

Applications of Integrated Photonic Spectrographs in astronomy

R. J. Harris[★] and J. R. Allington-Smith

Centre for Advanced Instrumentation, Physics Dept, Durham University, South Rd, Durham DH1 3LE

Accepted 2012 October 19. Received 2012 October 19; in original form 2012 July 19

ABSTRACT

One of the problems of producing instruments for extremely large telescopes (ELTs) is that their size (and hence cost) scales rapidly with telescope aperture. To try to break this relation alternative new technologies have been proposed, such as the use of the Integrated Photonic Spectrograph (IPS). Due to their diffraction-limited nature, the IPS is claimed to defeat the harsh scaling law applying to conventional instruments. In contrast to photonic applications, devices for astronomy are not usually used at the diffraction limit. Therefore, to retain throughput and spatial information, the IPS requires a photonic lantern (PL) to decompose the input multi-mode light into single modes. This is then fed into either numerous arrayed waveguide gratings (AWGs) or a conventional spectrograph.

We investigate the potential advantage of using an IPS instead of conventional monolithic optics for a variety of capabilities represented by existing instruments on 8 m telescopes and others planned for ELTs. To do this, we have constructed toy models of different versions of the IPS and calculated the relative instrument sizes and the number of detector pixels required. This allows us to quantify the relative size/cost advantage for instruments aimed at different science requirements. We show that a full IPS instrument is equivalent to an image slicer. Image slicing is a beneficial strategy for ELTs as previously demonstrated. However, the requirement to decompose the input light into individual modes imposes a redundancy in terms of the numbers of components and detector pixels in many cases which acts to cancel out the advantage of the small size of the photonic components. However, there are specific applications where an IPS gives a potential advantage which we describe. Furthermore, the IPS approach has the potential advantage of minimizing or eliminating bulk optics. We show that AWGs fed with multiple single-mode inputs from an PL require relatively bulky auxiliary optics and a 2D detector array which significantly increases the size of the instrument. A more attractive option is to combine the outputs of many AWGs so that a 1D detector can be used to greatly reduce the number of detector pixels required and provide efficient adaptation to the curved output focal surface.

Key words: instrumentation: spectrographs – techniques: imaging spectroscopy – techniques: spectroscopic.

1 INTRODUCTION

Spectroscopy is one of the most useful tools in astronomy, with applications in fields ranging from cosmology to exoplanet studies. Arguably, the most common form in astronomy is dispersive spectroscopy which uses a dispersive element to separate different wavelengths.

In its simplest form a dispersive spectrograph contains four components: slit to isolate the area to be dispersed, a collimator, a grating or prism to disperse the light and a camera and detector to record the intensity at each wavelength. In order to retain throughput the slit of the spectrograph is usually matched to the seeing of the telescope.

If this is not diffraction limited then the collimated beam must increase in size as the telescope grows in size in order to maintain the same resolution (Lee & Allington-Smith 2000). The immediate consequence of this is that the optics and disperser must also grow in size. This then leads physical problems such as stresses and flexure in the materials, along with the difficulties inherent in building these large monolithic structures.

In order to use the same design principles as existing instruments more exotic materials and construction processes are needed, which drives the costs of building the instruments up. Using the conventional approach, the cost of instruments scales with at least the square of the telescope aperture (Bland-Hawthorn & Horton 2006).

In order to reduce these problems the input to the spectrograph can be sliced by dividing the input field (the slit) into a number of thin slices. Each of these can then be fed into a spectrograph

[★]E-mail: r.j.harris@durham.ac.uk

or spectrographs. This technique is known as image slicing. Each of these spectrographs produces a spectrum for each slice. The resulting data is then reformatted into a 3D data cube with axes x , y , λ allowing the reconstruction of the initial image.

There are several different methods for image slicing detailed in the literature, from lenslet arrays feeding fibres to slicing mirrors (Allington-Smith 2007a). All have inherent advantages and disadvantages. The various methods have all been widely adopted in recent years and are used in various forms in instruments on the current generation of 8 m telescopes. Theoretical investigations into the scaling laws in relation to image slicing suggest that it will be an even more powerful tool to reduce instrument sizes and costs on the next generation of 30 m telescopes (Allington-Smith 2007b).

In this paper, we examine the potential applications of the Integrated Photonic Spectrograph (IPS), which shares some of the features of image slicers. Despite the physical differences they obey the same basic physical laws as conventional instruments.

The IPS devices take light from an input fibre which is usually matched to the seeing limit (as with conventional fibre-fed instruments) and so supports many modes. The light from this multi-mode fibre (MMF) is then split into a number of single-mode fibres (SMFs) by a photonic lantern (PL). At this point two options have been proposed (Bland-Hawthorn et al. 2010).

$P = 1$. The first requires a reformatting component (Thomson et al. 2011; Birks et al. 2012) to form a slit of SMFs which can then be dispersed by bulk optics. We shall call this the semiphotonic case.

$P = 2$. In the second, the SMFs are then fed into arrayed waveguide gratings (AWGs) which disperse the light into individual spectra, with one or more spectra per AWG (Bland-Hawthorn et al. 2010). These spectra have the advantage that they are in a linear format, so can be sampled using an array of fibres or a linear detector. This is our fully photonic case.

While the principles for both have been demonstrated (Cvetojevic et al. 2009, 2012; Leon-Saval, Betters & Bland-Hawthorn 2012), so far these have only used single or few modes from a single input, resulting in limitations in throughput and field. They are also not optimized in terms of size or the components used. This led us to investigate how complete instruments would perform compared to current instrumentation, with initial results suggesting that a fully photonic IPS would be best suited to small, diffraction-limited telescopes with small fields of view (Harris & Allington-Smith 2012).

In this paper, we determine the application areas where the IPS may have an advantage over conventional instrumentation. After noting the formal similarity between image slicing and photonic spectroscopy in Section 2, we consider the requirement for the field of view of the instrument in Section 3. Simplified models of the IPS are presented in Sections 4 and 5. The results of comparing conventional and IPS instruments are given in Section 6. In Section 7, we discuss different ways to reduce the number of AWGs and/or detector pixels, before presenting our conclusions in Section 8.

2 INSTRUMENT SIZE SCALE RELATIONSHIPS

It is often claimed that IPS violates the relationship between telescope diameter and spectral resolution:

$$R = \frac{m\rho\lambda W}{\chi D_T} = \frac{2 \tan \gamma D_{\text{col}}}{\chi D_T}, \quad (1)$$

where R is the resolution of the instrument, m is the diffraction order, ρ is the ruling density, λ is the wavelength, W is the length of intersection between the grating and collimated beam, χ is the angular slitwidth, D_T is the diameter of the telescope, γ is the blaze angle, ρ is the ruling density and D_{col} the diameter of the collimated beam.

This applies to a slit spectrograph using a diffraction grating as the dispersive element and shows that for a given resolution, angular slitwidth and blaze angle, the diameter of the collimated beam must increase in proportion to the telescope aperture, leading to a bigger instrument.

Unlike a conventional spectrograph the input to the IPS must be diffraction limited ($\lambda \approx \chi D_T$) due to its single-mode nature, so the resolution can be shown to be of the form

$$R = \frac{mN_{\text{wg}}}{C}, \quad (2)$$

where N_{wg} is the number of waveguides in the AWG or number of rulings on a conventional grating and C a factor to account for manufacturing errors (Lawrence et al. 2010). This has no dependence on telescope diameter so it would appear to break the relation. It must be noted though, this applies to a device operated at the diffraction limit of the telescope, not at the seeing limit as with equation (1).

To examine what happens when the input at the seeing limit, we consider the number of spatial modes in a conventional step-index fibre (Cheo 1990) which can be approximated as

$$M = \frac{V_{\text{fibre}}^2}{4}. \quad (3)$$

It is useful to remember that each spatial mode has two polarization states, though we do not include the factor here as each SMF accepts two polarizations. The associated V parameter is

$$V_{\text{fibre}} = \frac{\pi s \Theta}{\lambda}, \quad (4)$$

where s is the diameter of the fibre core (assumed equal to the slitwidth in equation 1) and Θ the numerical aperture at which the fibre is operated. This must be less than the limiting (i.e. maximum) numerical aperture of the fibre. Noting that

$$\Theta \approx \frac{1}{2F_T} \quad (5)$$

$$s = \chi f_T = \chi F_T D_T, \quad (6)$$

where F_T the telescope focal ratio and f_T is the telescope focal length, the number of modes is given by

$$M = \left(\frac{\pi \chi D_T}{4\lambda} \right)^2. \quad (7)$$

Therefore, it can be seen that for each sampling element the number of modes increases as the square of the telescope diameter in a similar way to the number of slices at the diffraction limit $(\chi D_T/1.22\lambda)^2$. This confirms that, to first order, photonic spectrographs are bound by the same scaling laws as conventional spectrographs. In what follows, we attempt to quantify areas where photonic spectrographs may confer an advantage, and suggest modifications which may allow photonic spectrographs to make a significant impact on future astronomical instrumentation.

3 THE INPUT FIELD AND SPATIAL MULTIPLEX

In the previous section, we calculated the number of modes per spatial sampling element (spaxel). As with diverse field spectroscopy

(Murray & Allington-Smith 2009), photonic spectrographs address a number of individual spaxels, which can be grouped (as in Integral Field Spectroscopy; IFS), or separate (as in Multi Object Spectroscopy; MOS). In order to fairly compare with conventional instrumentation we need to make sure that we sample the same number of spaxels (e.g. observe the same field).

A long slit can be thought of as a series of spaxels joined to form a rectangle of size $1 \times N$, where N is the total number of spaxels. In IFS, the field is equivalent to a series of slits (each composed of linked spaxels) joined so the total number of spaxels is $N = N_x N_y$, where N_x is the number of spaxels in the x direction and N_y in the y direction. MOS can be thought of as the same number of spaxels, N , distributed throughout the field of the telescope and brought together to form a long slit. See Fig. 2 for an illustration of this.

An important consideration is the sampling of the field. From Fig. 1 and equation (7) it can be seen that the number of modes produced is dependent on the overall size of the field, not the individual spaxel size (as the number of modes per spaxel is proportional to the square of the spaxel size). This means that the number of components for the IPS (and hence the approximate size of the instrument) required for the instrument will not depend on the sampling scale.

However, the amount of spatial information and throughput will depend on the sampling scale. Although it might appear best to reduce the spaxel size and the number of modes, this will reduce the coupling efficiency and throughput (Corbett & Allington-Smith 2006). At the other extreme the use of very large fibres would result in loss of spatial information. A balance must be found between throughput and spatial resolution. We do not investigate this fully here as it does not affect the total number of modes in the field or the required number of detector pixels. Thus, we choose to make our spaxel size equal to the full width at half-maximum (FWHM) of the seeing.

In order to calculate the scale length of the instrument we take the cube root of the volume of the instrument. We can calculate

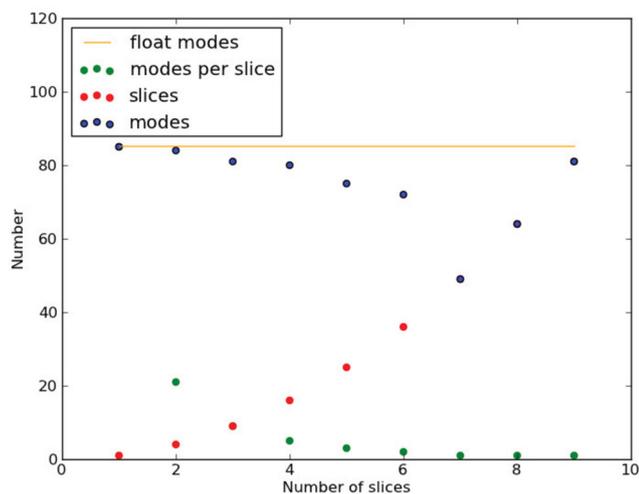


Figure 1. An example of the number of modes generated from a single spaxel on an 8 m telescope of 0.5 arcsec FWHM seeing at $\lambda = 1650$ nm. Fixing the size of the spaxel to the FWHM gives a single spaxel (here number of slices = 1); this spaxel is large and contains many modes. Slicing the spaxel produces smaller spaxel sizes, but larger numbers of them (the number of spaxels is the slices squared). This results in the same total number of modes in the area (the horizontal yellow line). The variation in the blue is due to the integerization of modes within individual slices. Note that the three final red points lie under the blue ones.

this from the number of spaxels in the total field passed to the spectrograph

$$\text{scale length} = \sqrt{NM(\lambda_{\min})^{P-1}L_xL_yL_z}, \quad (8)$$

where M is a function of the shortest wavelength in the spectrograph (λ_{\min}), in order to account for all spatial modes. L_x , L_y and L_z are the lengths of an individual component spectrographs or AWGs in the x , y and z direction, respectively, which will be defined in the next two sections. $P = 1$ and 2 represent the semiphotonic and photonic cases, respectively.

4 THE SEMIPHOTONIC IPS ($P = 1$)

The semiphotonic case involves taking an individual spaxel and using a re-arranged PL to form a diffraction-limited slit, which is then dispersed by bulk optics (Bland-Hawthorn et al. 2010).

4.1 Model geometry

Slicing the input of a spectrograph has already been examined theoretically in Allington-Smith (2007b). The paper took existing instruments and sliced the input, either adding the slices to the length of the slit or placing them into replica spectrographs. It showed that slicing could result in an instrument with a slightly smaller overall volume, though the instruments sliced to the diffraction limit were shown to be larger than their counterparts due to the extra components required. This is important to us as we showed that photonic spectroscopy is similar to image slicing to the diffraction limit in Section 2.

As conventional image slicing has already been examined, we restrict ourselves to examining only the IPS concept, which takes each individual spaxel (not a number of them) and separates it into a single spectrograph.

We will be using the modified model from Allington-Smith (2007b) described in Harris & Allington-Smith (2012) and adding this to our results. As the input to each spectrograph now depends on the number of modes per spaxel, we shall be setting the length of the slit to the number of modes (equation 7) instead of n_y in the previous papers.

4.2 Semiphotonic model limits and calibration

In order to calibrate the model we use the same method as Allington-Smith (2007b), with the $S = 1$ case oversizing using a multiplicative factor and the $S = 2$ case oversizing the spectrograph input beam. The scaling factors for our instruments can be found in Table 1.

Table 1. The scaling parameters for the semiphotonic versions of the conventional instruments. The scaling scenarios are described fully in Allington-Smith (2007b) and Harris & Allington-Smith (2012).

Instrument	S = 1		S = 2	
	a (m)	b	a (m)	b
GNIRS	0.1	2.1	0.46	1.1
CRIRES	0.1	2.2	0.61	1.1
NIFS (J)	0.1	7.0	0.86	1.1
SINFONI (H)	0.1	7.0	0.86	1.1
IRMS	0.1	8.0	2.06	1.1
IRIS	0.1	10.0	1.46	1.1

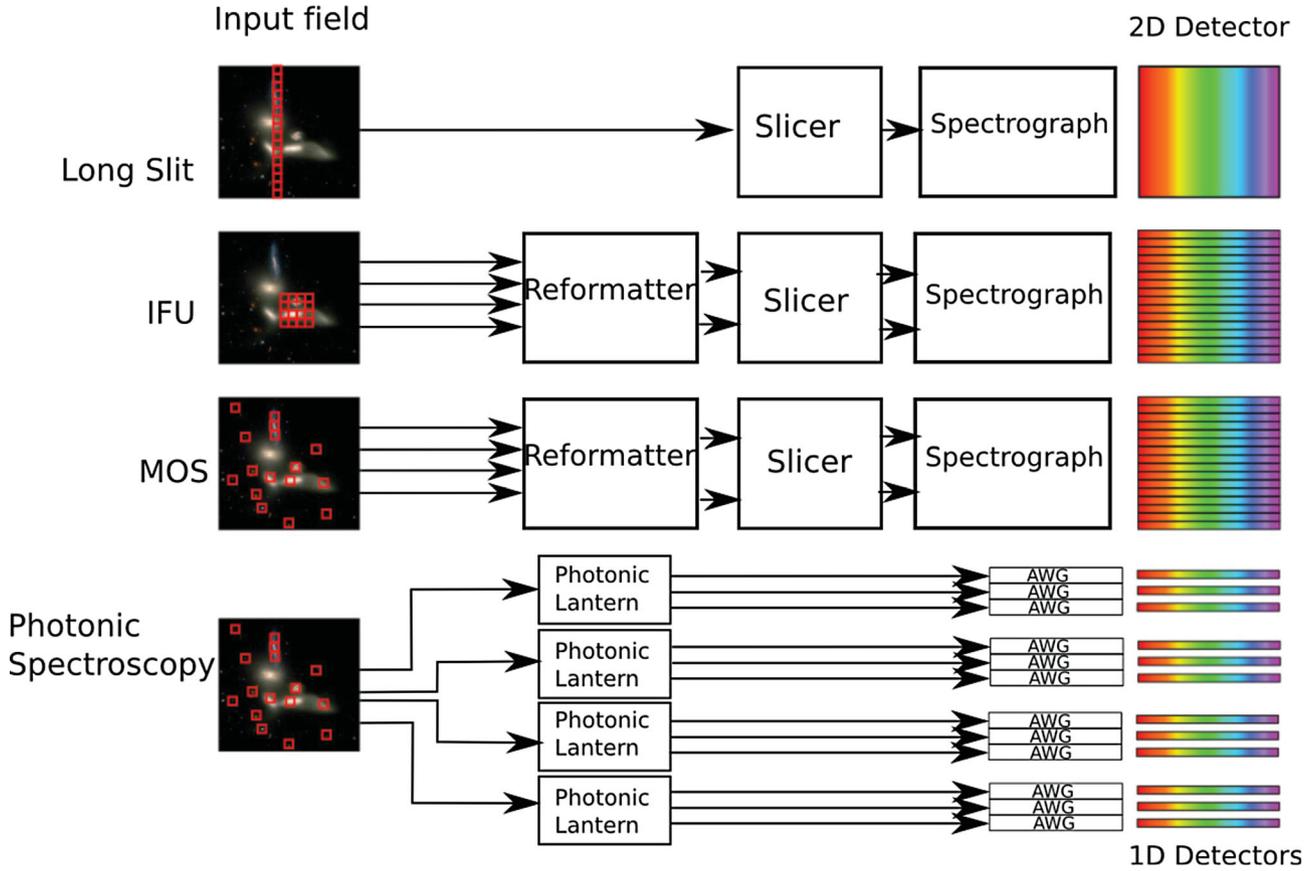


Figure 2. An illustration of conventional slicing and photonic slicing. All methods sample an area of the same size (e.g. the same number of spaxels). The three conventional methods reformat the input and disperse it, producing one spectrum per spaxel. The fully-photonic option takes each input spaxel and splits it into individual modes using an PL. Each of these modes is fed into an AWG to produce a spectrum. These then need to be recombined and summed to produce the spectrum for the spaxel. The semiphotonic option uses the same PL, but this is then reformatted into a long slit and fed into conventional spectrographs.

5 THE FULLY PHOTONIC IPS ($P = 2$)

For our fully photonic model, we shall concentrate on modelling the size of the AWGs, not the components that feed them. We shall include a factor for our detector sizes.

5.1 Model geometry

For the fully photonic model, we first need to consider the geometry of AWGs. These are available in many different variations, especially with respect to the geometry which generates the path difference between waveguides (e.g. S-bend, circular, horse-shoe). To keep the toy model simple, we have chosen a reflective AWG (Grave de Peralta et al. 2003, 2004) using a Rowlands circle arrangement for the free propagation region (FPR). We have removed the bend at the end of the waveguide array for simplicity. Because of this it looks almost identical to a conventional double-pass echelle spectrograph.

Using the definitions in Fig. 3, we arrive at the following equations for the size of the AWG model

$$L_x = (\max(D, E) + a_{\text{awg}})b_{\text{awg}} \quad (9)$$

$$L_y = (c_{\text{awg}} + w)b_{\text{awg}} \quad (10)$$

$$L_z = (a_{\text{awg}} + A + (N_{\text{wg}} - 1)\Delta L)b_{\text{awg}}. \quad (11)$$

Here A is the x -length of the FPR, ΔL the length difference between adjacent waveguides to achieve the required order for a given central wavelength (λ_c), D is the length containing the waveguides (analogous to the illuminated length of the grating in a standard echelle grating), E the x -length of the detecting surface and w is the waveguide diameter. The oversizing parameters a_{awg} , b_{awg} and c_{awg} parametrize the extra size required to implement a practical device.

First, we calculate the appropriate dispersion order, m , in terms of the free spectral range (FSR, $\Delta\lambda_{\text{FSR}}$) for an AWG,

$$m = \frac{\lambda_{\text{min}}}{\Delta\lambda_{\text{FSR}}}. \quad (12)$$

Setting $D = N_{\text{wg}}/\rho$, this can be combined with equation (2), where ρ is the density of waveguides (analogous to the ruling density of a conventional disperser) to give

$$D = \frac{CR}{m\rho}. \quad (13)$$

The physical extent of the FSR in an AWG is $X_{\text{FSR}} = (\lambda_{\text{min}}L_F\rho/n_s)$, where n_s is the refractive index of the slab and L_F the length of the free space propagation region. Combining with geometrical arguments gives

$$E = L_F \sin\left(\frac{\Psi}{2}\right) = L_F \sin\left(\frac{\lambda_{\text{min}}\rho}{n_s}\right), \quad (14)$$

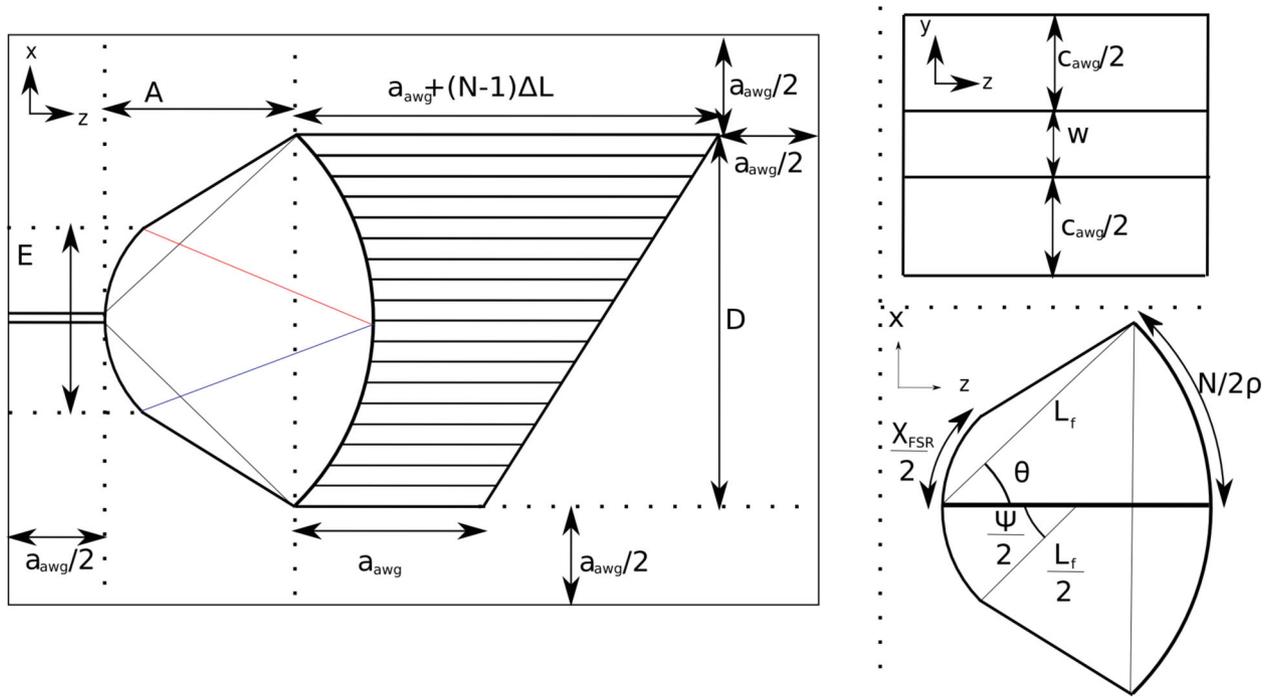


Figure 3. The fully photonic model. The left image shows the x - y view of the AWG, with the (left to right) input fibre, FPR and waveguides. The top-right image shows the side view of the AWG model, with the top and bottom cladding layers and the layer containing the waveguides in the centre. The bottom-right image is an enlargement of the FPR which is a Rowlands circle arrangement.

where Ψ is defined in Fig. 3. In order to calculate ΔL we make use of the equation for calculating the central wavelength of the AWG

$$\Delta L = \frac{\lambda_c m}{n_c}, \quad (15)$$

where n_c is the refractive index of the waveguides, the central operating wavelength is $\lambda_c = \lambda_{\text{min}} + \Delta\lambda_{\text{FSR}}/2$, so A can be calculated from geometry as

$$A = L_F \cos(\theta) = L_F \cos\left(\frac{N_{\text{wg}}}{2\rho L_F}\right), \quad (16)$$

where θ is defined in the figure and N_{wg} is calculated using equation (2). In order to calculate L_F , we make use of the fact that the imaging requires the number of detector pixels to be able to adequately sample at the resolution required (equivalent to sampling of the echelle model in Allington-Smith (2007b)). To do this, we take the dispersion relation

$$\left(\frac{\delta\lambda}{\delta x}\right) \simeq \left(\frac{d\lambda}{dx}\right) = \frac{n_s}{L_f m \rho}. \quad (17)$$

and combine it with equation (12), setting $\delta x = N_0 d_p$, where N_0 is the oversampling and d_p is the size of the pixels. We also take the equation for the spectral resolution $\delta\lambda = \lambda_{\text{min}}/R$. Minimizing to obtain the maximum L_F we find

$$L_F \geq \frac{n_s N_0 d_p R \Delta\lambda_{\text{FSR}}}{\rho \lambda_{\text{min}}^2}. \quad (18)$$

Finally, we can calculate the number of pixels we need for the required resolution

$$N_P = \frac{L_{\text{FSR}}}{N_0 d_p} = \frac{\lambda_c L_F \rho}{n_s d_p}. \quad (19)$$

5.2 Fully photonic model limitations and calibration

Astronomical spectrographs are usually designed to operate with a large FSR (typically several hundred nm). This is a problem for the IPS because conventional telecoms AWGs are designed with low FSR in order to deal with the discrete narrow-band input from the telecoms industry. For astronomy, single AWGs need to be re-designed to work in lower spectral orders by reducing the path difference between adjacent waveguides. This requires more waveguides to maintain the maximum theoretical resolution and an increase in L_F to maintain a practical one (see equation 18). This produces very large AWG dimensions which cannot be manufactured due to chip manufacturing size constraints (Lawrence et al. 2010).

We wish to avoid this problem and retain a fully integrated design with no external optics. As such we make use of the tandem AWG arrangement, where a primary AWG filters the light by wavelength into secondary AWGs, each encompassing a fraction of the original FSR (see Fig. 4 and Takada et al. 2001). This allows the individual component dimensions to be within manufacturing limits whilst allowing our full device to sample the correct FSR. It would also allow the AWG design process to remain similar to current specifications.

Since no full-scale AWG instruments currently exist, we are required to use a bottoms-up approach to estimate its size. To do so we take an existing AWG acquired from Gemfire Livingstone. We use its known parameters and adjust the model-produced dimensions until they match the real ones. In order to emulate Allington-Smith (2007b), we set the scaling parameters to two extremes.

S = 1. Minimize $b_{\text{awg}} = 1.1$, which yields a $a_{\text{awg}} = 10$ mm

S = 2. Minimize $a_{\text{awg}} = 0$ mm, which yields $b_{\text{awg}} = 2.8$.

For both scenarios, we keep $c_{\text{awg}} = 0.7$ mm as the device is planar so the height should not change.

For simplicity, we will not include the volume of the initial MMF bundle, the PL or the housing of the instrument. We will however

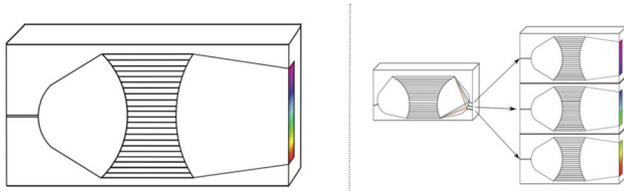


Figure 4. An illustration of the tandem and single AWG setup. The image on the left shows the conventional AWG dispersing the whole spectrum. The image on the right is the tandem configuration. The initial AWG (or other dispersive optic) splits the light by wavelength (here to $\Delta\lambda_{\text{FSR}}/3$) and feeds the second set of AWGs. This has the advantage that each individual AWG can be smaller, though it requires more AWGs, additional components (feeding fibres) and is subject to extra loss of throughput. Note that the length of the output must be the same in both cases in order for the detector to sample adequately.

include an estimate for the size of the detector. This value is calculated assuming the size of a typical detector sub-system including the cryostat. This is estimated as 10^{-7} m^3 per detector pixel.

6 RESULTS

To estimate the uncertainties within our model, we use both of our oversizing options for both models and also vary C between 1 (diffraction limited) and 4 (the initial results obtained in Cvetojevic et al. 2009) for our fully photonic instrument. This gives us two extremes and allows for the current maximum of $C = 1.6$ achieved in Cvetojevic et al. 2012.

To test our models, we choose two sets of instruments. First, those designed to represent current instrumentation on 8 m telescopes. Then, we test against instruments that have been designed for the Thirty Meter Telescope (TMT). We shall only investigate a single mode of operation for each instrument. This may be oversimplifying as most instruments are designed to operate at different resolutions and at different wavelengths by changing gratings or optics. It however keeps the model simple and IPS devices could be adapted to suit various purposes; this is discussed in the conclusions.

The first set of instruments are in current use on 8 m telescopes and are intended to represent generic instrument types. We use the parameters in Table 2 to calculate the volume of the instruments. The instruments we have chosen are as follows.

(i) Gemini Near-InfraRed Spectrograph (GNIRS) on Gemini-North. The instrument has an overall wavelength range of 1.0–5.4 μm , resolutions of between 1700 and 18 000, and slitwidths of between 0.1 and 1.0 arcsec. It has an imaging mode, a long-slit mode and originally an IFS mode (destroyed during maintenance

at the telescope). We will be comparing our photonic instrument to the long-slit configuration.

(ii) CRYogenic high-resolution InfraRed Echelle Spectrograph (CRIRES) is a high-resolution spectrograph on the Very Large Telescope (VLT). It is designed to operate between 1.0 and 5 μm , with a resolution of up to 10^3 . We have chosen it to illustrate a long-slit high-resolution spectrograph.

(iii) Near-Infrared Integral Field Spectrometer (NIFS) on Gemini-North is our first Integral Field Unit (IFU) instrument. It is designed to work with the Adaptive Optics (AO) system, over 0.9–2.4 μm .

(iv) Spectrograph for INtegral Field Observations in the Near Infrared (SINFONI) on the VLT is our second IFU instrument. It operates in the 1.1–2.45 μm range, again with AO.

We have also chosen two hybrid instruments proposed for the Thirty Meter Telescope (TMT; Simard et al. 2010). The Infra-Red Multi-object Spectrograph (IRMS) will employ 10 or more IFUs. Each one will have a 2×2 arcsec-squared field of view with a 50 per cent of the energy enclosed by 50 mas at wavelength 1 μm , resulting in 1600 spaxels per IFU. The Infrared Imaging Spectrograph (IRIS) has three IFU units, two of which will be lenslet arrays (for observing smaller fields) and one will be an image slicer (larger fields). Here, we model the slicer, which has 88 mirror facets, but will keep the best resolution possible with our AWG model. As with the 8 m instruments, the instrument scale lengths are fitted to values taken from the literature, see Table 2.

Table 3 shows the resulting parameters in the fully photonic case. The total number of AWGs required are shown in the second column; this number will be in the tens of thousands for 8 m instruments and the hundreds of thousands for the 30 m instruments. The large number of AWGs requires rigorous quality control to test the large number of individual components. This may be of advantage though, as the individual AWGs should be less prone to flexure and, due to their modular nature, are better suited to mass production and upgrades and expansion to suit cashflow. Note that the size predictions do not include provision for mounting hardware required to support the instrument components or to provide a suitable controlled environment.

The next eight columns show the different resulting normalized scale lengths of the instrument. Note that this is the scale length of the overall instrument, not the individual components. The first four are the scale lengths without provision for the detector size and show that the total size of the 8 m instruments will be of the same order as the conventional instrument. If the diffraction limit can be achieved, the resulting instruments are smaller for all scenarios, with the exception being NIFS with the $S = 2$ scaling. If the diffraction limit

Table 2. Table of input parameters for all instruments. Symbols except for N_{IFU} , the number of IFUs in the instrument, are explained in the text. The numbers are taken (and approximated from) Allington-Smith (2007b) for GNIRS; Kaeufl (2004) for CRIRES; McGregor et al. (2003) for NIFS; Eisenhauer et al. (2003) for SINFONI; Eikenberry et al. (2006) for IRMS and Larkin et al. (2010) for IRIS.

Instrument	χ (arcsec)	$N_y(N_x)[N_{\text{IFU}}]$	Total spaxels	R	λ_c (nm)	Δ_{FSR} (nm)	ρ (mm^{-1})	Vol (m^3)
GNIRS	0.3	330(1)	330	5900	1650	400	31.7	2.00
CRIRES	0.3	200(1)	200	100 000	1650	48	31.6	3.00
NIFS (J)	0.1	30(29)	870	6050	1250	600	600	2.75
SINFONI (H)	0.2	32(32)	1024	3000	1650	400	128.57	2.75
IRMS	0.1	40 (40)[10]	16 000	10 000	1200	400	128.57	16.00
IRIS	0.1	60(60)[1]	4000	8000	1200	400	310	55.00

Table 3. Table of the resulting scale lengths of the respective fully photonic instruments. The total number of modes in the whole instrument is shown first, with the next four columns showing the respective sizes for the model with no detector. This is followed by the model with detector. All of the scale lengths are normalized to the cube root of the volume in Table 2. The AWG model uses a waveguide separation of $\rho = 200 \text{ mm}^{-1}$.

Instrument	Total number of AWGs	Normalized scale length							
		No detector				With detector			
		$C = 1$		$C = 4$		$C = 1$		$C = 4$	
	$S = 1$	$S = 2$	$S = 1$	$S = 2$	$S = 1$	$S = 2$	$S = 1$	$S = 2$	
GNIRS	13 000	0.49	0.87	1.46	1.55	0.51	0.89	1.47	1.55
CRIRES	6300	0.50	0.90	1.19	1.32	0.60	1.09	1.22	1.42
NIFS	8900	0.79	1.45	1.66	1.93	0.80	1.46	1.66	1.93
SINFONI	28 000	0.39	0.64	1.36	1.39	0.39	0.66	1.36	1.39
IRMOS	520 000	1.69	3.08	3.53	4.09	1.71	3.11	3.53	4.11
IRIS	110 000	0.72	1.07	1.34	1.48	0.73	1.08	1.35	1.49

Table 4. Further information on the fully photonic model. The AWG model uses a waveguide separation of $\rho = 200 \text{ mm}^{-1}$.

Instrument	Modes per spaxel	Total detector pixels/(10^6)	Reference instrument pixels/(10^6)
GNIRS	39	60.69	1.05
CRIRES	31	47.31	2.10
NIFS	10	112.43	4.19
SINFONI	27	66.50	4.19
IRMOS	32	6263.44	83.89
IRIS	32	1127.42	16.78

cannot be achieved the instruments will have a scale length larger than the original instruments. The results for the 30 m instruments are similar to the 8 m ones, with the $S = 1$ scaled case of IRIS being slightly smaller and the rest being slightly larger.

The second four include the provision for detector and show similar results, though the scale lengths are increased slightly as expected. This shows though that the size of the additional detector pixels (discussed later) will not pose a significant size restriction on the instrument.

Table 4 shows the results corresponding to the number of modes per spaxel and hence the requirements in terms of detector pixels. The second column shows that all of the instruments will have around 30 modes per spaxel, with the exception of NIFS, which will have 10. This causes problems with oversampling in the fully photonic model due to each mode needing to be sampled using two detector pixels per resolution element (Nyquist sampling). There needs to be some way of combining the individual spectra to stop massive oversampling (shown in column 3). This will be discussed later, but will probably involve additional components, increasing the size of the instrument.

Table 5 shows the results from the semiphotonic model. The second column shows that the number of replica photonic instruments will be 10^3 – 10^4 . As such using the semiphotonic method will require mass production of the replica spectrographs which is not common in the astronomical instrumentation community. The alternative is to put many spaxels in the same spectrograph and would require a balance between redundancy and overlapped components for this version to work.

The next two columns show that using the semiphotonic model for the instrument will result in much larger instruments. This matches with the results of Allington-Smith (2007b), where as the input was sliced more the instrument tended to get bigger. As stated in Section 4, we are slicing the instrument to the diffraction limit in

Table 5. Results for the semiphotonic model. All of the scale lengths are normalized to the cube root of the volume in Table 2. The semiphotonic model uses ρ stated in Fig. 2.

Instrument	Number of replica spectrographs	Normalized scale length		Number of detector pixels ^a /(10^6)
		$S = 1$	$S = 2$	
GNIRS	330	4.30	5.30	1.05
CRIRES	200	3.06	4.66	2.10
NIFS	870	9.55	9.54	4.19
SINFONI	1024	8.62	9.36	4.19
IRMOS	16 000	26.07	32.48	83.89
IRIS	3600	8.49	9.20	16.78

^aThe number of detector pixels assumes that the modes in each spaxel can be reduced on to the detector appropriately, which may not be the case.

the spatial direction and then separating each spaxel into a separate spectrograph, which imposes huge redundancies. The number of required detector pixels are shown in the final column, for this we are assuming that all the detector separate modes can be combined on to a linear detector, which may not be possible.

7 MODIFICATIONS TO INTEGRATED PHOTONIC SPECTROGRAPHS

From the results already discussed, it is clear that IPSs in their current state offer little or no advantage in terms of size and detector pixels when compared with existing instruments on large telescopes or those planned for extremely large telescopes (ELTs). However, it is possible to envisage modifications to the fully photonic device which would make it possible to exploit the unique features of photonic spectrographs. One such scheme is already being studied (Cvetojevic et al. 2012).

7.1 Multiple-input arrayed waveguide gratings

So far we have restricted ourselves to one input per AWG (i.e. one mode per AWG), which as shown in previous sections requires many AWGs. Placing multiple inputs per AWG would reduce the dependence on equation (7), at the extreme eliminating it entirely if all the inputs could fit into a single AWG. In order to introduce these extra inputs additional fibres are placed at different positions on the input FPR. This introduces a path difference between each input with respect to the central one (see Figs 5 and 6). This path difference is carried through the system and results in the output

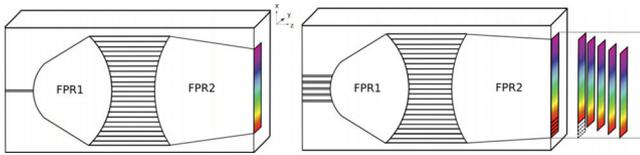


Figure 5. Cutaway diagram illustrating the difference between the single and multiple input versions of the AWG. To the left is the single input, which would make use of a linear detector array to sample the output spectrum. To the right the multiple input version. Here five inputs produce five separate spectra that overlap at the output of the AWG. This has a couple of implications, first the spectra would need to be cross-dispersed in order to be sampled and secondly, the end of the second FPR would need to be larger (though not the input which would remain the size of the waveguides). The path difference in the waveguides is not illustrated in this diagram for simplicity.

spectrum of each mode being shifted in the dispersion direction at the output. To remove the overlap between spectra, it will be necessary to introduce cross-dispersion.

The AWG also produces multiple diffraction orders (as with a conventional grating), so we also need to make sure that the inputs all lie within a region half that of the FSR of the central input. This is to stop the same wavelength from different orders lying in the same position in the linear output, resulting in the cross-dispersed spectra lying the same position on the detector. The refractive index change in fused silica is not great enough to disperse the light by polishing the AWG at an angle. In order to separate the spectra, the outputs need to be cross-dispersed using conventional optics (Cvetojevic et al. 2009). This means that each AWG will need a 2D detector, a dispersive element and collimating and camera lenses. Here, we look at the relative advantages of using cross-dispersion in the system.

7.2 Adding cross-dispersion

To cross-disperse, we need additional optics, which means that the device is no longer fully integrated, potentially making manufacture and maintenance more difficult, but reducing the number of AWGs required for the device. We construct a new toy model to see how the scale length of a cross-dispersed system (multiple inputs) compares to the one with linear arrays (single inputs). For this section, we have modified our fully photonic model so it is no longer reflective and the output of the second FPR outputs is linear (e.g. Fig. 5 and Lu et al. 2003). The first FPR is still in its original shape so as to allow the multiple inputs. Changing the model like this will affect the overall size of the instrument (due to the difference in AWG

design), but will still allow us to examine the relative sizes of the two scenarios.

We retain the single-input model for our comparison and use the length values calculated in previous sections. As such the scale length of the instrument (with all modes) is still $N_{\text{awg}}L_xL_yL_z$, where N_{awg} is the number of AWGs.

The calculation of the size of the multiple-input AWG option follows that for the cross-dispersion option requires the AWG, a collimator, a prism and then a camera in front of the detector (see Fig. 7). The equation for the volume of the system now becomes $N_{\text{awg}}L_xL_yL_z$, where $N_{\text{awg}} = N/N_i$, the total number of spaxels divided by the number of inputs per AWG and the dimensions being defined below.

We start by examining the output end of the AWG. For a single input, the x -length of the AWG system would be the same as described in equation (9), with E will now be X_{FSR} as we have flattened out the output. Adding extra inputs such as the ones illustrated in Fig. 7 will increase this x -length. The maximum distance between inputs must be less than X_{FSR} , to avoid the same wavelength in a different orders lying on the same position. For simplicity, we assume that evenly spaced inputs, which when combined with the previous condition, yield equation (26). We set the maximum number of inputs to be $X_{\text{FSR}}/D_{\text{input}}$, where D_{input} is the diameter of the input fibre (here set to 125 μm).

To calculate the y -length, we must consider how the spectra are to be cross-dispersed. We need to make sure that the output beam from the system is collimated. To do this, we make sure that the output angle of the collimator is small: $\theta_2 = w/2f_c < 0.01$, where θ_2 is the divergence in the collimator and f_c is the focal length of our collimator. As our system is diffraction limited, the diameter of our collimated beam is

$$D = \theta_1 f_c \approx \left(\frac{\lambda_{\text{max}}}{w} \right) f_c. \tag{20}$$

For cross-dispersion we use a prism, although a grating could also be used. We need to work out the required resolution of the prism, which is proportional to the number of inputs (e.g. as the number of inputs increases the FSR decreases by that factor). This gives

$$R_x = \frac{N_i \lambda_{\text{min}}}{\Delta \lambda_{\text{FSR}}}, \tag{21}$$

where R_x is the resolution of the cross-dispersed system. We can then combine this with the equation for the resolving power of a diffraction-limited prism (Foy & Foy 2002) to yield

$$t > R_x \left(\frac{d\lambda}{dn} \right) = \left(\frac{N_i \lambda_{\text{max}}}{\Delta \lambda_{\text{FSR}}} \right) \left(\frac{d\lambda}{dn} \right), \tag{22}$$

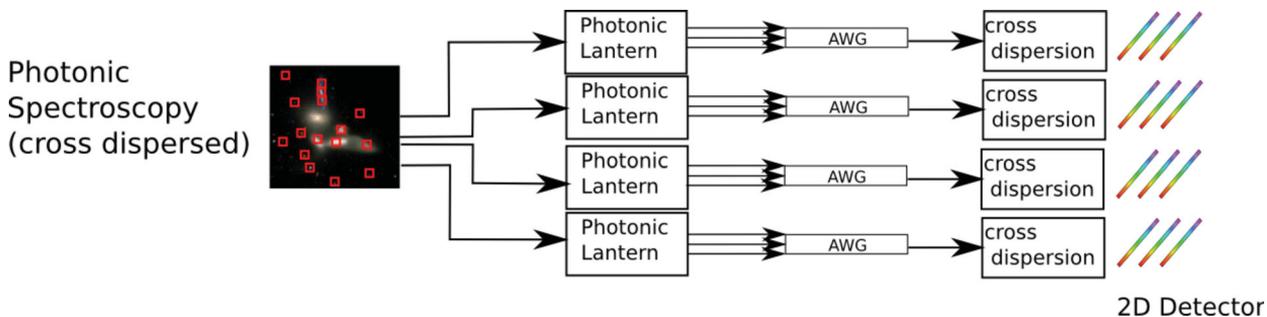


Figure 6. The multi-input model for the cross-dispersed system. Each spaxel from the input field is fed into an PL. The output SMFs are fed into AWGs, with multiple fibres in each AWG. The output from these AWGs is then cross-dispersed on to a 2D detector.

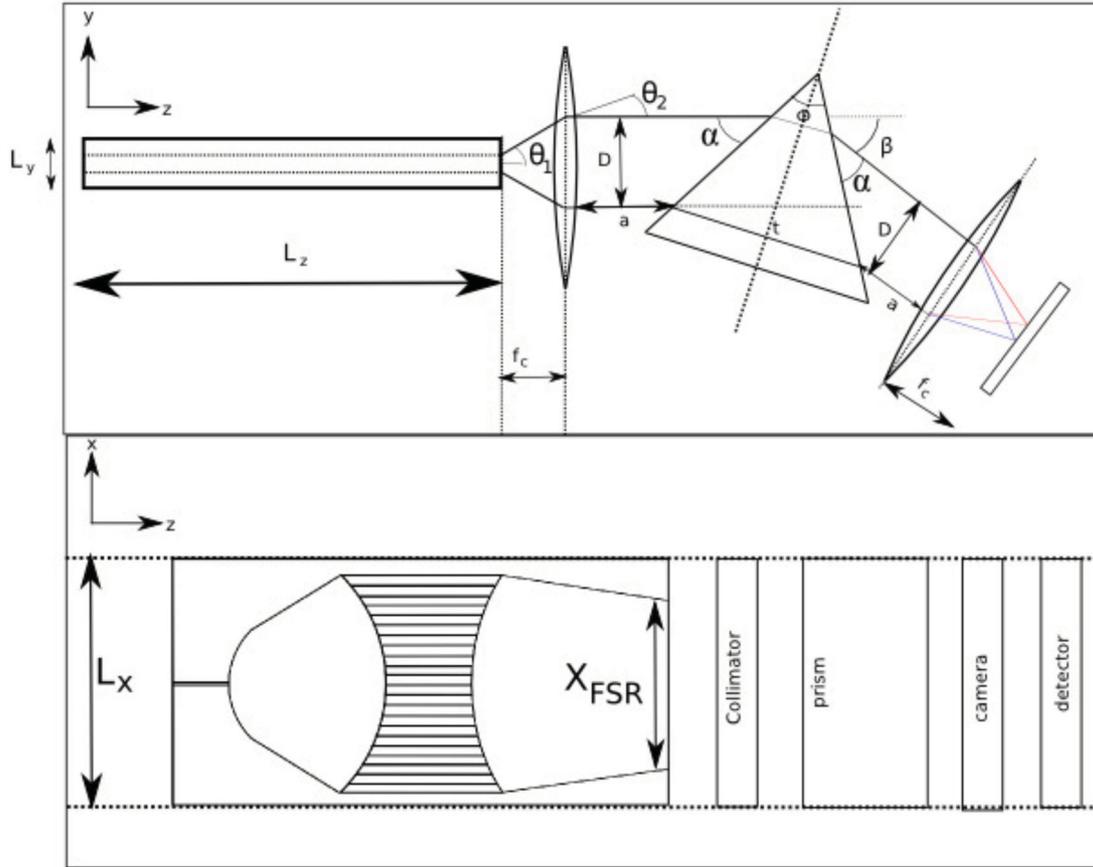


Figure 7. The new model for the AWG. This allows multiple inputs to the AWG and includes cross-dispersion at the output in order to separate overlapping spectra.

where t is the path difference between the upper and lower rays in the prism. In order to account for all wavelengths, we must use the maximum value of the material dispersion, $d\lambda/dn$ within the required wavelength range. The vertex angle of the prism is

$$\alpha = \arcsin\left(\frac{2D}{nt}\right), \quad (23)$$

where α is the angle the prism makes to the collimated beam and n is the refractive index of the prism. This allows the calculation of the vertex angle of the prism,

$$\phi = 2 \arcsin\left(\frac{t}{2D} \sin \alpha\right). \quad (24)$$

The output angle of the prism

$$\beta = \pi - \phi - 2\alpha. \quad (25)$$

We can then calculate L_y from Fig. 7, giving equation (27). Finally, from above and from the geometry in Fig. 7, we have

$$L_x = \text{MAX}\left(X_{\text{FSR}}\left(\frac{2N_i - 1}{N_i}\right), D\right) + 2a_{\text{awg}} \quad (26)$$

$$L_y = \frac{D}{2} + t \sin \delta + a \sin \beta + \text{MAX}\left(f_c \sin \beta, \frac{D}{2}\right) \quad (27)$$

$$L_z = L_z + f_{\text{col}} + a + t \cos \delta + (a + f_c) \cos \beta. \quad (28)$$

7.3 Results for cross-dispersed multiple-input AWG

We now run the simulation for all the instruments detailed in Section 6 using the model above and fused silica as the glass in our prism. We also set the maximum value of t to 30 cm, to represent sensible limits for the prism size.

By imposing our limit on t , we can see in Fig. 8 that the number of inputs per chip is limited to the tens for all the resulting graphs due to equation (22). The potential advantage of this is that all the resulting modes from a single spaxel could be fed into one AWG, meaning that each one could be isolated.

We can also see that though the instrument size decreases (particularly for NIFS, IRMS and IRIS) as more inputs are added, all instrument sizes will be much larger than the single-input version. Existing results have only put around 10 inputs on a chip and then cross-dispersed by the IRIS2 instrument (Cvetojevic et al. 2012), which fits with our results. There is no result for CRIRES as the prism would have to be too large to have sufficient resolution.

Not shown in the resulting graphs are the numbers of pixels required for the instruments, which would be of the same order or greater for this new setup.

It should be noted that we have used a prism in our example, which is usually used for lower resolution cross-dispersion. The alternative is to use a grating, though this would work in a similar way. Taking the equation for FSR and combining it with a diffraction-limited grating ($R_x = m\rho W$) yields

$$N_i = \rho W. \quad (29)$$

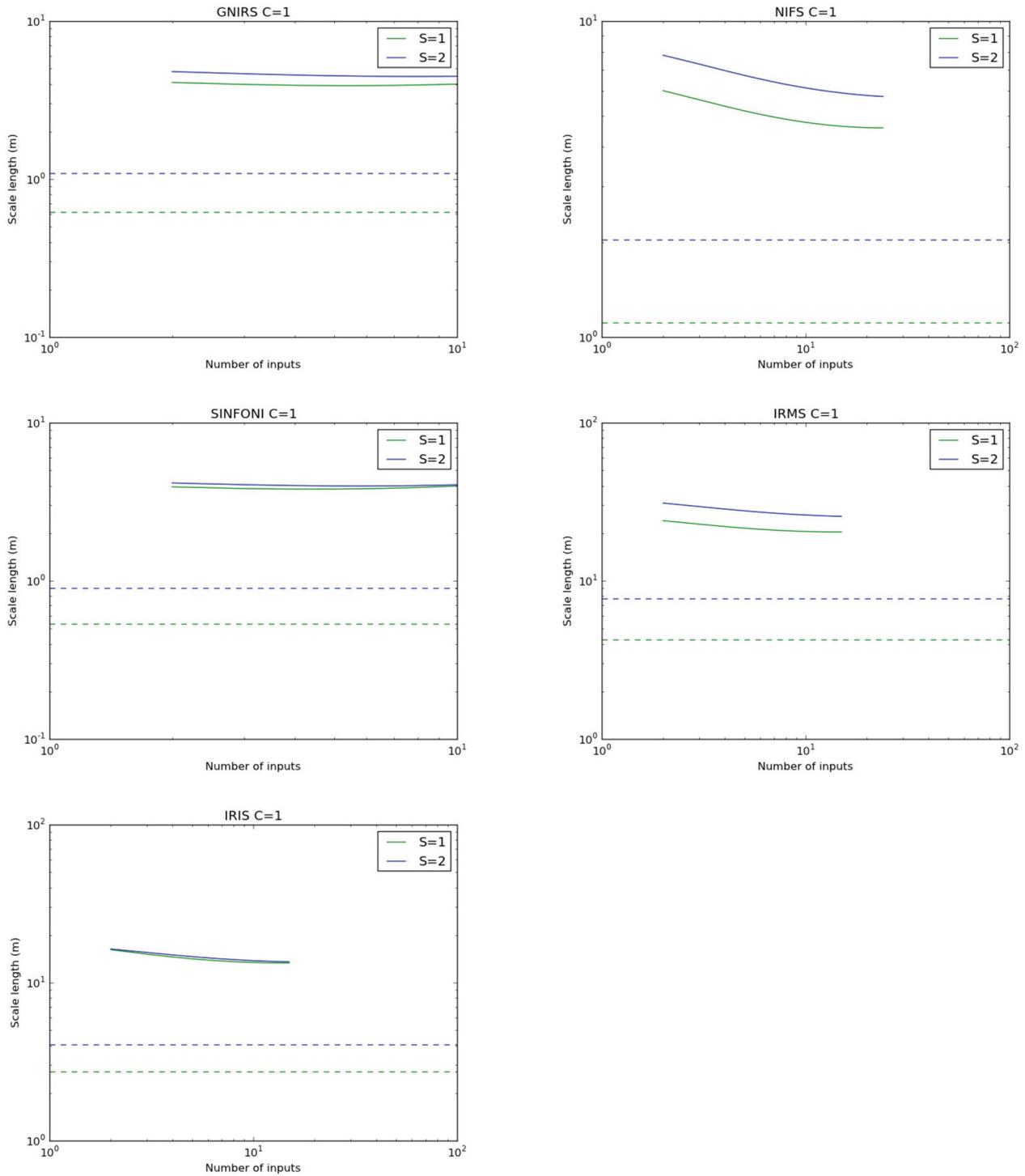


Figure 8. The resulting scale length due to varying the number of inputs to each AWG on each instrument. The different scaling cases are shown in green and blue with the dashed horizontal line indicating the scale length of the single input instrument. From the figure you can see that all the results will produce larger instruments than the single-input case. The result for CRIFES is omitted as no sensibly sized prism could be found with sufficient resolution to cross-disperse the outputs.

Showing the grating size (related to W) will increase as the number of inputs increases (given a maximum ruling density).

7.4 Other instrument options

In its present form, it is clear that trying to compete with large IFU-style instruments is not a viable option. As shown

in Harris & Allington-Smith (2012) the areas providing the greatest advantage would be small or diffraction-limited telescopes, preferably operating at longer wavelengths with instruments that only require a small field of view. There is potential for applications in Solar system science, planetary and stellar science and studies of individual stars in nearby galaxy populations.

Another option is to use multiple single-input AWGs, but to combine the output on to a single linear detector array to reduce the number of detector pixels by a factor equal to the number of modes per spaxel. This would greatly reduce the cost of the detector system and bring the benefit of adaptability of a 1D detector array to the output focal surface of the AWG. This would only be possible if the pixels had a large aspect ratio. This might incur a penalty in terms of extra detector noise, and the number of AWGs is not reduced. This option is currently under investigation.

A further option is to reduce the number of modes that are extracted from the input multi-mode light to produce an acceptable tradeoff between cost and performance defined as a combination of throughput, spectral resolution and field size. Options include.

(a) Restrict the number of SMFs output from the PL with a consequent loss of throughput. This may be acceptable because the population of excited modes is not likely to be uniform (Corbett & Allington-Smith 2006) but will reduce as a function of mode number to a cutoff value at high order. Thus, the overall performance of the system in terms of the product of cost and throughput may be acceptably high.

(b) Reduce the number of AWGs (and detector pixels) by making each work in a partly multi-mode (i.e. few mode) configuration so that the AWG disperses light which is not in a single mode. This may be acceptable if high resolving power is not required e.g. in a survey of faint, unresolved galaxies.

(c) Reduce the number of AWGs (and detector pixels) by decreasing the field of view. This directly trades off cost with field coverage. This is of relevance to applications requiring little spatial multiplex, e.g. single-object spectroscopy or spectroscopy with high-order AO such that the input image is already near the diffraction limit (Harris & Allington-Smith 2012).

8 CONCLUSION

We have examined the application of Integrated Photonic Spectrographs in astronomy and shown that an IPS is equivalent in function to an image slicer. We have shown that as the telescope diameter increases, the size of an IPS must also increase (provided that the slit is not matched to the diffraction limit) due to the increase in number of modes in the field (equivalent to the number of diffraction-limited slices). We have also shown that the number of modes in a field is independent of how the field is initially sampled [the size of the sampling element (spaxel) has no effect on the total number of modes in the field].

We modelled IPS instruments to compare them with conventional instruments on large telescopes and found that they require 10^3 – 10^5 AWGs or 10^3 – 10^4 replica spectrographs if bulk optics are used for instruments on 8 and 30 m telescopes. We found that fully photonic instruments were comparable in size to their conventional counterparts but only if the AWG was close to the diffraction limit. The semiphotonic instruments were found to be much larger, due to the redundancies of having multiple spectrographs.

We have also found that unless the input image is sampled near the diffraction limit, the number of component spectra in each spaxel is very high, requiring large numbers of pixels in the detector array. This is equivalent to oversampling the PSF and could also potentially increase detector noise in the instrument.

To combat the problem of size, we considered the effect of adding extra inputs to the AWG to reduce the number of AWGs required. However, the resulting instrument was of the same size

or larger. It also means that the problem of oversampling in the linear case remains unsolved and potentially will be worse since the spectra will need to have gaps between them to distinguish them.

We also examined other options for reducing the number of detector pixels and/or AWGs and concluded that instruments of photonic construction may be viable depending on the extent to which performance (including throughput, spectral resolution and spatial multiplex) can be traded against cost.

Even without these modifications or restrictions, there are some areas where IPSs may offer a significant advantage. These include spectroscopy of objects near the diffraction limit, e.g. single objects with high-order AO such as in exoplanet studies. Another is low-resolution multiplexed spectroscopy working in the few-mode limit.

ACKNOWLEDGMENTS

We gratefully acknowledge support from the Science and Technology Facilities council. We wish to thank the European Union for funding for the OPTICON Research Infrastructure for Optical/IR astronomy under Framework Programme 7. We are very grateful for the AWG supplied by Gemfire UK and their technical expertise. We also wish to thank Robert Thomson for many useful discussions into all things Photonic and Tom Shanks for inadvertently asking a poignant question. We would finally like to thank everyone else who has taken the time to discuss the workings of various components with us, particularly Jon Lawrence, Nick Cvetojevic and Nemanja Jovanovic.

REFERENCES

- Allington-Smith J., 2007a, in Kissler-Patig M., Walsh J. R., Roth M. M., eds, *Science Perspectives for 3D Spectroscopy 3D Instrumentation*. Springer-Verlag, Berlin, Heidelberg, p. 3
- Allington-Smith J. R., 2007b, *MNRAS*, 376, 1099
- Birks T. A., Mangan B. J., Díez A., Cruz J. L., Murphy D. F., 2012, *Opt. Express*, 20, 13996
- Bland-Hawthorn J., Horton A., 2006, *Proc. SPIE*, 6269, 21
- Bland-Hawthorn J. et al., 2010, *Proc. SPIE*, 7735, 77350N
- Cheo P. K., 1990, *Fibre Optics and Optoelectronics*. Prentice-Hall, Englewood Cliffs, NJ
- Corbett J., Allington-Smith J., 2006, *New Astron. Rev.*, 50, 313
- Cvetojevic N., Lawrence J. S., Ellis S. C., Bland-Hawthorn J., Haynes R., Horton A., 2009, *Opt. Express*, 17, 18643
- Cvetojevic N., Jovanovic N., Lawrence J., Withford M., Bland-Hawthorn J., 2012, *Opt. Express*, 20, 2062
- Eikenberry S. et al., 2006, *Proc. SPIE*, 6269, 62695W
- Eisenhauer F. et al., 2003, *Proc. SPIE*, 4841, p. 1548
- Foy R., Foy F., 2002, *Optics in Astrophysics*, 1st edn. Springer, Dordrecht, The Netherlands
- Grave de Peralta L., Bernussi A. A., Frisbie S., Gale R., Temkin H., 2003, *IEEE Photonics Technol. Lett.*, 15, 1398
- Grave de Peralta L., Bernussi A. A., Gorbounov V., Berg J. M., Temkin H., 2004, *IEEE J. Quant. Electron.*, 40, 1725
- Harris R. J., Allington-Smith J. R., 2012, *Proc. SPIE*, 8446, 84467J
- Kaeuf H. et al., 2004, *Proc. SPIE*, 5492, 1218
- Larkin J. et al., 2010, *Proc. SPIE*, 7735, 77356L
- Lawrence J., Bland-Hawthorn J., Cvetojevic N., Haynes R., Jovanovic N., 2010, *Proc. SPIE*, 7739, 77394I
- Lee D., Allington-Smith J. R., 2000, *MNRAS*, 312, 57
- Leon-Saval S. G., Betterts C. H., Bland-Hawthorn J., 2012, *Proc. SPIE*, 8450, 84501K
- Lu S., Wong W. H., Yan Y., Wang D., Yi D., Jin G., Avenue C., Kong H., 2003, *Opt. Quant. Electron.*, 35, 783

McGregor P., Hart J., Conroy P., 2003, Proc. SPIE, 4841, 1581
Murray G. J., Allington-Smith J. R., 2009, MNRAS, 399, 209
Simard L., Crampton D., Ellerbroek B., Boyer C., 2010, Proc. SPIE, 7735,
773523
Takada K., Abe M., Shibata T., Okamoto K., 2001, IEEE Photonics Technol.
Lett., 13, 577

Thomson R. R., Birks T. A., Leon-Saval S. G., Kar A. K., Bland-Hawthorn
J., 2011, Opt. Express, 19, 5698

This paper has been typeset from a \TeX/L\AA\TeX file prepared by the author.