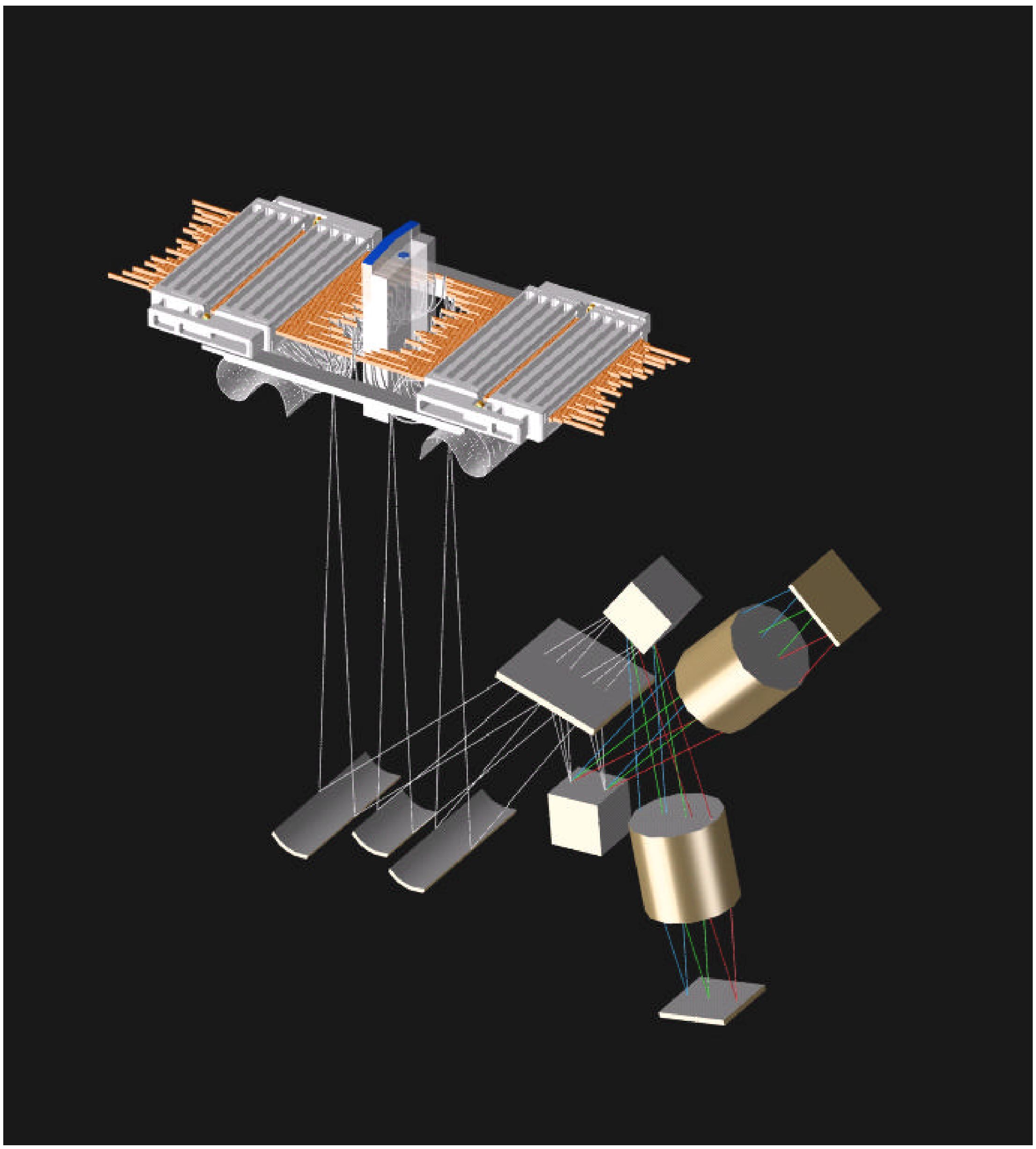
CANADIAN NGST NIR MOS/IFS CONCEPT STUDY

David Crampton and the Canadian NGST MOS/IFS team
E-mail: david.crampton@hia.nrc.ca



CANADIAN NGST MOS/IFS TEAM

National Research Council: D. Crampton, S. Roberts, G. Herriot

University of Toronto: S. Lilly

Simon Fraser University: S. Bakshi, A. Parameswaran, M. Syrzycki

University of Victoria: E. Steinbring, M. Nahon, I Sharf, B. Buckham, D. Erickson, J. Carretoro C. Abbott, R. Gaunt, G. Kabush, A. McIlwain, R. Vrooman

Camosun College Mechanical Technology students: D. Booey, A. Hill, K. Taylor, B. Townsend

EHR Optical Systems: E.H. Richardson

EMS Technologies: A.Bell, G. Buttner, C. Evans, D. Maule, N. Rowlands

Table of Contents

1.0 *Science*

1.1 Executive Summary

The Canadian MOS/IFS team studied a wide range of options for a conceptual design of a NIR Multi Object Spectrograph (MOS), an Image Slicer (IS), and a MOS/IS combination that will enable the scientific promise of NGST as exemplified by the Design Reference Mission (DRM). A low risk design for the MOS based on an elegant positioner for fibre optic slitlets combined with an equally elegant image slicer feeding a common spectrograph is proposed.

Detailed examination of the Design Reference Mission leads us to conclude that a wide range of spectral resolutions will be important for NGST. The lowest (R~100) resolutions will allow NGST to take advantage of the low relative background to obtain spectra of the faintest possible targets. The higher resolutions will allow the detailed analyses that are required to gain an understanding of the underlying physics. However, none of the programs require extremely high multiplexing capability, partly because of recent demonstrations that photometric redshifts are able to accomplish this part of high-ranked lower spectral resolution science. Hence we believe that a multiplexing capability of being able to observe ~100 objects simultaneously will satisfy most of the spectroscopic requirements for the NGST MOS instrument.

Our analyses also indicate that an image slicer delivering high spatial resolution with similar spectral resolutions will be an extremely powerful tool for detailed examination of individual objects from the nearby universe to the high redshift universe. A detailed simulation incorporating realistic telescope, instrument and detector parameters, and realistic projections of the high redshift universe, was constructed in order to investigate the critical design requirements for a NIR MOS. This simulation nicely demonstrates, for example, that precise slit widths are relatively unimportant for observations of high redshift galaxies: for most observations of interest the slit width has no effect on the noise and the effective spectral resolution is largely determined by their compact size. The simulation also permits us to investigate how best to carry out observing programs such as those given in the DRM. Both the simulation and our analyses indicate that the numbers of "interesting" objects per MOS field is typically 100 or less for most of the DRM projects. Hence the ability to accurately position slitlets on a 100 objects in a ~3 arcmin field is a strong requirement.

Three methods of positioning slitlets on ~100 targets in a 3' x 3' MOS field appear to be feasible for NGST. Two of these are mechanically-actuated movable slitlets basically formed by metal strips, and the other involves slitlets composed of two rows of microlenses and fibres. The latter form slitlets of dimension 0.4" x 3" which are transferred and reformatted by the fibres to form two long input slits for the spectrograph. Hence, one advantage of this solution is that spectra derived from targets distributed over a 4' x 3' area can be all lined up to optimize usage of the detector real estate. A second important advantage is that the spectrograph design is much simpler (and the resulting instrument much lighter) since each slit need only feed its own quasi-one dimensional collimator rather than a collimator capable of imaging the whole field. Our fibre positioner design uses only piezo actuators similar or identical to those being developed for NGST and allows slitlets to be positioned anywhere in the "MOS" field with a fill factor approaching 90%.

Our image slicer design will deliver essentially diffraction-limited images of a 2.5” x 4” object to the spectrograph in a slit 0.1” wide by 90” long. It consists of a solid block of an IR transmitting glass like calcium fluoride in which two arrays of mirrorlets and one array of lenslets are fabricated. Since the reflecting surfaces are internal and there are only two air-glass surfaces, it should have extremely high throughput and extremely low scattered light. Its performance is inherently athermal and stable since it is composed of a single block of glass. It can be designed such that it does not alter the locations of the focal planes of the telescope and spectrograph. However, if used in conjunction with the fibre positioner, it would likely be designed to transfer its reformatted NGST focal plane down closer to the location of the long spectrograph entrance slits formed by the fibres.

If the image slicer were to be located in the center of the input field of the spectrograph, then our fibre positioner could access two 1.5’ x 3’ fields on either side and the resulting three long slits could feed the same spectrograph. This would provide both a MOS and IFS capability in a single instrument. It is this combination which we propose for NGST NIR spectroscopy.

1.2 Science Capabilities – the need for spatially multiplexed faint object spectroscopy on NGST

That the NGST requires a spectroscopic capability in the near-infrared 1–5 µm waveband to perform its primary scientific mission of studying the high redshift Universe is beyond doubt. Less certain is which type of spectrograph, MOS, IFS or FTS, or combination thereof, is most suited to obtain the required data and, in the case of the MOS, which approach to the front-end multiplexing offers the best trade between scientific return and technical risk and cost. Answers to these questions depend on the relative scientific weighting attached to different regimes of spectral resolution, the desire to observe multiple objects simultaneously, or the requirement to cover a contiguous two-dimensional field.

A good fraction of the NGST science mission is aimed at understanding three related phenomena in the early Universe:

(a) The appearance of “first-light” in the Universe arising from the formation of the first generation of star-clusters or by accreting black-holes.
(b) The re-ionization of the inter-galactic medium (IGM) by these sources of ultraviolet radiation.
(c) The assembly of luminous and dark matter into large galaxies such as the Milky Way.

Understanding these phenomena requires a systematic approach to the study of galaxies and pre-galactic objects in the distant Universe. In 1994, the Dressler report laid out an approach to this problem that is still relevant today:

detection – identification – characterization – placement in context

While broad band measurements can detect galaxies at very high redshifts and, via spectroscopic redshifts, an estimate of the redshifts and broad evolutionary state obtained, detailed information on the physical conditions within the galaxies, measurements of their masses and metallicities, and an understanding of their environments can *only* be gained by spectroscopy. It is this latter information, going well *beyond* simple *x,y* location, redshift, luminosity and colour, which has been largely absent hitherto for galaxies at any cosmologically significant redshift. This absence has prevented our understanding of galaxy evolution from progressing beyond the level of simply describing the galaxy population at different cosmic epochs through crude distribution functions such as the luminosity function $\phi(L)$.

The NGST DRM lays out the case for spectroscopy in the near-infrared waveband in some detail and so we only briefly review this here with the aim of establishing relative priorities that then drive our choice of spectrograph concept.

1.2.1 Spectral Resolution Regimes

1.2.1.1 R=100 identification spectroscopy

For faint galaxies, the main use of spectroscopy at R~100 is in spectroscopic determination of redshifts. The identification of multiple spectral features in emission and absorption yields a redshift that is considerably more secure (and accurate) than those derived from much lower resolution "photometric" data at $R \leq 10$. The accuracy of the resulting redshifts are sufficient to identify the largest structures such as virialized clusters with $\sigma \geq 500$ kms^{-1} but not to characterize more typical environments. While some crude diagnostic information is available, such as $H\beta/H\alpha$, these are compromised by the resolution and the inability to resolve $H\alpha$ and the nearby N[II] lines, or to determine the effects of underlying stellar absorption underneath $H\beta$. Thus we envisage that spectroscopy on faint galaxies at this resolution will primarily be required for (a) confirmation of photometric redshift estimates obtained from lower resolution data for typical galaxies and (b) measurement of redshifts of faint objects of particular interest, such as those identified in X-ray, far-infrared and radio surveys.

R=100 spectroscopy is also required for identifying the damping wings that will be used to determine the neutral Hydrogen column density and thus to identify the redshift of re-ionization. Ideally this observation can be carried out on a few bright objects. However, such bright objects may well not exist at the required redshifts and it may be necessary to co-add the spectra of numerous faint very high redshift sources in order to see the signature of the un-ionized medium.

Finally, R~100 spectroscopy is required in the DRM to characterize the nature of supernovae.

Spectrograph implications:

We thus believe that the case for R~100 spectroscopy on NGST is extremely strong, especially since in this spectral regime NGST spectroscopy will certainly be background limited and thus the sensitivity gain over ground-based telescopes is maximised. It is also clear that a MOS multiplexing capability will be very important. However, many of the R~100 science drivers outlined above are for observations of single objects (e.g. identifications of particularly interesting objects, supernovae) or of objects of limited surface number density (bright extremely high redshift beacons for IGM studies), and thus we considered it an acceptable trade if our MOS concept did not achieve the maximum multiplexing gain ($n >> 100$) that the short spectra achieved at R~100 in principle allow. Probably the most important single requirement for the R=100 spectroscopy is to maximise the sensitivity since a main goal will be to push confirmation of photometric redshifts on selected "key" objects to the faintest possible levels.

1.2.1.2 R=300

As the resolution increases, the redshifts become accurate enough for useful kinematic information in systems with $\sigma \leq 500$ kms^{-1}. This is appropriate for identification of objects within small groups and thus for characterizing the large-scale environment of objects. However, it is still inadequate for sampling galaxy haloes and still suffers serious limitations in terms of emission line diagnostics. In the DRM, there is no specific requirement for R=300 spectroscopy and the environment science is achieved with higher resolution R=1000 spectroscopy required for these other objectives.

1.2.1.3 R=1000 diagnostic spectroscopy

In his review, Kennicutt (1998) identifies R=1000 as the ideal spectral resolution required for utilizing the multitude of emission line diagnostics that will be available with NGST spectroscopy in the near-infrared 1-5 μm waveband. In particular this resolution is sufficient to comfortably resolve Hα and [NII]6584 (R $\geq$ 400) Although dispersed spectra at this resolution are likely to be detector noise limited, the detectability of emission lines does not decrease with increasing spectral resolution until the lines become resolved (which will occur at R > 1000 for galactic systems).

As reviewed by Kennicutt (1998), flux measurements and flux ratios of emission lines lying between 3727 and 7000 Å can be used to determine a multitude of physical parameters that are highly relevant for studies of galaxy evolution, including metallicity, reddening, star-formation rate, and the nature of the ionizing source). The importance of these parameters in characterizing the evolutionary state of individual galaxies is obvious both in terms of understanding their current state, i.e. what they are doing at the epoch of observation, and in terms of understanding their past and future. In particular, metallicity measurements are extremely important in identifying whether star-forming galaxies are "primordial" and in limiting their possible descendants at later epochs. Basic metallicity information can be obtained through measurements of [OII]3727, [OIII]4959,5007, [NII]6583 and Balmer lines.

These analyses thus require spectrophotometric measurements over an octave of spectral range. Although the sensitivity gain of NGST over the ground-based telescopes is reduced in the detector-limited regime at R~1000 (assuming currently projected detectors) the difficulty of obtaining the required spectrophotometry of multiple features over the required wavelength range through the atmospheric windows and OH forest should not be underestimated. Not least, for all $z > 2.65$, H has $\lambda > 2.4$ μm and is effectively unobservable from the ground, while the key Hβ/[OIII] spectral region is lost for $z > 3.8$

Although ideally such information would be available at different spatial locations within a young galaxy, useful global measurements can be obtained by integrated spectroscopy (see Kobulnicky et al 1999).

In the DRM a large and systematic survey of faint galaxies at R = 1000 is envisaged. Such a survey would yield basic diagnostic information on all targeted galaxies, while the accurate redshifts would allow quantitative study of the environments of objects and the quantitative study of the development of clustering in both space and velocity. Most plausible cosmologies (e.g. $\Omega_0 = 0.2$, $H_0 = 65$ kms^{-1} Mpc^{-1}, with or without a Λ) have a comoving scale of approximately 2 Mpc arcmin^{-1} at high redshift (independent of redshift). Typical length scales of interest are 250 kpc (7 arcsec) for the turn-around radius of the Milky Way halo, 2 Mpc (60 arcsec) for the scale of the Local Group, and 8 Mpc (4 arcmin) for the correlation length of galaxy clustering at high and low redshifts (Giavalisco et al 1998, Adelberger et al 1998). Thus it is important that objects are observed over large contiguous regions of the sky. Furthermore, in order to gain a complete understanding of the assembly of galaxies, if possible every object at a given redshift within the scale of collapsing galaxy haloes and groups of galaxies should be observed.

Spectrograph implications:

The main requirement here is for efficient multiplexing so that as many objects can be observed as possible over angular scales of a few tens of arcsec (the extended halo of a forming massive galaxy) to a few arcminutes (the size of a group of galaxies and the observed correlation length at high redshifts). There is also a desire to fully sample the detected objects within a given redshift interval over these same spatial scales, so that the environmental context of individual objects and the processes occurring within them are understood. At R=1000 much of the spectroscopy need not use spatial information within a galaxy although this could be highly desirable, say in determining metallicity gradients in young disk galaxies.

1.2.1.4 R=5000 kinematic spectroscopy

R=5000 spectroscopy corresponds (at all redshifts) to a velocity resolution of 60 kms^{-1}. This is required to measure spatially resolved velocity fields within galaxies (e.g. rotation curves) and to estimate velocity dispersions within spatial unresolved systems, given the low velocities expected for galactic fragments at high redshift ($\sigma \sim 50$ kms^{-1}).

Determining the masses of galaxies and fragments is extremely important for two reasons. First, the theoretical paradigm for the origin of structure in the Universe is in terms of the growth of dark matter density fluctuations set up at very early times. On galactic scales, a hierarchical assembly of collapsed haloes on progressively larger scales leads to the build-up of large galaxies. Relating observed stellar systems to the dark matter haloes in which they reside therefore provides a key link to the theory and enables us to examine key questions such as the mass function of hales as a function of time and the role of physical mechanisms such as energy injection in regulating the movement of baryons into and out of dark matter haloes. The masses of objects also provide (with metallicity) major constraints on the possible later descendants of particular objects.

The dynamics of a system are best characterized by spatially resolved kinematic measurements and the interpretation of integrated spectra is problematical unless it is possible to be certain that the system is dynamically simple and symmetrical (e.g. elliptical galaxies or the HI rotation curves of disk galaxies). Many objects at high redshift are known to be very small (see e.g. Gardner and Satyapal 1999) and most of the extended objects have asymmetric, highly disturbed, morphologies. Three practical approaches to dynamical measurements are:

(a) Spatially resolved kinematics of nebula emission lines, where the motion can be ascribed to the gravitational potential of the system.
(b) Integrated measurements of stellar absorption lines for objects without emission lines. The Mg_b 5175 and Ca triplet features are frequently used. The 2.3 μm CO bandheads are attractive but lie beyond 5 μm for the redshifts of interest. These measurements require good S/N in the continuum.
(c) Integrated measurements of nebula emission lines. The interpretation of these are limited by the possibility of non-gravitational motion and uncertainty over which parts of the gravitational potential are being sampled, but this may be the only approach for the smallest and faintest objects.

We envisage that all three approaches will of necessity be employed. For all objects which are significantly extended, full 2-dimensional information will be very important – indeed most known galaxies at high redshift are so irregular that defining a "major axis" with any confidence at all that it would be dynamically significant will be very hard.

As noted above, the sensitivity to emission lines is not degraded as R increases, and we anticipate that much of the R ~ 5000 kinematic science with NGST will be undertaken on emission lines.

Spectrograph implications

Although some objects will be too small to make it worthwhile, the kinematic studies possible with R=3000 spectroscopy benefit the most from having full 2-d spatial information available and the case for an integral field type spectrograph is strongest at this resolution although 2-d information can of course be obtained by stepping multiple slits across multiple objects.

1.2.2 The need for large numbers of objects – MOS multiplexing parameters

A key parameter in considering MOS front-end concepts is the expected number density of targets. We assume for this purpose the Gardner and Satyapal (1999) model *L*-band counts, which extend to $L_{AB} = 34$. For the redshift distribution, we have assumed that the Fernandez-Soto et al (1999) redshift distribution between $25 < I_{AB} < 28$, which gives 25% of sources at $z > 2$ and 6% at $z > 4$.

In assessing the typical potential for multiplexing, we have considered the number density of targets. Following the DRM analyses, we have assumed extremely long integrations (24 hours) and taken the nominal S/N requirements from the DRM (usually 5-10 in the continuum). The range of L_{AB} in the third column then reflects the typical range of galaxy sizes.

Program	R	L_{AB} (max)	Nominal n arcmin^{-2}	ε	Effective n arcmin^{-2}	Comments
Identification	100	29.4–30.5	500–800	0.25	125–200	Assuming only $z > 2$ is interesting
Diagnostics	1000	25.4–27.0	100–250	0.25	25–60	ε from $z > 2$ fraction, but length of spectra already dictates a 1-d slit arrangement or at most two tiers of spectra.

Kinematics	5000	24.4–26.0	80–180	0.02	1.6–3.6	ε limited by limited range (i.e. *z* range) plus orientation effects.
Integral Field	1000	21.9–23.5	15–70	0.25	4–16	ε from $z > 2$ fraction
Integral Field	5000	20.9–22.5	7–30	0.08	0.6–2.4	ε from limited λ and z range

An important point is that the number densities of potential targets are generally high but not extremely high, except in the case of the R=100 spectroscopy where the faintness of the targets (and the short length of the spectra) would make very large numbers of targets accessible. As the resolution increases the number of potential targets decreases and the length of the spectra makes it more difficult to target all of the interesting objects at once. As will be discussed below, a movable slit design works efficiently up to a number density of 11 arcmin^{-2} – since that gives one object in about a 3.5 × 90 arcsec^2 region (i.e. the search area for each slit). Densities of ≤ 11 arcmin^{-2} are close to what is required for all of the above programs except the identification R = 100 program.

1.2.3 The need for two-dimensional information – IFS parameters

In our opinion the main driver for integral field spectroscopy is in studying the spatially resolved kinematics of high redshift galaxies. One concern is whether typical high redshift galaxies have bright enough extended structure to make such studies feasible – are there enough photons at R=3000 from the extended structures within the NGST PSF? Using the standard relations (Kennicutt 1999) between star-formation rate and Hα luminosity and ultraviolet continuum luminosity, and assuming a typical continuum $(I-L)_{AB} \sim 3$ (a present day Scd galaxy), it is possible to relate the NGST detection limits at 3.5 μm (i.e. Hα at $z = 4.3$) to an F814W surface brightness and then examine the images of $3 < z < 5$ galaxies in the HDF (at similar spatial resolution) to see what fraction lie above the detection threshold in the line and continuum.

The details are given in Appendix A1. The continuum detection level corresponds to a rest-frame $\mu_{AB}(R)$=14.3 mag arcsec^{-2} at z = 4 which is remarkably high and almost none of the known $3 < z < 5$ galaxies have spatially extended structure above this threshold. However, the equivalent detection threshold for line emission is a factor of ten or more fainter, and this is sufficient to bring much of the extended structure above the detectability threshold, as illustrated in Fig 1.1.

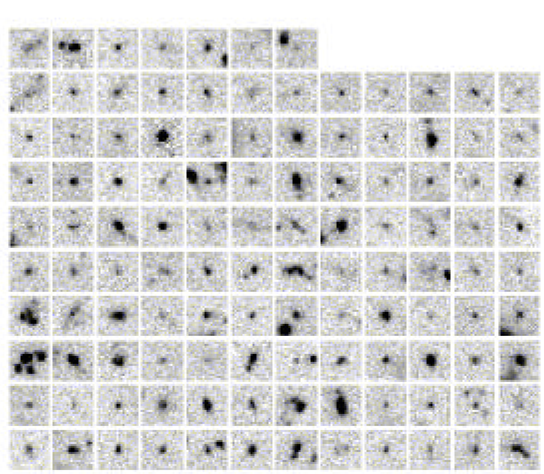


Fig. 1.1: Two mosaics of F814W images of 115 HDF galaxies with 3 < z <6, showing (top) all surface brightnesses, and (below) only those surface brightnesses equivalent to the Hα detection threshold. Individual boxes are 2 arcsec to a side.

1.2.4 Summary of Science Requirements leading to the proposed MOS/IFS concept

From the above considerations we identify the following capabilities as being essential to carrying out the scientific goals of the NGST. Whilst we view as all essential, we list these in decreasing order of importance.

(a) R=100 spatially-integrated spectroscopy at maximum sensitivity with multiplexing performance secondary.
(b) R=1000 spatially-integrated spectroscopy of multiple objects distributed over arcminute regions of sky with multiplexing performance paramount up to the expected number density of targets.
(c) R=3000-5000 spatially-resolved 2-d spectroscopy at the diffraction limit of NGST of individual objects extended on the scales of a few arcsec.
(d) R=1000 spatially resolved spectroscopy and R=3000-5000 spatially integrated spectroscopy.

This leads us to propose an integrated MOS/IFS design in which complementary front-ends achieve the MOS and IFS tasks and feed a common spectrograph capable of 100 < R < 5000. The MOS front-end was required to have a minimum of approximately 100 positionable apertures.

A detailed simulation tool was developed to investigate the critical design requirements for a NIR MOS and to gain a better appreciation of the practical problems which are likely to be encountered. The simulation incorporates realistic telescope, background, instrument and detector parameters, and conservative projections of the high redshift universe based on the deep HDF observations with NICMOS. This simulation tool nicely demonstrates, for example, that the slit width has little effect on the noise for most observations of high redshift galaxies and the effective spectral resolution is largely determined by their compact size. The simulation also permits us to investigate how best to carry out observing programs such as those given in the DRM. The results agree with the above analyses that the numbers of “interesting” objects per MOS field is typically 100 or less for most of the DRM projects. Details of the simulation are given in Appendix A2 and a web interface can be visited at http://astrowww.phys.uvic.ca/~steinb.

2.0 *Engineering*

2.1 Design Concept

Discussion of our MOS/IFS instrument concept is divided into descriptions of conceptual designs of the three major components:

(1) MOS multi-slit or fibre slitlet deployment concept
(2) Image slicer design
(3) Spectrograph concept

We admit at the outset that most of our effort was spent on development of a solution of the problem of how to reliably position slitlets on ~100 targets of interest distributed over a 3' x 3' field of NGST, and how to provide an appropriate integral field device for detailed studies of individual objects via 3D spectroscopy. Furthermore, unlike most other studies, we deliberately carried several options throughout our study, selecting our preferred solution near the end. Hence, our studies of the spectrograph and of the combination of all three components are much less complete. **Importantly, however, we did find a concept based on conventional (low risk) technology that combines multi-object and integral field spectroscopy in a single instrument that will enable the DRM science to be carried out efficiently and effectively.**

2.1.1 Fibre Slitlets and positioner concepts

2.1.1.1 NIR fibre properties

Optical fibres are extensively used for multi-object spectroscopy on ground-based telescopes, although most of the applications have been in the visible wavelength region (largely because of the small size of IR arrays and because sky background subtraction is more problematic). These systems are of some relevance, however, because they demonstrate optimal methods of achieving the highest possible efficiency and scientific benefit through the joint use of microlenses and fibres (e.g., Baudrand, Guinouard and Jocou 1998). *Infrared fibre* systems are now beginning to be incorporated into ground-based astronomy instruments and several companies supply fibres that transmit well in the 1 – 5 micron region (see Appendix B2). Studies to date, although admittedly not very extensive, indicate that both the optical and mechanical properties of these fibres at the 35K design temperature of NGST should not be substantially different from their room temperature performance (e.g., Levin et al. 1993 and verbal reports from several sources). Furthermore, there are several reports of fibres being considered or used for near-IR space missions (more details in Appendix B2), indicating that there are no known insurmountable problems associated with their use in space. For a long lifetime mission like NGST, however, extensive testing should be carried out, including testing in the expected radiation environment for possible degradation of their transmission and also of possible internal Cerenkov emission. Our instrument concept requires only relatively short lengths (~20cm) of fibres and also that the fibre bundles be ribbon shaped and bend relatively small amounts in only their "flat" direction. Thus it is expected that they will be not be overly stressed and that variations in their transmissive properties (that plague ground-based fibre systems) will be minimized. Once again, this expectation should be confirmed by actual tests. Special attention will have to be paid to the coupling of the microlenses and fibres and also of the slitlets to the probes in order to minimize stresses induced by cooling (and hence to minimize the so-called focal ratio degradation). Reports on cryogenic fibre devices used in ground-based applications are given, for example, for COHSI by Kenworthy, Parry & Ennico 1998, and for SINFONI by Tecza and Thatte 1998.

2.1.1.2 Fibre slitlet concept and actuation method

In order to access as much of the focal plane as possible, we propose to use miniature fibre slitlets composed of two rows of 15 fibres. The input microlenses on these fibres will be 0.2" square so that the overall size of the slitlets will be 3" by 0.4" with 0.2" spatial sampling of the target objects (see Figure 2.1). These slitlets will be affixed to the ends of probes which can be moved across the MOS field and be positioned on targets to milliarcsecond accuracy. The other ends of the fibre ribbons or bundles will be reformatted into 1 x 30 arrays of microlenses along the spectrograph input slit. Approximately 50 probes with attached slitlets can be inserted from each side of the 3 x 3' MOS field, leaving small (~0.5mm) gaps between the probes. Thus a total of ~100 slitlets can be positioned on targets in the field with an estimated fill factor of up to 90%.

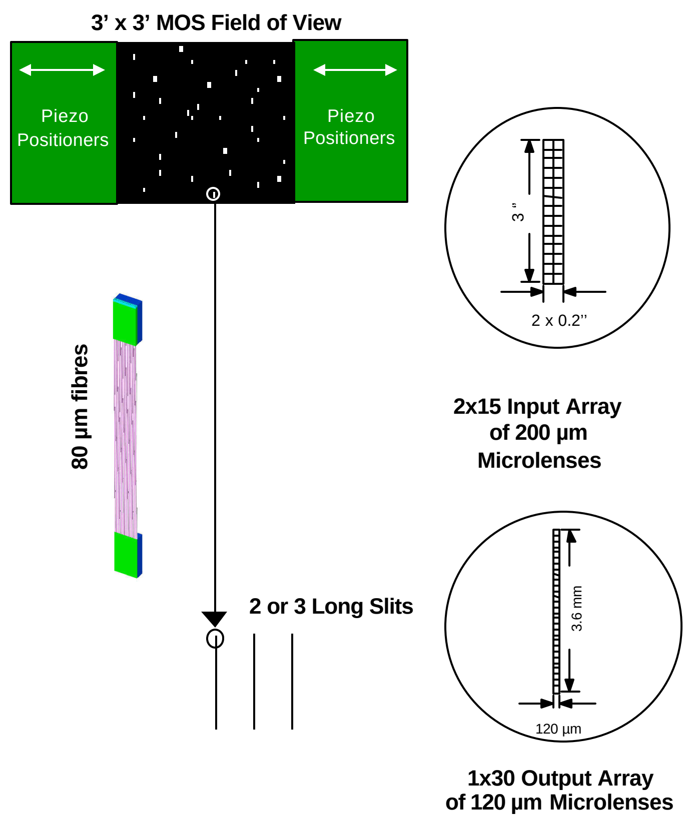


Figure 2.1: Schematic of fibre slitlet reformatting concept

Given that using fibre slitlets to transfer the light from target objects to the spectrograph slit appears to be feasible, the remaining problem is how to accurately and reliably position the slitlets on the objects. Clearly, most fibre positioning devices that have been used in ground-based applications are too complex and would not be suitable. We propose a simple concept in which rigid probes are used to move the fibre slitlets in one dimension across the focal plane. Translation of the probes is achieved using cryogenic piezo actuators that move all the probes simultaneously. Two clutch piezos are required for each probe and a shared piezo actuator moves two bars apart. Motion is achieved by co-ordinated clamping and unclamping as the bar moves back and forth relative to the fixed bar. The actuator concept is shown in Figure 2.2.

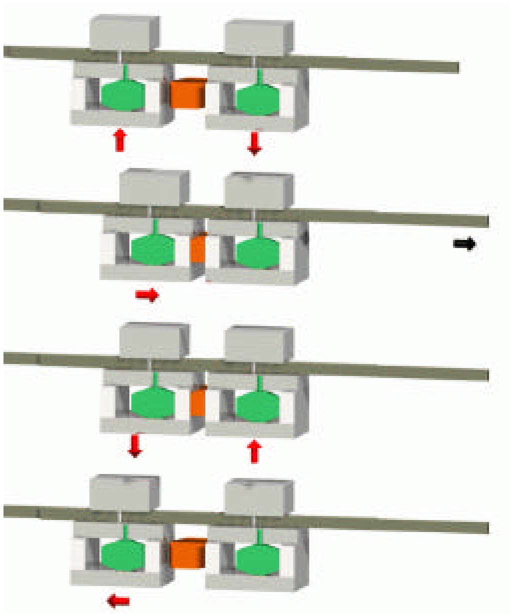


Figure 2.2: Actuation concept. The actuator on the right is attached to a fixed bar. By coordinated clamping and unclamping to the fixed and movable bars, the probe moves a step to the right in this diagram.

Since piezo actuators have very limited stroke or motion, movement of the probes across the focal plane is achieved by a large number of small steps. The motion actuator pushes or pulls one bar relative to the other and a spring is then used to reverse this motion. By clamping the probe to the movable bar during the relevant portion of the cycle, the probe can be moved in either direction and positioned to an estimated accuracy of a few microns with appropriate position sensing and feedback control. To provide redundancy of the motion, two or more motion actuators can be installed. An exploded view of the concept for the manipulator is shown in Figure 2.3. A prototype device was constructed which successfully demonstrated the concept at room temperature with three probes of which two were "active" and the other one fixed. Details of the proposed concept and several pictures of the prototype are given in Appendix B2.

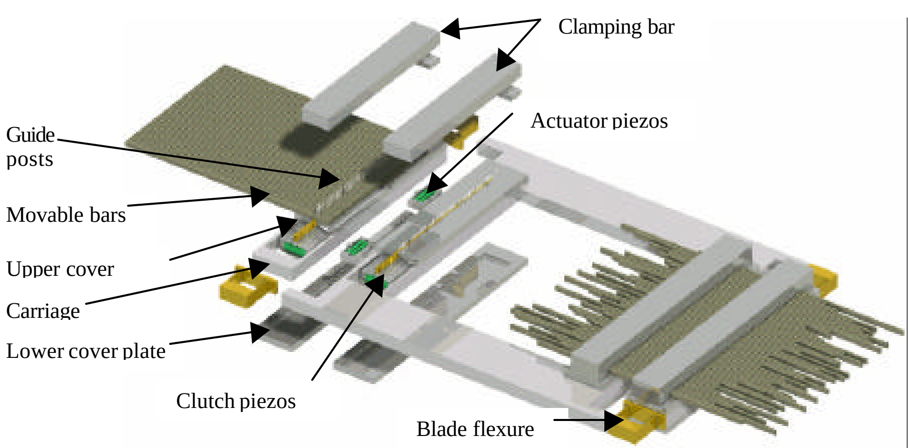


Figure 2.3: Exploded view of the conceptual fibre optic manipulator.

There are several obvious advantages of the proposed actuation concept:

- Uses only actuators similar or identical to those used elsewhere for NGST
- Simple, reliable concept
- Redundancy easily provided by extra actuators to provide motion
- Graceful failure modes for individual probes
- Low heat generation and power usage
- Conventional, low risk technology

2.1.1.3 Alternate multi-slit concepts

If, for any reason, fibres prove to be inappropriate for the NGST MOS, there are two additional methods of producing reconfigurable mechanical slits that appear to be practical. These involve either rolls of thin metal into which slits of two or more widths have been cut, or metal strips, the ends of which are slid together to form slits of variable width. Although various methods of actuation were investigated to provide the required motion as part of the studies of these two concepts, the same type of actuation as described above for the fibre slitlets could also be adapted to these mechanical slitlets. Details are given in Appendices B3-B5. Since the fibre slitlets appear to offer additional practical benefits, our discussion here is focussed on them.

A MEMS shutter device was also investigated as a possible concept for the MOS selection mechanism. Although our design yields almost ideal, variable width, apertures for NGST spectroscopy, the fill factor (percentage of field accessible) was less than ~30% and we were also concerned about residual radiation from the thermal actuators. Their performance at 35K also is unknown. For these reasons the mechanical or fibre slitlets appear to be more promising. Details of our MEMS studies are given in Appendix B6.

2.1.2 Image Slicer Concept

Our goal in designing an image slicer for NGST was to ensure that it would enable exploitation of the high spatial resolution of NGST for detailed investigation of single objects. With slices 0.1” wide, light from a 4” x 2.5” area could be redistributed along the equivalent of a 100” slit which in turn could map onto a 4K pixel detector and adequately sample the diffraction-limited (@2mu) PSF. Initially, in order that the image slicer could take the place of a regular spectrograph slit, a further goal was imposed to design the image slicer such that it could be used without refocusing either the telescope or spectrograph. As described in detail in Appendix B7, a very elegant solution was found. The image slicer consists of a solid block of glass (such as calcium fluoride) in which are machined, or to which are attached, two arrays of micromirrors and one lenslet array. The device is made confocal by folding the beam with totally internal reflections. A schematic is shown in Figure 2.4 and further details are given in Appendix B7 and in Richardson et al. 1999. Here we summarize the main features:

- Fabricated from a single glass block (e.g., calcium fluoride)
- Only two air/glass surfaces
- Uses two mirrorlet arrays and one lenslet array
- Diffraction limited at 2mu, panchromatic (1-5mu) performance

- High throughput, low scattered light
- Compact (135mm long)
- Stable, design insensitive to thermal effects

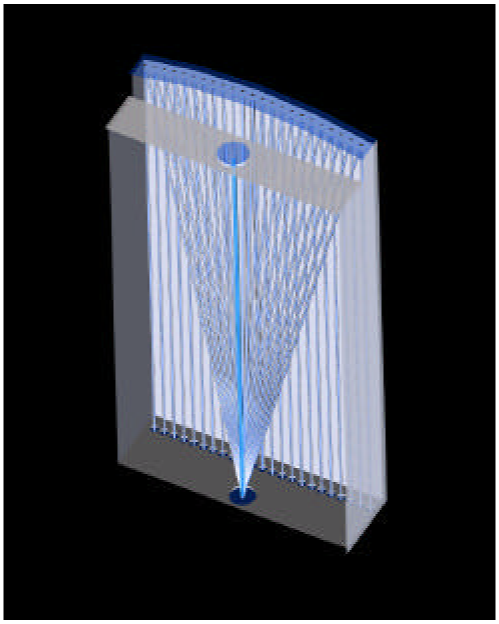


Figure 2.4 Image slicer concept. Light from the telescope enters the glass block at the top, reflects off a mirrorlet array at the bottom, returns to the top where it reflects off another (long) array of mirrorlets, and then exits through a microlens array at the bottom of the slicer.

2.1.3 Combined MOS/IFS

The image slicer and fibre positioner could share space in the telescope focal plane, producing three long slits at the entrance to a shared spectrograph. The concept is shown in a top (telescope) view in Figure 2.5.

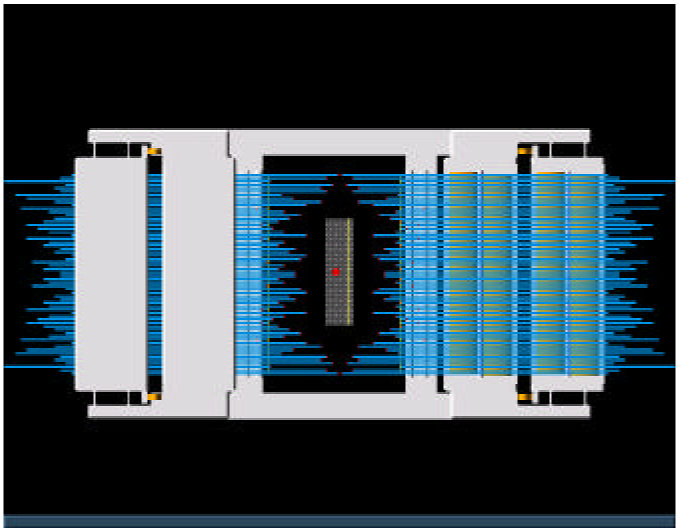


Figure 2.5 Cartoon showing the image slicer in the centre of the fibre positioner system discussed above.

The fibre positioner could patrol an arbitrarily large area (within reason!) since all the fibre bundles (or ribbons) lead to fixed long slits, i.e., increasing the patrol area for the fibre slitlets need not increase the size of the spectrograph input field. Hence, the MOS field could be enlarged to the same size as the imaging field, 4’ x 4’ without significantly increasing the cost. In the arrangement shown, objects in the centre of the field would be inaccessible to the fibre slitlets posing a limitation for some scientific targets (e.g., cluster studies). However, for most of the DRM studies the impact would be negligible, or could be alleviated by repointing the telescope.

2.1.4 Spectrograph Concept

Both our fibre concept and image slicer design produce a long slit format as entrance to the spectrograph. In this case, a simple, near-littrow spectrograph is possible. Our basic design requires only a collimator mirror, a curved grating and three lenses to achieve diffraction limited performance over the entire wavelength range from 1-5mu (see Figure 2.6). At a spectral resolution of R = 3000 the spectra would then extend over 8K, twice the length of our notional 4K detector. Ideally, a beamsplitter would be used to divide the wavelength range in two, feeding two such detectors. If cost or complexity prohibits this, then several gratings with appropriate order sorting filters would have to be installed on a turret to provide both a range of resolutions and a range of wavelength regions.

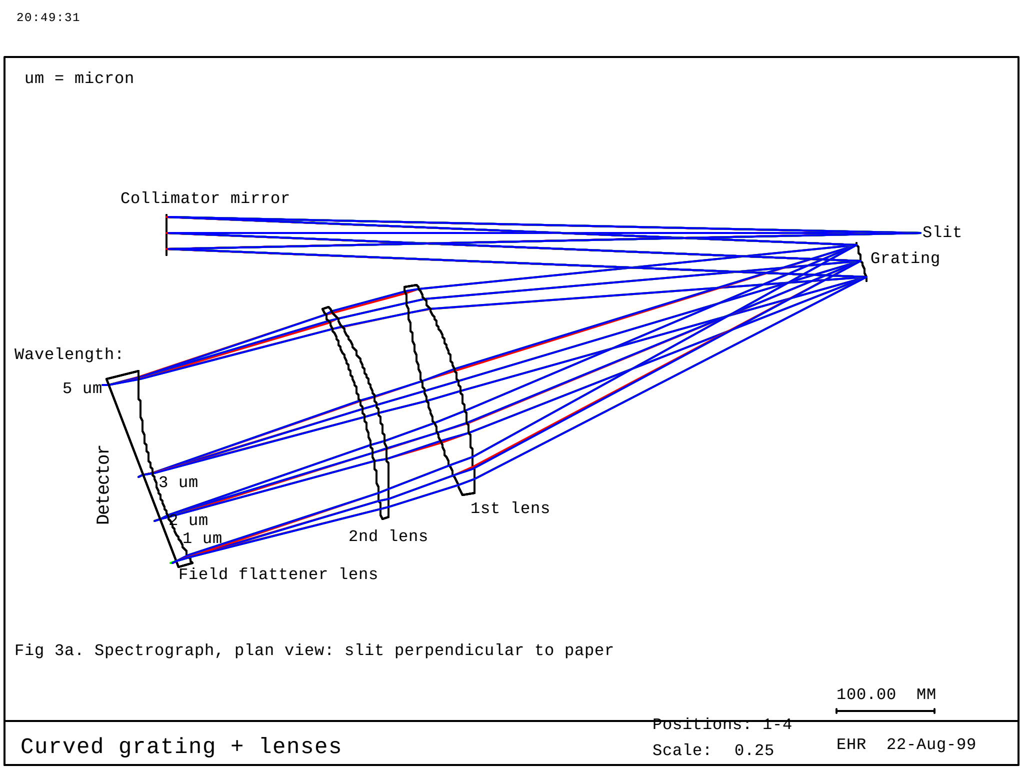


Figure 2.6: Basic spectrograph layout.

A sketch showing how the basic design could be incorporated into a combined MOS/IFS instrument is shown in Figure 2.7. One advantage of the long slit input format is that each collimator could be independently optimized to deliver diffraction limited images for each slit.

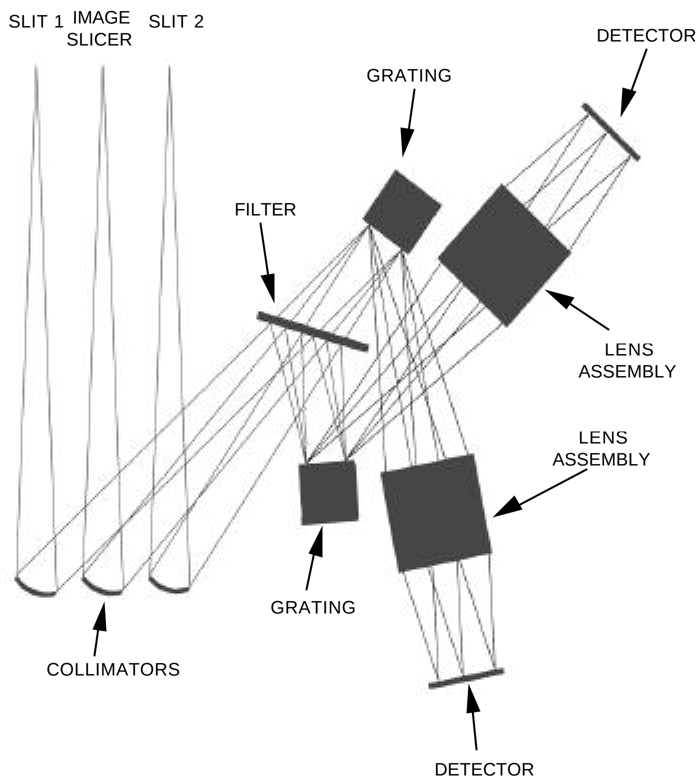


Figure 2.7: A schematic showing how input from several long slits could be fed into a single spectrograph and also how a dichroic could be used to split the spectral range into two parts, nominally from 1.2 – 2.4mu and 2.5 – 5mu.

2.1.5 Mechanical and Thermal Considerations

The overall design of the NIR MOS/IFS instrument is in a very rudimentary state. However, in this section we give results of some preliminary analyses.

2.1.5.1 Fibre Positioner Thermal Issues

As discussed above, the fibre positioner comprises ~100 fibre optic bundles, which can be individually positioned in the field of view to sample light from selected objects. The concept uses piezo actuators to drive the fiber bundles into the field of view according to a pre-selected coordinate table. In our notional design the patrol field was 3’ x 3’ or 180mm by 180mm. 100 fibre bundles, 50 per side are deployed with a 3.5mm pitch. Each fibre bundle is 3mm long and is comprised of two rows of 15 micro-lenses on the input end and one interlaced row of 30 micro-lenses at the output end. The positioning scheme outlined above requires 200 clutch piezo actuators (e.g., CEDRAT APA 100S) and a minimum of two but, for redundancy, more likely eight drive piezo actuators which could be of the same type. If they were operated at a frequency of 7.5Hz (4500 steps/10 minutes) then power dissipation estimates are:

Piezos:

- Power dissipation (low loss piezo) : 0.01W/actuator x 200 = 2 watts

Operating Cycle:

- 10 minutes per 24 hour period (assumed for worst case heat load)
- heat load ~ 1kJ/24 hr (assumes 200 actuators operate for only 5 min. on average from one position to next)
- average power dissipation over 24 hr period = 10mW

Heat will be generated by the positioner at a rate of 2W for approximately ten minutes immediately prior to each observation. There are no components active during the observation so for this time period (24 hours) heat generation will be zero. While 2 watts is a reasonable heat output in itself, the average dissipation over one observation period is only 10mW since the positioner is non-operational for most of the observation period. This suggests that the heat load is manageable and there are several thermal management strategies that can be applied:

(1) Absorb the heat load in a thermal reservoir during the ten minute operating period. Conduct the heat from the reservoir to the radiation panels during the observation period when the piezo actuators are inactive. This strategy may add 1 to 1.5 kg of mass to the unit (c_p for aluminum is 0.9 kJ/kg) and the reservoir would have to be well insulated to avoid thermal radiation towards the optical detectors.

(2) Use the average operating time for design purposes rather than the maximum operating time. On average, the fibers will only drive one quarter of the way across the field of view instead of halfway. Occasionally, the operating time would exceed the average by an amount large enough to force an increase in the cool down period before an observation could commence.

(3) Adopt a duty cycle that would limit the rate of heat dissipation and/or increase the cool down period before an observation could commence.

The optimum design solution will probably draw on all of the above.

2.1.5.2 Fibre positioner Structural Analyses

Only a few critical structural details have so far been investigated.

The stiffness of the carriage was examined to ensure the deflection of the carriage due to loading from the piezo actuator would be within the stroke limit for the actuator. The stroke limit is assumed to be 20µm at 30K. The carriage will deflect on the stationary and moving sides equally assuming equal stiffness on each side. In the worst case, assuming aluminum construction and a maximum deflection of 5µm, an 8 mm thick T-section, about 30mm wide with a 50mm web, would provide sufficient stiffness to resist loading from a row of 30 piezo actuators operating simultaneously. This is a practical size.

The deflection of the end of the fiber bundle drive rod normal to the field of view was examined. Assuming the rod is 2mm wide by 5 mm thick, loading on the rod is 19N (from manufacturer's specification) and the reaction points are 30 mm apart, the worst case deflection of the drive rod end due to loading from the piezo actuator would be about 0.18µm.

The fundamental resonance frequency for the drive rod would be greater than 500Hz. Coupling with the frequency of the piezo actuators should not occur. This calculation did not consider the mass of the fiber bundle at the end of the drive rod. This will tend to lower the fundamental frequency for the rod.

2.1.5.3 Other Mechanical and Thermal Issues

The image slicer is basically a monolithic glass block. There are no moving parts or heat generating components. The mechanical issues relating to the image splitter are mostly related to manufacturing techniques and mounting and alignment. There is a further requirement for a mechanism to block light into the image slicer when it is not being used (the fibre slitlets can be withdrawn to an obscuring position when the image slicer is being used). The detectors for the spectrograph and the grating exchange mechanism will dissipate some heat, however the level is low in comparison with the fibre positioner actuation mechanism.

2.1.5.4 Total mass estimates

The following mass estimates are based on the volumes and components illustrated in the figures. The volumes were assumed to have the density of aluminum (2700 kg/m^3).

Description	Mass (kg)	Comments
Spectrograph	21	
Positioner	9	without power electronics
Thermal sink	1	allowing 1°C temperature rise
Wiring Harness	5	for piezo actuators and detectors
Margin	15	30% contingency
Total	51	

2.1.6 SUMMARY: Proposed MOS/IFS Concept

Our specific proposal is for a combined MOS/IFS spectrograph instrument that combines our fibre slitlet positioner and solid glass image slicer. A graphical image of such a system is shown on page 1 of this proposal. Here we summarize its features:

MOS mode:

- ~100 fibre slitlets composed of 2 rows of 15 microlenses forming 0.4” x 3” slits
- Slitlets could be positioned anywhere in a 4’ X 3’ (or larger) field
- Simple, low risk fibre positioner

Image slicer:

- Diffraction limited images along 0.1” slices of a nominal 4” x 2.5” area of the target object
- Compact, stable, high throughput, low scattered light design

Spectrograph:

- Simple near-littrow arrangement possible
- R~ 3000 catadioptric design with flat focal surface

- Diffraction limited performance over 1-5mu range
- Alternate reflective camera designs being studied.

2.2 Technological Readiness

The MOS / IFS presented above is considered to be an approach which provides the scientific functionality demanded by the core of the DRM by use of technologies which significantly minimize risk and instrument complexity.

The solution proposed for the combined MOS / IFS represents a minimization in terms of mass and volume and maximum in terms of flexibility when compared to traditional solutions or those relying on “Breakthrough Technologies” such as Micro Mirror Arrays, Micro shutters or complex FTS / dispersive hybrids.

A limited amount of incremental development is however necessary to meet the challenges of the environmental or lifetime constraints of the NGST mission. It is recommended that the period prior to the start of the official phase A/B be used to allow this enabling technology development. Specific details of the recommendations are outlined below.

2.3 Development schedule and Integration & Test plan

Proposed Schedule

01/2000	Technology developments
04/2001	Phase A Start
05/2002	PDR STM Delivery
06/2003	CDR EM Delivery
06/2005	PFM Delivery

2.3.1 Pre Phase A

During the first year prior to the formal start of the Instrument Phase A it is proposed that a series of independent investigations be carried out on specific aspects of enabling technology. The aim of these studies will be to better understand the specific technologies and to allow early retirement of risks involved with the use of the required technologies in the extreme environment foreseen for NGST. This work will also allow more fully informed trade-offs to be made during the phase A.

It is recommended that radiation testing be performed on the microlenses and IR fibers and also on the candidate bulk materials for the Image Slicer to ensure their suitability for the radiation environment expected in the NGST location. Calcium fluoride has already been subject to radiation testing and provided of sufficient purity shows no degradation on exposure. (See for example Childs et al 1973).

It is recommended that several interwoven fiber bundles and microlenses be fabricated. These will be subject to mechanical and optical testing at temperatures down to 30K to allow for characterization. Similarly, it is recommended that an image slicer be fabricated, working with potential vendors to determine optimal methods and tradeoffs.

A mechanical test model similar to that developed by University of Victoria that incorporates the most promising piezo actuators and using appropriately designed mechanical supports for the fibre probes will be built and tested at cryogenic temperatures. Heat dissipation and other thermal issues associated with the actuators should also be studied.

The Pre Phase A will also include an overall system level study into the options available for the design of the rest of the instrument. This will encompass the spectrometer design, IFS material selection and instrument electronics issues.

2.3.2 Phase A/B

Assuming that the technology development work is carried out, then the instrument team will be well positioned to move forward in Phase A to complete the design trade-offs and select a baseline design.

It is proposed that rather than follow the simple STM, EM, PFM model philosophy suggested by the NASA proposal that a more conservative approach be taken with respect to the “Optical Head” (all the equipment to be housed on the cold side of NGST). It is recommended that either the EM or STM of the Optical Head be enhanced sufficiently that the qualification of the structural design prior to CDR. Thus will allow the Optical Head of the Flight Model to be acceptance tested only and thus subject to the minimum levels of stress possible. It is felt that this approach, whilst making the first two years of the project more difficult, will retire a significant amount of risk on the design of the Optical Head.

It is proposed that the EM electronics be limited to an electrically representative version only, ASICs being implemented by FPGAs etc. It is felt that because of the relatively benign environment on the warm side of the spacecraft that this is an acceptable risk.

2.3.3 Phase C/D

Given a successful Phase A/B, a two year Phase C/D to provide a PFM model electronics unit and FM Optical Head seems a reasonable proposition. Ultimately success will depend on good communications within the instrument team and spacecraft teams and effective definition of interfaces and software. This will be especially true if the proposed level of instrument integration within the ISIM come to fruition.

3.0 *Cost Estimate*

Provided by the Canadian Space Agency.

4.0 *References*

Adelberger, K.L., Steidel, C.C., Giavalisco, M., Dickinson, M., Pettini, M., Kellogg, M., 1998, ApJ, 505, 18.

Baudrand, J., Guinouard, I., & Jocou, L. 1998, Fiber Optics in Astronomy III, ASP Conf. Series, 152, 32.

Childs, B.G., Harvey, P.J., & Ritchie, H. 1973, Chalk River Report AECL-4454.

Gardner, J.P., and Satyapal, S., 1999, AAS, 194, 9106.

Giavalisco, M., Steidel, C.C., Adelberger, K.L., Dickinson, M.E., Pettini, M., Kellogg, M., 1998, ApJ, 503, 543.

Fernandez-Soto, A., Lanzetta, K., Yahil, A., 1999, ApJ, 513, 34.

Kennicutt, R.C., 1998, In "Proceedings of the 34th Liege International Astrophysics Colloquium "The NGST: Science Drivers and Technological Challenges", (ESA SP-429, October 1998), p81.

Kenworthy, M.A., Parry, I.R., & Ennico, K.A. 1998, Fiber Optics in Astronomy III, ASP Conf. Series, 152, 300.

Kobulnicky, H.A., Kennicutt, R.C., Pizagno, J.L., 1999, ApJ 514, 544.

Levin, K.H. et al. 1993, Fiber Optics in Astronomy II, ASP Conf. Series, 37, 295.

Richardson, E.H., Moore, A., Tillemar, T., & Crampton, D. 1999, 3D Spectroscopy, ASP Conf. Series, in press.

Tecza, M. & Thatte, N. 1998, Fiber Optics in Astronomy III, ASP Conf. Series, 152, 271.

List of Appendices

(These will be available at http://www.hia.nrc.ca/facilities/ngst/ngst_rpt.html

Appendix A1: NGST Integral Field Spectroscopy at High Spectral and Spatial Resolution by S. Lilly.

Appendix A2: Steinbring simulation report

Appendix B1: NGST MOS Slit Mask Design Study by B. Buckham, J.A. Carretero, D. Erickson, M. Nahon and I. Sharf (Initial survey of a broad range of options plus a survey of cryogenic actuators and material properties at low temperatures)

Appendix B2: Fibre Optic Positioning Device for the NGST Multi-Object Spectrograph by D. Erickson, B. Buckham. M. Nahon and I. Sharf.

Appendix B3: Mechanically Actuated Rolls of Slits Concept by Camosun College Mechanical Technology students.

Appendix B4: Mechanically Actuated Strip Method by University of Victoria Mech 400 students.

Appendix B5: Extensive appendices for Appendix B4.

Appendix B6: Surface Micromachined Reconfigurable Slit Mask by S. Bakshi, M. Parameswaran and M. Syrzycki.

Appendix B7: Solid Block Image Slicer for NGST by E.H Richardson.

Appendix B8: Moderate Resolution Spectrograph by E.H Richardson.