

On the Dispersive versus Arrhenius Temperature Activation of NBTI Time Evolution in Plasma Nitrided Gate Oxides: Measurements, Theory, and Implications

D. Varghese, D. Saha, S. Mahapatra, K. Ahmed^{#1}, F. Nouri^{#1}, and M. Alam^{#2}

Department of Electrical Engineering, IIT Bombay, Mumbai 400076, India (91-22-25720408 / souvik@ee.iitb.ac.in)

^{#1}Applied Materials, Inc., Santa Clara, CA; ^{#2}ECE Department, Purdue University, W. Lafayette, IN

Abstract

Negative Bias Temperature Instability (NBTI) is studied in p-MOSFETs having Decoupled Plasma Nitrided (DPN) gate oxides (EOT range of 12Å⁰ through 22Å⁰). Threshold voltage shift (ΔV_T) is shown to be primarily due to interface trap generation (ΔN_{IT}) and significant hole trapping (ΔN_{OT}) has not been observed. ΔV_T follows power-law time (t) dependence and Arrhenius temperature (T) activation.

Introduction

NBTI is a serious reliability concern for p-MOSFETs [1]. Oxynitrides (required for suppressing boron penetration and gate leakage) show worse NBTI than control oxides and has attracted much attention [2-5]. It is important to correctly *measure* and *extrapolate* t evolution of ΔV_T to accurately determine device lifetime. However, this extrapolation is complicated since (i) impact of delay time on measurement is not fully understood, (ii) ΔV_T origin (ΔN_{IT} that predicts tⁿ dependence or ΔN_{OT} that predicts log (t) dependence) is still unresolved, and (iii) the nature of temperature dependence (dispersive vs. Arrhenius activated) remains controversial. It is therefore necessary to resolve these inconsistencies for correct estimation of device lifetime.

Highlights of this work

It is shown that: (i) in addition to t-delay, measurement voltage ($V_{G,meas}$) and T strongly influence n, (ii) t-delay=0 measurements show power law t dependence for various EOT, N₂ dose, stress- V_G and T, (iii) recently reported large NBTI dispersion is an artifact of t-delay, (iv) for t-delay=0, short and long time degradation respectively are dispersive and Arrhenius activated (their relation is explained), (v) no ΔN_{OT} is observed, *only* ΔN_{IT} governs overall ΔV_T and hence can be explained by R-D model [10-12], and (vi) measured n along with activation energy (E_A) of diffusion uniquely identify neutral molecular H₂ as the diffusion species.

Impact of measurement delay

Fig.1 shows $\Delta V_T(t)$ ($= -(\Delta I_{dlin}/I_{dlin0}) * V_{GT0}$) from on-the-fly [6] and normal (with delay) Idlin measurements. t-delay increases n, and is known to be due to ΔV_T recovery during the stress “stopped” phase [6,7,11]. For identical t-delay, n also increases at lower $|V_{G,meas}|$ (Fig.1, LHS) and higher T (Fig.1, RHS). This is expected as ΔV_T recovery, believed due to passivation of broken Si- bonds [11,12] is larger at higher T and lower $|V_{G,meas}|$ [1,7,13]. Fig.2 (LHS) shows $\Delta N_{IT}(t)$ measured using charge pumping (C-P). Increase in n with higher t-delay and higher T is clearly seen. Fig.2 (RHS) shows $\Delta N_{IT}(t)$ simulated using R-D

model with and without t-delay for different T. For t-delay=0, n~1/6 (for all T) as expected for H₂ diffusion [11]. For t-delay>0, n increases at higher T, and is consistent with the measured trends. Fig.3 shows n (from Idlin and C-P) versus T for various t-delay. Increase in n at higher T can be clearly seen, which is more pronounced for larger t-delay. Further, as average $V_{G,meas}$ for C-P (~0, mean value of pulse) is lower than $|V_{G,meas}|$ for Idlin, n from C-P is much larger than that from Idlin for a given t-delay and T. These results are consistent with increased recovery at larger t-delay, less negative $V_{G,meas}$ and higher T. Since delay contaminates all measurements and results in uncertain values of n, only “on-the-fly” Idlin [6] should be used to quantify and interpret NBTI defect generation, as has been done below.

NBTI time evolution

Fig.4 shows $\Delta V_T(t)$ for various T, plotted in log-log and lin-log scales. It is obvious that NBTI follows power-law (not log) time dependence. When plotted in log-log scale, n shows a slight increase with T for short t, indicating weak dispersive diffusion [7,9]. However, the T-dependence of n completely disappears for longer t, and n becomes (slightly less than) 1/6 for all T. Fig.5 shows the T dependence of $\Delta V_T(t)$ for various EOT, N₂ dose, and stress- V_G . Long time degradation clearly shows an universal Arrhenius, power-law (n~1/6) behavior. Contrary to [8], the strong power-law time dependence suggests ΔV_T is governed by ΔN_{IT} and not ΔN_{OT} (at least for DPN oxynitrides for these EOT range). Furthermore, strong dispersion reported in literature [7,9] is most likely a consequence of measurement delay. Finally, we believe historically measured n (~1/5 – 1/3) for NBTI with uncertain t-delay can all be attributed to this universal n~1/6 behavior (for t-delay=0), which is explained below.

Explanation using “standard” R-D model

The Arrhenius T activated (longer time), power-law time evolution of ΔV_T (~ ΔN_{IT}) can be explained by R-D model [1,10-12]:

$$\Delta N_{IT} = (k_F \cdot N_0 / 2k_R)^{1/2} \cdot [D_0 \cdot \exp(-E_A/kT)t]^n,$$

where k_F , k_R are the Si-H dissociation and passivation rate constants, N_0 is total Si-H bond density, and $D=D_0 \cdot \exp(-E_A/kT)$ is the diffusivity of released H species. Fig. 6 shows that long t (varying T) data can be uniquely scaled along X-axis and Y-axis directions to universal relations, which yields E_A for D and ΔV_T respectively [1]. Obtained E_A (D) proves neutral H₂ as the diffusion species [14]. This is also consistent with n~1/6 power-law t dependence as predicted by R-D model for H₂ diffusion (Fig.2) [11]. Moreover, R-D

model suggests $E_A (\Delta V_T) = E_A (D) * n$ if $E_A (k_F) \sim E_A (k_R)$ [1], as also can be seen from Fig.6.

Interface-trap generation dominated NBTI

Fig.7 shows T dependence $\Delta V_T (t)$ for devices with same EOT but various N_2 dose. ΔV_T for both low and high N_2 dose (and their difference) show similar $E_A (\sim E_A (D) * n)$. This proves that increase in NBTI due to increased plasma N_2 dose is due to increase in ΔN_{IT} . To show the universality that ΔN_{IT} dominates ΔV_T , a large number of samples with different EOT and N_2 dose (including control oxides) were stressed at various T and oxide electric field (E_{OX}). $\Delta V_T (t)$ was obtained using on-the-fly Idlin. For few cases $\Delta N_{IT} (t)$ was obtained (separately) from C-P on larger EOT samples. The results are shown in Figs. 8 and 9 (control and plasma N_2 data plotted separately for clarity, and does not signify actual magnitude).

Fig.8 shows X-axis scaling factors of T dependent $\Delta V_T (t)$ and $\Delta N_{IT} (t)$ data. Very similar T activation (E_A ranges from 0.54eV through 0.62eV) of all data clearly proves that $\Delta V_T \sim \Delta N_{IT}$, and plasma N_2 and control oxides have very similar T dependent kinetics. Since $E_A (k_F) \sim E_A (k_R)$, T dependence of ΔN_{IT} is primarily due to that of D [1,9]. Moreover, that the diffusing species is always neutral molecular H_2 is also unequivocally established (consistent $E_A (D)$ and n values).

Fig.9 shows E_{OX} dependence of ΔV_T and ΔN_{IT} data. Since diffusion species is neutral H_2 , E_{OX} dependence is due to that of $(k_F/k_R)^{1/2}$ and has the form $A.E_{OX}^{1/2} \exp (B.E_{OX})$ [1,9]. Once again, qualitatively similar slope “B” suggests $\Delta V_T \sim \Delta N_{IT}$, and plasma N_2 and control oxides have similar E_{OX} dependent kinetics. We believe N_2 incorporation impacts “A”, possibly by reducing the Si-H bond dissociation energy [15]. In short, control oxides and DPN oxides are governed by very similar reaction and diffusion mechanics.

Dispersion in n

The transition from dispersive (short t) to Arrhenius (long t) T activation (Fig.4) can be interpreted rigorously, and is qualitatively explained using Fig.10. In the absence of deep trapping sites, H_2 would diffuse via shallow hopping sites (standard diffusion) and shows weak T dependence. ΔN_{IT} equals total released H_2 density (N_H) and $dN_{IT}/dt = dF_H/dx = D_0 d^2N_H/dx^2$ (F_H : H_2 flux).

In the presence of deep traps at a single energy level $E_A = E_A$, trapping and detrapping of H_2 (to and from this level) influences the T dependence of diffusion. If N_H^F, N_H^T are free and trapped H_2 density ($N_H = N_H^F + N_H^T$), then $N_H^F/N_0 = [N_H^T/N_T] \exp (-E_A/kT)$, if fast equilibrium is assumed. N_0, N_T are the density of shallow hopping sites and deep traps. Therefore, $dF_H/dx = D_0 d^2N_H^F/dx^2 = D_1 d^2N_H/dx^2$ (as only free H_2 can diffuse). Assuming $(N_0/N_T) \exp (-E_A/kT) \ll 1$, and by rearranging terms, $D_1 = D_0 (N_0/N_T) \exp (-E_A/kT)$. Hence, Arrhenius T activation is obtained. In such a case, N_{IT} time evolution can be explained by standard R-D model solution [1, 10-12] and the universal scaling scheme (Fig.6) will hold good [1,9].

In the presence of trap energy distribution up to infinity, trapping and detrapping of H_2 to and from multiple energy levels determine T dependence of diffusion. Emission time constants of these levels are $\tau = 1/v_0 \exp (E/kT)$, v_0 being the attempt to escape frequency. As t increases, more and more deep traps emit H_2 and as a result, the energy centroid of trapped $H_2, E_M (t)$, goes down in energy. Using exponential trap distribution as $g(E) = N_T/E_0 \exp (-E/E_0)$ and assuming fast equilibrium, $N_H^F/N_0 = [N_H^T/E_0 g(E_M)] \exp (-E_M(t)/kT)$, where $E_M (t) = kT \ln (v_0 t)$. Rearranging, $N_H^F = N_H [N_0/N_T] \exp [kT/E_0 \ln (v_0 t)]/(v_0 t)$, and $dF_H/dx = D_0 d^2N_H^F/dx^2 = D_2 d^2N_H/dx^2$, where $D_2 = D_0 (N_0/N_T) (1/v_0 t)^{[1 - kT/E_0]}$, which is a signature of dispersive diffusion [7,16].

For a real system with finite distribution of traps up to E_A , diffusion is dispersive till $t = t_0$, as long as $E_M(t)$ reduces with t. This get transformed to Arrhenius activated diffusion for $t > t_0$ as steady state is attained [$E_M(t_0) \sim E_A$]. The transition time is $t_0 = 1/v_0 [\exp (E_A/kT)]^{[1/(1-kT/E_0)]}$, and reduces at higher T. For $E_A = 0.58eV$, Fig.4 predicts $v_0 \sim 10^{12} s^{-1}$ and $E_0 = 80meV$, consistent with published values. (Accurate determination of t_0 can be influenced by non-equilibrium H to H_2 conversion, which can also affect short t slope).

Conclusion

NBTI time evolution is studied for plasma N_2 oxynitride p-MOSFETs. NBTI shows power-law time dependence for a wide range of EOT, N_2 dose, stress- V_G and T. It is shown that delay during measurements result in inaccurate power-law slope, which depends on delay time, measurement V_G and T. For zero delay measurements, short-time data shows dispersive but long-time data show Arrhenius T activation. It is clearly shown that strong long-time dispersion is an artifact of measurement-delay. A theory for the transition of dispersive to Arrhenius T activation is provided. NBTI is governed by interface trap generation, which follows R-D model. It is shown that DPN and control samples show very similar T and E_{OX} dependence of NBTI. Obtained n and E_A for diffusion suggests molecular H_2 diffusion governs the long-time evolution of NBTI behavior.

References

- [1] S. Mahapatra et al., p.105, IEDM 2004
- [2] Y. Mitani et al., p.509, IEDM 2002
- [3] Y. Mitani et al., p.117, IEDM 2004
- [4] V. Huard et al., p.40, IRPS 2004
- [5] H. Aono et al., p.23, IRPS 2004
- [6] S. Rangan et al., p.341, IEDM 2003
- [7] B. Kaczer et al., p.381, IRPS 2005
- [8] M. Dennis et al., p.109, IEDM 2004
- [9] M. Alam et al., Micro. Reliability, v.45, p.71, 2005
- [10] K. Jeppson et al., JAP, p. 2004, 1977
- [11] S. Chakravarthi et al., p.273, IRPS 2004
- [12] M. Alam, p.345, IEDM 2003
- [13] S. Mahapatra et al., SSDM 2005
- [14] M. L. Reed et al., JAP, p.5776, 1988
- [15] S. S. Tan et al., p.70, SSDM 2003
- [16] D. Monroe, Solid State Comm., p.435, 1986

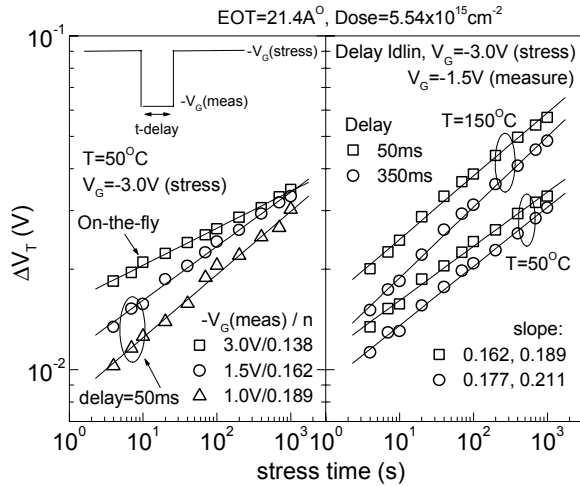


Fig. 1. ΔV_T (t) from Idlin measurements without and with delay. (LHS) Different measurement bias, (RHS) different measurement T. Slope increases for higher delay time, higher T and lower measurement V_G .

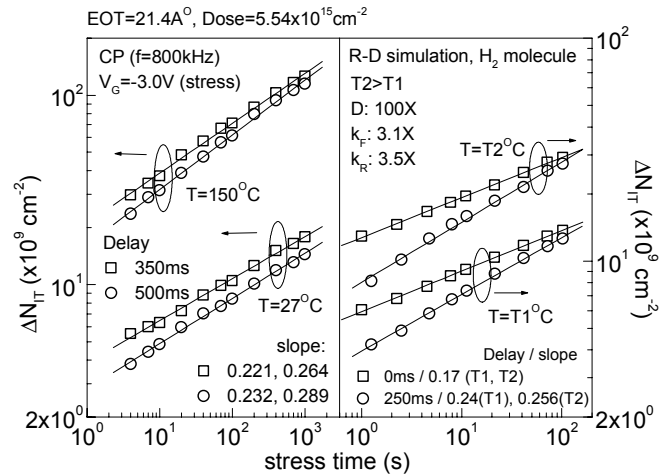


Fig. 2. (LHS) T dependence of ΔN_{IT} (t) measured using C-P for different t-delay. Increase in slope at larger T is more pronounced when delay is high. (RHS) Numerical solution of R-D model with and without delay. T-dependent slope observed for non-zero delay.

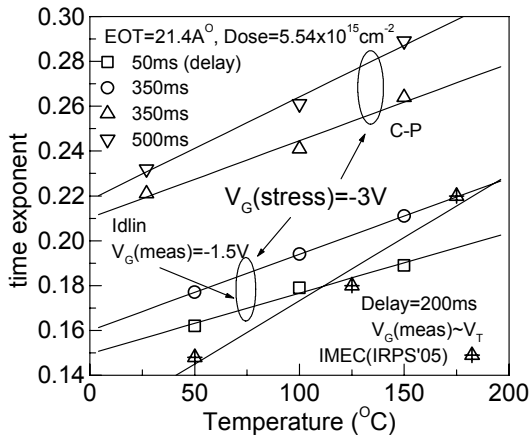


Fig. 3. T dependence of power-law time exponent from Idlin and C-P measurements. Higher dispersion (increase in n with T) is seen for higher measurement delay. C-P slopes are greater than Idlin slopes for same T and t-delay.

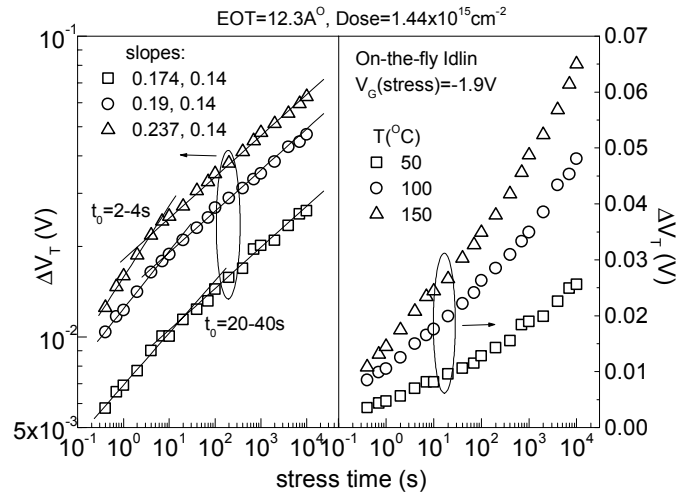


Fig. 4. T dependence of ΔV_T (t) measured using on-the-fly Idlin technique, showing power-law (not log) time dependence. Short time ($t < t_0$) data shows higher slope at higher T (dispersive). Long time data shows identical slope for all T (Arrhenius activation).

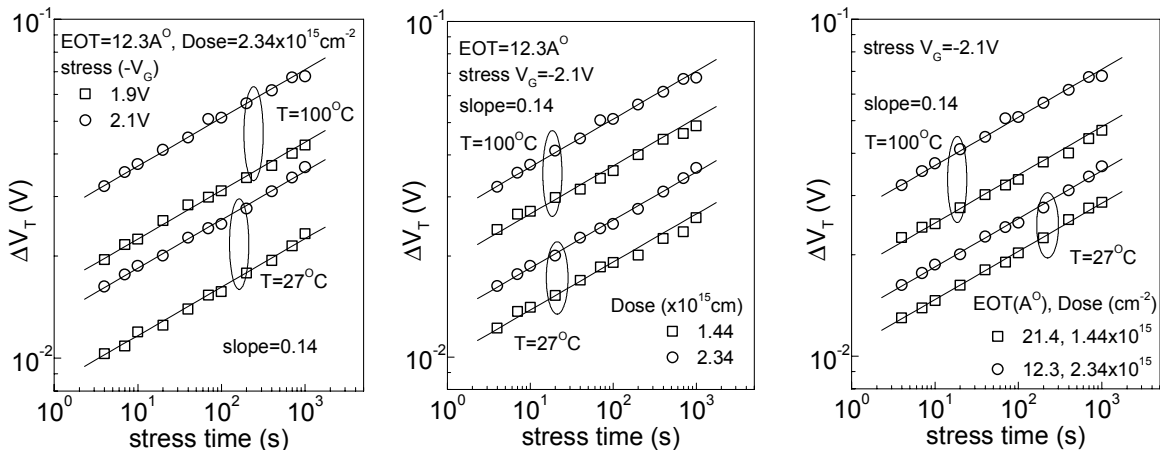


Fig. 5. T dependence of ΔV_T (t) for various stress- V_G , N_2 dose and EOT, measured using on-the-fly Idlin technique. Non-dispersive power law behavior is universally observed.

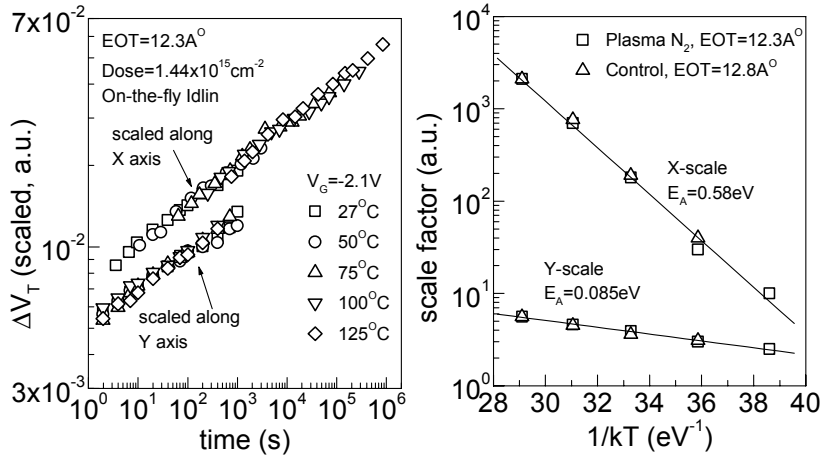


Fig.6. (LHS) Universal scaling scheme: $\Delta V_T(t)$ data for various T is scaled along X and Y-axis directions to universal relations. (RHS) T activation of X and Y-axis scale factors. Obtained E_A from X-axis scaling suggests molecular H_2 diffusion. E_A (Y-scale) = E_A (X-scale) * n as predicted by R-D model [1].

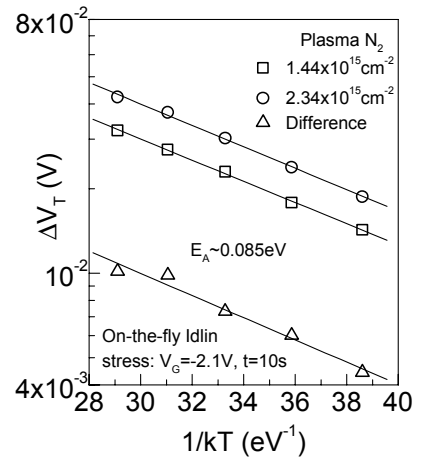


Fig.7. T activation of ΔV_T data for various N_2 dose. Lower and higher N_2 dose (and their difference) show identical E_A , similar to that of control oxide.

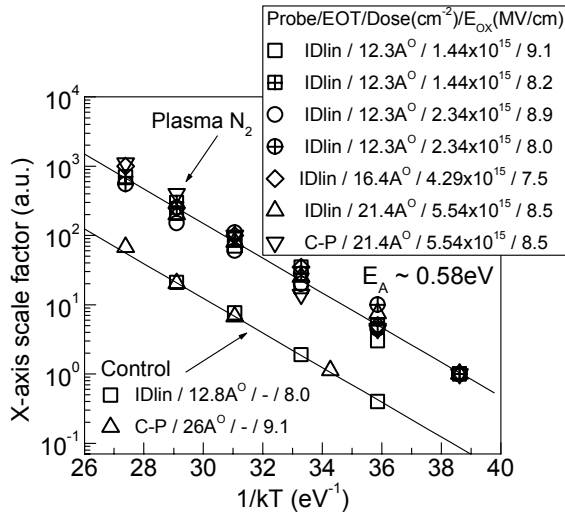


Fig.8. X-axis scaling (like Fig.6) factors for T dependent ΔV_T and ΔN_{IT} data, measured on films having different EOT and N_2 dose. Very similar T dependence is observed for all films. Line is best fit to control oxide data, and can be used as eye-guide for plasma oxynitride data.

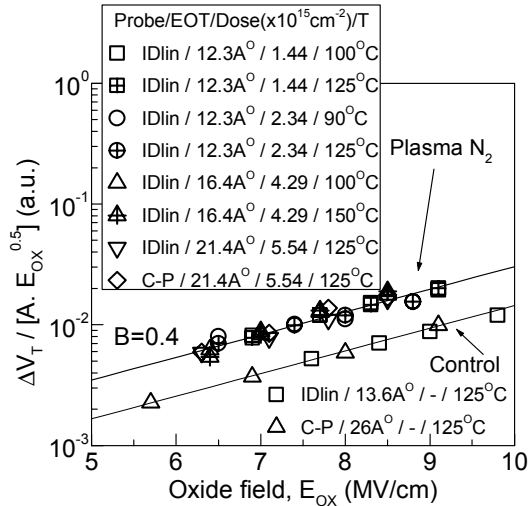


Fig.9. E_{OX} dependence of ΔV_T and ΔN_{IT} data, measured on films having different EOT and N_2 dose. Very similar E_{OX} dependence is observed for all films. Line is best fit to control oxide data, and can be used as eye-guide for plasma oxynitride data.

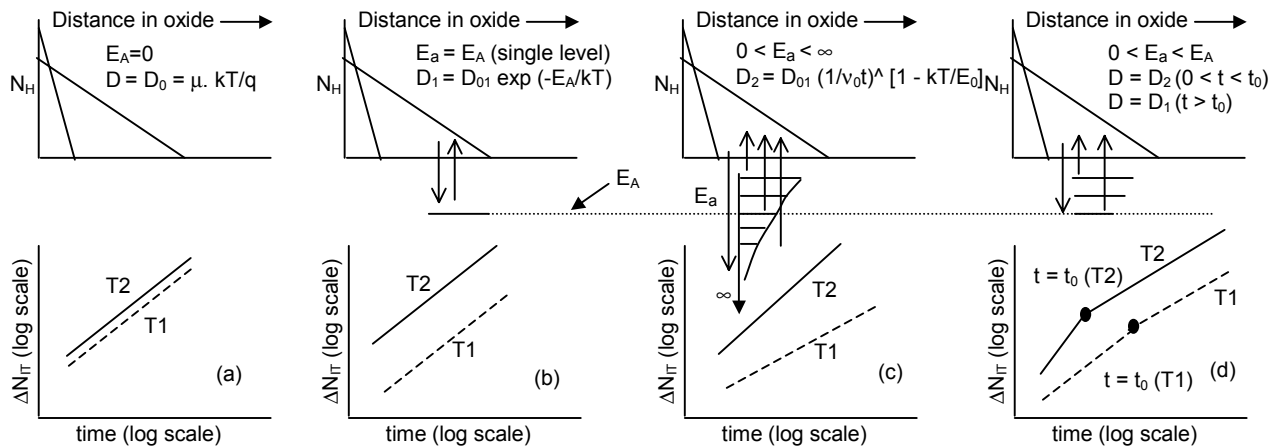


Fig.10. Schematic of H_2 diffusion process: (a) without trapping, (b) with single energy trap at E_A , (c) with distribution (shape factor E_0) of traps up to ∞ , and (d) with finite distribution of traps up to E_A . Resulting T dependence of ΔN_{IT} : (a) weak (Einstein relationship), (b) Arrhenius activated (c) dispersive, and (d) dispersive at short time ($t < t_0$), non-dispersive Arrhenius-like at longer time ($t > t_0$).