# **Supporting Information for:**

## Distributions of Conduction Electrons as Manifested in MAS-NMR of Gallium Nitride

James P. Yesinowski,<sup>\*†</sup> Andrew P. Purdy,<sup>†</sup> Huaqiang Wu,<sup>‡</sup> Michael G. Spencer<sup>‡</sup>, Janet Hunting,<sup>¶</sup> and Francis J. DiSalvo<sup>¶</sup>

<sup>†</sup>Chemistry Division, Naval Research Laboratory, Washington DC 20375; <sup>‡</sup>School of Electrical & Computer Engineering, Cornell University, Ithaca NY; <sup>¶</sup>Department of Chemistry and Chemical Biology, Cornell University, Ithaca NY 14853

## Calculation of Probability Density Function for Conduction Band Electron Concentration

The procedure for calculating the GaN:0.13%Ge conduction electron (carrier) probability density function shown in Figure 2(b) from the observed <sup>71</sup>Ga MAS-NMR Knight shift distribution will be discussed in greater detail here. The Knight shift distribution is shown in Fig. 1S, and corresponds to the mirror image of the spectrum in Fig. 2(a) with a new ppm scale. As the Knight shift (defined as the observed chemical



**Figure 1S:** <sup>71</sup>Ga Knight shift spectrum of GaN:0.13%Ge.

shift minus the "reference" shift of 329 ppm) approaches zero, there is an increase in intensity reflecting the small contribution from the minor broadened unperturbed peak at 329 ppm. In order to avoid possible ambiguities in interpreting such small shift differences, and to avoid too large a "spike" at small Knight shifts, we used a cut-off such that the Knight shift distribution only considers Knight shifts  $\geq$  26 ppm. This somewhat arbitrary cut-off eliminates only a small portion of the total integrated intensity of the spectrum and will not significantly affect the subsequent interpretation.

Having done so, we can make use of the theoretical relationship (subject to the assumptions given in ref. [6] of the main paper) that the Knight shift K is proportional to  $c^{1/3}$ , where c is the carrier concentration. To obtain a probability density function plot on a linear x-axis scale, where the y-axis corresponds to the relative probability of finding <sup>71</sup>Ga nuclei in a region having the conduction band electron (carrier) concentration given by the x-axis, we do the following on a point-by-point basis. The x-axis of Knight shift (in ppm) in Fig. 1S is replaced by  $aK^3$ , where the cube of the Knight shift is multiplied by a constant *a* having dimensions cm<sup>-3</sup> that reflects the exact (rather than proportional) numerical relationship between K and c. Since this "calibration" constant a is not obtainable directly from our experimental results, we estimate it by the procedure described in the following paragraph. The calculation of the proper y-axis requires us to consider that the Knight shift spectrum in Fig. 1S, like most NMR spectra, represents signal intensity along the y-axis vs. *uniformly-spaced* "bins" (in ppm units) along the xaxis. However, these points are no longer uniformly spaced when plotted as  $aK^3$ , being sparser at the higher carrier concentrations c. Since  $c = aK^3$ , it follows that dc = $3a(dK)K^2$ ; since  $K = (c/a)^{1/3}$  and  $K^2 = (c/a)^{2/3}$ , it follows that  $dc = 3a^{1/3}(dK)c^{2/3}$ . Thus, if we wish to connect the many unequally-spaced points in carrier concentration space to mimic a continuous smooth probability density function, we must first *divide* the NMR signal intensity by  $c^{2/3}$ .

To estimate the calibration constant *a*, we use the results of Corti et al in ref. [8] of the main paper for a highly-conductive GaN film sample and a resistive Mg:GaN film. The electron carrier concentration in the conductive sample was measured from Raman to be about  $5 \times 10^{19}$  cm<sup>-3</sup>. The <sup>71</sup>Ga  $T_1$  values (equivalent to 1/(2W) in their Figs. 7 and 8) were obtained at 9 T and 1.6 T (no field dependence is expected in the short correlation time limit). For the resistive Mg:GaN sample, relaxing solely by the quadrupolar mechanism, the  $T_1$  at our MAS-NMR temperature of 304 K is read off from Fig. 8 to be 4.3 s. The corresponding  $T_1$  due to the magnetic relaxation from conduction electrons alone for the conductive sample is read off from the extrapolated straight line of Fig. 7 to be 1.15 s. We then calculate the square root of the difference of these two rates, as  $(s^{-1/2})$ . By taking the <sup>71</sup>Ga shown in the Korringa-type plot of our Fig. 1(b), to be 0.93 MAS-NMR  $T_1$  values shown at selected shifts in the spectrum of GaN:0.13% Ge in Fig. 2(a) and plotting them as was done in Fig. 1(b) using a "reference" <sup>71</sup>Ga MAS-NMR  $T_1$ value of 3.51 s as measured for the h-GaN sample grown from a Na/K flux (spectrum in Fig. 1(a)), we obtain a best-fit straight line that enables one to predict, for the y-value of 0.93 (s<sup>-1/2</sup>) calculated above from the Corti et al. film results, a Knight shift of 78 ppm. (Because of the lower resolution of the Corti et al. results, they were not able to measure an accurate Knight shift, and their estimated value appears to be significantly in error visà-vis our results). We can then obtain the calibration constant  $a = 1.06 \times 10^{14}$  as the number that yields the carrier concentration measured by Corti et al. for a Knight shift K = 78 ppm.

The distribution of carrier concentrations shown in Fig. 2(b) of the main paper and expanded in the lower concentration region in Figure 2S is thus a properly-weighted probability density function (pdf). The area under the curve between any two x-values is thus proportional to the number of  $^{71}$ Ga nuclei in the sample (equivalently, since the Ga

atoms have essentially uniform density throughout the sample, to the relative amount of GaN) experiencing that range of carrier concentrations. The normalization of the y-axis is defined to give an integral under the entire pdf curve of unity. The *mean* carrier concentration of  $1.8 \times 10^{20}$  cm<sup>-3</sup> is defined to be that value about which the integrals on either side are equal to  $\frac{1}{2}$ . This also corresponds to the carrier concentration associated with a Knight shift that has a zero first moment in the Knight shift spectrum (i.e., the mean value of the Knight shift spectrum).



**Figure 2S:** Expanded region of probability density function for carrier concentration in GaN:0.13%Ge shown in Figure 2(b) of main paper.

It is clear that this distribution function is skewed towards larger values, although not to as great extent as the Knight shift spectrum itself, because of the intensity scaling required to change the x-coordinates as discussed above. A widely-used *family* of distribution functions for data exhibiting various degrees of this type of skew is the lognormal distribution function, which describes an impressively wide range of physical, financial, and social phenomena.<sup>1</sup> A variable is described as following a lognormal distribution. Such a variable can be supposed to arise from physical processes that involve many (multiplicative) products of random variables, rather than (additive) sums that yield a Gaussian distribution. We will not discuss mechanisms (such as Ge autocatalysis of crystal growth, etc.) that might possibly lead to such a distribution, since detailed mechanisms for lognormal behavior have usually not been provided. Instead, we will treat it as an empirical fitting, which might be suggestive if observed in other comparable situations.

There are a number of different parameterizations used to describe lognormal distributions expressed in the "linear" or "arithmetic" space of the x-axis of Fig. 2(b), and least-squares fitting can be unstable, so it is more straightforward to show the fit by plotting the y-axis vs. the *logarithm* of the carrier concentration, as is done in Fig. 2S. (Note that although base-10 logarithms were used here, natural logarithms are also

commonly used, and only change scaling factors in definitions). Note also that since the interval in this presentation is  $d(\log(c))$ , rather than the evenly-spaced d(c), the interpretation of such a plot is less straightforward than that of the pdf in linear space (Fig. 2(b)), although the characteristics such as mode (maximum), median, and mean have been well-worked out in terms of the equations used to describe lognormal distributions. The experimental data points (black) in Fig. 3S are shown over a range of



**Figure 3S:** Lognormal distribution fitting to a portion of the probability density function for carrier concentration in GaN:0.13% Ge shown in Figure 2(b) of main paper.

calculated carrier concentrations (in cm<sup>-3</sup>) from  $7x10^{18}$  (representing an upper estimate of the Mott transition described in the paper, and corresponding to a Knight shift of 40 ppm) to  $4.6x10^{21}$ . However, the Gaussian fitting was carried out for and plotted over the range of *c* values  $3x10^{19}$  (corresponding to a Knight shift of 66 ppm) to  $4.6x10^{21}$ . The experimental data over this range are seen to follow very well a lognormal distribution. The higher intensity observed on the left for lower concentrations may reflect a true bimodal distribution, or may possibly arise from other complicating factors. However, Fig. 3S is very misleading in terms of the relative amount of material contributing to this region, which is very small, and corresponds to the area of the region to the left of the "dip" in the pdf of Fig. 2S, which is very minor (ca. 4%) compared to the area under the majority of the curve.

It is noteworthy that the highest carrier concentrations in Fig. 2(b) exceed the  $4x10^{20}$  Ge/cm<sup>3</sup> limit seen in MBE-grown GaN films before separate Ge-rich phases were identified.<sup>2</sup> The different conditions of preparation may be responsible, or errors in the absolute calibration of the probability density function using combined NMR and Raman results may result in overestimates of the carrier concentration (but not the shape of the distribution), as indicated by the discussion in the paper. The low carrier concentration side of the distribution in Fig. 2S (to the left of the dip) appears to be close to the Mott transition.

## **Further Experimental Detail**

Trace element wet-chemical analysis for Ge was carried out by Complete Analysis Laboratories, Parsippany, NJ.

The conditions used to obtain the <sup>71</sup>Ga MAS-NMR spectrum of GaN:0.13%Ge shown in Figure 2(a) are also similar to those used for the measurements in Figure 1. A Bruker DMX-500 spectrometer with a <sup>71</sup>Ga Larmor frequency of 152.59 MHz at 11.7 T was used, with a filled 4 mm MAS rotor. The 90° pulse length for the central transition was 3.0  $\mu$ s, and a train of sixty 90° pulses asynchronously spaced at 2.01 ms intervals was used to saturate all transitions, followed by a saturation recovery period and then a rotor-synchronized 90°- $\tau$ -180°- $\tau$ -Acquire Hahn echo to acquire an echo that yielded an undistorted spectrum (flat baseline) within the 1.667 MHz wide spectral width. Fourier transformation starting from the echo peak was carried out after an exponential apodization corresponding to a linebroadening of 100 Hz (0.7 ppm) that minimized the convolution of different parts of the Knight shift spectrum. The spectrum in Fig. 2(a) took 8429 scans, or about 28 hours to acquire, with the recovery delay of 12 s resulting in a fully-relaxed and hence quantitative spectrum.

It is interesting that a peak at the unperturbed h-GaN shift, which may be broadened compared to the best sample (ref. [4] in the main paper) has been seen in many different doped samples (ref. [2b] in the main paper and unpublished results) and presumably arises from non-metallic GaN.

#### **References for Supporting Information**

- 1. J. Aitchison and J.A.C. Brown, "The Lognormal Distribution" Cambridge University Press, Department of Applied Economics Monographs 5, Cambridge, 1976.
- 2. Hageman, P.R.; Schaff, W.J.; Janinski, J.; Liliental-Weber, Z. J. Cryst. Growth 2004, 267, 123.