Theory

Calculation of electron flow in an imposed thermal gradient requires electron number density, average energy, and average velocity.

A.3.1 Electron Number Density

Electric current is proportional to electron number density. Although one or two electrons in the outermost shell in an atom are available to electric conduction, not all electrons in the conduction band are counted to electron number density. The electron number density at energy level from E to $E+dE$, $n(E)$, is obtained by the product of density of states, D, and Fermi-Dirac distribution function, f, as [Lun00]:

$$n(E) = D(E)f(E)dE \tag{A.9}$$

where E is the energy level. The electron number density, n, is calculated by integration at all energy levels as:

$$n = \int_0^\infty D(E)f(E)dE \qquad (A.10)$$

The density of states in 3D, 2D, and 1D are given, respectively, by:

$$D_{3D}(E) = \frac{\left(2m_e^*\right)^{3/2}}{2\pi^2 \hbar^3}\sqrt{\left(E - E_C\right)} \qquad (A.11)$$

$$D_{2D}(E) = \frac{m_e^*}{\pi \hbar^2} \qquad (A.12)$$

$$D_{1D}(E) = \frac{1}{\pi \hbar}\sqrt{\frac{2m_e^*}{\left(E - E_C\right)}} \qquad (A.13)$$

where m_e^* is the effective mass of electron; $\hbar$ is the Plank's constant divided by 2π; and E_C is the lowest energy level in the conduction band, which is set to zero in this Appendix. The differences in density of state among 3D, 2D, and 1D originate from quantum mechanical confinement of conduction electrons to 2D (plate) and 1D (line or nanowire).

The Fermi-Dirac distribution function, f, is given by:

$$f(E) = \frac{1}{1 + \exp\left(\dfrac{E - E_F}{k_b T}\right)} \qquad (A.14)$$

where E_F is the Fermi energy level, k_B is the Boltzmann constant, and T is the temperature. Since the value of E_C is set to zero, E_F for silicon is negative; whereas E_F for most of metal conductors, such as Al, Cu, is positive.

A.3.2 Average Energy

Using $D_{3D}(E)$, $D_{2D}(E)$, $D_{1D}(E)$, and $f(E)$, average energy of an electron can be given by:

$$E_{Avg,3D} = \frac{\int\limits_0^\infty E D_{3D}(E) f(E) dE}{\int\limits_0^\infty D_{3D}(E) f(E) dE} \qquad (A.15)$$

$$E_{Avg,2D} = \frac{\int\limits_0^\infty E D_{2D}(E) f(E) dE}{\int\limits_0^\infty D_{2D}(E) f(E) dE} \qquad (A.16)$$

$$E_{Avg,1D} = \frac{\int\limits_0^\infty E D_{1D}(E) f(E) dE}{\int\limits_0^\infty D_{1D}(E) f(E) dE} \qquad (A.17)$$

A.3.3 Average Velocity

An energy value of an electron can be simply converted to a velocity value, υ, as:

$$\upsilon = \sqrt{\frac{2E}{m_e^*}} \qquad (A.18)$$

Using $D_{3D}(E)$, $D_{2D}(E)$, $D_{1D}(E)$, and $f(E)$, average velocity of an electron can be given by:

$$\upsilon_{Avg,3D} = \frac{\int\limits_0^\infty \sqrt{\frac{2E}{m_e^*}} D_{3D}(E) f(E) dE}{\int\limits_0^\infty D_{3D}(E) f(E) dE} \qquad (A.19)$$

$$\upsilon_{Avg,2D} = \frac{\int\limits_0^\infty \sqrt{\frac{2E}{m_e^*}} D_{2D}(E) f(E) dE}{\int\limits_0^\infty D_{2D}(E) f(E) dE} \qquad (A.20)$$

$$\upsilon_{Avg,1D} = \frac{\int\limits_0^\infty \sqrt{\frac{2E}{m_e^*}} D_{1D}(E) f(E) dE}{\int\limits_0^\infty D_{1D}(E) f(E) dE} \qquad (A.21)$$

A.4 Thermoelectric Effect

Using Eq. (A.9-A.21), generated voltage values in an imposed temperature can be calculated. The Seebeck coefficient, S, is defined as [Lun13]:

$$S = \frac{\Delta V}{\Delta T} \tag{A.22}$$

where ΔV is the voltage difference and ΔT is the temperature difference.

A.4.1 Bulk or 3D Solid

The Seebeck coefficient of bulk or 3D solid material can be obtained by determining the change of E_F, ΔE_F, that satisfies the following ***energy balance equation*** using $D_{3D}(E)$:

$$E_{Avg,3D}\left(E_F,T_{Cold}\right)n_{3D}\left(E_F,T_{Cold}\right)=E_{Avg,3D}\left(E_F-\Delta E_F,T_{Hot}\right)n_{3D}\left(E_F-\Delta E_F,T_{Hot}\right) \tag{A.23}$$

In the case of most metals, E_F is positive. For Al, using E_F of 11.6 eV, T_{Cold} of 300 K, and T_{Hot} of 301 K, the calculated S is -0.94 μV/K, which is similar to the experimentally measured S of -1.8 μV/K [Kas05]. In the case of silicon, E_F is negative. For the n-type silicon, E_F ranges from 0 eV (heavy doping) to -0.55 eV (intrinsic Si). For Si, using E_F of -0.1 eV (n++ doping), T_{Cold} of 300 K, and T_{Hot} of 301 K, the calculated S is -551.3 μV/K, which is similar to the experimentally measured S of -200 μV/K to -500 μV/K.

A.4.2 Plate or 2D Solid

The Seebeck coefficient of plate or 2D solid material can be obtained by determining the change of E_F, ΔE_F, that satisfies the following ***momentum balance equation*** using $D_{2D}(E)$:

$$E_{Avg,2D}\left(E_F,T_{Cold}\right)n_{2D}\left(E_F,T_{Cold}\right)=E_{Avg,2D}\left(E_F-\Delta E_F,T_{Hot}\right)n_{2D}\left(E_F-\Delta E_F,T_{Hot}\right) \quad \text{(A.24)}$$

For Al, the calculated S is -0.31 µV/K, which is three times lower than that in 3D solid. For Si, the calculated S is -463.8 µV/K, which is very similar to that in 3D solid.

A.4.3 Nanowire or 1D Solid

The Seebeck coefficient of nanowire or 1D solid material can be obtained by determining the change of E_F, ΔE_F, that satisfies the following ***momentum balance equation*** using $D_{1D}(E)$:

$$E_{Avg,1D}\left(E_F,T_{Cold}\right)n_{1D}\left(E_F,T_{Cold}\right)=E_{Avg,1D}\left(E_F-\Delta E_F,T_{Hot}\right)n_{1D}\left(E_F-\Delta E_F,T_{Hot}\right) \quad \text{(A.25)}$$

For Al, the calculated S is 0.0 µV/K, which is many times lower than that in 3D solid. For Si, the calculated S is -420.5 µV/K, which is also very similar to that in 3D solid.

A.4.4 3D Solid versus 2D Solid

The differences between 3D and 3D solids are large when E_F is positive (as in Al) and small when E_F is negative (as in Si), as shown in Fig. A.1.
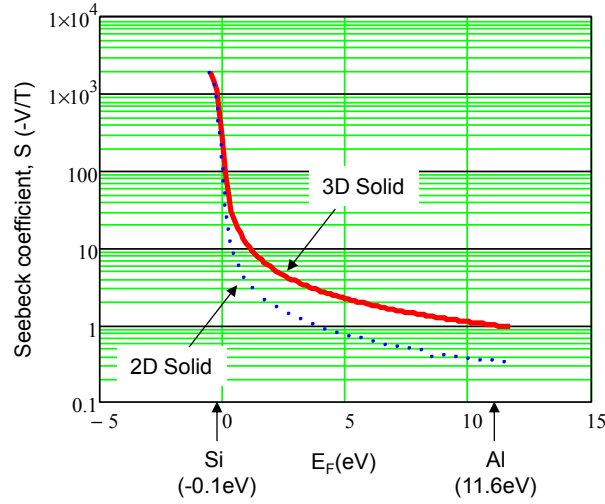
Figure A.1: Seebeck coefficient as a function of E_F for 3D solid and 2D solid.

In the case of gas molecules in multistage Knudsen vacuum pump, the pressure is lower at the cold end than at the hot end in the narrow channel; whereas the pressures are equal at both cold and hot end in the wide channel.

In the case of electronic transport, the voltage increases at the hot end in both 3D and 2D solids when the majority carrier is electron. This is because the electron in the hot end moves to the cold end, thereby resulting in electron depletion in the hot end. When the majority carrier is hole, the voltage increases in the cold end.

A.5. Potential Applications

A.5.1 Thermoelectric Voltage Using Multistage Design

For silicon, the differences in S between 3D and 2D are very small. Therefore, a serial cascade of narrow and wide channels provides only a small voltage.

For Al, the differences in S between 3D and 2D are large. Therefore, a serial cascade of narrow and wide channels combines voltage values at each stage. This combination can be achieved without using a serial cascade of differential materials. An

interesting thing is that a thin metal is analogous to a wide channel in a Knudsen pump; whereas a thick metal is analogous to a narrow channel in a Knudsen pump. A disadvantage of Al as a thermoelectric material is that the *S* value of Al is 2 orders lower than that of Si.

A.5.2 Thermoelectric Cooling

For comparison of cooling efficiency between various materials, the ZT figure of merit (FOM) is used. It is defined as [Lun13]:

$$ZT = \frac{S^2 \sigma_{el} T_C}{\kappa_{th}} \qquad (26)$$

where σ_{el} is the electric conductivity and κ_{th} is the thermal conductivity. Most efforts in research are focused on lowering κ_{th} while maintaining σ_{el}. Examples include silicon nanowire or bulk silicon plate with holes. These examples reduce thermal conductivity by confining phonon in small dimension or by inducing phonon scattering while maintaining the silicon electric conductivity. The thermal conductivity of crystalline solid could be reduced to the thermal conductivity of amorphous silicon dioxide thermal conductivity of 1.4 W/mK, which is close to theoretical limit in solid.

As shown in example calculations of Si, the difference in Seebeck coefficients of 3D and 2D solids is small. Therefore, if the thermal conductivity of a thin silicon film can be lowered, ZT FOM value is increased. For example, as in the high flow rate Knudsen pump (Chapter 6), vertically oriented (n-doped) polySi layers, which are trench refilled *in parallel*, could provide good electric conductivity. The thermal conductivity of a thin layer of polySi could be reduced due to phonon scattering and confinement, similar to nanowire. Parallel array is for high conductivity, but quantum confinement of density

of states as in $D_{2D}(E)$ needs to be considered for possible reduction in electric conductivity.

A.6 Conclusions

1) Thermoelectric effects are calculated assuming free electron gas model.

2) Momentum balance is used for calculating thermoelectric voltage in 2D and 1D solids.

3) Metal conductors, such as Al, can be used for obtaining thermoelectric voltage, using a cascade of thin and thick metal. Si cannot be used for such multistage design.

4) PolySi thin layers can be used for thermoelectric cooling, using parallel array. This needs more calculations for theoretical estimates.

REFERENCES

[Abq07] P. Abgrall and A-M. Gue, "Lab-on-chip Technologies: Making a Microfluidic Network and Coupling it into a Complete Microsystem - A Review," *Journal of Micromechanics and Microengineering*, vol. 17, no. 5, pp. R15-R49, 2007.

[Al06] A. Al-Azzawi, *Photonics: Principles and Practices*, CRC Press, pp. 336, 2006.

[Ale06] A.A. Alexeenko, S. F. Gimelshein, E. P. Muntz, and A. D. Ketsdever, "Kinetic Modeling of Temperature Driven Flows in Short Microchannels," *International Journal of Thermal Sciences*, vol. 45, no. 11, p 1045-51, Nov. 2006.

[An12] S. An, N.K. Gupta, and Y.B. Gianchandani, "A Monolithic 162-stage Two-part Knudsen Pump for High Compression Ratio," *Proceedings of the Solid State Sensors, Actuators and Microsystems Workshop (Hilton Head 2012)*, pp. 14-17, 2012.

[An13a] S. An, N.K. Gupta, and Y.B. Gianchandani, "A Si-Micromachined 162-Stage Two-Part Knudsen Pump for On-Chip Vacuum," *Journal of Microelectromechanical Systems*, to be published, 2013, DOI: 10.1109/JMEMS.2013.2281316.

[An13b] S. An and Y.B. Gianchandani, "A Dynamic Calibration Method for Pirani Gauges Embedded in Fluidic Networks," *Journal of Microelectromechanical Systems*, to be published, 2013, DOI: 10.1109/JMEMS.2013.2281319.

[An14] S. An, N.K. Gupta, and Y.B. Gianchandani, "Vacuum Sealing Using Atomic Layer Deposition of Al_2O_3 at 250°C," *Journal of Vacuum Science and Technology A*, vol. 32, pp. 01A101 (5 pages), 2014.

[Ant03] J. Antony, *Design of Experiments for Engineers and Scientist*, Elsevier, Chapter 2, 2003.

[Bee12] J.T.M. van Beek, and R. Puers, "A Review of MEMS Oscillators for Frequency Reference and Timing Applications," *Journal of Micromechanics and Microengineering*, vol. 22, no. 1, pp. 013001, 2012.

[Bel12] A. Bell, K. Pharas, W.D. Ehringer, and S. McNamara, "Body Temperature Powered Device for Dermal Wound Drug Delivery," *IEEE International Conference on Micro Electro Mechanical Systems (MEMS'12)*, pp. 930-932, 2012.

[Bel99] S. Bell, *A Beginner's Guide to Uncertainty of Measurement*, Crown, pp. 14, 1999.

[Bes12] A. Besharatian, K. Kumar, R.L. Peterson, L.P. Bernal, and K. Najafi, "A Scalable, Modular, Multi-Stage, Peristaltic, Electrostatic Gas Micro-Pump," *IEEE International Conference on Micro Electro Mechanical Systems (MEMS'12)*, pp. 1001-1004, 2012.

[Bes13] A. Besharatian, K. Kumar, R. L. Peterson, L. P. Bernal, and K. Najafi, "Valve-Only Pumping in Mechanical Gas Micropumps", *International Conference on Solid-State Sensors, Actuators and Microsystems (Transducers'13)*, pp. 2640-2643, 2013.

[Bru08] H. Bruus, *Theoretical Microfluidics*, Oxford University Press, New York, Chapter 3, 2008.

[Cab01] C. Cabuz, W.R. Herb, E.I. Cabuz, and T.L. Son, "The Dual Diaphragm Pump," *IEEE International Conference on Micro Electro Mechanical Systems (MEMS'01)*, pp. 519-522, 2001.

[Can06] R.N. Candler, M.A. Hopcroft, B. Kim, W-T. Park, R. Melamud, M. Agarwal, G. Yama, A. Partridge, M. Lutz, and T.W. Kenny, "Long-Term and Accelerated Life Testing of a Novel Single-Wafer Vacuum Encapsulation for MEMS Resonators," *Journal of Microelectromechanical Systems*, vol. 15, no. 6, pp. 1446-1456, 2006.

[Car09] P.F. Carcia, R.S. McLean, M.D. Groner, A.A. Dameron, and S.M. George, "Gas Diffusion Ultrabarriers on Polymer Substrates Using Al_2O_3 Atomic Layer Deposition and SiN Plasma-Enhanced Chemical Vapor Deposition," *Journal of Applied Physics*, vol. 106, no. 2, pp. 023533, 2009.

[Car11] M.R. Cardenas, I. Graur, P. Perrier, and J.G. Meolans, "Thermal Transpiration Flow: a Circular Cross-Section Microtube Submitted to a Temperature Gradient," *Physics of Fluids*, vol. 23, pp. 031702, 2011.

[Car12] P.F. Carcia, R.S. McLean, Z.G. Li, M.H. Reilly, and W.J. Marshall, "Permeability and Corrosion in ZrO_2/Al_2O_3 Nanolaminate and Al_2O_3 Thin Films Grown by Atomic Layer Deposition on Polymers," *Journal of Vacuum Science and Technology A*, vol. 30, n. 4, pp. 041515, 2012.

[Cha05] J. Chae, B.H. Stark, and K. Najafi, "A Micromachined Pirani Gauge with Dual Heat Sinks," *IEEE Transactions on Advanced Packaging*, vol. 28, no. 4, pp. 619-625, 2005.

[Che10] K. Cheung, L.F. Velasquez-Garcia, and A.I. Akinwande, "Chip-Scale Quadrupole Mass Filters for Portable Mass Spectrometry," *Journal of Microelectromechanical Systems*, vol. 19, no. 3, 2010.

[Chi06] M. Chiao and L. Lin, "Device-Level Hermetic Packaging of Microresonators by RTP Aluminum-to-Nitride Bonding," *Journal of Microelectromechanical Systems*, vol. 15, no. 3, pp. 515-522, 2006.

[Cho09] T.V. Chokshi, A. Ben-Yakar, and N. Chronis, "CO2 and Compressive Immobilization of C. Elegans On-Chip," *Lab on a Chip*, vol. 9, no. 1, pp. 151-157, 2009.

[Cho97] B.C.S. Chou and J-S. Shie, "An Innovative Pirani Pressure Sensor," *International Conference on Solid-State Sensors, Actuators and Microsystems (Transducers'97)*, pp. 1465-1468, 1997.

[Coc12] A.M. Coclite, F.D. Luca, and K.K. Gleason, "Mechanically Robust Silica-Like Coatings Deposited by Microwave Plasmas for Barrier Applications," *Journal of Vacuum Science and Technology A*, vol. 30, n. 6, pp. 061502, 2012.

[Col05] S. Colin, "Rarefaction and Compressibility Effects on Steady and Transient Gas Flows in Microchannels," *Microfluidics and Nanofluidics*, vol. 1, no. 3, pp. 268-279, 2005.

[Dar96] P. Dario, N. Croce, M.C. Carrozza, and G. Varallo, "A Fluid Handling System for a Chemical Microanalyzer," *Journal of Micromechanics and Microengineering*, vol. 6, pp. 95-98, 1996.

[Dic12] E. Dickey and W.A. Barrow, "High Rate Roll to Roll Atomic Layer Deposition, and its Application to Moisture Barriers on Polymer Films," *Journal of Vacuum Science and Technology A*, vol. 30, n. 2, pp. 021502, 2012.

[Dol05] A. Doll, M. Heinrichs, F. Goldschmidtboeing, H.J. Schrag, U.T. Hopt, and P. Woias, "A High Performance Bidirectional Micropump for a Novel Artificial Sphincter System," *Sensors and Actuators A: Physical*, vol. 14, pp. 445-453, 2005.

[Dom05] M. Doms, A. Bekesch, and J. Mueller, "A Microfabricated Pirani Pressure Sensor Operating Near Atmospheric Pressure," *Journal of Micromechanics and Microengineering*, vol. 15, no. 8, pp. 1504-1510, 2005.

[Dom07] M. Doms and J. Muller, "Design, Fabrication, and Characterization of a Micro Vapor-Jet Vacuum Pump," *ASME Journal of Fluids Engineering*, vol. 129, pp. 1339-1345, 2007.

[Dom11] M. Doms, A. Bekesch, and J. Mueller, "A Microfabricated Pirani Pressure Sensor Operating Near Atmospheric Pressure," *Journal of Micromechanics and Microengineering*, vol. 15, no. 8, pp. 1504-1510, 2011.

[DTR12] Defense Threat Reduction Agency (DTRA) solicitation website, https://www.fbo.gov/utils/view?id=73b4f6a6e141f3787dd6c3fc9b735411, accessed September, 2012.

[Ell00] R.E. Ellefson and A.P. Miiller, "Recommended Practice for Calibrating Vacuum Gauges of the Thermal Conductivity Type," *Journal of Vacuum Science and Technology A*, vol. 18, no. 5, pp. 2568-2577, 2000.

[Esa08] M. Esashi, "Wafer Level Packaging of MEMS," *Journal of Micromechanics and Microengineering*, vol. 18, no. 7, pp. 073001, 2008.

[Geo10] S.M. George, "Atomic Layer Deposition: An Overview," *Chemical Review*, vol. 110, no. 1, pp. 111-131, 2010.

[Ger95] T. Gerlach and H. Wurmus, "Working Principle and Performance of the Dynamic Micropump," *Sensors and Actuators A*, vol. 50, pp. 135-140, 1995.

[Gia06] Y.B. Gianchandani, C.G. Wilson, and J-S. Park, "Micromachined Pressure Sensors: Devices, Interface Circuits, and Performance Limits," *The MEMS Handbook 2nd Edition, MEMS Applications*, Edited by M. Gad-el-Hak, CRC, Chapter 3, 2006.

[Gri46] E.R. Grilly, W.J. Taylor, and H.L. Johnston, "Accommodation Coefficients for Heat Conduction Between Gas and Bright Platinum, for Nine Gases Between 80°K (or Their Boiling Points) and 380°K," *Journal of Chemical Physics*, vol. 14, no. 7, pp. 435-440, 1946.

[Guc84] H. Guckel, and D.W. Burns, "Planar Processed Polysilicon Sealed Cavities for Pressure Transducer Arrays", *IEEE International Electron Devices Meeting (IEDM '86)*, pp. 223-225, 1984.

[Guc89] H. Guckel and D.W. Burns, "Sealed Cavity Semiconductor Pressure Transducers and Method of Producing the Same," *U.S. Patent 4 853 669*, 1989.

[Guo09] X. Guo, D. Singh, J. Murthy, and A.A. Alexeenko, "Numerical Simulation of Gas-Phonon Coupling in Thermal Transpiration Flows," *Physical Reviews E*, vol. 80, pp. (046310) 1-10, 2009.

[Guo12] B. Guo, B. Wang, L. Wen, P. Helin, G. Claes, J.D. Coster, B. D. Bois, A. Verbist, R.V. Hoof, G. Vereecke, L. Haspeslagh, H.A.C. Tilmans, S. Decoutere, H. Osman, R. Puers, I.D. Wolf, S. Tanaka, S. Severi, and A. Witvrouw, "Poly-SiGe-Based MEMS Thin-Film Encapsulation," *Journal of Microelectromechanical Systems*, vol. 21, no. 1, pp. 110-120, 2012.

[Gup08] N.K. Gupta and Y.B. Gianchandani, "Thermal Transpiration in Zeolites: A Mechanism for Motionless Gas Pumps," *Applied Physics Letters*, vol. 93, no. 19, pp. 193511 (3 pages), 2008.

[Gup10] N.K. Gupta, "A Motionless Gas Micropump Using Thermal Transpiration in Bulk Nanoporous Materials", *Ph.D. Dissertation*, University of Michigan, 2010.

[Gup11a] N.K. Gupta and Y.B. Gianchandani, "Porous Ceramics for Multistage Knudsen Micropumps - Modeling Approach and Experimental Evaluation," *Journal of Micromechanics and Microengineering*, vol. 21, no. 9, pp. 095029 (14 pages), 2011.

[Gup11b] N.K. Gupta and Y.B. Gianchandani, "Thermal Transpiration in Mixed Cellulose Ester Membranes: Enabling Miniature, Motionless Gas Pumps," *Microporous and Mesoporous Materials*, vol. 142, no. 2-3, pp. 535-541, 2011.

[Gup12] N. Gupta, S. An, and Y. Gianchandani, "A Si-Micromachined 48-Stage Knudsen Pump for On-Chip Vacuum," *Journal of Micromechanics and Microengineering*, vol. 22, no. 10, pp. 105026 (8 pages), 2012.

[Han07] Y-L. Han and E.P. Muntz, "Experimental Investigation of Micro-Mesoscale Knudsen Compressor Performance at Low Pressures," *Journal of Vacuum Science and Technology B*, vol. 25, no. 3, pp. 703-714, 2007.

[Hau07] J.P. Hauschild, E. Wapelhorst, and J. Muller, "Mass Spectra Measured by a Fully Integrated MEMS Mass Spectrometer," *International Journal of Mass Spectrometry*, vol. 264, no. 1, pp. 53-60, 2007.

[Her87] A.W. van Herwaarden and P.M. Sarro, "Double-Beam Integrated Thermal Vacuum Sensor," *Journal of Vacuum Science and Technology A*, vol. 5, no. 4, pp. 2454-2457, 1987.

[Ike90] K. Ikeda, H. Kuwayama, T. Kobayashi, T. Watanabe, T. Nishikawa, T. Yoshida, and K. Harada, "Three-dimensional Micromachining of Silicon Pressure Sensor Integrating Resonant Strain Gauge on Diaphragm," *Sensors and Actuators A*, vol. 23, pp. 1007- 1010, 1990.

[Jia06] Y-B. Jiang, N. Liu, H. Gerung, J.L. Cecchi, and C.J. Brinker, "Nanometer-Thick Conformal Pore Sealing of Self-Assembled Mesoporous Silica by Plasma-Assisted Atomic Layer Deposition," *Journal of American Chemical Society*, vol. 128, n. 34, pp. 11018-11019, 2006.

[Jo05] S-B. Jo, M-W. Lee, S-G. Lee, E-H. Lee, S-G. Park, and B-H. O, "Characterization of a Modified Bosch-Type Process for Silicon Mold," *Journal of Vacuum Science and Technology A*, vol. 23, no. 4, pp. 905-910, 2005. doi: 10.1116/1.1943467.

[Jou08] K. Jousten, *Handbook of Vacuum Technology*, WILEY-VCH, Chapter 1-5 and pp. 921, 2008.

[Kac08] R. Kach and C. Kissling, "Diaphragm Pump," *US Patent* 7 373 872, 2008.

[Kad85] K. Kadoya, N. Matsunaga, and A. Nagahima, "Viscosity and Thermal Conductivity of Dry Air in the Gaseous Phase," *Journal of Physical and Chemical Reference Data*, vol. 14, no. 4, pp. 947-970, 1985.

[Kam98] K.P. Kamper, J. Dopper, W. Ehrfeld, and S. Oberbeck, "A Self-Filling Low-Cost Membrane Micropump," *IEEE International Conference on Micro Electro Mechanical Systems (MEMS'98)*, pp. 432-437, 1998.

[Kas05] S.O. Kasap, *Principles of Electronic Materials and Devices 3rd Edition*, McGraw-Hill, Chapter 4, 2001.

[Kav08] M. Kaviany, *Heat Transfer Physics*, Cambridge University Press, Chapter 2-3, 2008.

[Kes98] P.F. van Kessel, L.J. Hornbeck, R.E. Meier, and M.R. Douglass, "A MEMS-Based Projection Display," *Proceedings of the IEEE*, vol. 86, no. 8, pp. 1687-1704, 1998.

[Keu12] W. Keuning, P. van de Weijer, H. Lifka, W.M.M. Kessels, and M. Creatore, "Cathode Encapsulation of Organic Light Emitting Diodes by Atomic Layer Deposited Al_2O_3 Films and Al_2O_3/a-SiN$_x$:H Stacks," *Journal of Vacuum Science and Technology A*, vol. 30, n. 1, pp. 01A131, 2012.

[Kim07a] H. Kim, A.A. Astle, K. Najafi, L.P. Bernal, and P.D. Washabaugh, "A Fully Integrated High-Efficiency Peristaltic 18-Stage Gas Micropump with Active Microvalves," *IEEE International Conference on Micro Electro Mechanical Systems (MEMS'07)*, pp. 127-130, 2007.

[Kim07b] H. Kim, W. Steinecker, S. Reidy, G. Lambertus, A. Astle, K. Najafi, E. Zellers, L. Bernal, P. Washabaugh, and K. Wise, "Micropump-Driven High-Speed MEMS Gas Chromatography System," *International Conference on Solid-State Sensors, Actuators and Microsystems (Transducers '07),* pp. 1505-1508, 2007.

[Kim08] H. Kim, K. Najafi, and L. Bernal, "Gas Micropumps," *Comprehensive Microsystems Volume 2*, Edited by Y.B. Gianchandani, O. Tabata, and H. Zappe, Elsevier Publishers, pp. 273-299, 2008.

[Kim10] H. Kim, K. Najafi, and L.P. Bernal, "Helmholtz Resonance Based Micro Electrostatic Actuators for Compressible Gas Control: A Microjet Generator and a Gas Micro Pump," *Journal of Microelectronics and Electronic Packaging*, vol. 7, no. 1, pp. 1-9, 2010.

[Kim11] H.T. Kim, J.W. Park, and H. Kim, "All-Electric Peristaltic Vacuum Pump Utilizing Electromagnetic and Hydraulic Actuation with a Highly Flexible Latex Membrane," *International Conference on Solid-State Sensors, Actuators and Microsystems (Transducers'11)*, pp. 2454-2457, 2011.

[Knu09] M. Knudsen, "Eine Revision der Gleichgewichtsbedingung der Gase. Thermische Molekularstromung," *Annalen der Physik*, Leipzig, vol. 336, no. 1, pp. 205-229, 1909. (in German)

[Kos08] S. Kosuge and S. Takata, "Database for Flows of Binary Gas Mixtures through a Plane Microchannel," *European Journal of Mechanics B/Fluids*, vol. 27, pp. 444-465, 2008

[Kre06] E. Kreyszig, *Advanced Engineering Mathematics 9th Edition*, John Wiley & Sons, Inc., pp. 1049-1057, 2006.

[Laf98] J.M. Lafferty, *Foundations of Vacuum Science and Technology*, John Wiley & Sons, Inc., Chapter 1-2, 1998.

[Las04] D.J. Laser and J.G. Santiago, "A Review of Micropumps," *Journal of Micromechanics and Microengineering*, vol. 14, no. 6, pp. R35-R64, 2004.

[Lea03] A.M. Leach, A.R. Wheeler, R.N. Zare, "Flow Injection Analysis in a Microfluidic Format," *Analytical Chemistry*, vol. 75, no. 4, pp. 967-972, 2003.

[Lee11] Y.C. Lee, "Atomic Layer Deposition/Molecular Layer Deposition for Packaging and Interconnect of N/MEMS," *Proceedings of SPIE: Reliability, Packaging, Testing, and Characterization of MEMS/MOEMS and Nanodevices X*, vol. 7928, pp. 792802, 2011.

[Li08] Q. Li, J.F.L. Goosen, J.T.M. van Beek, F. van Keulen, K.L. Phan, and G.Q. Zhang, "Failure Analysis of a Thin-Film Nitride MEMS Package," *Microelectronics Reliability*, vol. 48, pp. 1557-1561, 2008.

[Lie05] M.A. Lieberman and A.J. Lichtenberg, *Principles of Plasma Discharges and Materials Processing 2nd Edition*, John Wiley & Sons, pp. 2, 2005.

[Lin10] C. Lin, Y. Zhang, A. Abdulagatov, R.Yang, S. George, and Y.C. Lee, "ALD-Enabled Hermetic Sealing for Polymer-Based Wafer Level Packaging of MEMS," *IEEE International Conference on Micro Electro Mechanical Systems (MEMS'10)*, pp. 528-531, 2010.

[Lin98] L. Lin, R.T. Howe, and A.P. Pisano, "Microelectromechanical Filters for Signal Processing," *Journal of Microelectromechanical Systems*, vol. 7, no. 3, pp. 286-294, 1998.

[Liu09] K. Liu, W. Zhang, W. Chen, K. Li, F. Dai, F. Cui, X. Wu, G. Ma, and Q. Xiao, "The Development of Micro-Gyroscope Technology," *Journal of Micromechanics and Microengineering*, vol. 19, no. 11, pp. 113001, 2009.

[Liu11] J. Liu, N.K. Gupta, K.D. Wise, Y.B. Gianchandani, and X. Fan, "Demonstration of Motionless Knudsen Pump Based Micro-Gas Chromatography Featuring Micro-Fabricated Columns and On-Column Detectors," *Lab on a Chip*, vol. 11, pp. 3487-3492, 2011.

[Liu99] C. Liu and Y.C. Tai, "Sealing of Micromachined Cavities Using Chemical Vapor Deposition Methods: Characterization and Optimization," *Journal of Microelectromechanical Systems*, vol. 8, no. 2, pp. 135-145, 1999.

[Loe34] L. Loeb, *The Kinetic Theory of Gases*, McGraw Hill, pp. 355-359, 1934.

[Lun00] M. Lundstrom, *Fundamentals of Carrier Transport 2nd Edition*, Oxford University Press, Chapter 1, 2000.

[Lun13] M. Lundstrom and C. Jeong, *Near-Equilibrium Transport Fundamentals and Applications*, World Scientific Publishing, Chapter 4-5, 2013.

[Man13] Manoharan lab's website, http://manoharan.seas.harvard.edu/siliconpc-doc.html, accessed September, 2013.

[Mar99] R. Marsili, "Lab-on-a-chip Poised to Revolutionize Sample Preparation," *R&D: Research and Development Kobe Steel Engineering Reports*, vol. 41, no. 2, pp. 34-40, Feb 1999.

[Mas07] N.D. Masters and W. Ye, "Octant Flux Splitting Information Preservation DSMC Method for Thermally Driven Flows," *Journal of Computational Physics*, vol. 226, no. 2, pp. 2044-2062, 2007.

[Mas91] C.H. Mastrangelo and R.S. Muller, "Microfabricated Thermal Absolute-Pressure Sensor with On-Chip Digital Front-End Processor," *IEEE Journal of Solid-State Circuits*, vol. 26, no. 12, pp. 1998-2007, 1991.

[McN05] S. McNamara and Y.B. Gianchandani, "On-Chip Vacuum Generated by a Micromachined Knudsen Pump," *Journal of Microelectromechanical Systems*, *vol*. 14, no. 4, pp. 741-746, 2005.

[Mey11] R. Meyknecht, "Micropump and Adhesive-Free Method for Joining Two Substrates," *US Patent* 8 043 073, 2011.

[Mit08] J. Mitchell, G.R. Lahiji, and K. Najafi, "An Improved Performance Poly-Si Pirani Vacuum Gauge Using Heat-Distributing Structural Supports," *Journal of Microelectromechanical Systems*, vol. 17, no. 1, pp. 93-102, 2008.

[Mun97] R.G. Munro, "Evaluated Material Properties for a Sintered α-Alumina," *Journal of American Ceramic Society*, vol. 80, no. 8, pp. 1919-1928, 1997.

[Mun02] E.P. Muntz, Y. Sone, K. Aoki, S. Vargo, and M. Young, "Performance Analysis and Optimization Considerations for a Knudsen Compressor in Transitional Flow," *Journal of Vacuum Science & Technology A*, vol. 20, no. 1, pp. 214-224, 2002.

[Mun06] E. P. Muntz, M. Young, and S. E. Vargo, "Microscale Vacuum Pumps," *The MEMS Handbook 2nd Edition, MEMS Applications*, Edited by M. Gad-el-Hak, CRC, Chapter 8, 2006.

[Nag98] H.T. nagel, S.S. Schiffman, and R. Gutierrez-Osuna, "How and why of electronic noses," *IEEE Spectrum*, vol. 35, no. 9, pp. 22-34, Sep 1998.

[Nam10] S-W. Nam, M-H. Lee, S-H. Lee, D-J. Lee, S.M. Rossnagel, and K-B. Kim, "Sub-10-nm Nanochannels by Self-Sealing and Self-Limiting Atomic Layer Deposition," *Nano Letters*, vol. 10, pp. 3324-3329, 2010.

[Ngu02] N-T. Nguyen, X. Huang, and T.K. Chuan, "MEMS-Micropumps: A review," *Transactions of ASME Journal of Fluids Engineering*, vol. 124, no. 22, pp. 384-392, 2002.

[Oxf13] Oxford ALD website, http://www.oxford-instruments.com/products/etching-deposition-and-growth/plasma-etch-deposition/atomic-layer-deposition, accessed March, 2013.

[Pha10] K. Pharas and S. McNamara, "Knudsen Pump Driven by a Thermoelectric Material," *Journal of Micromechanics and Microengineering*, vol. 20, no. 12, pp. 125032 (7 pages), 2010.

[Pot07] J.A. Potkay, G.R. Lambertus, R.D. Sacks, and K.D. Wise, "A Low-Power Pressure- and Temperature-Programmable Micro Gas Chromatography Column," *Journal of Microelectromechanical Systems*, vol. 16, no. 5, pp. 1071-1079, 2007.

[Puu05] R.L. Puurunen, "Surface Chemistry of Atomic Layer Deposition: A Case Study for the Trimethylaluminum/Water Process," *Journal of Applied Physics*, vol. 97, pp. 121301, 2005.

[Qu00] W. Qu, W. Wlodarski, and M. Austin, "Microfabrication and Reliability Study of Sapphire Based Ti/Pt-Electrodes for Thin-Film Gas Sensor Applications," *Microelectronics Journal*, vol. 31, no. 7, pp. 561-567, 2000.

[Red99] P.A. Redhead, "The Ultimate Vacuum," *Vacuum*, vol. 53, no. 1-2, pp. 137-149, 1999.

[Roj13] M. Rojas-Cardenas, I. Graur, P. Perrier, and J.G. Meolans, "Time-Dependent Experimental Analysis of a Thermal Transpiration Rarefied Gas Flow," *Physics of Fluids*, vol. 25, no. 7, pp. 072001, 2013.

[Sai10] M. Saidi and R.H. Abardeh, "Air Pressure Dependence of Natural-Convection Heat Transfer," *Proceedings of the World Congress on Engineering (WCE 2010)*, 2010.

[San11] F. Santagata, E. Iervolino, L. Mele, A.W. van Herwaarden, J.F. Creemer, and P.M. Sarro, "An Analytical Model and Verification for MEMS Pirani Gauges," *Journal of Micromechanics and Microengineering*, vol. 21, no. 11, pp. 115007, 2011.

[Sch02] C.G.J. Schabmueller, M. Koch, M.E. Mokhtari, A.G.R. Evans, A. Brunnschweiler, and H. Sehr, "Self-Aligning Gas/Liquid Micropump," *Journal of Micromechanics and Microengineering*, vol. 12, no. 4, pp. 420-424, 2002.

[Sch94] W.K. Schomburg, J. Vollmer, B. Bustgens, J. Fahrenberg, H. Hein, and W. Menz, "Microfluidic Components in LIGA Technique," *Journal of Micromechanics and Microengineering*, vol. 4, no. 4, pp. 186-191, 1994.

[Sen01] S.D. Senturia, *Microsystem Design*, Kluwer Academic Publishers, pp. 278-282, 2001.

[Sha96] F. Sharipov, "Rarefied Gas Flow through a Long Tube at any Temperature Ratio," *Journal of Vacuum Science and Technology A*, vol. 14, no. 4, pp. 2627-2635, 1996.

[Sha98] F. Sharipov and V. Seleznev, "Data on Internal Rarefied Gas Flows," *Journal of Physical and Chemical Reference Data*, vol. 27, no. 3, pp. 657-706, 1998.

[Sha99] F. Sharipov, "Non-Isothermal Gas Flow through Rectangular Microchannels," *Journal of Micromechanics and Microengineering*, vol. 9, pp. 394-401, 1999.

[Smi07] W.J. Smith, *Modern Optical Engineering 4th Edition*, McGraw-Hill Professional, pp. 14-16, 2007.

[Son13] J. Son, H. Kim, and H. Kim, "Pneumatic-Less High-Speed Vacuum Meso-Pump Driven by Programmable Hydraulics," *IEEE International Conference on Micro Electro Mechanical Systems (MEMS '13)*, pp. 584-587, 2013.

[Sta93] I. Stark, M. Stordeur, and F. Syrowatka, "Thermal Conductivity of Thin Amorphous Alumina Films," *Thin Solid Films*, vol. 226, pp. 185-190, 1993.

[Sta03] B.H. Stark, Y. Mei, C. Zhang, and K. Najafi, "A Doubly Anchored Surface Micromachined Pirani Gauge for Vacuum Package Characterization," *IEEE International Conference on Micro Electro Mechanical Systems (MEMS '03)*, pp. 506-509, 2003.

[Ste85] K. Stephan and A. Laesecke, "The Thermal conductivity of Fluid Air," *Journal of Physical and Chemical Reference Data*, vol. 14, no. 1, pp. 1227-1234, 1985.

[Ste93] E. Stemme and G Stemme, "A Valveless Diffuser/Nozzle-Based Fluid Pump," *Sensors and Actuators A: Physical*, vol. 39, pp. 159-167, 1993.

[Str96] R.L. Street, G.Z. Watters, and J.K. Vennard, *Elementary Fluid Mechanics*, John Wiley & Sons, pp. 125-141, 1996.

[Sun04] R. Sunier, P. Monajemi, F. Ayazi, T. Vancura, H. Baltes, and O. Brand, "Precise Release and Insulation Technology for Vertical Hall Sensors and Trench−Defined MEMS," *IEEE Sensors*, pp. 1442−1445, 2004.

[Sym09] R.R.A. Syms, "Advances in Microfabricated Mass Spectrometer," *Analytical and Bioanalytical Chemistry*, vol. 393, no. 2, pp. 427-429, 2009.

[Tan08] K. Tanaka, V.T. Dau, R. Sakamoto, T.X. Dinh, D.V. Dao, and S. Sugiyama, "Fabrication and Basic Characterization of a Piezoelectric Valveless Micro Jet Pump," *Japanese Journal of Applied Physics*, pp. 8615-8618, vol. 47, no. 11, 2008.

[Tas11] C. Tassetti, L. Duraffourg, J. Danel, T. Lagutère1, and F. Progent, "Development of a MEMS mass spectrometer based on TOF architecture for gas analysis," *8th Workshop on Harsh-Environment Mass Spectrometry, St. Pete Beach (USA)*, 2011.

[Ter79] S. Terry, J.H. Jerman, and J.B. Angell, "A Gas Chromatographic Air Analyzer Fabricated on a Silicon Wafer," *IEEE Transactions on Electron Devices*, vol. 26, pp. 1880-86, 1979.

[Tsa07] N-C. Tsai, and C-Y. Sue, "Review of MEMS Based Drug Delivery and Dosing Systems," *Sensors and Actuators A: Physical*, vol. 134, no. 2, pp.555-564, 2007.

[Tsu01] T. Tsuchiya, Y. Kageyama, H. Funabashi, and J. Sakata, "Polysilicon Vibrating Gyroscope Vacuum-Encapsulated in an On-Chip Micro Chamber," *Sensors and Actuators A*, vol. 90, pp. 49-55, 2001.

[Ubi51a] H.V. Ubisch, "On the Conduction of Heat in Rarefied Gases and its Manometric Application . I," *Applied Scientific Research*, vol. 2, no. 1, pp. 364-430, 1951.

[Ubi51b] H.V. Ubisch, "On the Conduction of Heat in Rarefied Gases and its Manometric Application . II," *Applied Scientific Research*, vol. 2, no. 1, pp. 403-430, 1951.

[Var99] S.E. Vargo, E.P. Muntz, G.R. Shiflett, and W.C. Tang, "Knudsen Compressor as a Micro- and Macroscale Vacuum Pump without Moving Parts or Fluids," *Journal of Vacuum Science & Technology A*, vol. 17, no. 4, pp. 2308-2313, 1999.

[Var01] S.E. Vargo and E.P. Muntz, "Initial Results from the First MEMS Fabricated Thermal Transpiration-Driven Vacuum Pump," AIP Conference Proceedings, no. 585, pp. 502-509, 2001.

[Vin75] W.G. Vincenti and C.H. Kruger, *Introduction to Physical Gas Dynamics*, pp. 12-15, 1975.

[Wan05] C.H. Wang, and G-B. Lee, "Automatic Bio-Sampling Chips Integrated With Micro-Pumps and Micro-Valves for Disease Detection," *Biosensors and Bioelectronics*, vol. 21, no. 3, pp. 419-425, 2005.

[Wan06] C.H. Wang and G-B. Lee, "Pneumatic Driven Peristaltic Micropumps Utilizing Serpentine-Shape Channels," *Journal of Micromechanics and Microengineering*, vol. 16, no. 2, pp. 341-348, 2006.

[Wan12] L. Wang, J.J. Travis, A.S. Cavanagh, X. Liu, S.P. Koenig, P.Y. Huang, S.M. George, and J.S. Bunch, "Ultrathin Oxide Films by Atomic Layer Deposition on Graphene," *Nano Letters*, vol. 12, pp. 3706-3710, 2012.

[Web10] L. Weber, "Micropump and Method for the Production Thereof," *US Patent* 7 771 176, 2010.

[Wen94] P.K. Weng and J. Shie, "Micro-Pirani Vacuum Gauge," *Review of Scientific Instruments*, vol. 65, no. 2, pp. 492-499, 1994.

[Wij00] W. Wijngaart, H. Andersson, and P. Enoksson, "The First Self-Priming and Bi-Directional Valve-Less Diffuser Micropump for Both Liquid and Gas," *IEEE International Conference on Micro Electro Mechanical Systems (MEMS'00)*, pp. 674-679, 2000.

[Wil04] S. Wilfert and C. Edelmann, "Miniaturized Vacuum Gauges," *Journal of Vacuum Science and Technology A*, vol. 22, no. 2, pp. 309-320, 2004.

[Wis04] K.D. Wise, K. Najafi, R.D. Sacks, and E.T. Zellers, "A Wireless Integrated Microsystem for Environmental Monitoring," IEEE International Solid State Circuits Conference, v. 1, pp. 434-537, 2004.

[Wil05] C.A. Wilson, R.K. Grubbs, and S.M. George, "Nucleation and Growth During Al_2O_3 Atomic Layer Deposition on Polymers," *Chemistry of Materials*, vol. 17, pp. 5625-5634, 2005.

[Wis08] K.D. Wise, A.M. Sodagar, Y. Yao, M.N. Gulari, G.E. Perlin, and K. Najafi, "Microelectrodes, Microelectronics, and Implantable Neural Microsystems," *Proceedings of the IEEE*, vol. 96, no. 7, pp. 1184-1202, 2008.

[Wu06] M.C. Wu, O. Solgaard, and J.E. Ford, "Optical MEMS for Lightwave Communication," *Journal of Lightwave Technology*, vol. 24, no. 12, pp. 4433-4454, 2006.

[Wu08] M-H. Wu, S.-B. Huang, Z. Cui, Z. Cui, and G.-B. Lee, "Development of

Perfusion –Based Micro 3-D Cell Culture Platform and its Application for High Throughput Drug Testing," *Sensors and Actuators B: Chemical*, vol. 129, no.1, pp. 231-240, 2008.

[Xia11] B. Xiao, T. Dong, E. Halvorsen, Z. Yang, Y. Zhang, N. Hoivik, D. Gu, N.M. Tran, and H. Jakobsen, "Integrated Micro Pirani Gauge Based Hermetic Package Monitoring for Uncooled VO*x* Bolometer FPAs," *Microsystem Technologies*, vol. 17, no. 1, pp. 115-125, 2011.

[Yon10] S. Yoneoka, C.S. Roper, R.N. Candler, S.A. Chandorkar, A.B. Graham, J. Provine, R. Maboudian, R.T. Howe, and T.W. Kenny, "Characterization of Encapsulated Micromechanical Resonators Sealed and Coated With Polycrystalline SiC," *Journal of Microelectromechanical Systems*, vol. 19, no. 2, pp. 357-366, 2010.

[You03] M. Young, Y. L. Han, E. P. Muntz, G. Shiflett, A. Ketsdever, A. Green, " Thermal Transpiration in Microsphere Membranes," *23rd International Symposium on Rarefied Gas Dynamics*, vol. 663,pp. 743-751, 2003.

[You05] M. Young, *Y.L. Han*, E.P. Muntz, and G. Shiflett, "Characterization and Optimization of a Radiantly Driven Multi-Stage Knudsen Compressor," *AIP Conference Proceedings*, no. 762, pp. 174-179, 2005.

[You99] R.M. Young, "Analysis of a Micromachine Based Vacuum Pump on a Chip Actuated by the Thermal Transpiration Effect," *Journal of Vacuum Science and Technology B*, vol. 17, no. 2, pp. 280-287, 1999.

[Zel04] E.T. Zellers, W.H. Steinecker, G.R. Lambertus, M. Agah, C.-J. Lu, H.K.L. Chan, J.A. Potkay, M.C. Oborny, J.M. Nichols, A. Astle, H.S. Kim, M.P. Rowe, J. Kim, L.W. da Silva, J. Zheng, J.J. Whiting, R.D. Sacks, S.W. Pang, M. Kaviany, P.L. Bergstrom, A.J. Matzger, Ç. Kurdak, L.P. Bernal, K. Najafi, and K.D. Wise, "A Versatile MEMS Gas Chromatograph for Determination of Environmental Vapor Mixture," *Proceedings of the Solid State Sensors, Actuators and Microsystems Workshop (Hilton Head 2004)*,pp. 61-66, 2004.

[Zha11] Y. Zhang, "ALD Enabled Wafer Level Polymer Packaging for MEMS," *PhD Dissertation,* University of Colorado, Boulder, 2011.

[Zho11a] H. Zhou, H.Q. Li, V. Sharma, and M.A. Schmidt, "A Single-stage Micromachined Vacuum Pump Achieving 164 torr Absolute Pressure," *IEEE International Conference on Micro Electro Mechanical Systems (MEMS'11)*, pp. 1095-1098, 2011.

[Zho11b] H. Zhou, "A Study of Micromachined Displacement Pumps for Vacuum Generation," *PhD Dissertation,* Massachusetts Institute of Technology, Massachusetts, 2011.