Nanodiamond Imaging: Methods and Apparatus

Alex Hegyi

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2014-186
http://www.eecs.berkeley.edu/Pubs/TechRpts/2014/EECS-2014-186.html

December 1, 2014

Nanodiamond Imaging: Methods and Apparatus

by

Alex Hegyi

Submitted to the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, in partial satisfaction of the requirements for the degree of **Master of Science.**

Readers:

_____

Professor Eli Yablonovitch
Research Advisor

_____

(Date)

****

_____

Professor Michael Lustig
Second Reader

_____

(Date)

Abstract: Nanodiamond imaging is a new molecular imaging modality that takes advantage of optically-detected electron spin resonance of nitrogen-vacancy defects in nanodiamonds to image the distribution of nanodiamonds, non-invasively within a living organism, with high spatial resolution and high sensitivity. Nanodiamonds can be tagged with a range of biologically active molecules such that they distribute within an organism according to the distribution of a biochemical target, such as by selectively concentrating in cancerous tissue. By imaging the spatial distribution of the nanodiamonds with a nanodiamond imaging system, one can thus discern information pertaining to various biological processes. In this project report, we describe the methods and apparatus that constitute nanodiamond imaging, both in terms of the current embodiment as well as in terms of hypothetical approaches that would offer improved performance.

I. Summary of invention

The invention "Nanodiamond Imaging" relates to a new kind of molecular imaging modality for imaging in one, two, or three dimensions that combines the sensitivity of optical detection with the resolving power of magnetic gradients. Molecular imaging refers to imaging the *in-vivo* distribution of a tracer or probe designed to target a particular biomolecule, as opposed to anatomical imaging, which captures only structural information. Applications such as monitoring gene expression patterns, diagnosis and longitudinal tracking of disease, stem cell tracking, and tumor detection are enabled by molecular imaging technologies [1]. Nanodiamond imaging complements existing and future methodologies for obtaining structural information about organisms such as magnetic-resonance imaging (MRI) and x-ray computed tomography (CT). It can function either as a stand-alone imaging modality or integrated into a hybrid molecular/anatomical imaging modality, such as nanodiamond/CT, for co-registration of anatomical and molecular information. Nanodiamond imaging will complement or improve upon other forms of molecular imaging, such as positron emission tomography (PET), single-photon emission computed tomography (SPECT), ultrasound, and optical fluorescence/bioluminescence imaging, due to its higher resolution and similarly high sensitivity [2]. Also, unlike PET, SPECT, and CT, nanodiamond imaging poses no risk of ionizing radiation.

Nanodiamond imaging derives its high levels of sensitivity and resolution from a tracer made of nanodiamonds containing the negatively charged nitrogen vacancy color center in diamond (NV), an atomic defect in the diamond lattice with many unusual and attractive properties. Among these properties is a paramagnetic ground state with long spin coherence and spin relaxation times. The spin state can be optically polarized with red light, detected via spin-dependent fluorescence intensity in the near-infrared (at wavelengths that easily penetrate biological tissue), and further manipulated via microwave and static magnetic fields. In addition to these physical properties, the nanodiamond host opens up an array of attractive biological applications, due to its non-toxicity and overall biocompatibility [3], easy bonding to other biomolecules [4], and time stability (allowing longitudinal studies).

The imaging works by defining a static magnetic field gradient with a field-free region, such as a field-free point or a field-free line. Microwaves and excitation light are applied to the sample to be imaged, and fluorescence from the sample is monitored as the field-free region is scanned across the sample. The microwaves induce a change in fluorescence from the nanodiamonds that are located in the field-free region. An image is formed by tracking these changes in fluorescence and indexing them to the instantaneous location of the field-free region.

Because the light used in imaging—both for excitation and fluorescence detection—must pass through the sample, the imaging modality will be most useful for small-

animal molecular imaging, where the required penetration depth is low. Sensitivities higher than 10 $nM \cdot mm^3 \cdot Hz^{-1/2}$ should be attainable. Resolutions are only limited by the strength of the magnetic gradient, and they should be finer than 100 μm with gradients as low as 5 T/m, potentially reaching single-cell resolution when imaging a portion of a small animal.

Although the targeted application of the invention is non-invasive preclinical molecular imaging, it is expected that the imaging principles described herein can be extended without modification to invasive imaging, non-invasive and invasive clinical molecular imaging of human subjects, and molecular imaging of smaller fields of view (such as within a cell) with up to single particle sensitivity and nanometer resolution, both *ex-vivo* and *in-vivo*.

II. Description of Imaging System

A schematic of one embodiment of the imaging system is shown in Figure 6. The components of the system work together to perform optically detected magnetic resonance imaging of NVs present in nanodiamonds within a sample that can be optically scattering. There is a magnetic subsystem to generate strong spatially-varying magnetic fields; an optoelectronic excitation and detection system; and a microwave excitation source.

A. Magnetics subsystem

The magnetics subsystem generates a magnetic field to map the positions of the NVs within the sample to changes in fluorescence as a function of magnetic field and/or microwave frequency. For example, consider the simple magnetic field shown in Figure 4. That is, let's say we want to generate a magnetic field $\boldsymbol{B}$ where $|\boldsymbol{B}| \rightarrow 0$ (a null, or a field-free region), and we want to shift this field relative to the sample. This field could be generated along the axis of a Maxwell pair of coils, which is two circular electromagnet coils located a distance apart on the same axis, with the current in one coil running in the opposite sense as the current in the other coil. The plotted field is then the axial component of the magnetic field along the coils' shared axis. The null occurs halfway in between the two coils, assuming the current through each coil is equal and opposite. However, the null can be moved by the application of a constant magnetic field supplied either by another set of coils (that is, with current in the same sense in both coils) or by the same set of coils. Moving the null across a point of NV-containing nanodiamonds and applying optical excitation and microwaves at 2.869 GHz, the changes in fluorescence caused by modulation of the microwaves are detected.

Depending on the imaging sequence performed, different magnetic fields may be required. They all have in common that they encode the nanodiamond position into observable variations in fluorescence intensity, in order to reconstruct an estimate of nanodiamond concentration within the sample.

1. Field-free point vs. Field-free line

There are two main magnetic field configurations that generate a null at some point while still satisfying the divergence-free condition for magnetic fields, $\nabla \cdot \boldsymbol{B} = 0$. These are the field-free point (FFP) and field-free line (FFL) configurations. In the FFP configuration, the null in the magnetic field exists at one point in space. Imaging takes place by scanning the point in a 3-dimensional raster, thus collecting 3-dimensional position information about the nanodiamond concentrations. In the FFL configuration, the null exists along a line. Imaging is done by moving the FFL across the sample, and the observed changes in fluorescence encode the nanodiamond positions as projected into a plane perpendicular to the FFL.

It is possible to generate all the required magnetic fields for both the FFL and FFP configurations using electromagnets. For an arrangement of six coils that would allow this, see Figure 5. Up to all six of the coils can participate in generating the FFP, whereas only four are needed to generate an FFL. For details of an imaging technique, called magnetic particle imaging, that uses a similar orientation of magnetic fields to generate a signal from particles in a field-free region, see [5]. By varying the amount of current going into each coil, the location of the FFP/FFL as well as the strength of the magnetic gradient at the FFP/FFL can be adjusted as desired. Ultimately there needs to be a uniquely defined field-free region within the sample volume.

Alternatively, a field-free region can be defined with the use of permanent magnets. If there is a limit to the amount of power available for the electromagnets, for example due to thermal constraints, it is more practical to generate the field-free region with permanent magnets and devote the entire power budget to the constant field required to shift the field-free region, rather than sharing it between generating and shifting the field-free region. Finally, it is possible to scan on one or more of the axes mechanically by placing the sample on a stage that moves relative to the FFL/FFP.

Note that it may be desired to have an adjustable magnetic gradient, which could be accomplished with permanent magnets by adjusting the distance from the magnets to the sample, or could also be implemented with only electromagnets. An adjustable gradient would allow for adjustable resolution imaging, so the field of view could be rapidly searched for nanodiamonds, for example by the method of bisection. That is, if a small isolated spot(s) of nanodiamonds is known to exist, the field of view can be divided into two voxels. The voxel producing the highest signal is further divided in two, and this process is repeated. In a field of view with N voxels, the scanning time to find the voxel yielding the maximum signal is reduced from N to $2 \log_2 N$.

B. Optical subsystem

The optical subsystem is comprised of excitation and detection stages.

1. Excitation

In the excitation stage, one or more sources of light, such as LEDs, lasers, etc., generate light in a range of wavelengths chosen to optimize depth penetration in tissue and optical absorption at the NV. For example, a high power red LED can generate several watts of light from 610-630 nm. A current source provides a stable source of power to the LED, minimizing intensity fluctuations that would otherwise contribute to system noise. Optical band-pass filters can be used in the excitation stage to remove tails from the LED emission or laser lines that overlap in wavelength with the NV fluorescence. Other optical elements, such as lenses, baffles, etc., can be added to relay the excitation light from the source to the sample. The light can be relayed to the sample via optical fiber, whereby the sample can be illuminated from the outside (non-invasively) or from the inside (invasively) such that the excitation source is closer to the NVs.

The excitation must be intense enough to effect a significant polarization of the NV spins. To fully polarize an NV spin requires absorption of three to four photons in a time less than the spin relaxation time of $T_1 \approx 1$ ms [6]. The signal strength depends linearly on spin polarization. Full spin polarization is not necessary; however. pulsed measurement schemes can be used to overcome the limitations of spin relaxation and thermal loading from high optical intensities. Or, the optical excitation spot can be scanned over the sample, such that the optical intensity is highest within the magnetic field-free region. Scanning the optical excitation, as opposed to continuous full-sample illumination, reduces the total optical power incident on the sample and thus the thermal load on the sample. It also reduces the non-modulated component of the fluorescence signal that contributes shot noise to the modulated component of the fluorescence signal, and it can potentially decrease the width of the system point-spread function by restricting the region of signal generation closer to the field-free region.

Furthermore, a fluid, such as water, can be run over the sample to provide the dual purpose of better index-matching to the optical excitation source as well as cooling the sample from the thermal load caused by the optical excitation (or maintaining a specific sample temperature).

2. Detection

The detection stage must record changes in fluorescence across the NV fluorescence band throughout the imaging procedure. Fluorescence from the sample is collected with lenses and passed through optical filters to selectively collect only NV fluorescence (and to specifically reject excitation light scattered from the sample). Any combination of optical elements may be used with the goal of maximizing light collection from the NV fluorescence and minimizing collection of all other light. Note that in a properly engineered system, the dominant source of noise will be shot noise from the NV fluorescence and background light. Background, which should be minimized as much as possible, consists of excitation light that makes it through the detection stage filters, as well as tissue autofluorescence and fluorescence from imaging system components.

The detector can be any sort of photoelectric detector, such as a photodiode operating in linear or avalanche mode, or a photomultiplier tube. It must introduce an acceptably low amount of system noise, its bandwidth must be great enough to encompass the measurement bandwidth, and the dynamic range must be large enough to measure the small signal modulation on the large DC background with high fidelity. A reverse-biased photodiode is ideal because it satisfies all three of these criteria. Although only a single pixel detector is necessary, a multi-element detection system could provide extra information about the location of the nanodiamonds that could be used during image reconstruction.

C. Electrical detection

The electrical subsystem comprises the front-end photodetector amplifier(s), as well as any relevant analog signal processing. It can be helpful to do some sort of narrow-band, synchronous detection. It is also possible at this stage to electronically remove additive noise on the signal caused by fluctuations in the DC background from amplitude variations in the excitation source, which may be necessary if the excitation source is a laser.

In Figure 6, the photodetector amplifier is a transimpedance amplifier (TIA).

1. Lock-in detection

Lock-in detection, or synchronous narrow-band detection, is used to maximize the SNR. In a CW measurement (i.e., excitation continuously on), one way to modulate the measurement is to amplitude modulate the microwaves with a low-frequency oscillator or LFO. This strategy will work if the microwave transition is not heavily saturated. Or, the microwaves can be chopped (i.e., square wave on/off modulation) to get to high modulation depths with a saturated microwave transition. Other elements of the system can also be modulated. For example, while the field-free region is being scanned, a small modulation can be applied to its location. The measured signal would be the derivative (with respect to position of the field-free region) of the signal obtained when modulating the microwaves.

The detected photocurrent is converted to a voltage and amplified by the transimpedance amplifier, and the resulting signal is fed to a lock-in amplifier, to synchronously detect the modulation in fluorescence. In Figure 6, this is shown schematically as multiplication of the output of the TIA with the LFO, and then passing the result through a low-pass filter.

2. Background “subtractive” reference detection

Because we are attempting to detect a small time-varying signal on a large DC background, minute fluctuations in the background will appear as a large noise component added to the time-varying signal. This additive noise will especially be a problem if a laser is used as the excitation source, because of the amplitude noise of lasers. Methods have been developed, such as [7], for subtraction of the additive noise component by using a reference photodetector and feedback circuit that automatically adjusts the gain of the subtraction circuit. This noise subtraction

technique can achieve shot-noise limited performance, and is still compatible with lock-in measurement techniques.

D. Microwave subsystem

The microwave subsystem consists of a microwave generator, possibly a modulator, a radiator, and any necessary tuning or matching circuitry. The microwave generator must generate microwaves in a band that includes at least the zero-magnetic-field spin transition of the NV at 2.869 GHz; high average and instantaneous powers and low phase noise are desired traits of a microwave generator. Many different kinds of modulation can be used—amplitude modulation of the microwaves or simple chopping (on/off) modulation of the microwaves with a PIN switch, or frequency modulation. Also, pulsed microwave schemes can be used as described in the next section, with either CW or pulsed optical schemes.

There must be some method of delivering the microwaves to the sample. A waveguide or coaxial cable may be used for this, with some kind of radiating structure or antenna at the end, such that the sample sees the far field, e.g. approximating a plane wave. Or, a small loop of wire, such that the sample sees the near field of the current in the wire, can be used. A resonant structure or cavity could be employed to enhance the microwave field, like a resonant coil of wire (even a single turn). Matching circuitry can be used to match the radiating or resonant element to the impedance of the waveguide or coax.

E. Pulsed measurement sequences

Pulsed measurement schemes increase either the signal-to-noise ratio or the number of measurement points acquired in a given time relative to CW measurements. Pulsing refers to concentrating the optical and/or microwave energies within a fraction of a measurement cycle and with a specified timing sequence, without increasing the average optical and/or microwave energies over the measurement cycle.

1. Continuous wave measurement

The simplest measurement scheme, also the easiest to implement, is the continuous wave (CW) excitation scheme, depicted in the timing diagram in Figure 7. Here, the excitation light is left on continuously, while the microwaves are modulated with a 50% duty factor (or amplitude modulated with a sine wave). The amount of spin polarization that occurs is a function of the excitation light intensity. Thus, the equilibrium spin polarization is the steady state between the polarization of the spins caused by optical pumping and the natural spin relaxation. At lower optical excitation intensities, where the spin polarization is linear in the excitation intensity, a two-fold increase in optical intensity thus results in a *four*-fold increase in signal, and only a $\sqrt{2}$-fold increase in noise. Since this is a synchronous measurement, the measurement phase is shown to alternate synchronously with the chopping of the microwaves. Following the timing diagram in Figure 7, the measurement is carried out as the field-free region is rastered across the sample.

2. Partial duty factor measurement

One of the ways to take advantage of the increase in signal at higher optical excitation intensities without excessive temperature rise of the sample (or to stay within legal/safe excitation intensity power limits, etc.) is to lower the duty cycle of the measurement, as in Figure 8. In the example above, a two-fold increase in optical excitation intensity induces a four-fold increase in signal (due to the increased spin polarization) and a $\sqrt{2}$-fold increase in noise. Such an increase in excitation intensity would be tolerated at a 50% duty cycle, which would decrease the signal by a factor of 2 and decrease the noise by a factor of $\sqrt{2}$. Overall, the SNR would increase by up to a factor of 2 by going to a 50% duty cycle at the same average excitation intensity.

3. Coherent microwave pulsed measurement

As shown in Figure 9, coherent microwave pulses allow measurements much faster than the spin relaxation time. A strong optical pulse pumps the NV spin polarization, followed by a weaker pulse to probe the NV fluorescence without inducing significant spin polarization. During the probe pulse, the NV spins within the field-free region are periodically inverted by a series of "π" pulses, and changes in fluorescence are measured synchronously with the spin inversions. This measurement technique allows much faster scanning of the field-free region than the previous measurement techniques allow: whereas before, the modulation (and thus the scanning rate in voxels/second) could not proceed at a rate faster than the spin relaxation time, now the modulation must proceed at a rate faster than the spin relaxation time (i.e. the spin must remain polarized for this technique to work).

Note that any combination of pulsed/cw excitation light as well as coherent and incoherent microwave excitation may be used. However, coherent microwave pulses require the ability to generate high instantaneous microwave powers, such that the microwave Rabi frequency is greater than the decoherence rate.

F. Imaging particles

The imaging particles used in this invention are nanodiamonds containing a concentration of nitrogen vacancy centers. Other particles can be used, but the desired traits are:

1) Some form of optically-detected magnetic resonance (such as optically induced spin polarization and fluorescence-based spin detection)
2) Long spin relaxation T_1 and spin coherence T_2 times
3) Near-infrared excitation and fluorescence
4) Non-toxicity/biosystem compatibility
5) Easily functionalizable
6) Stable
7) Easy to produce
8) Small zero-magnetic field splitting in the ground state

1. Nanodiamond

There are at least a couple of viable methods to produce NV-containing nanodiamonds, while tailoring the size to a particular application.

a) HPHT Synthesis

High-pressure-high-temperature synthesis is a common method of manufacture for synthetic diamond particles. Once diamond particles of the appropriate size are obtained, nitrogen vacancies must be created inside them, for example by ion implantation and annealing [8].

b) Detonation nanodiamond

Nanodiamond can also be created in a detonation process. Sintering of the detonation nanodiamond creates a very high concentration of NV centers [9]. The nanodiamond clusters can then be milled apart to the desired particle size.

c) Possible improvements to nanodiamond

A possible improvement to the nanodiamond would be to reduce crystal strain. Strain shifts apart the $m_s = \pm 1$ levels of the NVs even in zero-magnetic field, reducing the signal contrast obtainable from shifting apart the $m_s = \pm 1$ magnetically

2. Functionalization

The nanodiamonds or nanoparticles can be functionalized to attach to specific biomarkers [4]. Functionalization enables many applications for this imaging technique: tracking single cells or small groups of cells, quantitatively labeling the expression of various genes, targeting certain types of tumors, etc. Tumors can also be targeted and imaged using dynamic contrast enhancement (i.e., the enhanced permeability and retention effect) [10].

3. Nitrogen vacancy centers

Nitrogen vacancy centers are optically and magnetically active defects in a diamond crystal. They consist of two missing carbon atoms along a <111> direction, with one of the carbon atoms being replaced by a nitrogen atom. In this document, all mentions of nitrogen vacancy centers, NVs, etc., refer to the negatively-charged NV^-. The dangling bonds from the carbon atoms and nitrogen atom neighboring the vacancy, as well as an extra electron, join together to determine the energy levels of the NV center. The center has a net spin of S = 1 in the ground electronic state.

Both the optical and magnetic or spin properties of the NV are utilized in the present invention.

a) Optical Properties

The NV has an optical transition from state $^3A_2$ to state ^3E (see Figure 1). Although the zero-phonon line for this transition is at 637 nm, it can be excited non-resonantly, with the peak of excitation occurring around 560 nm. For the purposes of this invention, red light (~620 nm) is used for excitation, since it penetrates tissue more easily, although other wavelengths can be used. With more than 99% fluorescence quantum yield, the NV is an excellent fluorescent emitter. Its emission

peak is around 700 nm but emission spans the range from about 650 nm to about 800 nm, coinciding with the near-infrared window (absorption of hemoglobin toward the blue and absorption of water toward the infrared) in biological tissue.

b) Magnetic Properties

The electronic ground state of the NV is a spin triplet. There is a microwave transition between the triplet sublevels: spin sublevels $m_s = \pm 1$ lie above spin sublevel $m_s = 0$ by 2.869 GHz at zero magnetic field. An external magnetic field splits the $m_s = \pm 1$ energy levels according to the Hamiltonian

$$H = DS_z^2 - g\mu_B \boldsymbol{B} \cdot \boldsymbol{S}, \qquad (1)$$

where D is the zero field splitting ($D = 2\pi\hbar * 2.869$ GHz), $\boldsymbol{S}$ is the spin operator and S_z is its projection along the z-axis or axis from nitrogen to vacancy, μ_B is the Bohr magneton, g is the NV center's g-factor ($g \approx 2$), and $\boldsymbol{B}$ is the external magnetic field. At room temperature, the spin relaxation and spin coherence times are roughly 1.2 ms [6] and 1 μs [9].

c) Magneto-optic Properties

It is the combination of magnetic and optical properties that make the NV center unique and useful for this particular application, and there are two important aspects that combine these properties. The first is optically-induced spin polarization. Due to the level structure of the NV, upon optical cycling of the $^3A_2$ — ^3E transition, the spin population polarizes into the $m_s = 0$ state. This polarization occurs because of a preferential coupling of ^3E's $m_s = \pm 1$ states to the singlet state $^1A_1$, which couples preferentially to $^3A_2$'s $m_s = 0$ state, as in Figure 1. Just as important is the fluorescence-based spin-state detection. Because of the transitions shown in Figure 1, the $m_s = \pm 1$ state fluoresces less than the $m_s = 0$ state, by about 30%.

d) Possible improvements to NVs

The NVs used in the imaging system would benefit from an increased T_1 and T_2 time. The increased T_1 would allow for lower optical intensities to achieve the same steady-state spin polarization. The increased T_2 would contribute to narrower lineshapes, and would enable lower microwave powers and longer times for measurement schemes that relied on coherent use of microwaves.

III. Imaging Procedures

There are several imaging procedures of interest—somewhat briefly described in the section on creating magnetic fields—that all rely on generating a unique field-free region in space. These are the field-free line method (generating two-dimensional projections or three dimensional projection-reconstructions) and the field-free point method.

A. Field-free point scanning

One of the simplest imaging procedures to understand is the field-free point scan. In this technique, a field-free point, as described in the section on magnetic field generation (II.A), is rastered across the sample as one of the measurement schemes (pulsing or CW) is performed. As shown schematically in Figure 4, the signal generated by a point of nanodiamonds is highest when the field-free point overlaps with the point of nanodiamonds. It is possible to define a width of the field-free point as the distance the field-free point must move to trace out the full-width at half max of the signal peak caused by the point of nanodiamonds. Independent measurements of the signal can be taken on a 2-dimensional Cartesian grid (for a slice scan) or 3-dimensional Cartesian grid (for a full 3-d scan). The measurements as gridded constitute the raw image.

B. Field-free line (2-d projection)

Another imaging procedure is the field-free line procedure. Here, the field-free region is a line that projects through the sample. As in Figure 4, the magnitude of the signal is highest when the field-free region crosses a point of nanodiamonds. The line is rastered across the sample in a 2-dimensional grid. Because the line extends through the sample, any points of nanodiamond that are simultaneously along the line will be measured at the same grid point. This procedure thus yields a two-dimensional projection of the nanodiamond concentration.

1. Combination of 2-d projections into 3-d rendering

Several two-dimensional projections of the sample can be combined into a three-dimensional image, using techniques such as filtered back-projection or other iterative reconstruction methods. A multiplexing advantage results from acquiring three-dimensional data from 2-d projections relative to using the three-dimensional point scan technique. Due to the parallel acquisition of data at different depths using the projection technique, an $N \times N \times N$ image can be acquired N times faster with the projection technique than with the 3-d point-scan technique, assuming same signal to noise ratio. Or, in the same measurement time, an image taken with the projection technique will have a $\sqrt{N}$ higher SNR.

C. Signal processing

Once the raw data is acquired there are many different ways that it can be processed before display.

1. Raw data

The raw data consists of the fluorescence measurements derived from the measurement system (for example the output of the lock-in amplifier), indexed by the location of the field-free region. Physically, each measurement corresponds to the nanodiamond concentration within the sample at a certain point, after multiplication by a system matrix that may represent convolution by a point spread function or may also take into account position-sensitive dependencies of the location of the nanodiamonds.

2. Thresholded raw data

Because of the broad tails of the point-spread function, one of the ways to increase contrast of the resulting image is to threshold the raw data, such that the intensity of any image point below a threshold value is set to zero.

3. Constrained least squares deconvolution

A more quantitatively accurate method of image processing is to perform some kind of deconvolution of the image by the point-spread function or to solve the inverse problem of what nanodiamond concentration produced the resulting image. For example, assume that $\boldsymbol{Ax} + \boldsymbol{n} = \boldsymbol{b}$ is an equation that models the imaging system, where $\boldsymbol{x}$ is a vector representing the nanodiamond concentration over the sample volume, $\boldsymbol{A}$ is the system matrix that represents the transformation of the nanodiamond concentration into an image, $\boldsymbol{n}$ is noise added to the ideal image, and $\boldsymbol{b}$ is the measured data. The noise vector is a Gaussian random vector with known variance (assume the noise at each pixel or voxel is an independent identically distributed Gaussian). The maximum likelihood estimate of $\boldsymbol{x}$—let's call it $\boldsymbol{x}^*$—is then given by the solution to the equation: $\boldsymbol{x}^* = \text{argmin}_{\hat{\boldsymbol{x}}} |\boldsymbol{A}\hat{\boldsymbol{x}} - \boldsymbol{b}|_2$, where $|\cdot|_2$ represents the standard Euclidean (ℓ_2) norm. However, we must enforce the element-wise constraint on $\hat{\boldsymbol{x}}$ of $\hat{\boldsymbol{x}} \geq \boldsymbol{0}$, otherwise the solution will most likely include unphysical negative estimates of nanodiamond concentration.

When constructing the system matrix $\boldsymbol{A}$, many factors can be taken into account. A simple system matrix would assume that the system was shift invariant, essentially modeling a convolution. Underlying this assumption are the further assumptions that the excitation intensity, both optical and microwave, and the detection sensitivity, are shift invariant. In addition, for strict shift invariance the microwave field needs to remain perpendicular to the static magnetic field. This geometry can be accomplished by having the microwave field axis point in the same direction as the field-free line. Otherwise, the point-spread function may vary with the direction of the microwave field relative to the direction of the magnetic field at any given voxel.

A more complex system matrix would take into account spatial variations in excitation intensity and detection sensitivity, and would calculate the observed lineshape at each pixel/voxel by solving the Hamiltonian for the nitrogen vacancy center. A distribution of nanodiamonds within a pixel or voxel can be assumed to include nanodiamonds in all orientations with varying strains. The Hamiltonian of an NV in a static magnetic field, taking strain into account, and written in the basis of the NV axis as the z-axis, is

$$H = DS_z^2 + E\left(S_x^2 - S_y^2\right) - g\mu_B \boldsymbol{B} \cdot \boldsymbol{S}, \qquad (2)$$

with E being the strain parameter and the other terms defined after Eq. (1). First, the energy levels of the Hamiltonian are solved for a particular strain E and particular crystal orientation relative to the magnetic field. The corresponding states can be ordered in energy: $|0\rangle$, $|1\rangle$, and $|2\rangle$, where $|0\rangle$ is composed primarily of

the $m_s = 0$ spin state of the NV and $|1\rangle$ and $|2\rangle$ are composed of superpositions of the $m_s = \pm 1$ states. Next, we calculate the Rabi frequency for the $|0\rangle \rightarrow |1\rangle$ and $|0\rangle \rightarrow |2\rangle$ transitions, given a microwave field $\boldsymbol{B_1}$ oscillating at a frequency ω. We seek these transitions because the optical excitation pumps the system into state $|0\rangle$, and the Rabi frequency is a measure of the extent to which the microwaves will induce a $|0\rangle \rightarrow |1\rangle$ or $|0\rangle \rightarrow |2\rangle$ transition. The Rabi frequencies will be a function of strain E, microwave frequency ω, and orientation of the NV's axis (θ, ϕ) relative to the lab or magnetic field frame. We thus have:

$$\begin{aligned} \Omega_1(\omega, E, \theta, \phi) &= \frac{g\mu_B}{2\hbar} \langle 0 | \boldsymbol{B_1} \cdot \boldsymbol{S} | 1 \rangle, \text{ and} \\ \Omega_2(\omega, E, \theta, \phi) &= \frac{g\mu_B}{2\hbar} \langle 0 | \mathbf{B_1} \cdot \mathbf{S} | 2 \rangle. \end{aligned} \tag{3}$$

We define E_1 as the $|0\rangle \rightarrow |1\rangle$ transition energy, and $\omega_1 = E_1/\hbar$. Similarly, we define E_2 and ω_2. The optically detected magnetic resonance lineshapes for the two transitions are

$$\begin{aligned} h_1(\omega, E, \theta, \phi) &= \frac{4\Omega_1^2 T_1^* T_2}{1 + (\omega - \omega_1)^2 T_2^2 + 4\Omega_1^2 T_1^* T_2} \text{ and} \\ h_2(\omega, E, \theta, \phi) &= \frac{4\Omega_2^2 T_1^* T_2}{1 + (\omega - \omega_2)^2 T_2^2 + 4\Omega_2^2 T_1^* T_2}, \end{aligned} \tag{4}$$

i.e., the Bloch equations for a two-level system [11] where $T_1^* = \left(T_1^{-1} + T_{\text{pump}}^{-1}\right)^{-1}$, and T_{pump} refers to the optical spin polarization time. Because the NV is a three-level system, the $|0\rangle \rightarrow |1\rangle$ and $|0\rangle \rightarrow |2\rangle$ transitions are not independent and must be combined. A combination which seems to match observations well is $h = \max(h_1, h_2)$. To calculate the final lineshape we take the following average:

$$h(\omega) = \frac{1}{4\pi} \int_{\Omega} \int_{E} h(\omega, E, \theta, \phi) p_E(E) dE \, d\Omega, \tag{5}$$

where $p_E(E)dE$ represents the probability of finding a nanodiamond with a strain in the range $[E, E + dE]$, and $\int_{\Omega} d\Omega/4\pi$ represents the collective contribution of nanodiamonds pointing at all angles. Given a source point $\boldsymbol{r'}$ of nanodiamonds and a location $\boldsymbol{r}$ of the magnetic field-free region, this function is evaluated at $2\pi \cdot 2.869$ GHz to yield one point in the system transfer function $h(\boldsymbol{r}; \boldsymbol{r'})$, or equivalently, one element of the system matrix $\boldsymbol{A}$.

It is also possible to determine the lineshape empirically. The system matrix can be directly measured by putting a test point of nanodiamonds at one voxel or pixel location, and recording a full image by rastering the field-free point. This measurement is repeated by recording a full image with the test point of nanodiamonds moved to each voxel or pixel location. Another possibility to determine the lineshape empirically is to measure the system response with the test

point of nanodiamonds at one or more locations, and then fit a lineshape to the measurements. The width of the observed lineshape $h(\omega)$ scales approximately linearly with applied magnetic field, or equivalently, with distance from the field-free region. Thus, the amplitude of the lineshape at its center should scale inversely with distance from the field-free region, i.e. $h(\boldsymbol{r};\boldsymbol{r}') \propto 1/|\boldsymbol{r}-\boldsymbol{r}'|$. However, this function diverges; broadening (for example from strain) limits the maximum amplitude. A more accurate fit can be achieved with the addition of a broadening term and a constant offset. Thus, a good fitting function for the response of the system with field-free region at point $\boldsymbol{r}$ and source point at point $\boldsymbol{r'}$ is

$$h(\boldsymbol{r};\boldsymbol{r}') = \frac{b}{\sqrt{(\boldsymbol{r}-\boldsymbol{r}')^2 + a^2}} + c, \qquad (6)$$

where in general a, b, and c can be functions of $\boldsymbol{r'}$ (or $\boldsymbol{r}$) and can be interpolated from the measurement of the system response at a small number of source points. If we discretize the position vector into M locations and the source vector into N locations, i.e. $\boldsymbol{r} \in \{\boldsymbol{r_i}\}, i \in 1 \ldots M$ and $\boldsymbol{r}' \in \{\boldsymbol{r_j'}\}, j \in 1 \ldots N$ then the system matrix is simply $A_{ij} = h(\boldsymbol{r_i};\boldsymbol{r_j'})$. Note that Wiener deconvolution and other Fourier-domain techniques can be applied when the point-spread function is shift-invariant.

4. Regularization techniques

Regularization techniques are used to limit the high-frequency noise-gain inherent in any deconvolution or system inversion. Typically, regularization is added by modifying the objective function to be minimized when doing the least-squares data fitting as described above. One type of regularization, known as Tikhonov regularization, attempts to find a minimum norm solution to $\boldsymbol{x}^* = \text{argmin}_{\hat{\boldsymbol{x}}}|\boldsymbol{A}\hat{\boldsymbol{x}} - \boldsymbol{b}|_2$ by modifying the problem to $\boldsymbol{x}^* = \text{argmin}_{\hat{\boldsymbol{x}}}(|\boldsymbol{A}\hat{\boldsymbol{x}} - \boldsymbol{b}|_2 + \beta|\hat{\boldsymbol{x}}|_2)$ where β is a positive real number. One way to interpret this minimization is that the extra term $\beta|\hat{\boldsymbol{x}}|_2$ reflects the higher probability of lower norm solutions. More generally, we can include a term $|\boldsymbol{B}\hat{\boldsymbol{x}}|_2$ where $\boldsymbol{B}$ is a matrix that can be used to weight the probabilities of $\hat{\boldsymbol{x}}$ according to the singular vectors and singular values of $\boldsymbol{B}$. For example, $\boldsymbol{B}$ can be structured to minimize the presence of sharp "edges" in $\hat{\boldsymbol{x}}$.

IV. Imaging System Performance

A. Potential resolution

The native resolution of the imaging system is determined by the width of the point-spread function, which is in turn determined by the strength of the magnetic gradient and the amount of strain present in the nanodiamonds. With the nanodiamonds and the 1 T/m gradient used in the prototyped system, the resolution (or half-width at half-max of the point-spread function) is about 500 microns. It is expected that sub-100 μm resolution will be easily achievable over preclinically-relevant fields of view.

B. Potential maximum sensitivity

The sensitivity of the current prototype system (with ~1 W/cm^2 optical excitation and a 1 cm^3 field of view) is approximately 300 μM·mm^3·Hz$^{-1/2}$, that is, in units of [molar concentration of carbon atoms] × [1 mm^3 of voxel volume] × [1 s of measurement time]$^{1/2}$. Given foreseen improvements of going to a 10% pulsed duty cycle, using high-NV-concentration sintered detonation nanodiamonds, and a coherent microwave pulsing scheme, we can expect sensitivities approaching 10 nM·mm^3·Hz$^{-1/2}$. This would allow a single (30 μm)^3 cell, labeled with a 40 μM concentration of nanodiamond, to be tracked with 1 mm precision at 1 s intervals. Note that if only a single labeled cell is within the field of view, it can be tracked without having to scan the field-free region over the entire field of view, by implementing a feedback system that keeps the field-free region at the location of the cell, with feedback on the signal magnitude. Similarly, a small number of cells can be tracked without having to scan over the whole sample volume, if the positions and velocities are measured from scan to scan and used to update the rough scan locations for the next scan step.

Here we estimate the time for a typical scan. Assume that a 4 cm × 2 cm × 2 cm volume is to be imaged with (1 mm)^3 resolution using the 2d projection reconstruction method. Assuming a measurement time of 1 minute, the measurement time is 3.75 ms per voxel, multiplied by 20 (2 cm/1 mm) due to multiplexed measurement. The noise in each voxel would be less than 40 nM. Due to the decrease of optical intensity with depth (both of the excitation light and the received fluorescence), and the decrease in spin polarization with optical intensity, we can expect that the signal will decay as $e^{-3\delta z}$, where δ is the effective optical attenuation in the tissue [12] (assume δ ≈ 1 cm^{-1}) and z is the depth into the tissue. Thus, the center of the volume (1 cm from the surface) will be imaged with only 5% of the sensitivity as the surface. If a mouse is injected with 1 μmol (12 μg) of nanodiamonds (no toxic effects shown—see [3]), and 10% of voxels are filled, voxels on the surface of the mouse will be imaged with an SNR of ~15,000, and those at the center of the mouse will be imaged with an SNR of ~800.

V. References

[1] R. A. de Kemp, F. H. Epstein, C. Catana, B. M. W. Tsui, and E. L. Ritman, "Small-animal molecular imaging methods.," *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, vol. 51 Suppl 1, no. 1, p. 18S-32S, May 2010.

[2] S. R. Meikle, P. Kench, M. Kassiou, and R. B. Banati, "Small animal SPECT and its place in the matrix of molecular imaging technologies.," *Physics in medicine and biology*, vol. 50, no. 22, pp. R45-61, Nov. 2005.

[3] E. K. Chow et al., “Nanodiamond therapeutic delivery agents mediate enhanced chemoresistant tumor treatment.,” *Science translational medicine*, vol. 3, no. 73, p. 73ra21, Mar. 2011.

[4] J.-I. Chao et al., “Nanometer-sized diamond particle as a probe for biolabeling.,” *Biophysical journal*, vol. 93, no. 6, pp. 2199-208, Sep. 2007.

[5] P. W. Goodwill and S. M. Conolly, “Magnetic Particle Imaging Devices and Methods,” U.S. Patent US 2011/02214382011.

[6] D. Redman, S. Brown, R. Sands, and S. Rand, “Spin dynamics and electronic states of N-V centers in diamond by EPR and four-wave-mixing spectroscopy,” *Physical Review Letters*, vol. 67, no. 24, pp. 3420-3423, Dec. 1991.

[7] P. C. D. Hobbs, “Ultrasensitive laser measurements without tears,” *Applied Optics*, vol. 36, no. 4, p. 903, Feb. 1997.

[8] Y.-R. Chang et al., “Mass production and dynamic imaging of fluorescent nanodiamonds.,” *Nature nanotechnology*, vol. 3, no. 5, pp. 284-8, May 2008.

[9] P. G. Baranov et al., “Enormously high concentrations of fluorescent nitrogen-vacancy centers fabricated by sintering of detonation nanodiamonds.,” *Small (Weinheim an der Bergstrasse, Germany)*, vol. 7, no. 11, pp. 1533-7, Jun. 2011.

[10] H. Maeda, J. Wu, T. Sawa, Y. Matsumura, and K. Hori, “Tumor vascular permeability and the EPR effect in macromolecular therapeutics: a review,” *Journal of Controlled Release*, vol. 65, no. 1-2, pp. 271-284, Mar. 2000.

[11] A. Yariv, *Quantum Electronics*. Wiley, 1989, p. 676.

[12] V. Tuchin, *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis*, 1st ed. SPIE Publications, 2000, p. 9.

VI. Figures

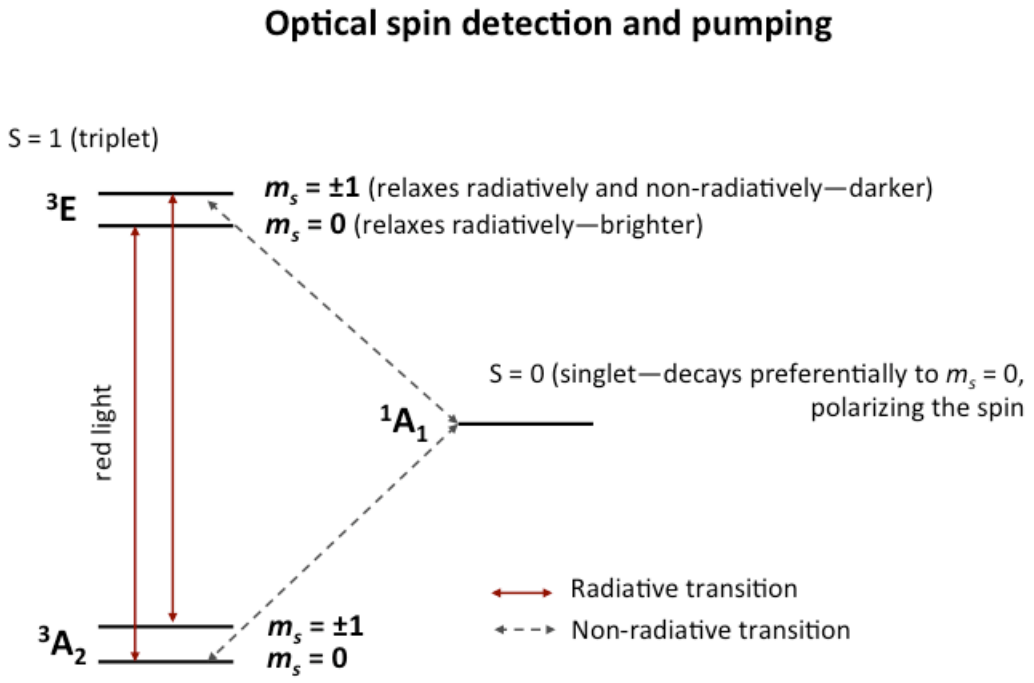

Figure 1. Optical spin detection and pumping of the NV level structure. The ground electronic state ($^3A_2$) can be excited via an optical transition to the excited electronic state (^3E), either resonantly (at 637 nm) or non-resonantly, with absorption peaking in the green around 560 nm. The $m_s = 0$ spin substate of ^3E will relax radiatively to the $m_s = 0$ substate of $^3A_2$, whereas the $m_s = \pm 1$ spin substate of ^3E will relax both radiatively to the $m_s = \pm 1$ substate of $^3A_2$ and non-radiatively to the $^1A_1$ (singlet) state. Thus, the spin state (whether $m_s = 0$ or $m_s = \pm 1$) can be determined based on the quantity of fluorescence from the NV. Furthermore, once in the singlet state, the NV will relax preferentially to the $m_s = 0$ spin substate. Thus, the NV spin becomes polarized into $m_s = 0$ upon continuous optical cycling.

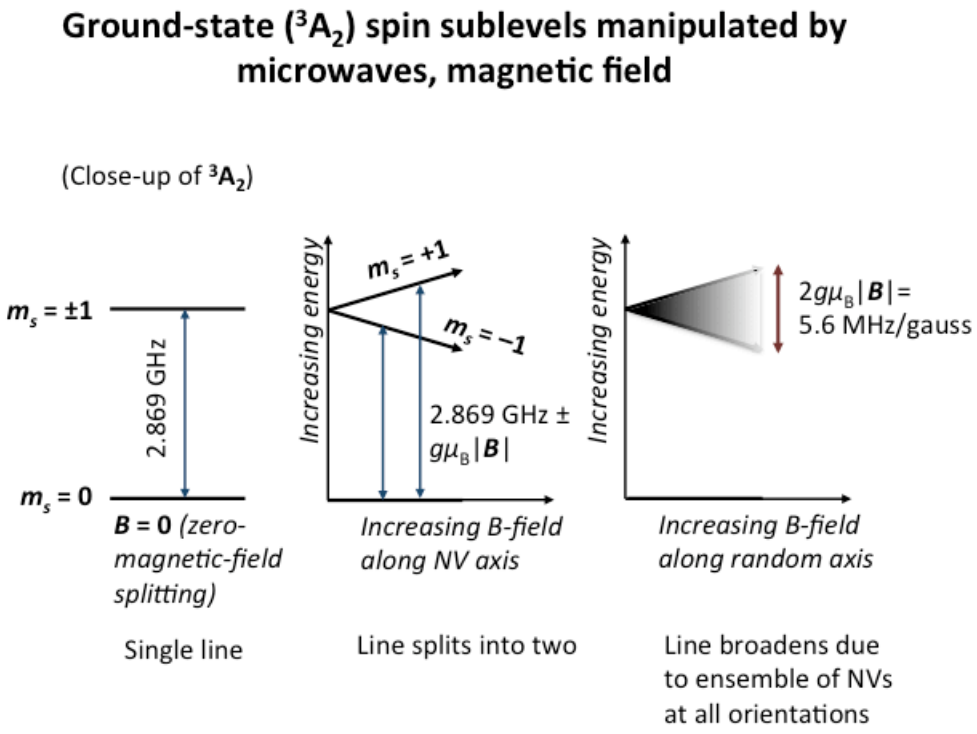

Figure 2. Close-up of how the spin sublevels of the ground electronic state shift with applied magnetic field. If the field is along the NV axis, the $m_s = +1$ and $m_s = -1$ have a maximal splitting. If one measures the fluorescence of the NV while applying a magnetic field and sweeping the frequency of a microwave source, one will observe dips in fluorescence at two transition frequencies corresponding to the transitions between the $m_s = 0$ and $m_s = +1$ states, as well as the $m_s = 0$ and $m_s = -1$ states. When there is an ensemble of NVs pointing in random directions, the splitting of each NV will depend on the projection of the applied magnetic field along the NV's axis, and the combined effect of the ensemble is a single line that broadens with applied magnetic field.

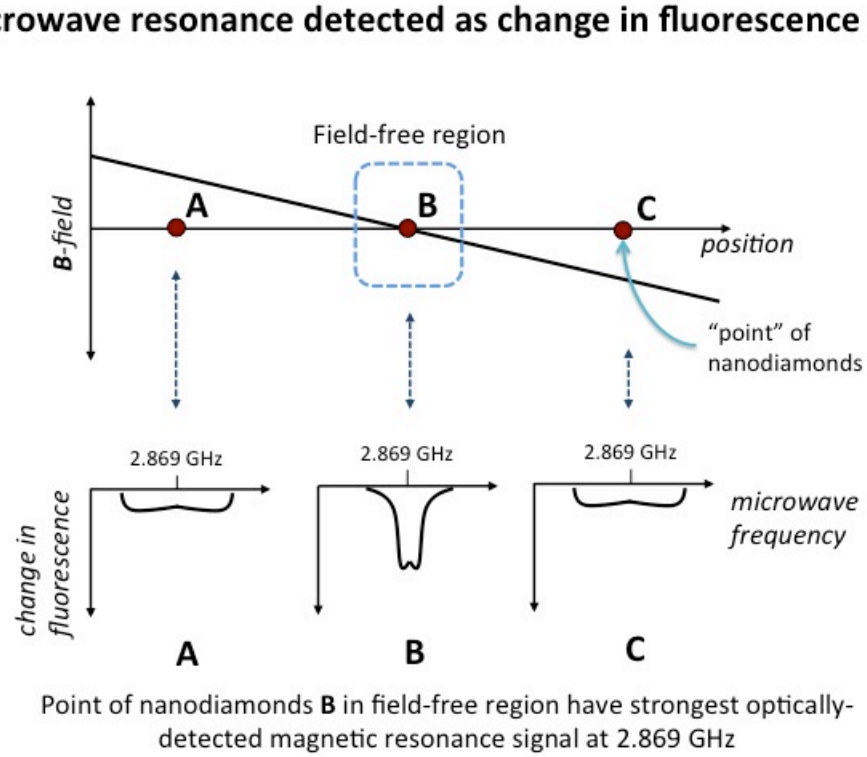

Figure 3. The top shows a possible one-dimensional magnetic field. This field could exist along the axis of a Maxwell coil pair (two circular coils that share the same axis but are spaced a distance apart to maximize the quadrupole moment). Points of nanodiamonds located at **A**, **B**, and **C** yield the corresponding optically-detected magnetic resonance signals. At the center frequency of 2.869 GHz, the signal is strongest when the point of nanodiamonds is within the field-free region.

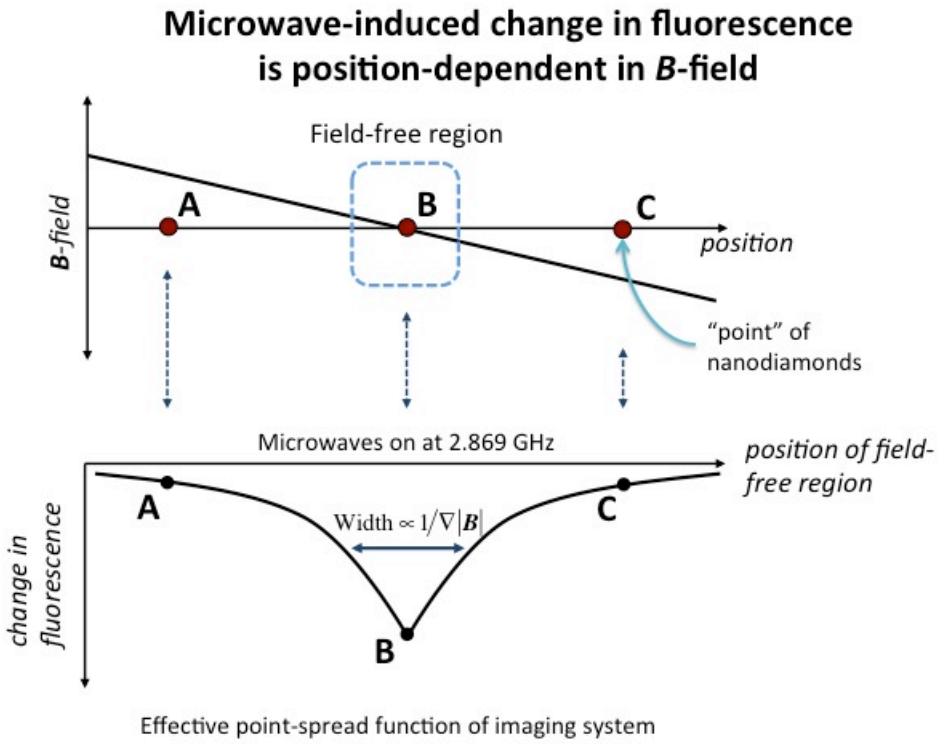

Figure 4. The same field as in Figure 3 is generated. The field is moved relative to a point of nanodiamonds. When the nanodiamonds are at **B**, or the null of the magnetic field, the change in fluorescence caused by 2.869 GHz microwaves, as shown on the bottom axis, is greatest. When the nanodiamonds are moved away from the central null, as in positions **A** and **C**, the change in fluorescence relative to baseline is less. Baseline is considered the observed fluorescence with no microwave field applied. Moving the null of the magnetic field relative to a point of nanodiamonds is one way to measure the imaging system's point-spread function.

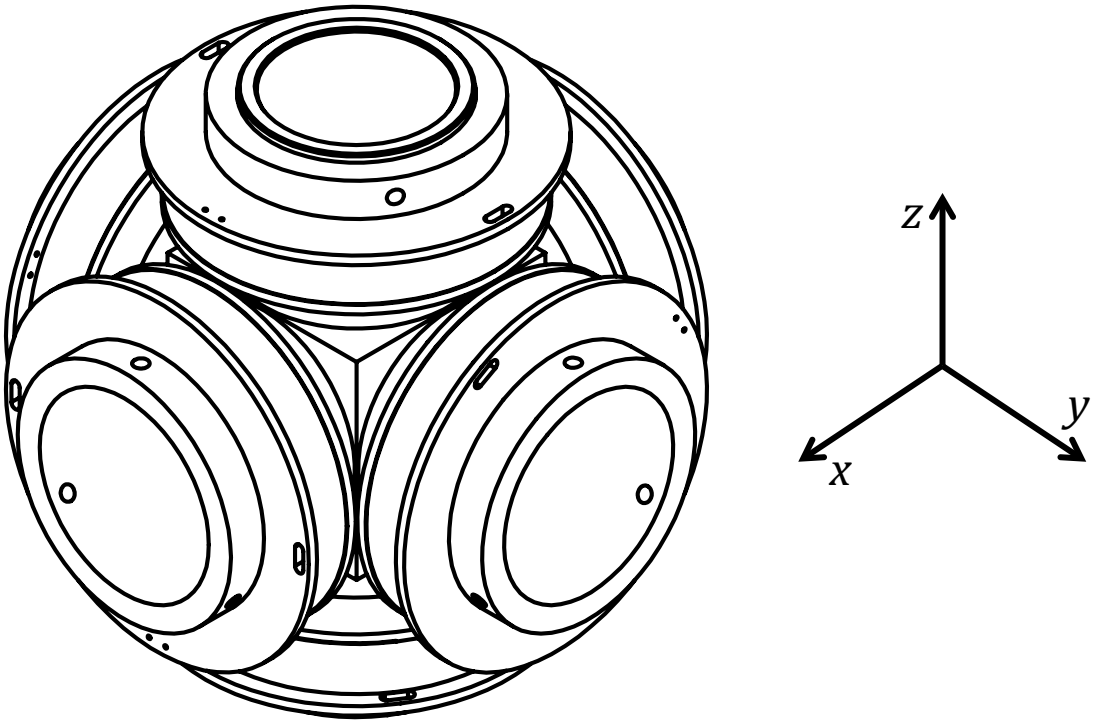

Figure 5. An arrangement of electromagnet coils with two coils on each Cartesian axis.

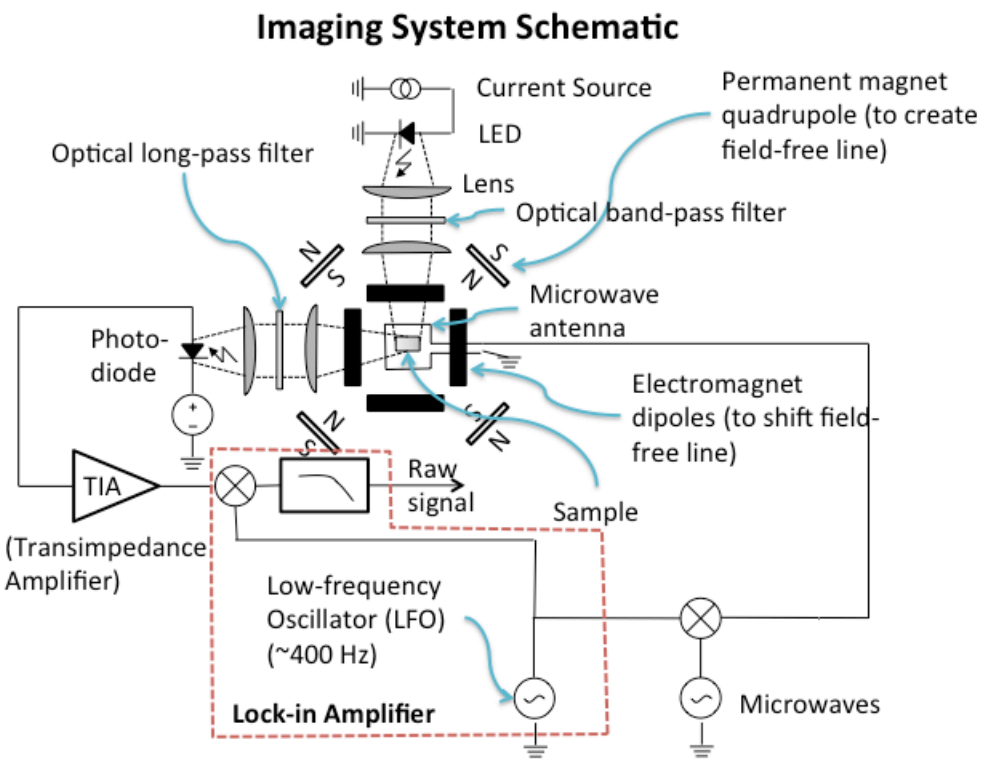

Figure 6. A schematic overview of one implementation of the imaging system. Light from an LED is band-pass–filtered to clearly define its spectrum, and then it is focused onto the sample. Fluorescence is separated from scattered light via a long-pass filter and is focused onto a photodiode. Microwaves at 2.869 GHz are applied to the sample with an antenna. The microwaves are amplitude-modulated, and changes in fluorescence are measured synchronously with the microwave modulation as various magnetic field patterns are applied.

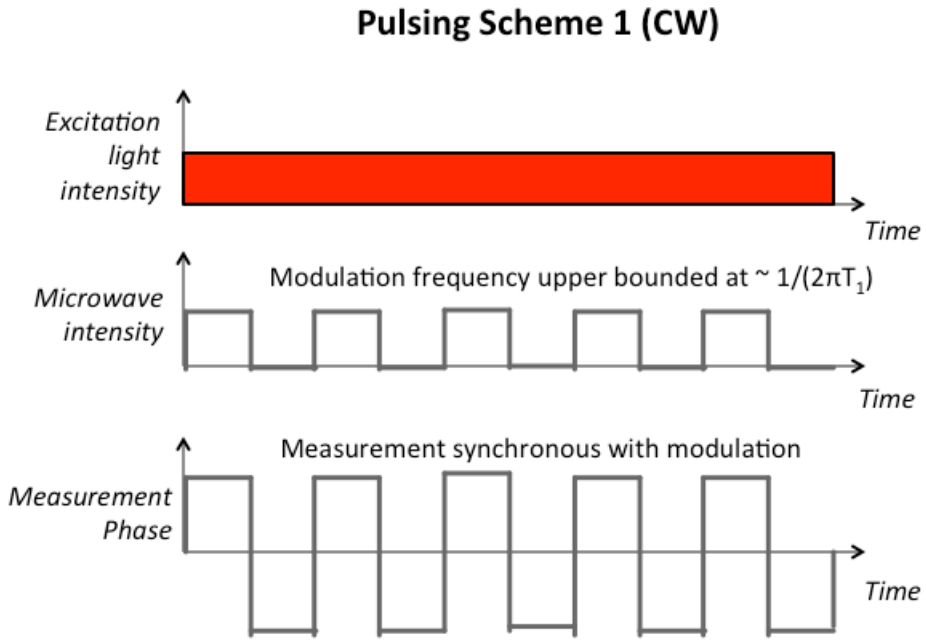

Figure 7. The easiest measurement scheme to implement is not technically a pulsed scheme. Rather, it is a continuous (CW) measurement scheme where the excitation is on continuously, and the microwaves are chopped at 50% duty cycle or amplitude modulated with a sine wave, creating periodic dips in the NV fluorescence from the NVs present in the field-free region. The reversal in measurement phase that is synchronous with the microwave chopping is meant to indicate synchronous detection.

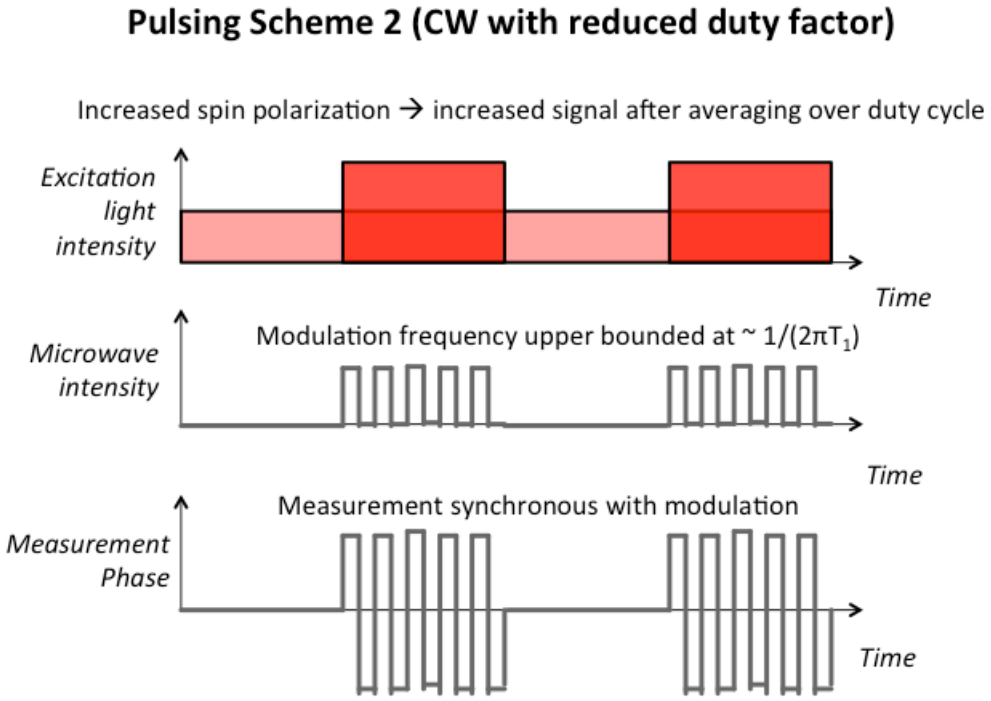

Figure 8. This pulsing scheme is similar to the CW pulsing scheme except it is run at a duty factor less than 100%, allowing for higher optical intensities. For example, reducing the duty cycle to 50% allows the same amount of light to be collected at twice the excitation intensity. Thus the noise for a given measurement is the same, but the signal can be up to twice as great because of enhanced spin polarization at higher optical intensities.

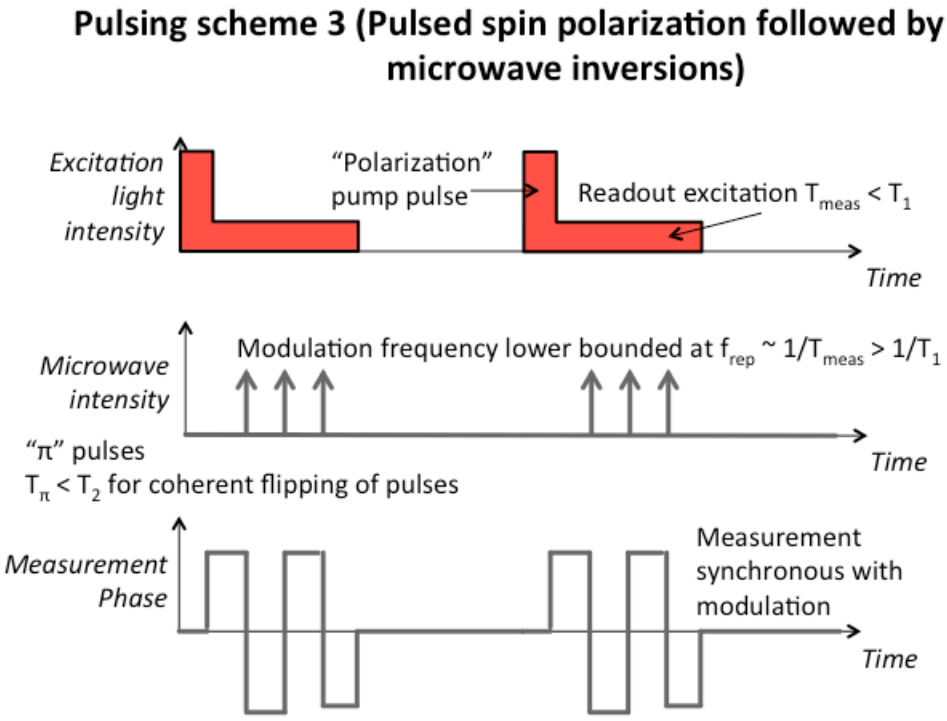

Figure 9. A pulsed scheme that utilizes coherent microwave pulsing in order to achieve a measurement speed-up. In the previous two pulsing techniques, the spins have to relax each time the microwaves are switched on and off. Thus, the speed of the measurement scheme is limited by the spin relaxation time. Here, a dynamic measurement is performed where first the spins are polarized using a brief intense pulse of light, and then the fluorescence is monitored at a lower excitation intensity while the spins are periodically inverted with a series of microwave "π" pulses. Due to the low duty factor of the microwaves, it is possible to have a high enough instantaneous microwave intensity such that the time of a "π" pulse is less than the decoherence time T_2 of the spins, while still maintaining low average microwave power. The excitation light is dimmed during the measurement in order not to repump the spin while it is inverted. Signal can be up to double the previous two techniques, because we are inverting the spin population from the brighter state to the darker states, rather than just equalizing the spin populations.