8.3 Sparse Sampling in Microscopy

Kurt Larson[1], Hyrum Anderson[2], and Jason Wheeler[1]

A fundamental goal of microscopy is capturing images, which in the digital age generally means representing the variable spatial properties of a specimen as an array of pixels with different values. The operator of a microscope chooses pixel spacing – the true distance of on a specimen represented in a pixel – as a parameter of the image collection process. On many microscopes pixel spacing is derived from other user-selected parameters such as field of view and pixels per image. The relationships between signal structure and sufficient sampling were worked out by Nyquist, Shannon, and others [1, 2]. For electron microscopes in which the pixel spacing is greater than the electron beam spot size, or optical microscopy above the diffraction limit, the pixel spacing is equivalent to half the dimension of reliably detectable features. In practice, the operator often chooses a pixel spacing that is significantly smaller than half the minimum dimension, as many pixels per structural feature reduces the impact of noise, improves visual interpretation, and significantly improves algorithms for tasks such as segmentation that may be used for automated image interpretation.

Due to variable structures in most specimens and user-selected spatial oversampling, most images collected in microscopy have pixel values that are correlated spatially. Generally, neighboring pixels are likely to have similar intensity values, although other spatial correlations exist. Such images are sparse or compressible – they can be represented as a linear combination of elements in a basis in fewer values than the number of pixels. Often, the most compact sparse representation may be a transformation of the original image (for example, a gradient image). For compressible images, the Nyquist and Shannon signal detection criteria may not be necessary to completely describe the signal. Sparse sampling and reconstruction, also known as compressive sensing (CS), was mathematically formalized by Candes, Donoho, and others in the mid-2000's [3-5]. CS describes the situations and methods for which a sparse representation of the domain of interest can be exploited, and provides formal performance guarantees for some situations. One goal of a CS microscope can be the acquisition of the specimen structure in fewer measurements than the number of pixels in the image. The opportunities are broader, however. For the designer, CS provides an alternative framework for thinking about sensing systems. The potential benefits of CS in any particular application vary depending on the intent and creativity of the designer. For domains and systems which have constraints on the sampling process – such as one or more of cost, power, size, bandwidth, dose, speed, etc. – design utilizing CS may provide surprising or even radically disruptive performance compared to traditional approaches. The well-known single-pixel camera, in addition to being a demonstration of compressive sensing, could be an alternative imaging design in wavelengths for which detector arrays are difficult to fabricate[6]. The fact that CS is computationally simple to enables an on-orbit reprogramming of the Herschel Space Observatory for

[1] Sandia National Laboratories. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly-owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
[2] FireEye Technologies.

transmission of high-bandwidth data without lossy compression artifacts, while at the same time improving resolution[7].

In this section, we consider the collection of sparse samples in electron microscopy, either by modification of the sampling methods utilized on existing microscopes, or with new microscopes that are specifically designed and optimized for collection of sparse samples. Electron microscopy is well suited to compressive or sparse sampling due to the difficulty of building electron microscopes that can accurately record more than one electron signal at a time.

8.3.1 Motivations for Sparse Sampling in Electron Microscopy

Electron microscope images are usually smooth and are often compressed via block-DCT or wavelet compression schemes using JPEG or JPEG-2000 standards, respectively, while still maintaining high image fidelity. To assess image sparsity of typical electron microscopy images, we gathered 1022 electron microscopy images (SEM, TEM and E-SEM) from the public domain Dartmouth gallery[8]. The images are of a variety of biological, geological, and materials specimens. To standardize analysis, we excised the center 512 x 512 pixels of each image to remove banners and rescaled images to [0 1] grayscale values. For each 512 x 512 image, we computed the sparsity K by counting the number of coefficients in the block-DCT domain (32 x 32 blocks) that cumulatively accounted for at least 99.75% of the total coefficient energy. A histogram of the results is shown in Figure 2, demonstrating that most of these images are highly compressible even when losing only 0.25% of the image energy.

Sparse sampling in electron microscopy has been considered for dose reduction, improving 3D reconstructions and accelerating data acquisition. In optical microscopy, sparse sampling biological fluorescence microscopes [9] and holographic sparse sampling microscopes have been developed[10, 11].

8.3.2 Sparse Sampling and Reconstruction

CS allows signal acquisition and reconstruction using few measurements. Whereas the Shannon-Nyquist condition provides recoverability guarantees for band limited signals provided a sufficiently dense sampling, CS theory provides recoverability guarantees for sparse or compressible signals, provided that a sufficient number of non-adaptive measurements are taken. Romberg and Willett (this volume) discuss reconstruction of signals from sparse data.

In CS, a measurement vector y is acquired according to the linear model

$$y = Ax + n,$$

where A is an $M \times N$ matrix (e.g., M measurements of an N pixel image) that can be decomposed into $A = \Phi\Psi$, and n is additive noise. The "sensing matrix" Φ represents the measurements employed in the system and Ψ is a compression basis or dictionary suitable to the domain (such as wavelets or block-DCT). Thus, Ψx is the signal or image decomposed into a vector x with only K significant coefficients. For example, x could represent the wavelet coefficients of an image with the wavelets captured in the basis Ψ; or x could, itself, represent a sparse image (e.g., a "starfield") with $\Psi = I$ (no compression basis). In compressed sensing, the measurements are non-adaptive, in that A does not depend on x in

any way. In the context of electron microscopy, the sensing matrix Φ is some set of illumination patterns that can be projected onto the sample by the electron microscope, where the signal for each pattern, measured by an electron detector, is collected as a single entry in the measurement vector *y* .

Since $M \ll N$, the transformation from x to y is a dimensionality reduction or compression, that in general, loses information. However, since only K elements of x are nonzero, recovering x from (noiseless) measurements y could, in theory, be achieved through NP-hard combinatoric search for the support of x using only $M \geq 2K$ appropriately defined measurements. Indeed, if x is a vector with only K nonzero elements, then a brute force solution of the non-convex problem

$$\min_x \|x\|_0 \quad s.t. \quad Ax = y$$

could recover x, where $\|x\|_0$ counts the number of nonzero elements in x. However, that problem is computationally intractable even for very small N.

Although many alternative methods have been proposed to reconstruct the signal x from measurements y, basis pursuit—a computationally tractable ℓ_1 inverse problem that is essentially a convex relaxation of the combinatoric recovery problem—is the most common approach that enjoys broad theoretical guarantees. Basis pursuit for noiseless measurement solves

$$\min_x \|x\|_1 \quad \text{subject to} \quad Ax = y,$$

where $\|x\|_1 = \sum_{i=1}^{N} |x_i|$, and the constraint is replaced by $\|Ax - y\|_2^2 \leq \epsilon^2$ for noisy measurements. A myriad of efficient methods exist to solve basis pursuit in polynomial time.

Ideally, compressive sensing measurements are 'wholistic' or global and each measurement represents a response that interrogates a broad spatial area (for 2D reconstructions) or volumetric area (for 3D reconstructions). Compressive sensing theory does not automatically promise a benefit if the number of conventional measurements is simply reduced. The greatest benefit from CS could derive from collecting measurements of the specimen in non-traditional ways. If less-than-ideal benefit is acceptable, however, any measurement method can potentially be used. A direct approach to trying compressive sensing in microscopy is to simply choose a sparsifying basis and collect samples using the normal measurement methods, just fewer of them. In scanning methods such as SEM and AFM where each pixel is individually measured, by this approach the measurement matrix is simply a subset of the identity matrix in which many rows representing individual pixel measurements are removed (not measured). This yields a 'starfield' sampling pattern. Simply reducing the number of point samples might be expected to yield less promising results than alternative multi-point sampling strategies could.

8.3.2.1 Sparse Sampling in Transmission Electron Microscopy

Sparse sampling in electron microscopy has been most extensively reported in transmission electron microscopy (TEM). TEMs direct a beam of nominally 100-300 keV electrons through a thin specimen and measure the signal emitted from the side opposite the illumination. In ordinary TEM, an array of detectors collects an image from a wide-field beam. For sparse sampling, variations of scanning TEM (STEM) are typically used. In STEM, the electron probe is scanned across the specimen, and the detector

measurement is recorded as a function of probe location. Many modes of operation have been devised for TEM. Single scans are performed to explore two-dimensional material structures at sub-nm resolution, usually with the beam axis aligned with a crystallographic axis. To perform three-dimensional tomography, many 2D scans are completed with various rotations and tilts of the specimen. Goris et al. mention some of the modes of TEM that can be used for tomographic reconstruction[12]. Each traditional 2D tilt-rotation measurement is wholistic or global and suited for CS in the sense that it integrates across the thickness of the sample. In traditional weighted backprojection reconstruction as well as sparse reconstruction, the set of tilt-rotation STEM measurements provides the constraint for the inferred 3D structure. Compressive sensing has been simulated and experimentally demonstrated in TEM to recover 2D and 3D structure. The earliest explorations emphasized better recovery of structure with fewer scans. Reduction in electron dose for sensitive specimens is also identified as a potential benefit and has been explored.

CS may have been first applied to electron microscopy in 2007 for cryo-TEM which seeks to reconstruct the 3D structure of a from one 2D image containing many specimen copies in different orientations[13]. A novel L1 reconstruction method that converges using a 'sliding mode' technique was developed to reconstruct virus structure in experimental and simulated data. In simulated data reconstruction was shown for fewer views. Reconstructions were better than those acquired using conventional weighted back-projection. Subsequent developments in simulated cryo-TEM data include a technique for improved SNR and resolution by promoting sparsity in the wavelet domain and exploiting Toeplitz properties of the forward-projection and back-projection operators[14, 15].

Veeraraghaven et al. [16]used compressive sensing computed tomography in serial section TEM to simultaneously achieve z-resolution of 5 nm and reduce the number of tilted sample images from tens or hundreds to about five. 30 consecutive, 50-nm thick sections of fly larva brain were imaged at 4 nm x,y resolution for five tilts. Utilizing prior knowledge in the domain of neural tissue, a sparsifying basis was constructed using planar local basis functions in a variety of orientations as well as a basis functions to represent uniform intensity and blob image primitives. Segmentation of neurites was improved compared non-tomographic methods for the cost of four extra tilt-rotation images, and the speed of acquisition was much faster than traditional computed tomography which uses many more tilt-rotation images, or focused-ion beam (FIB) sectioning. Hu et al. [17] used an unsupervised learning method to infer a domain-specific sparsifying three-dimensional basis from a high-resolution training set collected on a focused ion beam (FIB) tool. Using the learned basis, they demonstrate on 15 slices of fly larvae neural tissue that the number of STEM tomographic slices could be reduced to five while maintaining high resolution using a sparse reconstruction technique. A drawback was noted that the images require very precise alignment prior to sparse reconstruction. A number of potential mathematical enhancements of the reconstruction process were noted.

Goris et al. [18] used a total variation minimization sparse reconstruction technique and compared it to the simultaneous iterative reconstruction technique (SIRT) algorithm for tomographic reconstruction of dispersed Ag nanoparticles, PbSe-CdSe core shell particles, and a Si needle with Pb inclusions. The three specimens were imaged in different TEM microscopes with different techniques. For all three specimens, the TVM sparse reconstruction suffered less from the missing wedge of projections due to

the limited orientations possible in a TEM system, required fewer projections, yielded better segmentation, and demonstrated higher detection probability for small features. Subsequently, using experimental nanoparticle data and synthetic phantoms, others demonstrate reduced elongation in the direction of the missing wedge, improved segmentation, improved resolution or detection of small features, reduced artifacts, and 3D reconstruction with fewer samples, compared to SIRT and weighted back-projection [19, 20].

Motivated principally to reduce the dose for specimens that might otherwise be altered or destroyed, Binev et al. [21] develop the application of compressed sensing to TEM for 2D scanning and 3D tomographic recovery, and validate concepts via numerical simulations. Stevens et al. [22] use Bayesian dictionary learning to infer a domain-specific basis for STEM, which is used to recover effective images from approximately 5% of the pixels (20X) reduction. Beche et al [23] implement a fast shutter mechanism in a 2D scanning TEM to enable arbitrary, random illumination of pixels without modifying the scanning mechanism. Images with 50% reduction in samples (50% dose reduction) maintained image quality across all specimens tested; images with 80% reduction had variable results depending on specimen and imaging factors.

Although compressive sensing may yield the greatest benefit with alternative illumination strategies, we are unaware of simulation or experiments for TEM in modifying the beam (for example, utilizing STEM with a beam with a spot size broader than one pixel) or utilizing multiple beams in 2D or 3D TEM to implement sparse sampling, although it has been conceived as a potential approach to reducing dose in TEM sampling [21].

8.3.3 Sparse Sampling in Scanning Surface Microscopy

The traditional SEM or AFM samples each pixel independently to generate an image. As most images are sparse in some basis due to image structure, it is natural to consider whether sparse sampling methods could be applied to these instruments. It is fairly easy to conceptualize and model sparse sampling techniques that could capture sufficient information for reliable compressive recovery in fewer samples for data acquisition with a reduced dose. For example, one could randomly sample a subset of the desired pixels and utilize sparse reconstruction to recover the image. A high-speed shutter would be useful to precisely select pixels and exposure times, such as has been recently created for TEM [24] and are commercially available for SEMs.

To speed up imaging, however, an alternative faster than to scanning every pixel equally is necessary. Dynamic processes in scan electronics preclude arbitrarily sampling individual pixels at high speed. Simple robust methods to recover an accurate image with fewer measurements, such as iterative sub-sampling, in practice yield a process that takes more time than sampling all of the pixels. A second and important consideration is that compressive sensing favors measurements of a global nature – rather than a reduced set of point measurements.

In both SEM and AFM, electrical lags in both electrostatic and electromagnetic components exist and the beam or tip does not come to rest at a commanded position until some finite time after the command. In the case of AFM, the tip assembly also has physical inertia. SEM and AFM typically compensate for

these dynamical lags by moving slowly or by allowing the system to come into stable motion prior to imaging; in SEM this occurs beyond the edges of the desired image. Essentially, extra pixels outside of the desired image area are scanned and thrown away until the beam speed becomes constant. A consistent image is created by replicating identical scan commands for each row of an image, some of which is discarded. Since the dynamical motion is highly dependent on the specific piece part values of amplifiers and other electrical components, each SEM or AFM needs to be calibrated whenever hardware is replaced so that electrical control signals can be correlated to real dimensions. One implication of the electrical and mechanical dynamics is, in most scans, even after the transients have died out, the probe is constantly lagging the continuously changing commanded positions, and the measurements represented as a pixel in an image are actually collected by a probe in motion sampling a linear section of the specimen. Electrical systems with electromagnetic components cannot avoid transient and steady-state dynamical responses. Power supplies with fast switching and large voltage range for use in electrostatic scanning SEMs are possible, however they are expensive and not normally used, so even electrostatically-scanning SEMs have lag dynamics in scanning. When these system dynamics are taken into consideration, many concepts for collecting data sparsely in existing real microscopes are likely to be no faster, or even slower than sampling every single pixel at the desired signal-to-noise ratio. This may be acceptable for exploratory R&D or dose reduction applications, but is a significant challenge in applications emphasizing speed.

8.3.4 Sparse Sampling in Atomic Force Microscopy

Atomic force microscopy (AFM) uses a sensing probe on a cantilever arm to detect variations in the interaction of the probe and specimen as a function of probe location. The sensing probe and property measured have many variations [25], which affect the rate of data acquisition. Interest in compressive sensing in AFM is aimed at increasing speed of acquisition or reduced sample interaction to reduce damage. Monitoring the progress of in-situ surface or biochemical reactions emphasizes speed of acquisition or revisit rates.

Song et al. developed a sparse sampling procedure and reconstruction method for AFM which uses a fixed sampling pattern in which the tip traces a square spiral [26]. The sampling ratio (sampled pixels to total pixels) is not specified. They use a total-variation norm to reconstruct the image, which shows possible aliasing due to the highly structured sampling methodology. Because their sampling process is highly structured and has appreciable gaps, the method is not generalizable to all samples.

Andersson and Pao randomly chose 20% of the pixels from a complete AFM data set to simulate random-pixel sampling in AFM [27], and also tried a concentric-square sampling method similar to that used by Song et al. [26] Pure L1 reconstruction was poor for both methods. Much better results were obtained using total variation minimization and Delaunay triangulation. The authors conclude that compressive sensing could be used advantageously in AFM when sample interaction needs to be minimized, or for speedup when very slow measurements such as force mapping is being conducted.

A fast and arbitrarily controlled (or random) probe motion and image inversion technique has been developed which could be applicable to AFM [28]. This method would require inversion of accurate

models of AFM tip motion such as those developed by Chang et al. for use in a feedback-controlled variable scanning process [29].

8.3.5 Sparse Sampling in Scanning Electron Microscopy

Anderson et al. demonstrate a sparse imaging method on SEM, in which the electron probe measured a subset of locations on the specimen with varying illumination times [28]. The dynamics of the scanning process were modeled and compensated for to allow a compressed sensing reconstruction of the image. Each detector measurement encodes an integrated response over many pixels as the beam moves transiently to commanded locations. By this method, recovery of an image similar to the normal SEM image was possible in about 1/3 of the total imaging time. In addition to speed, a dose-minimization motivation is applicable to use of compressive sensing in SEM, in which certain biological or dielectric materials may be damaged or exhibit charging artifacts if the dose rate is too high.

8.3.6 Compressed Sensing in Multi-Beam Electron Microscopes

Volumetric and wide-area imaging at resolution of a few nm is being sought in neuroscience and materials science. In neuroscience, imaging the approximately 500 mm^3 volume of a mouse brain [30] with 4 nm pixel pitch and 30 nm tissue slices [31] would generate 10^{18} voxels. The Intelligence Advanced Research Projects Activity (IARPA) Circuit Analysis Tools program sought a 400X increase in the speed of imaging a 1 cm^2 area at 2.75 nm resolution (13.2 trillion pixels) for microcircuit analysis [32], which was not realized by the program [33]. The IARPA Rapid Analysis of Various Emerging Nanoelectronics (RAVEN) program seeks to image all layers of a 1 cm^2-area microcircuit with 14 nm minimum feature size within 25 days [34]. In microcircuit fabrication, which must keep pace with an ever-shrinking minimum feature size, the capability to routinely inspect large areas with ~nm resolution, which is well beyond all optical techniques, is becoming necessary for process development and defect detection. Statistical characterizations of structural heterogeneity benefit greatly from wide-area imaging and 3D reconstructions, e.g., natural materials such as geologic samples, or engineered materials after failure, chemical degradation, or aging. These applications demonstrate that near-nm resolution imaging at sustained rates over large areas is scientifically and commercially valuable. For many applications, imaging at sustained rates of hundreds of millions of pixels per second could be sufficient; in the case of neuroscience, sustained imaging at many billions of pixels per second is necessary to even conceive that large volumes of brain tissue could be imaged.

Collecting surface images of samples by SEM is ultimately limited in speed by the single-pixel-at-a-time nature of the scanning process. Even as detectors become faster, and system engineering innovations reduce time required for beam dynamics and stage movement, the maximum attainable speed at a given resolution is limited by one or more of (1) the physics of material contrast, (2) electron repulsion and lens aberrations which limit beam current in the column for a given spot size, and (3) sample dose tolerances limiting the beam intensity per unit area on the sample. The fastest scanning electron microscopes operate in specialized applications such as microelectronics inspections at approximately 100 million pixels per second. Compared to the necessary speeds, the single beam systems are not fast enough to collect the desired quantities of data in a reasonable amount of time. Parallel collection of pixels is necessary.

NOTE FIGURES AND TABLES APPENDED AT END OF MANUSCRIPT DRAFT

In the microcircuit fabrication domain, parallel imaging can be implemented in a footprint the size of the wafer, so widely separated scanning systems can be built which simplifies parallel imaging. In biology and materials science, with unique small samples, the parallel imaging must be achieved in a much smaller area. Development of electron microscopes that can image in parallel in a small area is difficult. In response to an incident beam, electrons scatter from the sample by elastic and inelastic processes with a broad range of energies. The broad range of electron energies creates challenging conditions for imaging electron optics. For this reason, standard scanning electron microscopes use a simple potential field in the chamber to attract low-energy, inelastically-scattered (secondary) electrons to a single detector, one pixel at a time. If used for multiple illumination beams, this detection paradigm mixes the electrons emitted from the multiple incident beams. Wide-field or parallel imaging optics utilizing complex lens assemblies have been devised for specialized applications as in LEEM [35, 36]. The Zeiss MultiSEM uses innovative illumination and imaging optics to scan many single-beam SEM images in parallel at landing energies less than 3 keV [37, 38]. Due to the complexity of their imaging optics, the LEEM and MultiSEM have significant constraints on electron landing energy which makes these systems specialized rather than general-purpose.

The 61-beam Zeiss MultiSEM 505 collects 1.22 billion samples per second allocated as approximately 200 electrons per pixel per sample (~672 pA x 50 ns) with a minimum pixel spacing of 3 nm over a field of view about 100 um wide [39, 40]. Some speciments would require more than one sample per pixel to generate sufficient material contrast in the collected image. Stage moves and other non-imaging processes occur between imaging intervals. With two samples per pixel and assuming 1 second for non-imaging process such as a stage move between imaging intervals, the sustained rate for large-area imaging using the MultiSEM in the neighborhood of 200 million pixels per second. Clearly, it is important to reduce the 'down-time' for stage moves. Given a suitable characterization of the amplifier and electrostatic plate beam dynamics in the MultiSEM, a single-beam compressive sensing technique could be attempted with it [28, 41]. Since the scanning time in the current MultiSEM is a relatively small fraction of the elapsed time, using compressive sensing in the current system would not significantly increase its effective speed. Other benefits might be possible such as improved image quality for a given exposure time or mitigating sample charging by using an unstructured illumination pattern. If the stage move time in the MultiSEM becomes significantly smaller, in some domains the use of compressive sensing on it could conceivably increase its effective rate into the billions of pixels per second range.

The difficulties associated with using multiple detectors to image in parallel makes surface electron microscopy very appealing for compressive sensing approaches, which generally utilize one or a few detectors, each of which aggregates information sampled globally rather than locally. Illuminating an area of many μm^2 with a series of electron patterns that satisfy the global nature desired for sparse sampling is feasible, as various multi-beam electron illumination columns for mask writing or lithography applications have been developed in which illumination patterns with essentially arbitrary structure over the field-of-view can be generated. One uses MEMs apertures and deflection electrodes to create 262,144 electron beams with 20 nm or 10 nm spot size within an 82 μm x 82 μm write field [42]. The beams are individually programmable on/off, and the overall system is capable of addressing the write location of the beams within 0.1 nm. Another uses MEMs electron mirror/trap devices to create an

array of 248 x 4096 adjacent beams (1,015,808 total) that can be switched on/off individually with beam location precision of about one nm [43]. Some of the technologies developed in the multi-electron beam lithography and photomask R&D could potentially be adapted or repurposed for compressive sensing electron microscopy.

Anderson et al. [44] model compressed sensing recovery in SEM with a single detector when multiple electron probes (beams) simultaneously illuminate the specimen. For a 50% samples-to-pixels ratio using simulated data with ~5% sparsity, binary (on-off) electron beam illumination patterns yielded minimum reconstruction error for sparse beam-pixel ratios of 0.1% to 6%; above and below these values, reconstruction error increased rapidly. For Bernoulli measurements with -1,+1 illumination, which in electron microscopy requires a method to estimate the mean response of the specimen, reconstruction error was low for all fill ratios between 0.25% and 50%. Simulated reconstruction by sampling images collected in SEM suggested that non-ideal noise in real electron microscopes may be an important limiter in compressive sensing. The authors also noted that the methodology could have introduced more noise than would occur in a real multi-beam microscope and that a less-than-optimal reconstruction method was used, both of which suggest the results are potentially pessimistic. Experiments in a multi-beam instrument were identified as necessary to fully understand the effect of stochastic sample interactions and detector responses in compressive sensing.

8.3.7 Theoretical Analysis of an Electron Column for a Multi-Beam CSEM

Here we develop a theoretical analysis of a hypothetical compressive sensing electron microscope to develop understanding of potential performance and challenges. First, the desired maximum diameter of the electron beam at the sample, also known as the spot size, is defined as 5nm. Collimation is the angular spread of electrons in the beam and is defined as the spot size divided by the focal length of the objective lens. A value of 5 mm focal length is chosen for this analysis. This leads to a beam collimation requirement of 1×10^{-6} radians (1 micro-radian).

The system concept uses an array of apertures to create a pattern of beamlets from an incident flood beam that is created by an electron gun and collimation lens. An array of Micro-Electro-Mechanical (MEMs) devices over about 2 mm x 2 mm (4 mm^2) is assumed as a design starting point and seems feasible to construct, as similar devices have been developed for electron lithography [42]. A beam waist is a position in a microscope where the beam has a minimum dimension and angular divergence. The 4 mm^2 area of the aperture array is selected as a beam waist from which an emittance value could be derived. Emittance E is defined as $E^2 = A\omega$, where A is the area of the beam waist and ω is the solid angle of the beam, which is approximately the collimation squared. Therefore, the emittance of the flood beam entering the aperture array is 2×10^{-9} A/m^2sr. The beam current I is related to brightness and emittance by $I = BE^2$, where brightness is a property of the electron emitter. Schottky field-emission emitters have a brightness of about 5×10^{12} A/m^2sr. Thus, the beam current entering the 4 mm^2 aperture array is about 20 uA.

To continue the analysis, an assumption about how the aperture array divides the incident collimated flood beam into beamlets is necessary. We assume the aperture array passes 10% of the incident

electrons as some unspecified number of beams. A design process iterating between electron column models and sparse reconstruction analysis would be necessary to determine the optimal number of beams for a real system. For the time being, the additional optics to create patterns and demagnify the beams onto a sample are ignored, which would also require significant iterative modeling to verify resolution and other performance attributes. Given these assumptions, about 2 uA of beam current could be used to create "sensing function" patterns for sparse sampling.

2 uA of beam current is equivalent to approximately 12,500 electrons per nanosecond. At ordinary detector quantum efficiencies (DQE), acceptable SNR for many imaging applications can be generated from between 10^3 and 10^4 incident electrons per pixel. If 10^4 electrons per pixel are budgeted to generate SNR or contrast, a conservatively large value, the theoretical data rate for the multi-beam compressive microscope is 1.25 billion pixels per second. Specimens that require fewer electrons per pixel to generate sufficient material contrast, or a system that has a higher DQE, could achieve proportionately higher pixel rates. Furthermore, as illustrated in Figure 1, many surface microscopy images are sparse, so an additional increase in the pixel rate could potentially be achieved because compressive sensing can enable collection of fewer samples than pixels. This first-order estimate of a potential data rate in the billions of pixels per second range for a multi-beam compressive sensing electron microscope exceeds the current state of the art for most applications. Further investigation of the possible designs and engineering challenges of a multi-beam compressive sensing electron microscope seem warranted.

8.3.8 Potential Embodiments of a Multi-Beam CSEM

In this section we explore potential embodiments of a multi-beam CSEM. As has been noted, multi-beam electron lithography systems have developed a variety of technologies that could be used in multi-beam compressive sensing [42, 43]. In lithography, a common approach to creating the exposure patterns is to assign each beam a fixed location in the array of beams, and turn the beam off (e.g. with an electrostatic deflection into a blanking plate or absorption) if exposure is not necessary at that location in the current exposure period. In areas where the lithographic pattern is sparse, most beams are off; where it is dense, most of the beams are on. The beam current or number of patches being exposed at any one time is job and location specific and is not a quantity of principal concern. The bandwidth or switching speed of the patterning system is related to switching beams on and off.

A similar approach could be taken for a compressive sensing microscope [45]. A contiguous array of many on|off switchable beams could be constructed, with each beam mapping to one location in the image to be reconstructed. Such an array could expose any number of pixels or patches in a contiguous region of the sample, depending on the operator's choice of which beams are on or off. A sequence of patterns could be specified as the sensing matrix. This sort of microscope could potentially borrow technology for many devices from electron beam lithography systems.

Many possible designs for compressive sensing microscopes could be made with on|off switching, and undoubtedly multi-beam compressive sensing could be tested and thoroughly explored in an on | off system. A application interest in minimizing acquisition time emphasizes maximizing beam current to the sample, and raises electronics bandwidth management as a challenging engineering issue. In an

on|off system, there are as many potential beams as pixels, and 50% (or more) of the beams are off during any sensing pattern illumination. One perspective of such a system is that the available electronics bandwidth is used to reduce beam current by interrupting beams that have already been formed. It thus seems like a sub-optimal approach to utilizing the available beam current and bandwidth. Rather than switching beams on or off, could a system be made to switch beams 'here' or 'there'? In other words, could the electronics bandwidth be used to modify illumination patterns without affecting the beam current to the sample? Approximately half of all pixels (or fewer) are not illuminated in any given sensing pattern. Instead of absorbing or deflecting 50% or more of the formed beams, a more efficient and therefore faster microscope might result if each beam could be steered in some way to select among two (or more) pixels (or patches), because in this manner many more pixels than beams would be measured in any sequence of patterns. This section explores the potential for compressive sensing microscopes with multiple steerable beams, or more precisely, beams that are switchable among a pre-determined set of locations. Each beam is always 'on' and beam current to the specimen is constant or, in other words, maximized to drive speed up; what varies among the sensing patterns is the location each beam is illuminating. We devise two potential methods and explore the compressive sampling and reconstruction, engineering challenges, and potential speed of each.

8.3.8.1 Concept for Steerable Beams

Larson et al. [45, 46] explore using steerable beams to create the patterned illumination of the measurement matrix. The switching mechanism would be located in the electron column between the electron gun and the objective lens. It would use a system of microelectronic devices with two purposes. First, the devices create an array of beams from a collimated flood beam emerging from the electron gun. Second, the devices deflect (steer) the beams to their desired position on the sample. To be consistent with the preceding theoretical evaluation, the total area of the apertures is about 10% (or potentially more) of the array area, and the size of the apertures is inversely proportional to their number. The area not utilized for apertures would be used for circuitry such as steering electrodes, switching logic, routing, and power. The devices would need to be mechanically, thermally, and electrically stable in the environment of an electron column.

Each device or aperture would have x-axis and y-axis circuitry to control deflection in orthogonal directions. Conceptually, the circuitry design could be simple. Each aperture is divided into four quadrants: -X, +X, -Y, +Y. Voltage differential in the X elements provide X-axis deflection; similarly for Y. Grounding all elements allows a beam to pass undeflected. By applying voltage steps to the control elements, each beam can be deflected by pixel increments along the X and Y axes. With two voltages in each axis (e.g., [0, +1]), each beam could address four different pixels. With three voltages [-1, 0, +1], each could address nine pixels. Thus, each beam can be 'addressed' to a number of pixels that is the product of the available X and Y voltage increments. The number of pixels that can be patterned is some multiple of the number of beams. Depending on the number of pixel offsets allocated to each beam and the spacing of the switching apertures, a variety of exposure patterns and beam/pixel ratios can be created including ones in which pixels can be 'exposed' by more than one beam at a time. The bandwidth is utilized to select a particular pattern rather than to reduce beam current. The bandwidth per pattern required by the switching mechanism is directly related to the number of beams, the

number of pixel offsets allowed for each beam, and the electron exposure required to obtain a sufficient measurement from each pattern. We explore two varieties of this configuration here.

8.3.8.2 Array of Correlated Steerable Beams

To explore sparse sampling with a high ratio of beams to pixels, a square array is postulated with nominally 100 x 100 switching apertures, for ten thousand total beams. Given 2 μA total beam current to the sample from 10% of the area of a 4 mm^2 array, each beam aperture has a diameter of about 7 μm and passes 200 pA current. The circuit complexity and bandwidth required to independently switch this many beams in a small area is likely prohibitive, so a simplification is used. Each X-axis beam control electrode is linked to its neighbors along each row so that they all receive the same deflection offset command, and each Y-axis control electrode is similarly linked along each column[45]. The trade-off for this significant simplification of circuitry is that the offsets along each row and along each column are correlated. All beams in a row are offset by the same number of pixels in the orthogonal Y direction, and all beams in a column are offset by the same number of pixels in X. Correlation in the sampling matrix results from this concept, which ought to reduce the efficiency of sparse sampling and require more samples for image reconstruction. The engineering benefits are that control circuitry can be placed external to the array, and the switching bandwidth is significantly reduced. This idea is conceptually similar to Toeplitz or circulant approaches that can be used to reduce control bandwidth and which can be nearly as good as random sampling [47].

The deflection array and subsequent optics could be configured so the beams have non-overlapping address space. For example, if each control line has two settings [0,+1], the beam centers (neutral position) would occur on every other pixel position and a 10,000 beam array could create patterns over forty thousand pixels, for ¼ beam-to-pixel fill ratio. Three setting controls would have beam centers on every third pixel index and interrogate ninety thousand pixels for 1/9 fill, and so forth. The no-overlap configuration minimizes bandwidth and prevents a pixel from being illuminated by multiple beams simultaneously. A potential disadvantage of the no-overlap configuration is only one pixel within a single beam's address area can be illuminated at a time, which means that the sensing patterns cannot directly test the correlation of some adjacent pixels. This approach also makes the system more vulnerable to defective beamlets, which may manifest as 'always off' or 'unsteerable'.

Alternatively, the deflection array can be configured to allow the address space of beams to overlap. The possible fill ratios are as in the no-overlap case, however the increased addressing allows for any two adjacent pixels to be illuminated in a sensing pattern as well as the condition of multiple beams illuminating one pixel at the same time. The bandwidth required for the overlap configuration is higher than the no-overlap case. A hardware deflection array built to accommodate an overlap configuration could be limited in software to run in the no-overlap mode.

This configuration of beam controls and signals can be described precisely. The ith 2-dimensional $U \times V$ electron sensing pattern P_i can be written as the outer product of a vector that defines the rows $r_i \in \mathbf{R}^U$, and another that defines the columns $c_i \in \mathbf{R}^V$, as $R_i = r_i c_i^T$. This scenario is related to previous work in separable imaging operators for compressed sensing [48, 49]. Sensing patterns generated by one potential configuration of a 21x21 array of beams that addresses a 49x49 array of pixels is shown in

Figure 2. A parameter-less implementation of basis pursuit called SPGL1 [50, 51] evaluates this sampling for randomly-generated images with 5%, 10%, and 15% sparsity over a range of sampling ratios (Figure 3).

8.3.8.3 Array of Individually Steerable Beams

Previously, it has been shown that sparse matrices can perform theoretically as well as dense Gaussian ensembles for compressed sensing recovery[52], when the degree (fill ratio) of the sensing matrix was very small. Sparse measurement matrices offer an advantage of efficient image recovery, since each iteration of the process becomes a simple multiplication by a sparse matrix. For recovering images that are canonically sparse, (e.g., starfield images), very efficient methods exist that are based on expander graphs [53].

The beam steering apparatus could be configured to create a very sparse fill ratio, on the order of 1 beam per thousand pixels. Assuming a system in which 1000 beams illuminate sensing functions over a million pixels, each beam needs a minimum of $\sqrt{1000} \sim 32$ addresses in X and Y. More addresses would allow for overlapping fields of illumination. In terms of bandwidth, five bits per beam for each X and Y control axis requires 32 different voltages and yields $(2^5)^2 = 1024$ pixel addresses per beam; six bits requires 64 voltages yielding 4096 positions. The potential performance of this system is discussed in the following section.

8.3.8.4 Managing the Electron Budget

The allocation of electrons among beams and compressive samples is a key design consideration of a fast compressive sensing electron microscope. For a given source and column design, electrons are the resource in scarce supply, and utilizing them efficiently is the key to achieving quality images in the shortest possible time.

There are several factors that bear consideration. First, since the system has multiple beams and electrons are few and discrete, an initial question is how many electrons must pass through each aperture to ensure that the sample is sufficiently interrogated by each beam in a sample? The time and location of an electron being emitted from a source is essentially a random process, and thus the aperture that an individual electron passes through is unknowable.

For the system concept described in Section 8.3.7, the emission of electrons from a Schottky source was modeled as a Poisson process, in which electrons are emitted at inexact time intervals in a normal distribution over the facet of the source. Thus, each aperture hole has an equivalent chance of passing the emitted electron. If the number of electrons is approximately equal to the number of holes, the statistics work out such that many holes will pass one electron, some will pass more than one, and some will pass zero. To assess the number of emitted electrons necessary to achieve statistically-equivalent illumination per beam, sparse image reconstruction accuracy as a function of the average number of electrons per aperture hole was assessed through computer simulation, while also varying the beam-to-pixel fill ratio (Figure 4). A sparse fill ratio was significantly superior to dense, and an average value of 40 or more electrons per hole was the “knee in the curve” above which additional illumination provided slight to negligible improvement in reconstruction accuracy for this source of noise. This method of

quantifying noise is independent of the beam current in the system and the specimen being imaged and is thus a useful representation of the minimum number of electrons that must be transmitted to the sample, per measurement per beam, in a multi-beam CSEM. Since reduced-dose imaging is a motivation for compressive sensing, particularly in TEM, increased exploration of the minimum realizable electron dose per compressive sample would be useful for several applications.

Is there any benefit in going above the minimum electron budget? Stochastic processes in the interaction of sample and beams and detectors make this an issue that may be best answered experimentally. However, some insight from SEMs might be applicable. Many SEMs have a fixed sampling rate, 10 to 20 million samples per second on many modern machines. In a SEM with 200 pA beam current and 10 million samples per second, each measurement uses about 125 electrons. SNR in an image pixel is controlled by operator selections of beam current (electrons per sample) and averaging more than one sample into the reported pixel value. In the case of sample averaging, the SNR increases as the square root of the number of samples. Excellent SNR in the recorded image (aka material contrast) is often built up from many poor measurements.

In a multi-beam CSEM, responses from many pixels are aggregated into one compressive measurement. An open question is the necessary SNR of the individual pixel responses. Each pixel is measured many times across the set of compressive measurements. For example, a multi-beam CSEM with 1000 beams, a million pixel image area, and 100,000 sparse measurements (measurements/pixels = 0.10) will represent each pixel in about 100 compressive measurements. If on the order of 100 electrons per pixel is allocated for each compressive measument, each pixel's value is represented by approximately 10,000 electrons across the entire collection. Fundamentally, if few electrons per pixel are sufficient in each compressive sample, then fast collection of the compressive samples can be achieved. Conversely, if many electrons are required per pixel per compressive sample, fast imaging becomes challenging.

8.3.9 Multi-Beam CSEM Speed Estimates

Estimates of speed are developed in this section for a variety of potential embodiments. In Section 8.3.7, a conceptual column design for a multi-beam CSEM was developed suggesting that approximately 12,500 electrons per nanosecond could be delivered to the sample if distributed over many beams. Accepting the critical caveat that scientific uncertainty and engineering challenges exist, which are discussed in a following section, and assuming they can be solved, the necessary speed-related parameters are principally the electron budget per pixel per sample – which as previously discussed could be as low as about 40, but not likely lower – and the number of compressive samples to be collected, which depends theoretically on the sparsity of the specimen being imaged and practically on the optimization of the sampling matrices and image reconstruction methods. In the scenarios presented in Table 1, 100 electrons per beamlet per sample and a 10% ratio of measurements to pixels are assumed. Both of these are unverified in an actual multi-beam CSEM. Speed estimates are linear in total electron budget, electrons per beam per sample, and the compression ratio.

Scenario 1 through 4 explore dense measurement matrices. Scenarios 1 and 2 explore the effect of [on|off] vs steerable switching. Scenario 1 has 10,000 [on|off] switchable beamlets and a 10,000 pixel image, with 50% of the beamlets on at any one time, and half of the electron budget diverted to a

collection plate. Scenario 2 has 5000 steerable beamlets over two positions, and maintains the full electron budget on the sample for all patterns. Scenario 2 is twice as fast as Scenario 1. Scenario 2 and 3 show that for a given pixel-to-beam ratio, the system with fewer beams is faster. Scenario 3 recovers a 20,000 pixel image, and Scenario 4 recovers a 160,000 pixel image. Comparing Scenarios 3 and 4 shows that, all other factors being equal, the systems have the same realizable speed. In fact, Scenarios 1, 3, and 4 each have 10,000 beams and are equally fast.

Scenarios 5 through 8 explore sparse measurement matrices. Scenarios 5 and 6 show that, like Scenarios 3 and 4, the ratio of the number of pixels in the image to the number of beams in the sensing system has no effect on the achievable speed. Scenario 6 and 7 show that for images with the same number of pixels, the system with fewer beams will have a faster rate. Scenarios 7 and 8 also show that for a given image size to beam ratio, the system with fewer beams is faster.

The 'Switch Rate, MHz' column explores the engineering complexity of the Scenario, indicating the detector rate required and the maximum pattern switch rate for the system. This is an indicator of the bandwidth required by the system. Extremely careful electronics design or even fundamental advances in microelectronics would be required to achieve the performance some of these scenarios demand. The 'nA per Beam' indicates the physics complexity of the Scenario; in most electron microscopes, the resolution is inversely proportional to beam current.

The 'MP/sec Formula' column indicates the value calculated by the formula A/BCD where A is the electron budget per second (12.5×10^{12}), B is the number of beams in the Scenario, C is the electrons per beamlet per sample (100), and D is the sparsity, or ratio of samples to pixels, 0.1. The units cancel

$$\frac{\frac{e}{sec}}{beam\,\frac{e}{beam * sample}\,\frac{sample}{pixel}} = \frac{\frac{e}{sec}}{\frac{e}{pixel}} = \frac{pixel}{sec}$$

to a rate of pixel/sec. The values calculated by this formula match those displayed in the 'MP per second' column. The electron budget (beam current) in the numerator indicates an unsurprising direct linear correlation with speed. Sparsity in the denominator is also expected, and shows that the microscope has an effective increase in speed, just by reducing the number of compressive measurements such as might occur over time with R&D into image reconstruction methods and improved sparse representations. The electron budget per beam per sample in the denominator is also not surprising. The electron dose per beam per sample has a minimum realizable value of around 40, and as discussed previously the necessary value is an open scientific question for compressive sensing in electron microscopy (it is also dependent on the specific detector system: electron capture and detector efficiency).

The somewhat non-intuitive aspects of the speed formula is that the speed of a multi-beam electron microscope implementing sparse sampling is inversely proportional to the number of beams and does not depend on the beam-to-pixel ratio. The key insight is the available beam current, in the numerator, ought to be allocated to as few beams as possible which is physics-limited for a domain of application.

In a sense, the multi-beam compressive sensing system optimized for speed would use as much beam current per beam as would be used in the single beam of a SEM highly optimized for speed.

The intermediate term is the ratio of total electron budget to the average number of electrons each pixel receives in a compressive sample. For a fixed number of electrons per beam per measurement, the only variable in the denominator is the number of beams. Thus, the number of beams in the illumination pattern is a primary control on the number of times a pixel is sampled in a fixed number of compressive measurements, or in other words, the number of incident electrons each pixel receives in a given measurement matrix.

Conveniently, the number of pixels per image and the methods for selecting pixels in any sensing pattern are free variables that can be chosen optimally for the engineering constraints of the system or the mathematical constraints of the compressive sensing issues in a given domain of interest.

8.3.10 Scientific Challenges in a Multi-Beam CSEM

A variety of multi-beam electron microscopes exist for lithography applications, proving that various and essentially arbitrary illumination patterns can be created using multiple electron beams. In general, compressive reconstructions are more sensitive to noise in the measurement process than traditional imaging methods. The principal scientific challenges in a multi-beam CSEM are the extents to which sources of measurement uncertainty in the system dominate the compressed measurements. As compressive sensing in electron microscopy is lightly explored, the topics introduced in this section are intended to outline areas that would benefit from research and development, rather than to raise insurmountable objections to the concept.

A finite and relatively small number of electrons are used in electron microscopy to measure responses of the specimen. A well-optimized system is essentially signal-to-noise dominated in which the elapsed time to acquire an image is principally related to the number of electrons per pixel necessary to acquire the desired image quality. In a multi-beam system in which an array of apertures is used to create many beams from one source, each electron emitted can pass through at most one aperture hole. An immediate uncertainty is the number of electrons necessary to create sufficiently uniform illumination among the beams created by the aperture array. This uncertainty was evaluated by simulation and becomes negligible with 40 or more electrons per aperture per sample. Alternative methods for creating multiple beams, such as an array of sources or multiple illumination of a photocathode create variable emission that could be problematic in compressive sensing.

An associated uncertainty, already discussed, is the extent to which imprecise vs. precise measurement of each pixel value is necessary in each compressive sample. As in SEM and TEM, the secondary and backscatter electron response of the sample to the incident beam is a principal source of uncertainty and the principal driver of long illumination time to increase SNR, which improves as the square root of the illumination duration. An open scientific question is the necessary SNR or dwell in a CSEM collecting sparse samples of a global nature. This is related to CS theory but also to stochastic and noise processes in an electron microscope system. Clearly, if each pixel can be ‘undersampled’ in compressive measurements, samples can be collected faster.

While not noise per se, beam deformation and spreading due to lens aberrations and electron repulsion require careful consideration. For example, to reduce electron repulsion effects, multi-beam systems being developed for lithography applications are often operated at electron acceleration voltages higher than the 0.5-30 keV typically used in surface electron microscopy [42, 43]. Methods for simulating these quantities in complex microscopes are mature and commercially available [54, 55]. For a given total beam current, electron repulsion would tend to increase as the number of beams decrease.

Clearly, a fast CSEM requires an efficient detector system which collects a large percentage of the electrons emitted from the sample. Chamber response is the emission of electrons (indirect BSE and SE3, etc.) from surfaces struck by BSE emitted by the sample. Since it derives directly from BSE emitted by the sample, it can be considered as an indicator or amplification of primary BSE signal [56]. Chamber response electrons would be measurement noise in a CSEM only if the specific application requires collection of SE2 exclusively. In principal, detectors are a final potential source of noise however most detectors in common use are relatively clean compared to the variability of specimen response.

The uncertainties and sources of noise discussed above are all common to existing electron microscopes. A real compressive sensing electron microscope may have additional uncertainties in the precise shapes and positions of the beams in the global illumination patterns, deriving from many sources. First, the shape of each beam will to some extent be unknown and varying across the array due to the discrete, stochastic nature of electrons as well as various engineering imperfections and optical aberrations in the illumination electron column. This will be true even in the case of long illumination times, careful calibration, and the potential use of anti-stigmation micro-optics on each beam.

The optimal size of the beams on the sample is an area for further study. The use of beams that are smaller than the intended pixels may help in reducing the effects of uncertainty of beam shape and position. Using beams larger than pixels, which would sample patches of pixels simultaneously while maintaining pixel precision in location, may have help average out noise in pixel-domain sampling [57]. However, uncertainty in beam location is an explanation offered for degraded reconstruction in experiments on a SEM [28], which could occur for patch sampling in a multi-beam CSEM.

Another source of sampling matrix uncertainty is imperfect knowledge of the sample. Uneven topography will distort the sensing pattern. Sample tilt will induce skew distortions. The application of a CSEM favors flat samples in mountings that provide high precision orientation orthogonal to the axis of illumination.

A further source of sampling matrix uncertainty is charging in the sample. Under many imaging conditions, a sample illuminated by an electron beam develops an imbalanced charge near the incident beam. Surface charge (of either polarity) causes the incident beam to deflect from the intended position, and can affect the capture efficiency of detectors. Since a multi-beam CSEM envisions sampling globally over an area perhaps many microns wide, differential surface charge could develop which would cause spatially-variable distortion of the measurement patterns and possibly spatially-variable detection efficiency of the emitted electrons. The magnitude of the charging-induced distortions would be unknowable and potentially progressive over time depending on the illumination

process and sample properties. Spatially-variable, time-progressive distortions would be a significant challenge in reconstruction. Specimens or specimen preparations resistant to charging such as carbon coating or active charge suppression methods such as gas injection or complementary illumination would be desirable for surface imaging applications of compressive sensing. Since sample charging is a local phenomenon, it could be that sparse measurement matrices such as 1/1000 beam-to-pixel ratio would allow charge imbalance to dissipate as the beam landing locations jump around irregularly.

The devices that create the sensing patterns, like many multi-beam electron lithography systems, would probably use some sort of switchable or configurable microelectronics in the illumination column. The devices in this system are subject to degradation and possible failure, which would be another source of measurement matrix uncertainty. An aperture could become contaminated, affecting illumination and beam shape. Of potentially greater impact would be the failure of a device such that a beam that is supposed to vary in its location or intensity becomes stuck in one position. For long service life between major maintenance, the favored system would be resilient in device function, provide run-time knowledge of device state, and allow any of several beams to illuminate a location in the sensing matrix, so that failure of one device does not leave gaps in areal coverage over a series of patterns.

Methods to address image recovery in the presence of uncertain deviations from intended sensing patterns and other uncertainties have been considered [58]. It is likely that a CSEM would require reconstruction methods that accommodate some measurement matrix uncertainty.

In summary, the principal scientific questions for electron microscopy with sparse sampling and compressive reconstruction center on the impact of system noise and uncertainty on the compressive reconstruction process. It seems possible that experiments using one or more of the various existing multi-beam electron systems, adapted from electron beam lithography, could be conducted to explore and characterize the specific effects of each of these uncertainties. As these experiments would intend to answer issues of a mostly scientific nature, they could be conducted at much slower speed and much lower resolution than would be desired in an operational CSEM. A well designed and executed demonstration of compressive sensing in a multi-beam electron microscope, exploring and answering the issues outlined in this section, would be an important milestone.

8.3.11 Engineering Challenges in a Multi-Beam CSEM

A scientific demonstration of compressive sensing in a multi-beam system could be undertaken at any resolution and speed, and thus any existing multi-beam system is a potential candidate for compressive sensing experiments that could significantly reduce scientific uncertainty. A compressive sensing microscope capable of sustained operation collecting imagery faster than existing alternatives at near-nanometer resolution is another matter altogether. Some SEMs operate at near SNR-limited rates over large mosaics, and an emergent parallel scanning SEM could potentially be optimized to collect data over sustained rates more closely to its peak speed of greater than 1 billion samples per second. For a CSEM to go faster, significant engineering challenges exist in many of the necessary subsystems. This section identifies and briefly discusses some of the anticipated engineering issues that would need to be solved in a CSEM that is dramatically faster than alternative approaches. As in the scientific challenges section, the intent here is to open avenues of potential research rather than to raise objections to the

possibility of a fast CSEM. Only a few people have examined the vast engineering landscape of a fast CSEM and feasible solutions may already exist for some or many of these challenges.

A simple but vital issue is focus and stability. How would a multi-beam CSEM be focused, and how would the focus be kept stable for a long period of time so that the sustainable data rate is high? Most likely, a scanning mode or setup mode would be necessary, in which the compressive sampling and reconstruction is bypassed to enable a slower and more direct imaging process. With fast enough reconstruction (potentially in hardware), an iterative scheme could be attempted.

Like the electron beam lithography industry, a multi-beam CSEM would likely incorporate micro-optic devices to create, shape, and control beams. Two principal engineering challenges are micro-optic bandwidth and fabrication. Bandwidth is the product of the number of independent controls (bits) times the switching speed of the array, yielding bits per second. It is not possible to route hundreds of thousands or millions of time-synchronized lines, each one carrying one of several possible voltages, into micro-optic devices laid out over a few square mm. Either only a few thousand (at most) deflection elements can be built, or some method of reducing routing bandwidth is necessary. A steerable beam concept might utilize on the order of a thousand independent deflection devices. To explore bandwidth, assume each deflection device requires X and Y deflection controls with one deflection voltage per control setting. If each beam has just two voltage settings (0 and +1), a thousand independent beams requires 4000 bits per pattern. A sampling frequency of 20 million hertz (50 ns dwell per sample) is common in existing SEMs. Such a CSEM would have a bandwidth of 80 billion bits per second. A different steerable beam concept might create a million beams arranged in a 1000 x 1000 array with some sort of bandwidth-reducing control system as discussed previously. A million beam array configured thus would use a similar number of bits per sample as the array of a thousand independent beams. 80 billion bits per second is large for MEMs devices and thus bandwidth to the deflection array of micro-optic devices is a significant challenge. A vast worldwide community desires increased microelectronics bandwidth, however, so a current challenge becomes increasingly tractable over time.

The voltage controls on the micro-optics will require either CMOS logic integrated proximal to the devices, or a pattern switching method that is very low bandwidth. CMOS integration with MEMs structures has been performed in electron microscopes [42] however some embodiments of the CSEM could require MEMs-CMOS integration at a complexity not yet attempted. A bandwidth-reduction solution that reduced randomness in the patterns of beams would theoretically require more compressive samples and hence represent an engineering trade-off between switching bandwidth and microscope speed. Similar to bandwidth, in this area the CSEM benefits from innovation in the broader microelectronics industry.

A second area of engineering challenges is in the objective lens of the microscope, which demagnifies the beam pattern from the micron-scale dimensions in the electron column to the nanometer scale dimensions necessary on the sample. As mentioned previously, approximately 2 uA of beam current could pass through the objective lens as 1000 or more individual beams. An IMAGE computer model of parallel beams passing through an existing electromagnetic objective lens maintained spot sizes on the sample of several nanometers suggesting that this requirement is reasonable [46]. However, contrary

evidence is that the electron beam lithography industry shifts to high acceleration voltage (~50 keV) to reduce electron repulsion interactions.

A thorough canvassing of available detectors and their suitability for compressive sensing applications has not been conducted. Potential issues include detector placement, dynamic range, sensitivity, and speed. A detector placed at the bottom of the polepiece is sometimes used in SEMs, and would present a large capture area for CS measurements. The optimal detector configuration and location would be determined in formal design. As the compressive measurement patterns are created by many beams striking the sample simultaneously with high total current, there will be a large electron flux. The detector may need a dynamic range and sensitivity to be able to distinguish relatively small variability in signals. If relatively few beams and few electrons per beam per sample are sufficient, the sampling rate may exceed 100 million samples per second.

Alternative approaches to potential detector challenges may be possible. Noting that undersaturation compresses dynamic range and detector speed is inversely proportional to bit depth, sampling strategies and sparse recovery conditions have been developed for saturated responses and as few as 1 bit of detector depth [59, 60]. Independent of their potential as a path forward for detector challenges, these alternative strategies point to the value of 'contrarian thinking' in designing a CSEM.

In compressive sensing, the measurements are somewhat independent of the back-end reconstruction. During the design and prototyping phases, reconstruction algorithms simply need to be 'good enough' to provide quick iteration and useful feedback to designers, and evidence that the proposed system is competitive or economically justifiable. As system development proceeds, custom, domain-specific sparsifying representations can be developed, as has already been done in TEM. Improved reconstruction methods occur as domain-specific constraints are identified and mathematical challenges are worked out. These developments can only result in improved image quality and higher speed by requiring fewer measurements. Thus, relative to the scientific uncertainties and engineering challenges in a multi-beam CSEM, robust methods for image reconstruction appears as a lesser risk.

Compressive reconstruction can be computationally complex. Significant computational resources may be necessary to invert the compressive measurements into usable images. Since a myriad of approaches are possible, including custom hardware, the issues of computational complexity are regarded as an engineering challenge. In the case of using compressive sensing to collect large quantities of data, the engineering challenge of compressive reconstruction could be insignificant compared to the overall challenge of storing and automatically processing terabytes to petabytes or, someday, exabytes of information to a compact representation of the desired information.

8.3.12 Conclusions

Sparse sampling and compressive reconstruction offer an alternative design space for electron microscopy, which is largely unexplored. The first applications of CS in electron microscopy used existing electron microscopes with reduced sampling and sparse reconstruction. This is most active in TEM, which integrate information tomographically in each measurement which is highly suited to CS. Identified benefits of CS in TEM include reduced number of tilt-rotation samples required, reduced

overall dose to the specimen, improved resolution of the z-dimension (depth) in the reconstructed volume, and reduced impact of the 'missing wedge' (inaccessible tilt-rotations) in reconstructed images. Fewer explorations of sparse sampling and reconstruction exist for AFM and SEM, which do not intrinsically measure integrative samples.

Engineering modification of existing electron microscopy technology for sparse sampling has been explored in a fast electrostatic beam shutter in TEM, alternative sampling patterns in AFM with sparse reconstruction, star-field sampling in SEM, and modeling and inverting SEM beam dynamics enabling knowledge of the multi-pixel sensing function associated with each measurement. The modification of beam shape in existing electron microscopes has not been explored.

Sparse sampling and reconstruction in multi-beam electron microscopes has been lightly explored from theoretical and conceptual design perspectives, chiefly motivated by the desire for fast collection of large quantities of imagery. The difficulty of wide-field imaging in electron microscopy makes multi-beam illumination with single-detector response a natural fit for compressive sensing and motivates development. The existence and advanced technological development of multi-beam electron lithography microscopes demonstrates the ability to create precise, repeatable, and essentially arbitrary electron beam patterns on a specimen. Furthermore, many of the beam-forming and control technologies developed for lithography appear highly suited to multi-beam compressive sensing. A conceptual design suggests that multi-beam electron microscopy with sparse sampling and reconstruction could collect upwards of billions of pixels per second, due to parallel illumination allowing high total beam current. Principal scientific challenges relate to uncertainties in noise processes and their effects in sparse reconstruction. Of the many engineering challenges, a principal concern is the micro-electronics bandwidth to create the sensing patterns. The theoretical/conceptual analysis suggests that a multi-beam CSEM optimized for speed would have as few beams as possible with maximal current in each beam subject to resolution and other issues, rather than fewer beams with lower beam current.

8.3.13 References

1. Nyquist, H., *Certain topics in telegraph transmission theory.* Transactions of the A.I.E.E., 1928(2): p. 27.
2. Shannon, C.E., *Communication in the presence of noise.* Proceedings of the Institute of Radio Engineers, 1949. **37**(1): p. 11.
3. Candès, E.J., *Compressive Sampling.* Proceedings of the International Congress of Mathematicians, Madrid, Spain, 2006. **III**: p. 20.
4. Candès, E.J., J.K. Romberg, and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements.* Communications on Pure and Applied Mathematics, 2006. **59**(8): p. 1207-1223.
5. Donoho, D.L., *Compressed sensing.* IEEE Transactions on Information Theory, 2006. **52**(4): p. 1289-1306.
6. Duarte, M., et al., *Single-Pixel Imaging via Compressive Sampling.* IEEE Signal Processing Magazine, 2008. **25**: p. 83-91.

7. Barbey, N., et al., *Feasibility and performances of compressed sensing and sparse map-making with Herschel/PACS data.* Astronomy & Astrophysics, 2011. **527**: p. A102.
8. *Dartmouth College Electron Microscope Facility*. [cited 2016 March 4, 2016]; Available from: http://www.dartmouth.edu/~emlab/gallery/.
9. Studer, V., et al., *Compressive fluorescence microscopy for biological and hyperspectral imaging.* Proceedings of the National Academy of Sciences, 2012. **109**(26): p. E1679-E1687.
10. Brady, D.J., et al., *Compressive Holography.* Optics Express, 2009. **17**(15): p. 13040-13049.
11. Marim, M.M., et al., *Compressed sensing with off-axis frequency-shifting holography.* Optics letters, 2010. **35**(6): p. 871-873.
12. Goris, B., et al., *Electron tomography based on a total variation minimization reconstruction technique.* Ultramicroscopy, 2012. **113**: p. 120-130.
13. Kim, M.W., et al. *Cryo-electron microscopy single particle reconstruction of virus particles using compressed sensing theory*. 2007.
14. Vonesch, C., et al. *Fast wavelet-based single-particle reconstruction in Cryo-EM*. in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*. 2011. IEEE.
15. Wang, L., Y. Shkolnisky, and A. Singer, *A Fourier-Based Approach for Iterative 3D Reconstruction from Cryo-EM Images.* arXiv:1307.5824V1, 2013.
16. Veeraraghavan, A., et al. *Increasing depth resolution of electron microscopy of neural circuits using sparse tomographic reconstruction*. in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. 2010. IEEE.
17. Hu, T., et al., *Super-resolution using sparse representations over learned dictionaries: Reconstruction of brain structure using electron microscopy.* arXiv preprint arXiv:1210.0564, 2012.
18. Goris, B., et al., *Electron tomography based on a total variation minimization reconstruction technique.* Ultramicroscopy, 2012. **113**: p. 120-130.
19. Leary, R., et al., *Compressed sensing electron tomography.* Ultramicroscopy, 2013. **131**: p. 70-91.
20. Thomas, J.M., et al., *A new approach to the investigation of nanoparticles: Electron tomography with compressed sensing.* Journal of colloid and interface science, 2013. **392**: p. 7-14.
21. Binev, P., et al., *Compressed sensing and electron microscopy*. 2012: Springer.
22. Stevens, A., et al., *The potential for Bayesian compressive sensing to significantly reduce electron dose in high-resolution STEM images.* Microscopy, 2014. **63**(1): p. 41-51.
23. Béché, A., et al., *Development of a fast electromagnetic shutter for compressive sensing imaging in scanning transmission electron microscopy.* arXiv preprint arXiv:1509.06656, 2015.
24. Béché, A., et al., *Development of a fast electromagnetic beam blanker for compressed sensing in scanning transmission electron microscopy.* Applied Physics Letters, 2016. **108**(9): p. 093103.
25. Abramovitch, D.Y., et al. *A tutorial on the mechanisms, dynamics, and control of atomic force microscopes*. in *American Control Conference, 2007. ACC'07*. 2007. IEEE.
26. Song, B., et al. *Video rate atomic force microscopy (AFM) imaging using compressive sensing*. in *Nanotechnology (IEEE-NANO), 2011 11th IEEE Conference on*. 2011. IEEE.
27. Andersson, S.B. and L.Y. Pao. *Non-raster sampling in atomic force microscopy: A compressed sensing approach*. in *American Control Conference (ACC), 2012*. 2012. IEEE.
28. Anderson, H.S., et al. *Sparse imaging for fast electron microscopy*. in *IS&T/SPIE Electronic Imaging*. 2013. International Society for Optics and Photonics.
29. Chang, P.I., et al., *Local raster scanning for high-speed imaging of biopolymers in atomic force microscopy.* Review of scientific instruments, 2011. **82**(6): p. 063703.
30. Vincent, T.J., et al., *Longitudinal brain size measurements in APP/PS1 transgenic mice.* Magnetic Resonance Insights, 2010. **4**: p. 19.

31. Lichtman, J.W. and W. Denk, *The big and the small: challenges of imaging the brain's circuits.* Science, 2011. **334**(6056): p. 618-623.
32. Intelligence Advanced Research Projects Agency, S.a.S.O.O., *Circuit Analysis Tools (CAT) IARPA Broad Agency Announcement*. 2009.
33. Intelligence Advanced Research Projects Agency, S.a.S.O.O., *Rapid Analysis of Various Emerging Nanoelectronics (RAVEN) Proposer's Day*. 2015.
34. Intelligence Advanced Research Projects Agency, S.a.S.O.O., *Rapid Analysis of Various Emerging Nanoelectronics (RAVEN)*. 2015.
35. Bauer, E., *Low energy electron microscopy.* Reports on Progress in Physics, 1994. **57**(9): p. 895.
36. Telieps, W. and E. Bauer, *An analytical reflection and emission UHV surface electron microscope.* Ultramicroscopy, 1985. **17**(1): p. 57-65.
37. Kemen, T., et al. *Further advancing the throughput of a multibeam SEM*. in *SPIE Advanced Lithography*. 2015. International Society for Optics and Photonics.
38. Eberle, A., et al., *High-resolution, high-throughput imaging with a multibeam scanning electron microscope.* Journal of microscopy, 2015. **259**(2): p. 114-120.
39. Michael, J.R., et al., *High-Throughput SEM via Multi-Beam SEM: Applications in Materials Science.* Microscopy and Microanalysis, 2015. **21**(S3): p. 697-698.
40. Malloy, M., et al. *Massively parallel E-beam inspection: enabling next-generation patterned defect inspection for wafer and mask manufacturing*. in *SPIE Advanced Lithography*. 2015. International Society for Optics and Photonics.
41. Anderson, H.S., et al., *Sparse sampling and reconstruction for electron and scanning probe microscope imaging*, in *US Patent and Trademark Office*. 2015, Sandia Corporation: United States.
42. Klein, C., H. Loeschner, and E. Platzgummer, *50-keV electron multibeam mask writer for the 11-nm HP node: first results of the proof-of-concept electron multibeam mask exposure tool.* Journal of Micro/Nanolithography, MEMS, and MOEMS, 2012. **11**(3): p. 031402-1-031402-7.
43. McCord, M.A., et al. *REBL: design progress toward 16 nm half-pitch maskless projection electron beam lithography*. in *SPIE Advanced Lithography*. 2012. International Society for Optics and Photonics.
44. Anderson, H.S., J.W. Wheeler, and K.W. Larson, *63. Compressed Sensing for Fast Electron Microscopy*, in *TMS 2014 Supplemental Proceedings*. 2014, John Wiley & Sons: Hoboken, NJ, USA.
45. Larson, K.W., H.S. Anderson, and J.W. Wheeler, *Fast electron microscopy via compressive sensing*, U.P.a.T. Office, Editor. 2014, Sandia Corporation: US.
46. Larson, K.W., et al., *An Electron Microscope for Compressive Sensing with Flood Beam Illumination*. 2014, Sandia National Laboratories: Albuquerque, NM.
47. Yin, W., et al. *Practical compressive sensing with Toeplitz and circulant matrices*. in *Visual Communications and Image Processing 2010*. 2010. International Society for Optics and Photonics.
48. Rivenson, Y. and A. Stern, *Compressed imaging with a separable sensing operator.* Signal Processing Letters, IEEE, 2009. **16**(6): p. 449-452.
49. Duarte, M.F. and R.G. Baraniuk, *Kronecker compressive sensing.* Image Processing, IEEE Transactions on, 2012. **21**(2): p. 494-504.
50. van den Berg, E. and M.P. Friedlander, *SPGL1: A solver for large-scale sparse reconstruction*. 2009.
51. Van Den Berg, E. and M.P. Friedlander, *Probing the Pareto frontier for basis pursuit solutions.* SIAM Journal on Scientific Computing, 2008. **31**(2): p. 890-912.

52. Gilbert, A. and P. Indyk. *Sparse recovery using sparse matrices*. 2010. Institute of Electrical and Electronics Engineers.
53. Jafarpour, S., et al., *Efficient and robust compressed sensing using optimized expander graphs.* Information Theory, IEEE Transactions on, 2009. **55**(9): p. 4299-4308.
54. *SIMION*.
55. *Munro's Electron Beam Software.*
56. Goldstein, J., et al., *Scanning electron microscopy and x-ray microanalysis.* ISBN. Vol. 306472929. 2003, United States: Springer. 9780306472923.
57. Chartrand, R. 2015.
58. Parker, J.T., V. Cevher, and P. Schniter. *Compressive sensing under matrix uncertainties: An approximate message passing approach*. in *Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on*. 2011. IEEE.
59. Laska, J.N., et al., *Democracy in action: Quantization, saturation, and compressive sensing.* Applied and Computational Harmonic Analysis, 2011. **31**(3): p. 429-443.
60. Jacques, L., et al., *Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors.* Information Theory, IEEE Transactions on, 2013. **59**(4): p. 2082-2102.

NOTE FIGURES AND TABLES APPENDED AT END OF MANUSCRIPT DRAFT

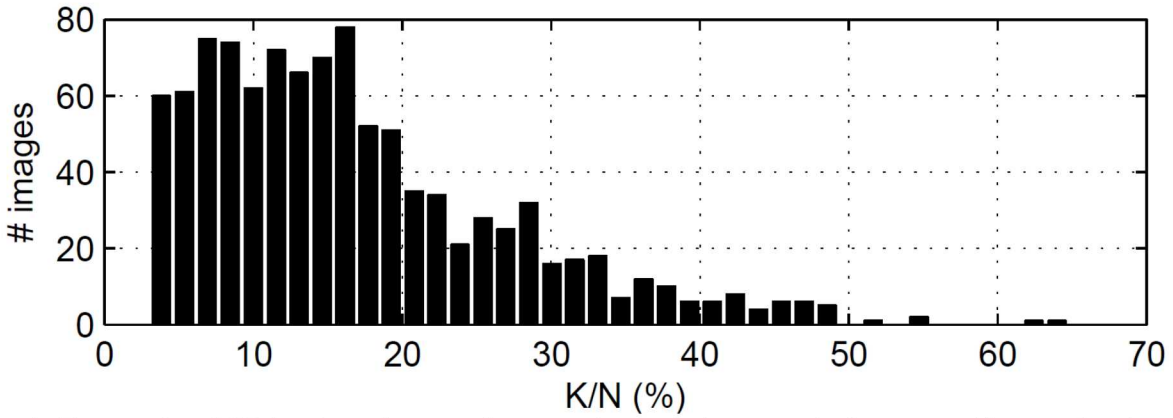

Figure 1. The sparsity of 1022 various electron microscopy images, where sparsity is measured by counting the number of block-DCT coefficients K that account for at least 99.75% of the image energy and dividing by the number N of pixels. The average sparsity is 17%, with half of all images less than 15% sparse, and three-quarters less than 20% sparse.

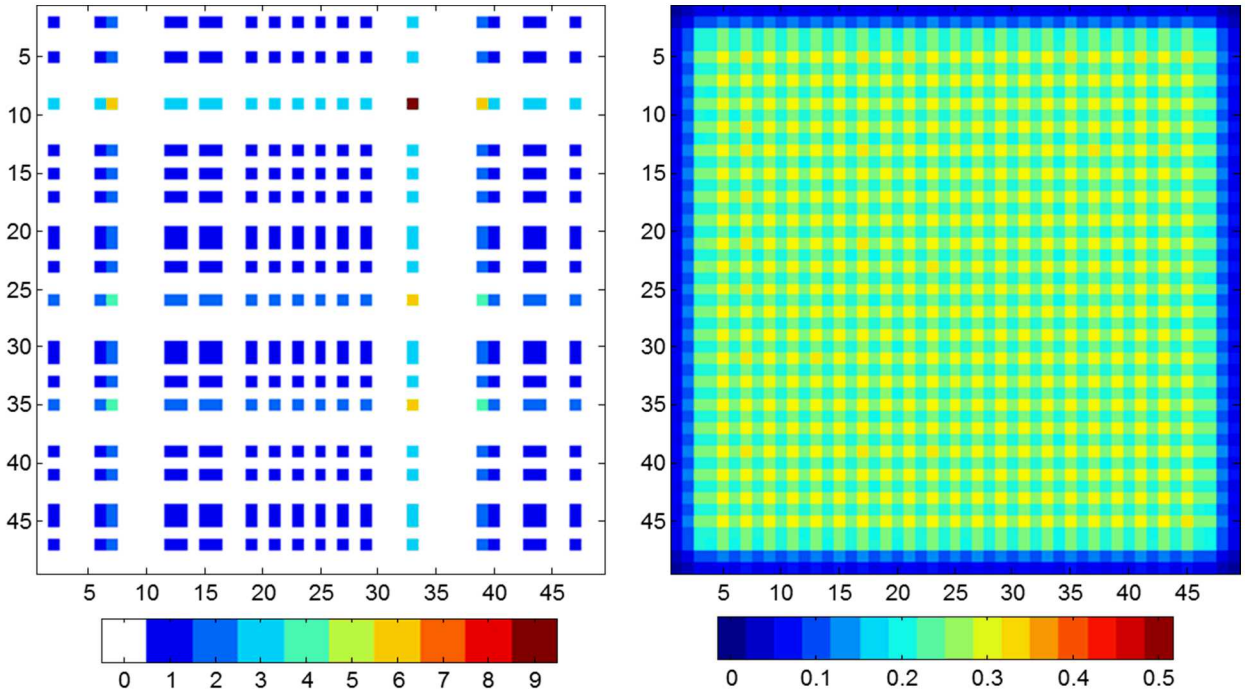

Figure 2. (left) One realization of a sampling pattern for an array of 47x47 beams, with a two-pixel offset between beam centers and a -2,-1,0,+1,+2 steerable offset randomly selected and applied uniformly to each row and column. The beam-to-pixel ratio in this configuration is 25% and a pixel is illuminated by 0 to 9 beams in any realization, as indicated by the color bar. (right) The average illumination per pixel in a realization if offsets are selected uniformly from the distribution.

NOTE FIGURES AND TABLES APPENDED AT END OF MANUSCRIPT DRAFT

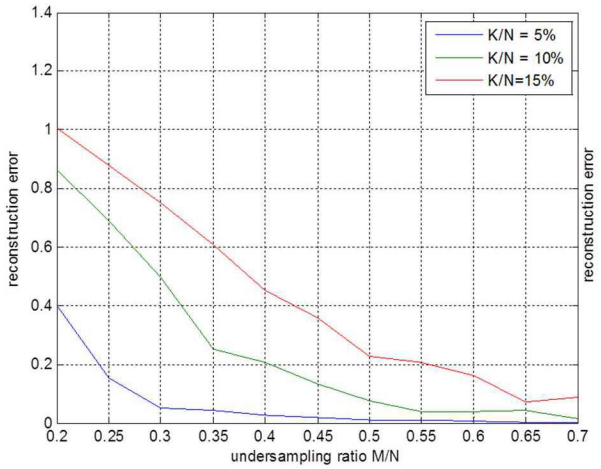

Figure 3. Reconstruction results for an array of correlated beams configured as in Figure 2. Acceptable reconstruction error of less than 0.1 is achieved for images with 5% or lower sparsity above approximately 0.3 undersampling ratio.

NOTE FIGURES AND TABLES APPENDED AT END OF MANUSCRIPT DRAFT

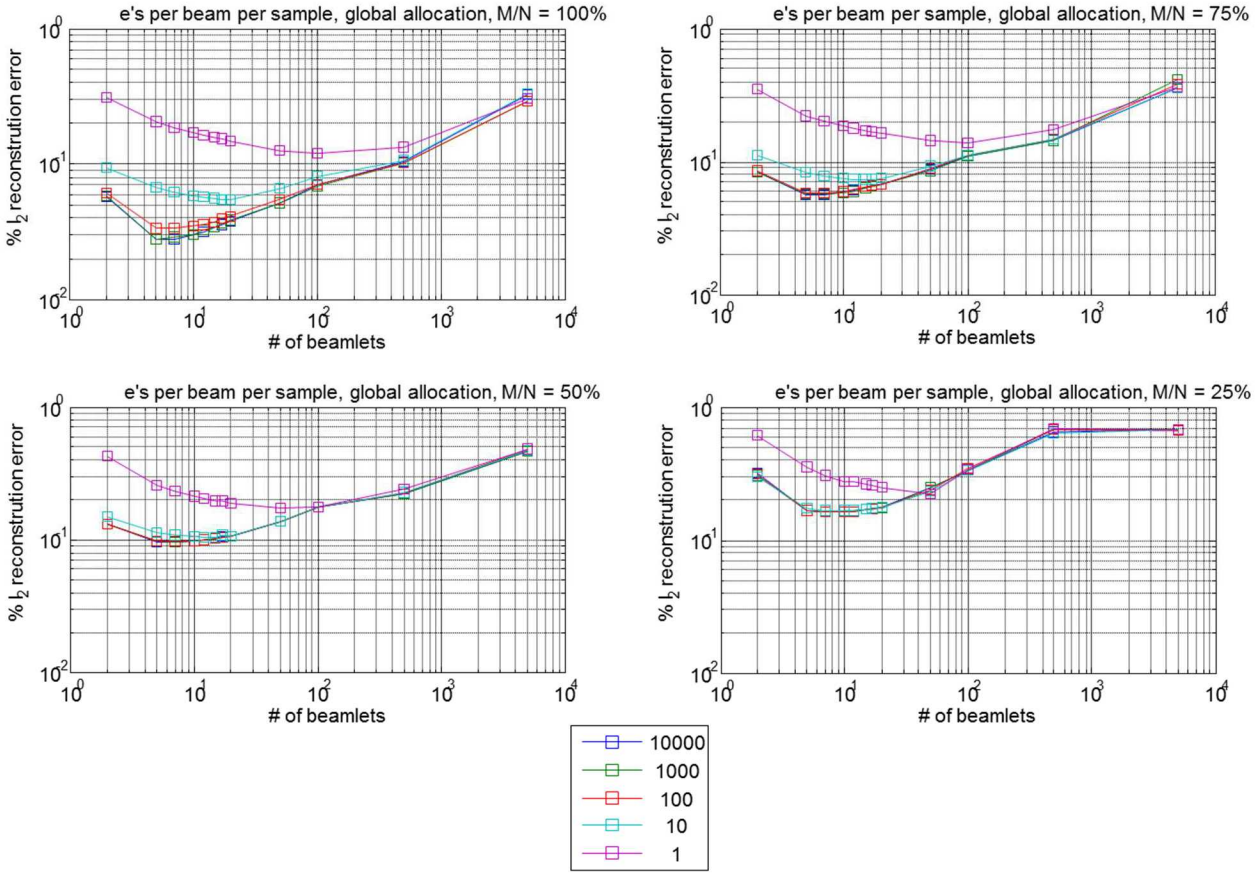

Figure 4: Reconstruction error as a function of the number of beams (beamlets) in a hypothetical CSEM for a set of 100x100 pixel images. Colored lines indicate the electrons per beam per sample. The four subplots test sampling ratio (samples to pixels). Global allocation means that each beam was assigned a random location in the image without restriction. Global allocation was uniformly slightly better than a local allocation scheme in which each beam was restricted to a 4x4 pixel patch (results not shown). Reconstruction error is smaller for sparse measurement matrices (few beams) rather than dense. 10 or fewer electrons per beam yield significantly poorer reconstruction, and 100 or more electrons per beam per sample provide indistinguishable results. The 'knee in the curve', or threshold electron budget, is about 40 electrons per beam per compressive sample.

NOTE FIGURES AND TABLES APPENDED AT END OF MANUSCRIPT DRAFT

Scenario	# Beams	Pixel to Beam Ratio	µs per image	MP per second	Pattern Switch Rate, MHz	nA per Beam	MP per second Formula
1	10000	1 [on|off]	80	125	12.5	0.2	125
2	5000	2	40	250	25	0.4	250
3	10000	2	160	125	12.5	0.2	125
4	10000	16	1280	125	12.5	0.2	125
5	1000	1000	800	1250	125	2	1250
6	1000	500	400	1250	125	2	1250
7	500	1000	200	2500	250	4	2500
8	250	1000	50	5000	500	8	5000

TABLE 1. Calculations of speed for a hypothetical multi-beam compressive sensing electron microscope in which 2 nA of beam current (12500 electrons per nanosecond) are distributed among the indicated number of beams. The exposure is assumed to be 100 electrons per beam(pixel) per compressive sample. The Pixel to Beam ratio is simply the size of the recovered image divided by the number of beams used to collect it. The µs per image indicates the total time to collect samples necessary for reconstruction, assuming no switching latency. The MP per second is the data rate for a single image. The pattern switch rate indicates the rate of compressive sample collection. The nA per beam indicates the beam current per beam. The Formula and Scenarios are discussed in the text.