Crystalline nucleation in undercooled liquids: A Bayesian data-analysis approach for a nonhomogeneous Poisson process

A. Filipponi,[1,*] A. Di Cicco,[2] and E. Principi[3]

[1]*Dipartimento di Scienze Fisiche e Chimiche, Università degli Studi dell'Aquila, I-67100 Coppito, L'Aquila, Italy*

[2]*CNISM, Scuola di Scienze e Tecnologie, Sezione di Fisica, I-62032 Camerino, Macerata, Italy*

[3]*Elettra-Sincrotrone Trieste S.C.p.A., Strada Statale 14, I-34149 Basovizza, Trieste, Italy*

(Received 2 September 2012; published 6 December 2012)

A Bayesian data-analysis approach to data sets of maximum undercooling temperatures recorded in repeated melting-cooling cycles of high-purity samples is proposed. The crystallization phenomenon is described in terms of a nonhomogeneous Poisson process driven by a temperature-dependent sample nucleation rate $J(T)$. The method was extensively tested by computer simulations and applied to real data for undercooled liquid Ge. It proved to be particularly useful in the case of scarce data sets where the usage of binned data would degrade the available experimental information.

DOI: 10.1103/PhysRevE.86.066701 PACS number(s): 05.10.−a, 64.60.Q−, 02.50.−r

I. INTRODUCTION

Crystalline nucleation in metastable undercooled liquids is a fascinating phenomenon involving the ability of the system to explore the entire available phase space and fall into the stable equilibrium state. A basic understanding of the phenomenon is given by the classical nucleation theory (CNT) [1–5], based on reasonable assumptions for the free energy associated with the formation of a crystalline nucleus in the surrounding metastable liquid (involving volume and surface terms estimated from macroscopic thermodynamic quantities) and a kinetic model predicting a metastable stationary distribution of crystalline cluster sizes. From the above assumptions [3] the stationary state nucleation rate $I(T)$ is given by the product of a weakly temperature- (T-) dependent prefactor by an exponential term with a strong T dependence:

$$I(T) \simeq N_{at} k^{+}(n^{*}) Z e^{-\beta \Delta G(n^{*})}, \tag{1}$$

where $\beta = 1/(k_B T)$ with k_B the Boltzmann constant, n^{*} is the number of atoms in critical nuclei, for which a maximum of the Gibbs free energy $\Delta G(n^{*})$ occurs, and $k^{+}(n^{*})$ is the atomic attachment rate to these nuclei. Z is the Zeldovich factor resulting from the saddle point integration over the nuclei size and N_{at} is the number of atoms (or potential nucleation centers for heterogeneous nucleation) in the system. By assuming a linear dependence for the volume Gibbs free energy difference between solid and liquid phases around the melting point T_m and adopting the negentropic model to estimate the interface energy [6–8] $\sigma = \alpha T \Delta S_f / (N_A V_m^2)^{1/3}$ (where $\Delta S_f = \lambda_m / T_m$ is the molar entropy of fusion, N_A the Avogadro number, and V_m the molar volume), the resulting leading T dependence in the exponential term is

$$e^{-\beta \Delta G(n^{*})} \approx \exp\left(-\frac{16\pi}{3} \alpha^3 \, \frac{1}{N_A k_B} \, \frac{\lambda_m}{T_m} \, \frac{T^2}{(T_m - T)^2} \, f(\theta) \right). \tag{2}$$

The dimensionless constant α is associated with the degree of ordering in the liquid in the neighborhood of the interface with the crystal nucleus and depends on the crystal structure. The additional factor $f(\theta) = (2 - 3\cos\theta + \cos^3\theta)/4$ represents the ratio between the volumes of the spherical sector and the entire sphere arising in heterogeneous nucleation [9] on foreign substrates with contact angle θ.

The limits of the theory have been critically assessed [4,5] and in recent times a renewed interest was brought about by computer simulation work [10–14], in particular on tetrahedral liquids [15–17], directly relevant to the present investigation, or other systems [18] where a a metastable fluid-fluid phase transition occurs. A complete understanding of this phenomenon is limited by the absence of any experimental technique able to probe what is actually going on at the microscopic level in the regions of the fluid where nucleation actually takes place. This is because any structural probe of the metastable fluid will be sensitive to average bulk structural properties dominated by uninteresting configurations, while only a negligible fraction of the sample atoms are involved in the successful (overcritical) nucleation process.

Much progress has been performed in relatively recent times with the development of suitable materials-processing techniques including containerless environments based on drop tubes [19–24] or levitation (electromagnetic [25], aerodynamic [26], electrostatic [27–29]) or specific techniques involving supported samples in environments which do not favor heterogeneous nucleation effects [30–34]. With the above methods several metallic and insulating liquids can be brought into a reproducible metastable undercooled state and subject to various measurements of thermophysical or microscopic properties. The undercooled liquid state terminates with a nucleation event leading to macroscopic crystallization and resulting in a sudden temperature increase of the specimen to the equilibrium melting temperature T_m (recalescence). For micrometric to millimetric sized samples, recalescence occurs on the microsecond to millisecond time scale, and the onset of crystallization is marked by a characteristic "flash" (easily observable with a pyrometer for high-melting-point substances). The lowest temperature T reached by the sample prior to recalescence is itself the measurement that can be performed at the lowest possible temperature in an undercooled liquid and contains important information on the nucleation

*Corresponding author: adriano.filipponi@aquila.infn.it

 ©2012 American Physical Society

rate. Owing to the stochastic nature of the phenomenon, T is a random variable depending on the nucleation rate and thermal history of the sample. This paper is actually focused on the understanding of this stochastic phenomenon and on the possible data-analysis methods. The subject is not new and several investigations of maximum undercooling temperature data sets have been performed [31–38], developing and exploiting the basic formalism introduced by Skripov [39]. In the present paper a Bayesian approach that turns out to be particularly appropriate for the case of scarce data sets is proposed and applied to a Ge recalescence experiment.

The paper is organized as follows: The main assumptions behind the description of the crystallization events in terms of a nonhomogeneous Poisson process will be reviewed in Sec. II. The data-analysis inversion problem is addressed in Sec. III. The Bayesian approach to the data analysis is illustrated in Sec. IV. Finally in Sec. V we present a preliminary application of the method to real data sets.

II. NUCLEATION STATISTICS

We will assume that the crystalline nucleation process for a given sample is described by a sample nucleation rate $J(T)$ dependent on the instantaneous temperature T. In the ideal case $J(T)$ corresponds to the lower possible limit associated with homogeneous nucleation processes and results proportional to the sample mass m: $J(T) = mI_h(T)$, where $I_h(T)$ is the homogeneous nucleation rate per unit mass. In real cases $J(T)$ will be typically dominated by heterogeneous nucleation processes driven by impurity effects and will still scale with m for given impurity concentrations. The assumption of the existence of a $J(T)$ implies that the temperature changes are slow compared to the time needed for the sample to reach a metastable stationary state, which is quite reasonable. The sample thermal history $T(t)$ will be considered a monotonically decreasing function of time [$\dot{T}(t) < 0$] going from a maximum heating temperature $T_{\max} > T_m$, through T_m (the specimen melting point) at $t = t_0$, and decreasing towards the undercooling regime $T < T_m$. The probability $P_l(t)$ that the sample remains liquid at time t is described by a nonhomogeneous Poisson process since the rate is dependent on time through the parameter T and is given by

$$P_l(t) = \exp\left(-\int_{t_0}^{t} J(T(t'))dt'\right), \tag{3}$$

where the integral in the exponential assumes stochastic independence among the nucleation probability in successive infinitesimal time slices dt' and runs from t_0: the time from which $J > 0$. This probability can be conveniently expressed as a function of T using the thermal history $T(t)$ as a substitution variable in the integral [$T(t_0) = T_m$, $T(t) = T$, and $dT = \dot{T}dt' = -(-\dot{T})dt'$]. Then the probability that the sample remains liquid when it has reached temperature $T < T_m$ is

$$P_l(T) = \exp\left[-\int_{T}^{T_m} \frac{J(T')}{C_r(T')}dT'\right], \tag{4}$$

where the minus sign was explicitly introduced to make $-\dot{T}(t(T')) = C_r(T') > 0$ (i.e., the positive sample cooling rate as a function of the dummy integration variable T') and the integration boundaries accordingly exchanged to express $P_l(T)$ as the negative exponential of the forward integration of a positive quantity. The sample nucleation rate vanishes, $J(T) = 0$, for $T \geqslant T_m$ and is expected to display a monotonic steep increase when the temperature T decreases below T_m. For this reason, upon decreasing T this probability [starting at $P_l(T_m) = 1$] will initially remain $P_l(T \lesssim T_m) \simeq 1$ but, around some typical maximum undecooling temperature T_u, will display a rapid drop to 0. So the maximum undercooling temperature in repeated cooling processes with rate $C_r(T) > 0$, obtained by any equivalent thermal history $T(t)$ with arbitrary time origin, will be a random variable T described by the cumulative distribution function $P_l(T)$. The probability density of this random variable is obtained from

$$\begin{aligned} g(T) &= \frac{dP_l(T)}{dT} = \frac{J(T)}{C_r(T)}P_l(T) \\ &= \frac{J(T)}{C_r(T)}\exp\left[-\int_{T}^{T_m}\frac{J(T')}{C_r(T')}dT'\right], \end{aligned} \tag{5}$$

and is expected to be a narrow distribution peaked around a typical undercooling temperature T_u.

Equations (4) and (5) are the solution for the problem of finding the distribution of the random variable maximum undercooling temperature for a given sample nucleation rate $J(T)$ and cooling rate $C_r(T)$. This distribution will depend on the sample mass m (and impurity concentration) and $C_r(T)$ and, provided a functional expression for $J(T)$ is given, can be calculated analytically or numerically. It is also possible to generate a sample of independent random numbers distributed according to Eq. (5) by extracting a uniform random number x in the interval $[0,1)$ and solving for T the expression $x = P_l(T)$, i.e., inverting Eq. (4).

III. DATA ANALYSIS AND INVERSION

The theoretical treatment of the nonhomogeneous Poisson process illustrated in Sec. II provides a solution for the direct problem of finding the distribution of the undercooling temperature random variable. The experimental information that can be obtained from repeated melting and undercooling experiments is a set $\{T_i\}$ of temperature values for this random variable. The objective of any data-analysis approach is to derive $J(T)$, or any related information, from this data set, corresponding to the inverse problem to that solved in Sec. II. Owing to the stochastic nature of the phenomenon a statistical analysis is required.

An obvious approach consists in realizing that $J(T)$ appears in the integrand of Eq. (4) with the T variable in the lower boundary of the integral. By differentiating the natural logarithm of $P_l(T)$ with respect to T, the integrand function at T is directly obtained, that is,

$$J(T) = \frac{d\ln[P_l(T)]}{dT}C_r(T) = \frac{1}{P_l}\frac{dP_l}{dT}C_r(T). \tag{6}$$

Equivalently this result is obtained by inverting Eq. (5). The application of the above expression requires knowledge of the random variable distribution and in particular of the ratio between probability density and cumulative distributions involved in the logarithmic derivative. Clearly the experiment

will provide a statistical approximation to the above functions represented by suitable histograms of the data set, and this will introduce statistical uncertainties. It is also clear that the set of recalescence temperatures will be confined within a narrow temperature interval and the retrievable information on $J(T)$ will be limited to this interval with the highest information content in the temperature region of maximum data density. This approach has been previously exploited by Perepezko and co-workers in various investigations [31–33] and independently in a previous work from our group [34]. Expressions quoted in Refs. [31–33] can be recognized as equivalent implementations of Eq. (6).

This approach, which is certainly known and exploited in different scientific contexts, as originally quoted, has the great advantage of being model independent in the sense that no assumptions are made on $J(T)$ and the resulting functional shape with uncertainty bounds comes directly from the application of Eq. (6) and related expressions. On the other hand, this direct inversion formula suffers from practical limitations associated with the statistical uncertainty. As a matter of fact, obtaining a numerous and homogeneous set of recalescence temperatures for a given sample is not an easy job, whatever experimental technique is used. Techniques using supported samples are more affected by possible changes in the sample contamination [31–34], and the succession of maximum undercooling temperatures has to be critically inspected in order to extract a subset of successive values obtained under comparable sample conditions. Typical sets recorded with levitation techniques [37,38] are also limited to a few hundred repetitions. Drop tube experiments [19–21,24] performed on high-purity specimens typically allow a single trial per sample. So very often the only pieces of available information on the undercooling ability of a well-characterized specimen are scarce data sets, and the statistical analysis embodied in Eq. (6) requiring data binning has some practical limitations. Lastly a further drawback of this approach is that the resulting estimate for $J(T)$ is given by a sequence of values with uncertainty, but no constraint is introduced to account for the expected smooth functional dependence of $J(T)$.

IV. A BAYESIAN APPROACH

In the attempt to provide an alternative data-analysis scheme for these experiments we explored the possibility of adopting a Bayesian statistical approach [40] following current trends in many scientific fields. The way this approach was implemented starts from the observation that the narrow distribution of maximum undercooling temperatures is able to provide information on $J(T)$ only in a limited temperature range. So it is possible to find approximations for $J(T)$ in the temperature range of interest (depending only on a few parameters, indicated as A,B in the following) that can be used to predict the resulting distributions.

In the Bayesian approach the probability density of the model parameters, with the condition that a given data set of undercooling temperatures $\{T_i\}$ is obtained, is given by

$$f(A,B|\{T_i\}) = \frac{\phi(\{T_i\}|A,B)\, p(A,B)}{\int dA\, dB\, \phi(\{T_i\}|A,B)\, p(A,B)}, \qquad (7)$$

where $p(A,B)$ represents the prior probability distribution for the parameters and $\phi(\{T_i\}|A,B)$ is the likelihood function representing the probability density for the data set $\{T_i\}$ given the values for the model parameters. The denominator accounts for the correct normalization of the distribution. In the absence of any prior information on the parameters, $p(A,B)$ can be regarded as uniform in the allowed parameter range, and in practice omitted in Eq. (7) apart from the resulting integration boundaries in the denominator. In the case of stochastically independent undercooling temperatures T_i, which is reasonable if they are obtained in subsequent melting-undercooling cycles under stable sample conditions, the likelihood function is given by the product of the probability densities calculated for the various observed temperatures. Therefore the resulting probability density in the parameter space is proportional to the likelihood function regarded as a function of the model parameters:

$$f(A,B|\{T_i\}) \propto \phi(\{T_i\}|A,B) = \prod_i g(T_i|A,B). \qquad (8)$$

In Eq. (8) the normalization denominator was omitted for simplicity. The probability densities $g(T_i|A,B)$ are given by Eq. (5), once $J(T)$ has been suitably parametrized. It should be emphasized here that within a given functional model the undercooling temperature distribution $g(T|A,B)$ is insensitive to the specific choice of the parameters (A,B), and various functional combinations giving the same $J(T)$ can be used. However, when these expressions are evaluated as a function of the parameters (A,B), such as in Eq. (8), this choice matters as well as the assumption of a uniform prior which is strictly associated with the parameter definition. So the parameter choice and the assumptions on their prior distribution are important elements of the Bayesian approach contributing to the shape of the distribution.

Suitable parametrizations for $J(T)$ are certainly those suggested by the CNT and they can be used to calculate Eq. (5) using a numerical approach.

A. Empirical models

For practical purposes it is convenient to use empirical expressions leading to analytically solvable integrals. The simplest expression is

$$J(T) = Ae^{-B(T-T_0)}, \qquad (9)$$

where T_0 is a reference undercooling temperature and $A = J(T_0)$ and $B = -\frac{d\ln J(T)}{dT}|_{T_0}$. Equation (9) is not intended to represent $J(T)$ in a wide temperature interval but only to model its functional shape in a narrow interval in the spirit of a Taylor expansion of its logarithm around T_0. The rationale is that typical undercooling temperatures are expected to occur in a narrow interval where the information on $J(T)$ can be obtained. In the same spirit $C_r(T)$ will be approximated by its value at T_0, $C_r = C_r(T_0)$. With this trivial model we have

$$P_l(T) = \exp\left(-\frac{A}{BC_r}[e^{-B(T-T_0)} - e^{-B(T_m-T_0)}]\right), \qquad (10)$$

and the resulting probability density is

$$g(T|A,B) = \frac{A}{C_r} e^{-B(T-T_0)} P_l(T) = \frac{A e^{-B(T-T_0)}}{C_r} \exp\left(-\frac{A}{BC_r}[e^{-B(T-T_0)} - e^{-B(T_m-T_0)}]\right). \quad (11)$$

A sample of random temperatures distributed according to Eq. (11) can be obtained from uniformly distributed pseudo-random numbers from computer routines $x \in U[0:1)$ using the inverse of Eq. (10):

$$T_r = T_0 - \frac{1}{B} \ln\left(-\frac{BC_r}{A} \ln x + e^{-B(T_m-T_0)}\right). \quad (12)$$

In all the above expressions the term $e^{-B(T_m-T_0)}$ accounts for the fact that in Eq. (9) $J(T)$ does not strictly vanish at T_m. For realistic parameter values it gives a negligible contribution.

A more realistic model for $J(T)$ should contain an exponential temperature dependence in agreement with the CNT prediction. A very general elementary integrable expression is

$$J(T) = A f'(T) e^{-Bf(T)}. \quad (13)$$

In the framework of the negentropic model of the interface energy [6–8], we assume

$$f(T) = \phi\left(\frac{T^2}{(T_m - T)^2} - \psi\right), \quad f'(T) = \frac{2\phi T T_m}{(T_m - T)^3}, \quad (14)$$

where ϕ and ψ are constants that can be conveniently determined to maintain for A and B the same meaning as in model (9) with respect to the reference temperature T_0. We find

$$\phi = \frac{(T_m - T_0)^3}{2T_0 T_m}\left(1 + \frac{1}{B}\frac{T_m + 2T_0}{T_0(T_m - T_0)}\right) \quad (15)$$

and

$$\psi = \frac{T_0^2}{(T_m - T_0)^2} - \frac{1}{\phi B} \ln f'(T_0). \quad (16)$$

An elementary integration yields

$$P_l(T) = \exp\left(-\frac{A}{BC_r} e^{-Bf(T)}\right) \quad (17)$$

and correspondingly

$$g(T|A,B) = \frac{A}{C_r} f'(T) e^{-Bf(T)} \exp\left(-\frac{A}{BC_r} e^{-Bf(T)}\right). \quad (18)$$

The random sample of recalescence temperatures can be obtained with the transformation

$$T_r = \frac{T_m}{1 + \left[\psi - \frac{1}{\phi B} \ln\left(-\frac{BC_r}{A} \ln x\right)\right]^{-1/2}}. \quad (19)$$

The functional factor $f'(T)$ in Eq. (13) has the only purpose of allowing for an elementary integration in Eq. (4); its contribution to the overall shape of $J(T)$ is small because it is not in the exponential term. A similar idea was already adopted in previous investigations [35–38] where an exact CNT expression was approximately integrated assuming a small prefactor contribution; here an exact integration is performed on an approximate expression. Clearly, further investigation may suggest the adoption of improved functional models.

The two models here introduced are useful to illustrate the method and will be proven to be adequate. The models appear to depend on A, B, and T_0, but only two parameters are actually independent since the same functional form can be obtained using a different reference temperature T_0 by suitably changing the values of A and B. This is obvious for model (9) since a T_0 translation simply involves a redefinition of A (B is constant), but a similar argument applies also to Eq. (13). Namely, with ϕB and $\phi A \exp(\phi\psi B)$ constants the same T dependence in both preexponential and exponential terms is obtained.

B. Numerical simulations

In order to assess the validity of the data-analysis approach and to compare models (9) and (13) we performed a numerical simulation. We used parameters roughly appropriate for the case of heterogeneous nucleation of Ge, using $T_m = 1200$ K, $A = 1$ Hz, and $B = 0.1$ K^{-1}, at $T_0 = 1000$ K and $C_r = 10$ K/s. Random samples of undercooling temperatures were generated using (12) and (19) and the related histograms are compared with the predicted distributions in Fig. 1. The undercooling temperature distribution shows the characteristic asymmetric shape due to the exponential increase of $J(T)$ at lower temperature. The two models yield slightly different distributions, with Eq. (13) resulting in the narrower $g(T)$ peak owing to the greater nucleation rate variation. The histograms were obtained for samples of 100 000 simulated undercooling temperatures to show the close agreement with the predicted distributions. For realistic data sets of a few hundreds of undercooling temperatures the two models give less distinguishable results within the statistical uncertainty, especially for larger and more realistic B values, leading to narrower distributions. The model functions $J(T)$ and the resulting estimated values from Eq. (6) are compared in Fig. 2. Model (9) is tangent to Eq. (13) at T_0 and can be regarded as its first-order Taylor expansion, which is expected to be sufficiently accurate in a narrow temperature interval around T_0. Model (13) can be regarded as a further approximation to $\ln[J(T)]$ with a curvature specified according to the CNT predictions and the correct limiting behavior for $T \to T_m$.

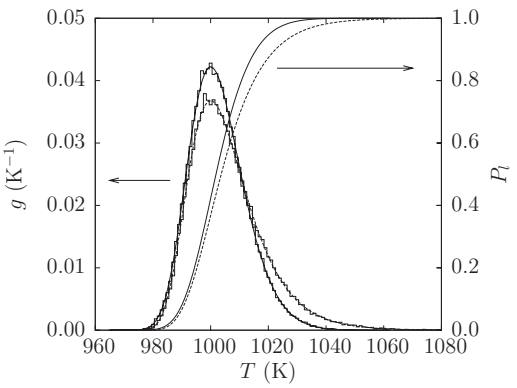

FIG. 1. Simulated undercooling temperature distributions using model (9), dashed line, and (13), solid line, with the same parameters $T_m = 1200$ K, $A = 1$ Hz, and $B = 0.1$ K^{-1}, at $T_0 = 1000$ K and $C_r = 10$ K/s.

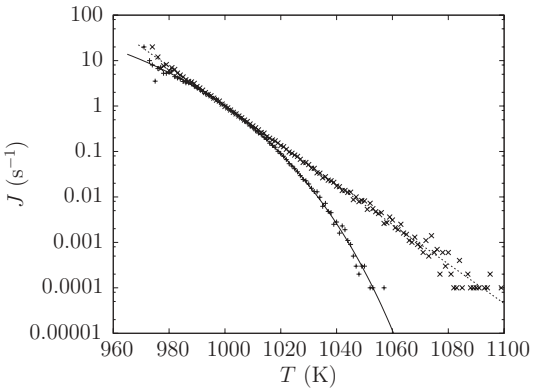

FIG. 2. Sample nucleation rate $J(T)$ on a semilogarithmic plot associated with models (9), dashed straight line, and (13), solid curve. The noisy data points refer to the application of the inversion formula (6) to the data sets of $N = 100\,000$ undercooling temperatures binned every 1 K.

Samples with 100 000 data points allow the model functional shapes to be retrieved; obviously this will not be the case for scarce data sets.

C. Likelihood function simulations

In order to implement the Bayesian data-analysis approach the probability density in parameter space, which is proportional to the likelihood function in the case of a uniform prior, has to be evaluated according to Eq. (8) using the product of the appropriate number of functions of the type (18) associated with the undercooling temperature data set. In this analysis the reference temperature T_0 is unknown and will be indicated as the parameter T_u, representing a typical undercooling reference temperature around which the information on $J(T)$ can be retrieved. The actual choice for the parameters and their prior distribution becomes important here. We decided to use $\ln(A)$ instead of A as the parameter since $\ln(A)$ is in the exponent of $J(T)$, similarly to B, and this choice results in more symmetrical distributions. Owing to the nonlinear relationship, a uniform prior with constant $u(\ln(A), B) = u$ in $\ln(A)$ does not correspond to a uniform prior in A, since $p(A,B) \propto A^{-1}$ because $u\, d(\ln A) = p(A)dA$. In a similar manner the numerator of Eq. (7) transforms as

$$\phi(\{T_i\}|A,B)A^{-1}dA = \phi(\{T_i\}|e^{\ln A},B)d(\ln A). \qquad (20)$$

As a consequence, to represent the likelihood function it is possible to use a product of Eqs. (11) or (18), maintaining the A parameter with the dimensions of a rate, but plotting them on a $\ln A$ scale. This corresponds to the substitution $A = \exp(\ln A)$, leading to the correct probability density accounting for the assumed prior distribution. Any possible assumption on the prior distribution will eventually become less important for large data sets. For scarce data sets the Bayesian approach forces one to define any assumption on the prior.

In order to understand the principle of application of this method it is instructive to visualize the likelihood function (8) associated with simulated samples of undercooling temperatures. The probability density in the two-parameter $(\ln A, B)$ space can be easily plotted in three dimensions (3D). For this purpose we adopted model (13) with $C_r = 10$ K/s, $T_m =$ 1200 K, $A = 1$ Hz, and $B = 0.1$ K^{-1}, at $T_0 = 1000$ K, and simulated two data sets of $N = 20$ and $N = 100$ undercooling temperatures. The 3D shape of the $f(\ln A, B|\{T_i\})$ distributions (apart from normalization) and the effect of the choice of T_u in the data analysis are illustrated in Fig. 3 reporting two $\prod_i g(T_i|A,B)$ functions for $N = 20$ for slightly different T_u. The likelihood functions are peaked around the most probable $(\ln A, B)$ couple for the given T_u. The $\ln(A)$ and B random variables are not stochastically independent: $f(\ln A, B|\{T_i\})$ cannot be written as the product of two independent functions of $\ln A$ and B, and in general is an asymmetric peak. Confidence intervals for the $(\ln A, B)$ couple can be in principle obtained using a suitable contour $f(\ln A, B|\{T_i\}) = \text{const}$, which encloses the desired percentage (confidence level) of the corresponding probability. This is, however, not easy for the presence of asymmetry and the normalization requirement. A Gaussian approximation for the peak shape, however, becomes progressively more accurate as N increases and may be adopted under certain conditions. For small N finding the correct normalization for $f(\ln A, B|\{T_i\})$ is anyway a difficult task.

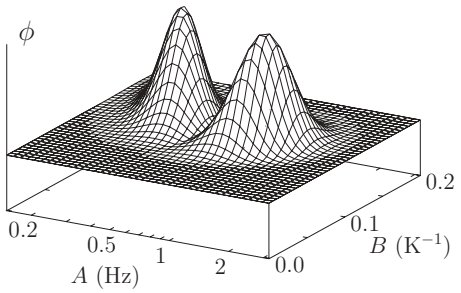

FIG. 3. Likelihood functions in the $(\ln A, B)$ parameter space associated with a simulated data set of 20 undercooling temperatures (with $T_m = 1200$ K, $A = 1$ Hz, and $B = 0.1$ K^{-1}, at $T_0 = 1000$ K) for $T_u = 998$ K (surface on the right side) and $T_u = 1008$ K (surface on the left side). The probability densities are not normalized.

In general the $\ln A$ and B random variables are also correlated, but this correlation depends on the choice of T_u. This effect is clearly visible in Fig. 4, reporting a sequence of contour plots for two data sets with $N = 20$ and $N = 100$ and the same sequence of T_u values. The change in the tilt of the main axes of the contour plots in going from "high" T_u (negative correlation) to "low" T_u (positive correlation) is evident. This is explained by the fact that interception of a given data set from lower temperatures can be equivalently obtained with larger or smaller $\ln A$ and simultaneously larger or smaller negative slopes. The opposite occurs when the reference temperature is greater then those in the data set. From the above considerations it appears quite natural to choose the best T_u value for a given data set in such a way that the correlation among the $\ln A$ and B parameters is negligible. This will allow one to express the confidence interval in the simplest way as a couple of optimal values with error bars (and no correlation coefficient). The parameter T_u therefore assumes the meaning of a reference temperature where the data set $\{T_i\}$ allows one to determine independently the coefficients of the expansion for the logarithm of the sample nucleation rate, $\ln A$ and B. The comparison between the $N = 20$ and

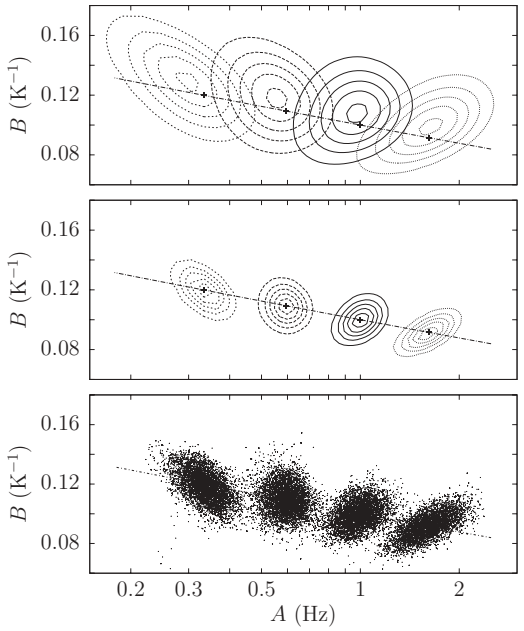

FIG. 4. Contour plots of the likelihood functions in the $(\ln A, B)$ parameter space associated with a simulated data set of 20 undercooling temperatures (top panel) and 100 undercooling temperatures (middle panel). The four sets of contours correspond to $T_u = 995, 1000, 1005, 1010$ K from right to left. The bottom panel reports a scatter plot of the Metropolis Monte Carlo simulation on the $N = 100$ data set; see Sec. IV D.

$N = 100$ data sets in Fig. 4 is also useful to appreciate the corresponding reduction of uncertainty in parameter space. Equivalent contours appear to be shrunk by a factor $\sqrt{5}$ as expected from elementary statistical arguments.

D. Metropolis Monte Carlo sampling

The graphical method illustrated in the previous section is not particularly useful for the application of the Bayesian approach to real (or simulated) data sets to solve both the problems of normalizing $f(\ln A, B|\{T_i\})$ and finding the appropriate T_u to obtain uncorrelated variables. Both these aspects can be elegantly tackled using a numerical approach based on the Metropolis Monte Carlo method (MMC) [41]. In this approach a driven random walk in the $(\ln A, B)$ parameter space is generated from a starting point $(\ln A_0, B_0)$ by tossing trial uniform random displacements in suitable ranges $\pm\delta \ln A$ and $\pm\delta B$, refusing the move, i.e., maintaining the previous position, with probability $1 - p$ if

$$p = \frac{f(\ln A_{\rm trial}, B_{\rm trial}|\{T_i\})}{f(\ln A_{\rm old}, B_{\rm old}|\{T_i\})} < 1, \qquad (21)$$

and accepting the move in all other cases. The random walk will first approach the region of maximum $f(\ln A, B|\{T_i\})$ (equilibration), thus performing a parameter fitting, and will subsequently be trapped around the likelihood function peak, simulating the allowed statistical fluctuations in parameter space. The theory of this Markov process assures that after equilibration the random walk sequence provides an unbiased sampling of the $f(\ln A, B|\{T_i\})$ probability density. The absence of normalization requirements for f in the acceptance rule makes this method efficiently applicable in all those cases where a probability density normalization is difficult to calculate. The resulting $f(\ln A, B|\{T_i\})$ is obtained by a statistical analysis of the random walk (after equilibration) and is approximated by the corresponding histogram, suitably normalized. The application of this MMC method to such a simple two-parameter space is a trivial computational task. The maximum of the likelihood function can be reached rapidly from any starting point and the process can run for several millions of steps to achieve the required statistical accuracy. The $C_{\ln A,B} = \langle(\ln A - \ln\bar{A})(B - \bar{B})\rangle$ correlation can be computed. If this parameter is not sufficiently small T_u can be changed and another run performed till the uncorrelated parameter conditions is satisfied.

To illustrate this method a MMC random walk was generated by changing T_u every 5000 steps. The corresponding couples of $(\ln A, B)$ values obtained for the $N = 100$ simulated data set are reported in a scatter plot representation in the bottom panel of Fig. 4. It can be appreciated that the scatter plot, after equilibration, reproduces the shapes of the distributions of the corresponding contours of the likelihood function reported in the middle panel of the same figure. The dependence of the correlation $C_{\ln A,B}$ on T_u was investigated in an extended run performed scanning T_u every 1 K using $N_e = 500$ equilibration steps and acquiring statistics for the successive $N_s = 50\,000$ steps. The results are reported in Fig. 5, showing the T_u dependence of the dimensionless correlation coefficient $\rho_{\ln A,B} = C_{\ln A,B}/(\sigma_{\ln A}\sigma_B)$, normalized to the standard deviations of the respective variables $\sigma_{\ln A}$ and σ_B. In the same figure the simulated data set is reported as marks at the corresponding temperatures in the T_u scale on the $\rho = 0$ line. The large variation of $\rho_{\ln A,B}$ obtained on scanning T_u through the temperature range of the data set and the principle for choosing the optimal T_u for $\rho_{\ln A,B} \approx 0$ can be appreciated. This optimal value is close to but does not coincide with the average undercooling temperature, and the numerical simulation approach is an effective means to optimize this choice. In a standard data-analysis scheme the numerical procedure is obviously implemented with a more efficient zero-search algorithm starting from the average $\{T_i\}$ and there is no need to scan T_u as was done here for illustrative purposes. After the first trial run the displacement ranges can

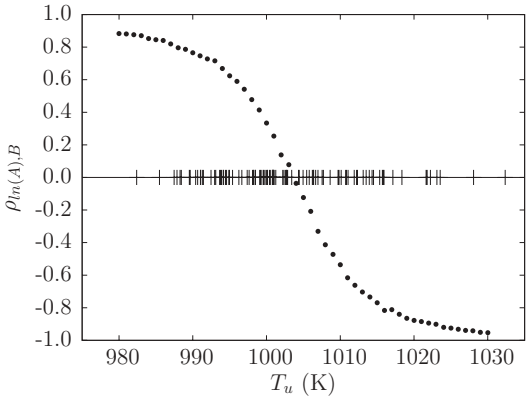

FIG. 5. Dimensionless correlation coefficient $\rho_{\ln A,B}$ as a function of T_u for the simulated data set of 100 undercooling temperatures marked as thin vertical segments on the $\rho = 0$ line.

be conveniently set to $\delta \ln A = \xi \sigma_{\ln A}$ and $\delta B = \xi \sigma_B$ with ξ (a dimensionless constant) chosen to maximize the diffusion in the parameter space. We adopted $\xi = 1.5$, achieving an acceptance rate around 57%. The statistical uncertainty in $\rho_{\ln A,B}$ depends on the number of MMC steps, and in the case of $|\rho_{\ln A,B}| \lesssim 0.2$ the standard deviation in successive equivalent runs is found to be about $\sqrt{\nu/N_s}$, where the empirical factor ν (found to be $\approx$6.7) accounts for the correlated nature of the random walk used to evaluate the statistics. The apparent noise in Fig. 5 is due to this effect for $N_s = 50\,000$. In practice by using $N_s = 100\,000$ $\rho_{\ln A,B}$ is determined with a statistical accuracy in the 0.01 range. In these conditions the numerical statistical accuracy for the (practically uncorrelated) $\ln A$ and B parameters is much smaller than their uncertainty represented by $\sigma_{\ln A}$ and σ_B.

Because of the two-dimensional nature of the parameter space the joint confidence intervals for a given confidence level c are represented by suitable contours of the probability density. In the assumption of an uncorrelated two-dimensional Gaussian distribution the confidence interval sections are extended by $\eta = \sqrt{-2\ln(1-c)}$ times the standard deviations of the corresponding variables, and the corresponding contour is located at $\exp(-\eta^2/2) = 1 - c$ of the value at the maximum. Thus the joint confidence interval drawn at 1σ contains only $c \simeq 39.3\%$ of the probability. In order to include 95% of the probability ($c = 0.95$), $\eta \simeq 2.448$.

E. Methods and models performances

In order to understand the advantage of the Bayesian approach over the direct binning approach, both data-analysis methods were applied to the $N = 100$ simulated data set. The results are both compatible with the adopted $J(T)$ model and are compared in Fig. 6. The binning method was implemented by subdividing the sample into ten temperature intervals containing ten data points each. There is of course a certain arbitrariness in this choice that affects the results in conjunction with the sample statistical fluctuations. The Bayesian approach instead exploits the assumption of the existence of an analytical $J(T)$ function to explain the entire data set and the limited number of estimated parameters is reflected in an overall smaller statistical uncertainty. The binning approach becomes progressively less applicable or useful for less numerous samples. The Bayesian analysis is instead in principle applicable even if only a single maximum undercooling temperature is available. On the other hand it is clear that the main limitation of the Bayesian method is the requirement for a model assumption. It should be pointed out here that it is not necessary to assume the validity of the CNT or a specific $J(T)$ model for an extended T range. It is sufficient to assume that $\ln[J(T)]$ is a continuous and derivable function in the T range of interest. Clearly, if $J(T)$ is expected to display an extremely discontinuous T dependence the models considered in this paper require improvement, but this is very unlikely.

An indication of the effects and limitations associated with model assumptions can be obtained by investigating the ability of model (9) to approximate data simulated according to model (13). The contour plots of the likelihood function for model (9) are compared with the equivalent quantities for model (13) in Fig. 7 for the usual $N = 100$ data set. The best B value

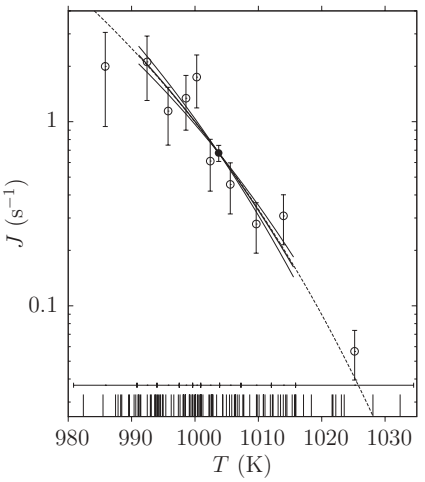

FIG. 6. Comparison of the performance of the Bayesian approach with the binning method. The adopted $J(T)$ model is reported as a short-dashed curve. The $N = 100$ simulated data set is marked with thin vertical lines at the bottom. The horizontal line with small ticks indicates the adopted binning. The ○ data points with error bars, reported at the central bin T, refer to the statistical evaluation according to Eq. (6) and [34]. The Bayesian analysis provides an estimate of A (• with error bar) and B at T_u, determining the shape of $J(T)$ according to model (13) (thick solid curve) with an uncertainty fan associated with $\pm\sigma_B$ (thin solid curves), traced over the region of the central 90th percentile of the simulated data. All errors refer to one standard deviation.

for model (9) is roughly constant with T_u at variance with the case of model (13). This is explained by the fact that in this functional model a change in T_0 does not affect B but only A. The different slopes between the sequences of centroids should not concern the reader. The real comparison should be made on the central profiles corresponding to $\rho \approx 0$. In this case model (9) leads to a B value overestimated by 12% and a comparable A value with respect to model (13), used to simulate the data. The effect is explained by the extension of the undercooling data into a region of a steep decrease of $J(T)$, as shown in Fig. 2, that model (9) attempts to fit with a slightly larger

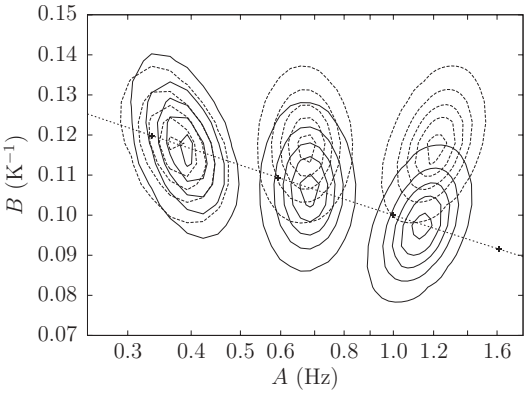

FIG. 7. Contour plots of the likelihood function drawn at levels 0.1, 0.3, 0.5, 0.7, and 0.9 of its maximum for model (13), solid curves, and (9), dashed curves, for the optimal T_u (central), $T_u + 5$ K (left), and $T_u - 5$ K (right). The outer curves include approximately (in the Gaussian approximation) 90% of the probability.

B value. For these $N = 100$ samples the difference between the two models is comparable with the parameter uncertainty. In other words, the systematic error made by assuming the "wrong" model (9) is of the order of the statistical uncertainty. In the case of less numerous or narrower $\{T_i\}$ distributions the performances of the two models are comparable. In real applications the issue of the systematic uncertainty associated with the assumption of an incorrect empirical model is clearly important and not easy to assess. It can be argued, however, that by assuming model (13), which accounts for the correct curvature and asymptotic behavior of the exponential term predicted by the CNT, the systematic errors are anyway smaller than the differences with the analysis based on model (9), with zero curvature, and generally will be smaller than the statistical uncertainty.

V. APPLICATIONS TO UNDERCOOLED LIQUID Ge

An extended experimental program has been undertaken over recent years to investigate the nucleation properties of supported Ge samples. The interest in elemental Ge stems from the existing competition between covalent and metallic bonding possibly leading to a liquid-liquid phase transition line between high- and low-density liquids (HDL-LDL) in the undercooled range, with potential effects on the crystalline nucleation barrier, as recently highlighted in a computer simulation work on silicon [17]. The specific motivation of this research was to reproduce on millimeter sized Ge droplets equivalent experimental conditions to those encountered in supported micrometer sized powder Ge samples dispersed in boron nitride (BN) probed by x-ray absorption temperature scans [30], which are suitable to reach a high degree of undercooling (of even 300 K). For this purpose we used a high-vacuum chamber, a BN crucible placed on a resistive heater, and a pyrometric temperature probe coupled to a PC-based acquisition system as previously described [34]. Over 20 zone-refined 99.9999% Ge samples in a mass range $1 \lesssim m \lesssim 500$ mg were processed by carefully melting the samples, avoiding overheating above the evaporation threshold, and allowing the chamber to outgas for about 1 h to recover high-vacuum conditions under continuous turbomolecular pumping. The sample was successively subjected to repeated melting-cooling cycles (typically heating for 30 s and switching off the heater for 20 s) with continuous monitoring at 2000 Hz rate of the surface temperature of the sample. A number of recalescence events from a few tens to a few hundreds were recorded for each sample, according to the sample performance. Examples of sequences of maximum undercooling temperatures obtained in successive melting-cooling cycles for two representative Ge samples of different masses are reported in Fig. 8. The cycle history of the sample performance is conceptually equivalent to similar results reported by other groups [32,33]. Here the contamination effects are severe and the samples often display changes in the degree of possible undercooling. In particular, in many cases the sample is observed to improve the undercooling ability after the first few melting cycles. We believe that this is due to the progressive removal of the native oxide of Ge which decomposes and is pumped out under high-temperature high-vacuum conditions. This process occurs in the first sample

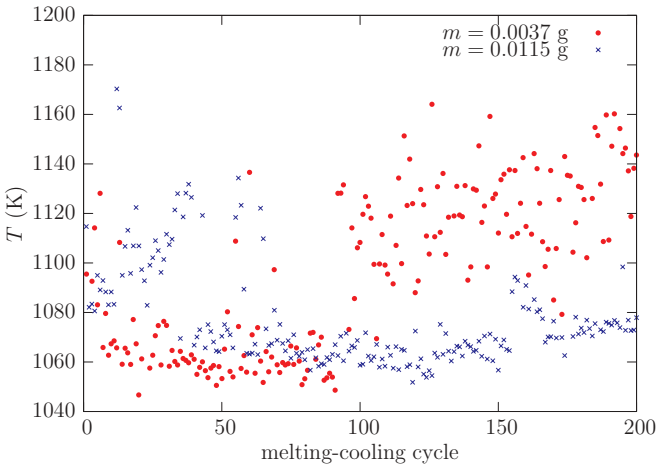

FIG. 8. (Color online) Sequence of maximum undercooling temperatures obtained in successive melting-cooling cycles for two representative Ge samples.

melting cycle and normally continues in the next cycles. We usually observed that after this first stage the sample reaches stable undercooling capabilities, indicating the achievement of the best stationary purity conditions we could obtain. After a few tens to a few hundred cycles the undercooling sample properties are subject to degradation in an irreversible manner. We attribute this effect to an irreversible sample contamination from heater and/or crucible sources.

In order to extract a sequence of undercooling temperatures associated with stationary contamination conditions, the performance histories (similar to Fig. 8) have to be carefully inspected to find homogeneous contiguous subsequences. Although strictly valid for Gaussian distributions, the statistical t test for consistency between the mean values of two subsamples can be applied to determine the maximum acceptable extension of homogeneous subsequences. Obviously the effect of adding or eliminating data points on a single subsequence can be directly verified. Subsequences associated with both cleaner and contaminated conditions can often be extracted from a single sample run. Each subsequence can be processed with the Bayesian data-analysis method proposed in this paper and the corresponding sample nucleation rate can be determined. The list of samples considered in the present investigation and the properties of the corresponding subsequences are reported in Table I. The sequences of maximum undercooling temperatures are shown in Fig. 9.

The result of the Bayesian analysis for the sample nucleation rates $J(T)$ is reported in Fig. 10. It should be emphasized that the graphical representation of the uncertainty in both $\ln(A)$ and B refers to one standard deviation only. For a 95% confidence level the confidence interval should be extended to about 2.45σ. The results are spread over the entire diagram range, also owing to the difference in the sample masses. Dividing $J(T)$ by the corresponding sample masses, the resulting substance nucleation rate data (per unit mass), reported in Fig. 11, display a greater compatibility among selected subsets.

Theoretical predictions for the nucleation rate can be obtained in the framework of the classical nucleation theory for both homogeneous and heterogeneous nucleation. The leading

TABLE I. List of samples, labeled with lower-case letters, investigated in the present analysis. The corresponding masses m, cooling rates C_r, and model parameters resulting from the Bayesian analysis of the sequence of $N_{\rm rec}$ events are also reported with statistical errors indicated in parentheses.

Sample	m (mg)	C_r (K/s)	$N_{\rm rec}$	T_u (K)	A (Hz)	B (K^{-1})
a	44.2	22.5	77	1059.2	2.25(26)	0.158(16)
b	11.5	20.0	268	1076.8	3.40(21)	0.275(13)
c	3.7	22.5	75	1056.1	2.98(35)	0.233(23)
d	432.0	12.2	71	1112.1	0.519(26)	0.049(8)
e	90.6	16.2	95	1071.6	1.21(12)	0.118(11)
f	184.1	14.5	137	1077.7	0.88(7)	0.084(6)
g	8.4	24.3	39	1061.0	3.6(6)	0.259(16)
h	1.3	23.2	8	1048.6	3.5(1.3)	0.28(8)

exponential term was calculated according to Eq. (2) using $\lambda = 36.940$ kJ/mol and $\alpha = 0.537$ [42], a value which is believed appropriate for the diamond structure and found consistently slightly lower than that for closer packed structures. Even if it is an ideal limiting case, it is instructive to estimate the attachment rate in Eq. (1) for the homogeneous nucleation limit using the approximation [3]

$$k^+(n^*) \simeq \frac{6D}{\delta^2} 4(n^*)^{2/3}, \tag{22}$$

where $D = D(T)$ is the Ge self-diffusion coefficient, δ represents a typical diffusion distance (approximated by the cubic root of the atomic volume), and $4(n^*)^{2/3}$ is an estimate of the number of possible attachment sites on the surface of the nucleus. Owing to a partial cancellation between $(n^*)^{2/3}$ and the Zeldovich factor $Z = \sqrt{\beta C/(2\pi)}$, where $C = -\frac{\partial^2 \Delta G(n)}{\partial n^2}|_{n^*}$, the resulting prefactor is

$$k^+(n^*)Z \simeq 4.5 \frac{D(T)}{\delta} \left(\frac{\alpha \lambda_m}{\left(N_A V_m^2\right)^{1/3} k_B T_m} \right)^{1/2}, \tag{23}$$

and its T dependence appears dominated by $D(T)$. This latter physical quantity can be estimated from the results of *ab initio* molecular dynamics simulations [43] interpolated by

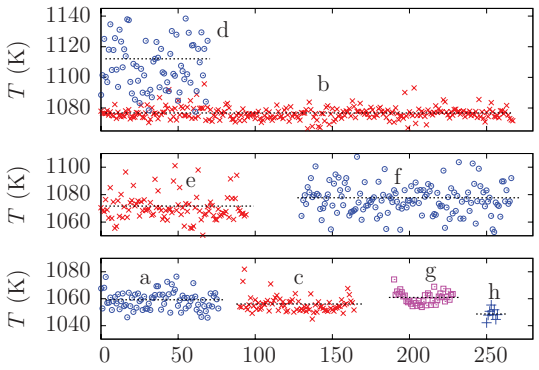

FIG. 9. (Color online) Sequences of maximum undercooling temperatures used for the present analysis collected under stable conditions for the corresponding samples listed in Table I. The horizontal dotted lines refer to the optimal T_u values.

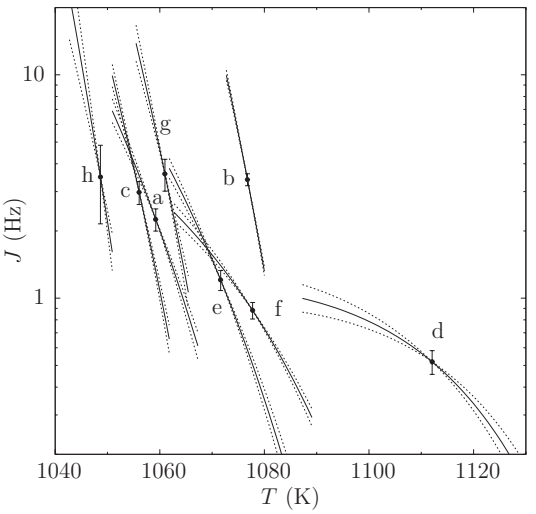

FIG. 10. Sample nucleation rates resulting from the Bayesian analysis of representative data sets. The • symbols are located at the corresponding reference temperatures T_u and A values on a logarithmic scale. The error bar is one standard deviation. The solid curve is the most probable sample nucleation rate adopting model (13) and the dashed curves represent the fan associated with $\pm\sigma_B$. The curves are traced over the region of the central 90th percentile of the observed undercooling temperatures.

an Arrhenius empirical function

$$D(T) = 2.0 \times 10^{-7} e^{3700/T\ (\rm K)}\ (\rm m^2/s), \tag{24}$$

in reasonable agreement with recent quasielastic neutron scattering measurements [44] in the stable liquid T range.

The nucleation rate computed according to the CNT for the homogeneous case is, as expected, several orders of magnitude

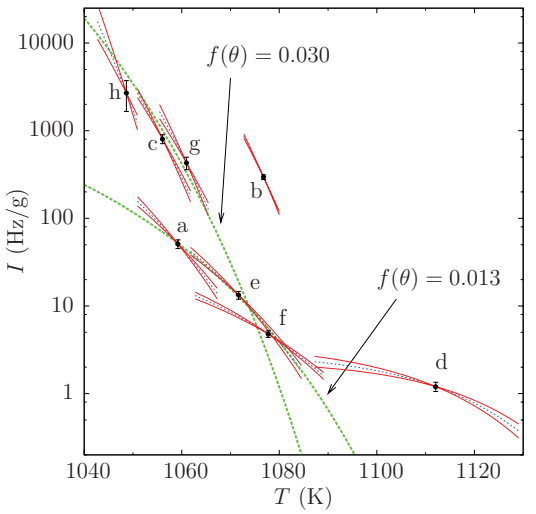

FIG. 11. (Color online) Ge nucleation rates resulting from the Bayesian analysis of representative data sets reporting the data of Fig. 10 divided by the corresponding sample mass. The data should ideally coalesce into a single master curve. The departure from this scenario indicates a variability in the concentration of impurities acting as heterogeneous nucleants. Representative estimates of the heterogeneous nucleation rate for different values of $f(\theta)$ are reported as the labeled dotted curves (green).

lower than the range of Fig. 11. Estimates of the heterogeneous nucleation rate can, however, be obtained, according to Eq. (2), by introducing the corrective $f(\theta)$ factor, θ being the contact angle, in the exponential term and reducing the prefactor by several orders of magnitude to match the observed range (this should reflect the number of active impurities which is much lower than the number of atoms in the system). Two representative curves are reported in Fig. 11, which roughly match the results (samples h, c, and g) for $f(\theta) = 0.030$ ($\theta \simeq 37.5^\circ$) and (samples a, e, and possibly f) for $f(\theta) = 0.013$ ($\theta \simeq 30^\circ$). The data set b is still compatible with $f(\theta) = 0.030$ but with a much higher impurity concentration. Sample d is apparently associated with an even lower $f(\theta)$.

The scenario that emerges from these results is that the limited Ge undercooling ability is due to the presence of impurities and dominated by heterogeneous nucleation phenomena. The analysis indicates the presence of at least two major impurity substrates characterized by the above, relatively small, contact angles. In light of the recent findings [17] information about the T dependence of the Ge nucleation rate may provide an insight into the nucleation kinetics and barrier, possibly resulting from the interplay between HDL and LDL, the latter acting as a heterogeneous substrate for the tetrahedral crystal nucleation also in the presence of foreign impurities. The present results are however still limited in T range and maximum undercooling temperature. Further insight can be obtained by a more extended analysis involving other samples, other subsequences, other crucible and substrate materials. Such an analysis is however in progress and beyond the scope of this article.

VI. CONCLUSIONS

In this paper we have proposed a Bayesian data-analysis approach applicable to data sets of maximum undercooling temperatures reached by undercooled liquid samples subject to continuous radiation cooling prior to recalescence and crystallization. The assumptions are that the temperature random variable is the result of a nonhomogeneous Poisson process with a temperature-dependent sample nucleation rate $J(T)$. The adopted model $J(T)$ function mimics the functional form predicted by the classical nucleation theory, and the Bayesian approach is able to retrieve optimal values and confidence intervals for the parameters representing the central value and temperature slope of the logarithm of $J(T)$. The method has been extensively tested against simulated data with success. The Bayesian approach provides the shape of the two-dimensional probability density of the parameters compatible with the given data set, from which suitable confidence intervals can be determined. The implementation of a Metropolis Monte Carlo simulation in parameter space is very effective for parameter fitting, to optimize the reference temperature which minimizes parameter correlation, and to determine the confidence intervals associated with statistical uncertainty.

The method was applied to milligram weight undercooled liquid Ge droplets supported by BN crucibles melted using a resistive heater. This environment introduces foreign impurities acting as heterogeneous nucleants which limit the undercooling ability of the specimen to a maximum of about 170 K. The impurity concentration displays a strong variability among samples and as a function of the processing history. The model functional form for $J(T)$ of Eq. (13), owing to the correct exponential T dependence and in spite of its *ad hoc* prefactor, was found to mimic reliably the nucleation rate expressions suggested by the CNT. The results of the analysis were interpreted in terms of heterogeneous nucleation processes involving at least two impurity substrates characterized by different contact angles. These results provide a solid benchmark for the validation of the method.

The scientific relevance of this investigation resides in the possibility to use BN as a confinement environment for micrometric sized Ge droplet samples suitable to achieve a deep undercooling and usable for several x-ray measurements, and also at high pressure. Understanding the undercooling behavior of the Ge/BN system is therefore a prerequisite for the correct interpretation of the sample properties and behavior. For the present methodological paper, however, the selected experiment represents a challenging case of samples in a mass range covering nearly three orders of magnitude in a contaminated environment. The method was able anyway to provide useful information and should be therefore applicable [45] in a straightforward manner to experiments performed in containerless environments and/or cleaner conditions, similar to those currently performed by several groups [31–38]. Further applications, possible improvements of the functional models, and comparisons with different approaches will be greatly encouraged by the present findings.

[1] D. Turnbull and J. C. Fisher, J. Chem. Phys. **17**, 71 (1949).
[2] D. Turnbull, J. Appl. Phys. **21**, 1022 (1950).
[3] K. F. Kelton, A. L. Greer, and C. V. Thompson, J. Chem. Phys. **79**, 6261 (1983).
[4] K. F. Kelton, in *Solid State Physics*, edited by H. Ehrenreich and D. Turnbull, Vol. 45 (Academic, New York, 1991), pp. 75–177.
[5] D. Oxtoby, J. Phys.: Condens. Matter **4**, 7627 (1992).
[6] F. Spaepen, Acta Metall. **23**, 729 (1975).
[7] F. Spaepen and R. B. Meyer, Scr. Metall. **10**, 257 (1976).
[8] F. Spaepen, Mater. Sci. Eng., A **178**, 15 (1994).
[9] D. Turnbull, J. Chem. Phys. **18**, 198 (1950).
[10] P. R. ten Wolde and D. Frenkel, Phys. Chem. Chem. Phys. **1**, 2191 (1999).
[11] S. Auer and D. Frenkel, Nature (London) **409**, 1020 (2001).
[12] S. Auer and D. Frenkel, J. Chem. Phys. **120**, 3015 (2004).
[13] A. Cacciuto, S. Auer, and D. Frenkel, Phys. Rev. Lett. **93**, 166105 (2004).
[14] S. E. M. Lundrigan and I. Saika-Voivod, J. Chem. Phys. **131**, 104503 (2009).
[15] T. Li, D. Donadio, L. M. Ghiringhelli, and G. Galli, Nat. Mater. **8**, 726 (2009).

[16] T. Li, D. Donadio, and G. Galli, J. Chem. Phys. **131**, 224519 (2009).
[17] C. Desgranges and J. Delhommelle, J. Am. Chem. Soc. **133**, 2872 (2011).
[18] L. Xu, S. V. Buldyrev, H. E. Stanley, and G. Franzese, Phys. Rev. Lett. **109**, 095702 (2012).
[19] W. H. Hofmeister, M. B. Robinson, and R. J. Bayuzick, Appl. Phys. Lett. **49**, 1342 (1986).
[20] B. Vinet, L. Cortella, and J. J. Favier, Appl. Phys. Lett. **58**, 97 (1991).
[21] L. Cortella, B. Vinet, P. J. Desré, A. Pasturel, A. T. Paxton, and M. van Schilfgaarde, Phys. Rev. Lett. **70**, 1469 (1993).
[22] A. L. Greer, Mater. Sci. Eng., A **178**, 113 (1994).
[23] B. Vinet and L. Cortella, Mater. Sci. Eng., A **178**, 125 (1994).
[24] L. Cortella and B. Vinet, Philos. Mag. B **71**, 11 (1995).
[25] D. M. Herlach, Annu. Rev. Mater. Sci. **21**, 23 (1991).
[26] S. Krishnan, J. J. Felten, J. E. Rix, J. K. Richard Weber, P. C. Nordine, M. A. Beno, S. Ansell, and D. L. Price, Rev. Sci. Instrum. **68**, 3512 (1997).
[27] W.-K. Rhim, S. K. Chung, D. Barber, K. F. Man, G. Gutt, and A. Rulison, Rev. Sci. Instrum. **64**, 2961 (1993).
[28] A. J. Rulison and W.-K. Rhim, Phys. Chem. Liq. **30**, 169 (1995).
[29] A. J. Rulison, J. L. Watkins, and B. Zambrano, Rev. Sci. Instrum. **68**, 2856 (1997).
[30] A. Filipponi, M. Borowski, P. W. Loeffen, S. De Panfilis, A. Di Cicco, F. Sperandini, M. Minicucci, and M. Giorgetti, J. Phys.: Condens. Matter **10**, 235 (1998).
[31] M. J. Uttormark, J. W. Zanter, and J. H. Perepezko, J. Cryst. Growth **177**, 258 (1997).
[32] G. Wilde, J. L. Sebright, and J. H. Perepezko, Acta Mater. **54**, 4759 (2006).
[33] G. Wilde, C. Santhaweesuk, J. L. Sebright, J. Bokeloh, and J. H. Perepezko, J. Phys.: Condens. Matter **21**, 464113 (2009).
[34] A. Filipponi and M. Malvestuto, Meas. Sci. Technol. **14**, 875 (2003).
[35] C. W. Morton, W. H. Hofmeister, R. J. Bayuzick, and M. B. Robinson, Mater. Sci. Eng., A **178**, 209 (1994).
[36] W. H. Hofmeister, C. W. Morton, and R. J. Bayuzick, Acta Mater. **46**, 1903 (1998).
[37] T. Schenk, D. Holland-Moritz, W. Bender, and D. M. Herlach, J. Non-Cryst. Solids **250–252**, 694 (1999).
[38] S. Klein, D. Holland-Moritz, and D. M. Herlach, Phys. Rev. B **80**, 212202 (2009).
[39] V. P. Skripov, in *Current Topics in Material Science, Crystal Growth and Materials*, edited by E. Kaldis and H. Scheel, Vol. 2 (North-Holland, Amsterdam, 1977), p. 23.
[40] G. D'Agostini, Rep. Prog. Phys. **66**, 1383 (2003).
[41] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[42] D. Li and D. M. Herlach, Europhys. Lett. **34**, 423 (1996).
[43] G. Kresse and J. Hafner, Phys. Rev. B **49**, 14251 (1994).
[44] S. M. Chathoth, B. Damaschke, T. Unruh, and K. Samwer, Appl. Phys. Lett. **94**, 221906 (2009).
[45] The data-analysis C code is freely available upon request from the corresponding author.