Supplemental Material for "Vacancy-induced low-energy states in undoped graphene"

Sambuddha Sanyal,[1] Kedar Damle,[2] and Olexei I. Motrunich[3]

[1]*International Center for Theoretical Sciences, Tata Institute of Fundamental Research, Bengaluru 560089, India*

[2]*Department of Theoretical Physics, Tata Institute of Fundamental Research, Mumbai 400005, India*

[3]*Department of Physics, California Institute of Technology, Pasadena, California 91125, USA*

Abstract

In this Supplemental Material, we present additional numerical evidence and analytical arguments which support the key findings described in the main text.

PACS numbers: 71.23.-k;73.22.Pr;71.23.An;72.15.Rn

I. ADDITIONAL NUMERICAL EVIDENCE

A. Other concentrations

Figure 1 displays $N_L(\Gamma)$ for $n_v = 0.05$, for the three largest sizes studied. The corresponding extrapolation to the thermodynamic limit is shown in Fig. 2. The corresponding results for $n_v = 0.075$ are displayed in Fig. 3. In all these figures, we focus on $\Gamma < \Gamma_g^*(L_{\rm min})$, where $L_{\rm min}$ is the smallest size for which data is displayed, and Γ_g^* is read off from the peak in the histograms of Γ_g shown in Fig. 4 and Fig. 5. The corresponding histograms for $n_v = 0.0625$ and $n_v = 0.1$ are displayed in Fig. 6

As is clear from these results for $n_v = 0.05$ and $n_v = 0.075$, $N_L(\Gamma)$ is found to fit well to a power-law form $N_{\rm Dyson} \equiv c\Gamma^{-y}$ up to a fairly well-defined and readily-identified crossover scale $\Gamma_c(L) \equiv \log_{10}(t/|\epsilon_c(L)|)$. However, beyond Γ_c, the asymptotic fall-off is clearly faster than a power law. While the increase of $\Gamma_c(L)$ with L is more significant at the smallest concentration studied ($n_v = 0.05$), it is nevertheless clear that $\Gamma_c(L)$ does saturate to a finite value even in this case. This is clear from the fact that the extrapolated thermodynamic density of states $N(\Gamma)$ (Fig. 2) also displays the same crossover seen in the finite-size data. In the large-Γ regime beyond this crossover, the modified Gade-Wegner form $N_{GW}(\Gamma) \equiv a\Gamma^{1/3}e^{-b\Gamma^{2/3}}$ is seen to provide a very good fit of the data for both these concentrations. The corresponding values of $\Gamma_c(L)$ and Γ_c, and of the best fit values of y, provide us additional points that fill in the curves shown in Fig. 4 and Fig. 6 of the main text, which display the n_v dependence of y and Γ_c, and the close relationship between $l_c \equiv N(\Gamma_c)^{-1/2}$ and $l_w \equiv w^{-1/2}$. Finally, we re-emphasize a point made already in the main text: Our computational constraints prevent us from accessing the thermodynamic limit for the much larger values of Γ at which we expect to see the same crossover for the lowest concentration $n_v = 0.02$.

B. Extrapolations

Since states at any finite Γ (*i.e.*, away from the band center $\epsilon = 0$) in such particle-hole symmetric hopping problems are not critical, one expects the leading corrections to the thermodynamic limit $N(\Gamma)$ at any finite Γ to be regular rather than singular, similar to the finite-size corrections expected in noncritical phases of matter (*i.e.*, away from critical

points or critical lines). Guided by this rationale, the thermodynamic limit $N(\Gamma)$ is obtained from $N_L(\Gamma)$ by performing a polynomial extrapolation in $1/L$ (note that we expect that the leading finite-size corrections are $\mathcal{O}(1/L)$ rather than $\sim \exp(-L/\xi)$ because of "surface" contributions associated with the semi-open boundary conditions we employ). Since we are careful to only use large enough sizes for which almost every sample has at least one zero mode ($f_L \approx 1$), our finite-size data is already rather close to the thermodynamic limit, leading to a rather small secular drift with increasing L. In most cases, given the size of our error bars relative to the magnitude of this secular drift with L, the inclusion of the next-order term c/L^2 only results in an over-interpretation of statistical fluctuations. Therefore, a simple linear (in $1/L$) extrapolation $a + b/L$ has been used in most cases.

We have also tested the stability of this extrapolation procedure to the inclusion of data at larger sizes. For the representative case of $n_v = 0.0625$, this is shown Fig. 7, which is devoted to a comparison of the thermodynamic limit obtained in the main text using sizes $L = 200, 180, 160$ with two other alternatives: A linear extrapolation from three sizes $L = 220, 200, 180$, and a linear extrapolation from four sizes $L = 220, 200, 180, 160$. As is clear from this figure, all three extrapolations (*i.e.*, the one used in the main text as well as the other alternatives which use data at a larger size) yield extrapolated values that lie within the error-bars of each other. Further, there is no systematic trend that suggests that any one of these extrapolations yields a consistently higher or lower value of $N(\Gamma)$ at all Γ. Details of all three extrapolations, for each value of Γ, are also shown as a separate multi-page figure (Fig. 20) placed at the end of this supplemental section for ease of inspection. Some examples of extrapolations used to arrive at $N(\Gamma)$ from data for $N_L(\Gamma)$ at other concentrations are also shown in Figs. 8, 9, 10, and 11. From this careful and detailed study, we conclude that our approach indeed allows us to reliably obtain the thermodynamic limit curve $N(\Gamma)$.

II. FURTHER ANALYSIS OF ZERO MODES

Our data for f_L, the probability that an $L \times L$ sample has at least one zero mode, is shown in Fig. 12. Clearly, f_L tends to 1 as $L \to \infty$, as already mentioned in the main text. This is consistent with the analytical argument in the main text, which also provides a simple rigorous lower bound for the density of zero modes. The 4-triangle zero mode used in this

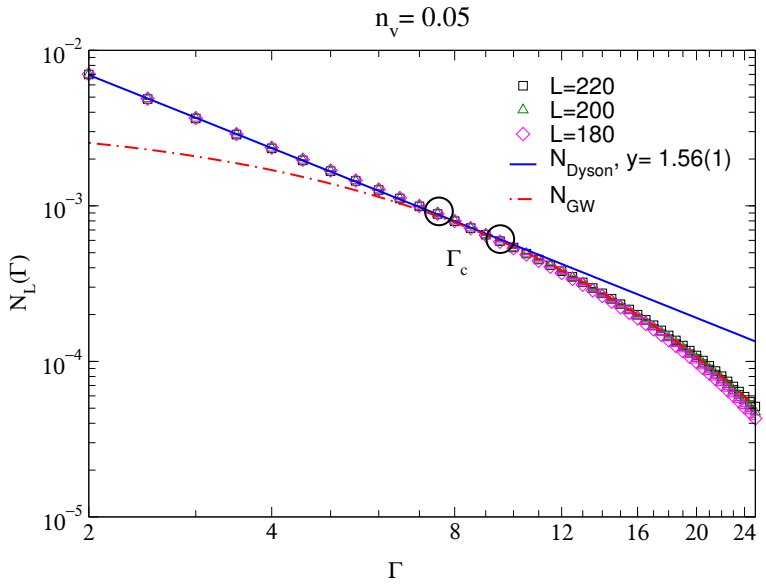


FIG. 1: $N_L(\Gamma)$ at the three largest values of L studied for $n_v = 0.05$. Circles demarcate the crossover region centered at the crossover scale Γ_c. Data for $\Gamma \lesssim \Gamma_c$ fits well to power-law form $N_{\rm Dyson}(\Gamma)$ (see text for details) with the value of y indicated in the figure, while the large-Γ regime fits well to the modified Gade-Wegner form $N_{\rm GW}(\Gamma)$.

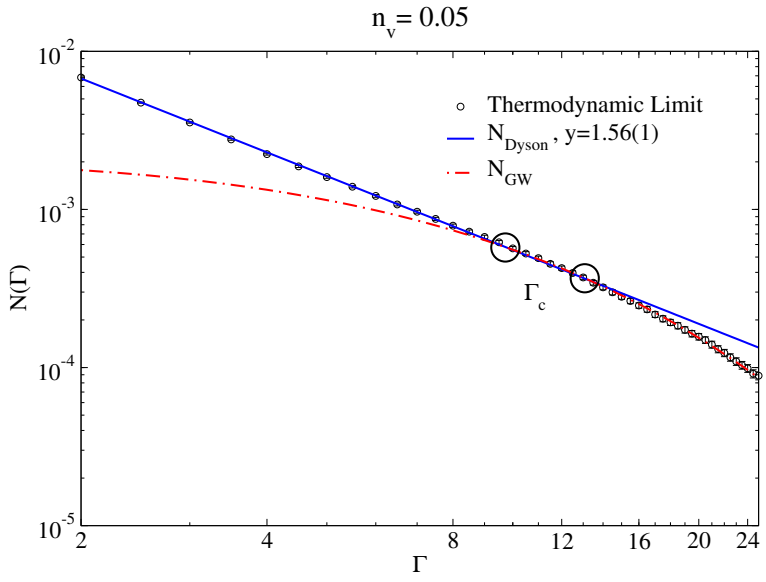


FIG. 2: $N(\Gamma)$, the extrapolation to the thermodynamic limit of the finite-size data from the previous figure. Again, circles demarcate the crossover region centered at the crossover scale Γ_c. Data for $\Gamma \lesssim \Gamma_c$ fits well to power-law form $N_{\rm Dyson}(\Gamma)$ with the value of y indicated in the figure, while the large-Γ regime fits well to the modified Gade-Wegner form $N_{\rm GW}(\Gamma)$.

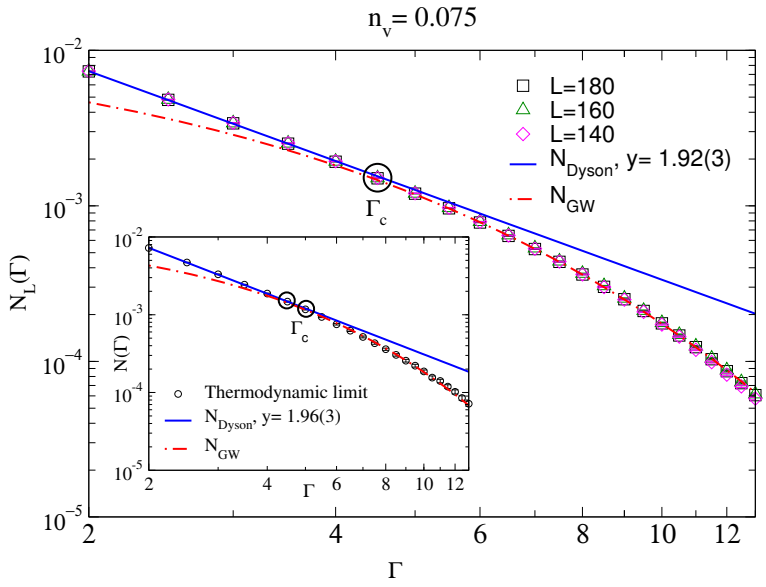


FIG. 3: $N_L(\Gamma)$ at the three largest values of L studied for $n_v = 0.075$. Circles demarcate the crossover region centered at the crossover scale Γ_c. Data for $\Gamma \lesssim \Gamma_c$ fits well to power-law form $N_{\rm Dyson}(\Gamma)$ with the value of y indicated in the figure, while the large-Γ regime fits well to the modified Gade-Wegner form $N_{\rm GW}(\Gamma)$. Inset shows the extrapolation to the thermodynamic limit, in which the same crossover is clearly visible.

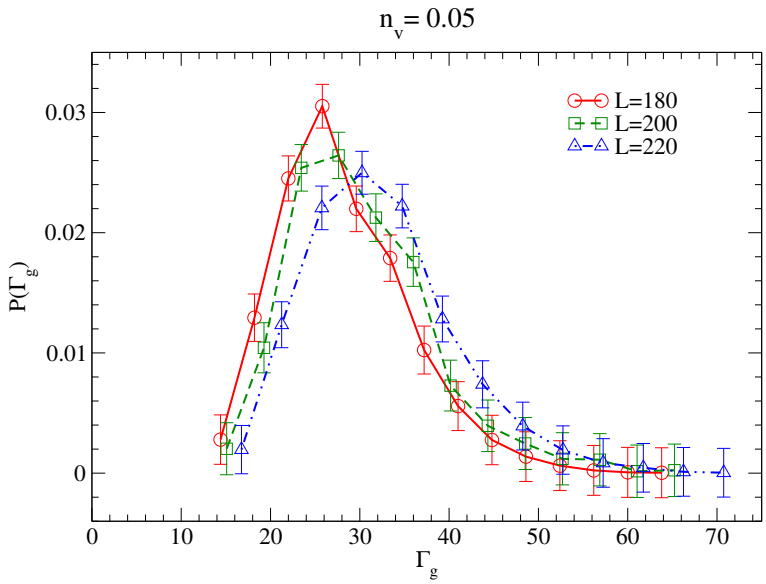


FIG. 4: Histogram of Γ_g, corresponding to the lowest nonzero gap for the three largest sizes studied at $n_v = 0.05$.

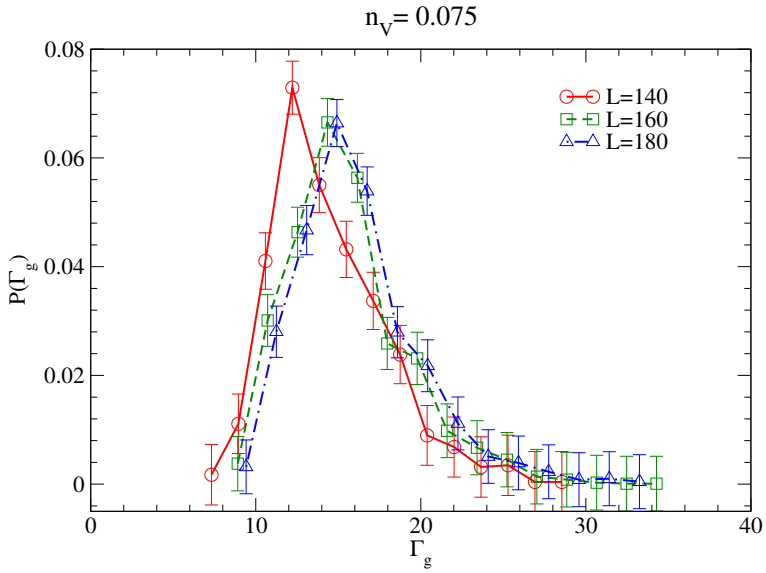


FIG. 5: Histogram of Γ_g, corresponding to the lowest nonzero gap for the three largest sizes studied at $n_v = 0.075$.

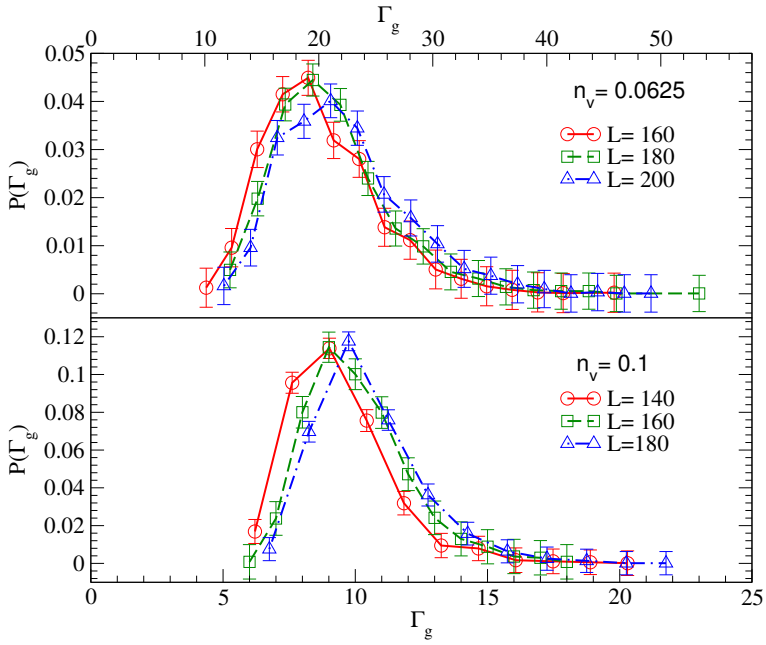


FIG. 6: Histograms of Γ_g at the three largest values of L studied for $n_v = 0.0625$ and $n_v = 0.1$.

argument is the first term in an infinite series in n_v, with higher powers of n_v arising from bigger patterns consisting of a larger number of impurities in specific locations relative to each other. In Fig. 13 and Fig. 14, we show a few examples of zero mode constructions that contribute to this series. However, as already noted in the main text, terms in this series do

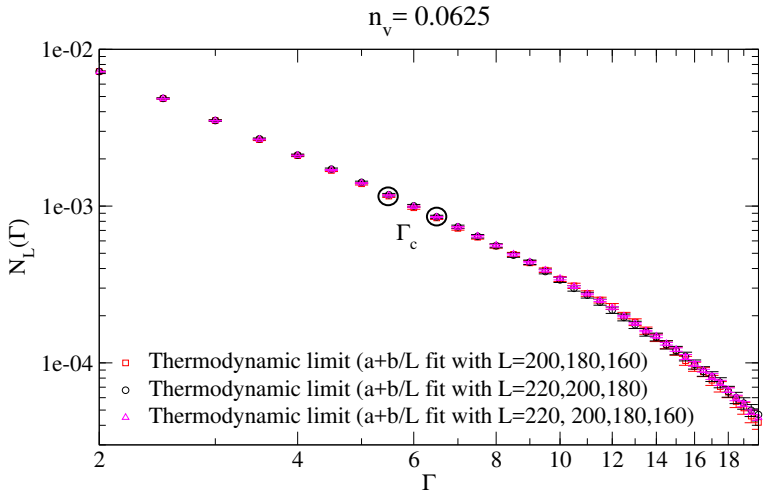


FIG. 7: Three different extrapolations yield results for the thermodynamic limit $N(\Gamma)$ that fall within the error bars of each other, confirming the reliability and stability of our procedure to obtain the thermodynamic limit for the representative case of $n_v = 0.0625$.

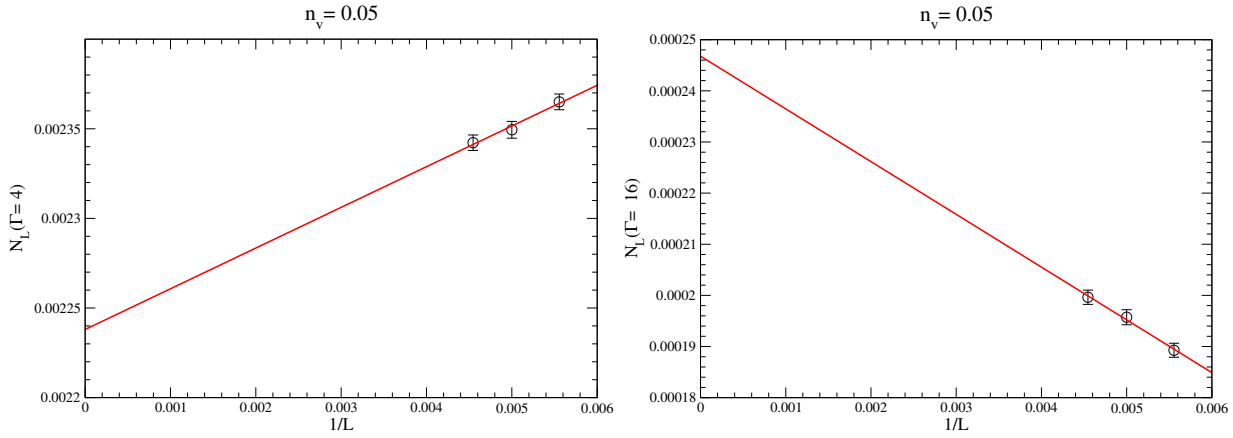


FIG. 8: Examples of extrapolation of $N_L(\Gamma)$ to the thermodynamic limit at $n_v = 0.05$. For this concentration, $\Gamma_c \approx 11$ (see Fig. 4 in the main text), and the left panel illustrates the extrapolation for $\Gamma < \Gamma_c$, while the right panel is for $\Gamma > \Gamma_c$. Note in particular that our extrapolation for $\Gamma > \Gamma_c$ is very likely an overstimate, so one can be fairly confident that $N(\Gamma)$ in the thermodynamic limit drops below $N_{\rm Dyson}(\Gamma)$, ruling out a fit to this form for $\Gamma > \Gamma_c$.

not give the dominant contribution to w at the not-too-small values of n_v studied by us in this work. Indeed, we have explicitly measured the density of 4-triangles and checked that it is significantly smaller than the density of zero modes for all n_v at which we have computed w (including $n_v = 0.05$). Additionally, we have enumerated all possible clusters of fewer

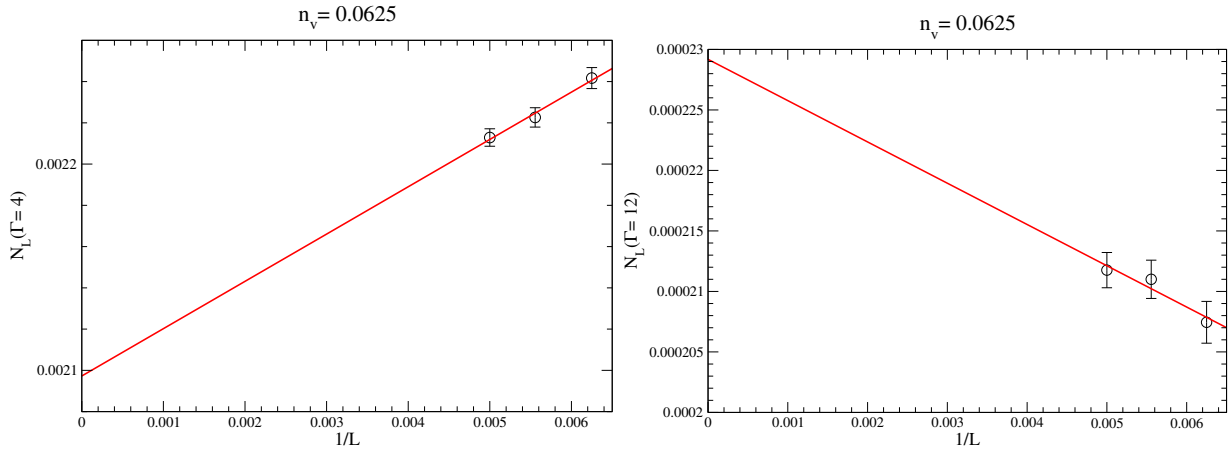


FIG. 9: Examples of extrapolation of $N_L(\Gamma)$ to the thermodynamic limit at $n_v = 0.0625$. For this concentration, $\Gamma_c \approx 7$ (see Fig. 4 in the main text), and the left panel illustrates the extrapolation for $\Gamma < \Gamma_c$, while the right panel is for $\Gamma > \Gamma_c$. Note in particular that our extrapolation for $\Gamma > \Gamma_c$ is very likely an overstimate, so one can be fairly confident that $N(\Gamma)$ in the thermodynamic limit drops below $N_{\rm Dyson}(\Gamma)$, ruling out a fit to this form for $\Gamma > \Gamma_c$. Extrapolations at other values of Γ, as well as extrapolations including a larger size ($L = 220$) are shown in Fig. 20.

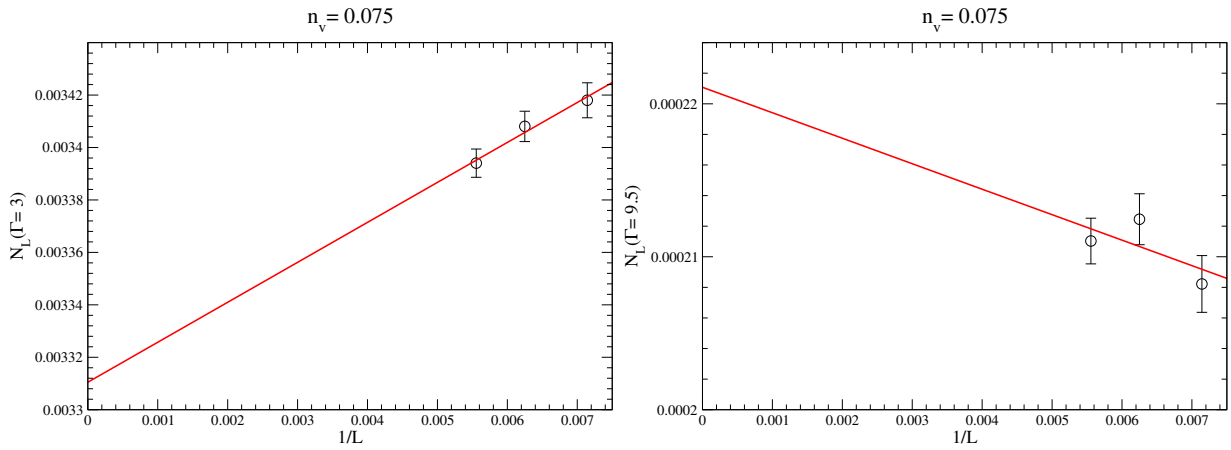


FIG. 10: Examples of extrapolation of $N_L(\Gamma)$ to the thermodynamic limit at $n_v = 0.075$. For this concentration, $\Gamma_c \approx 5$ (see Fig. 4 in the main text), and the left panel illustrates the extrapolation for $\Gamma < \Gamma_c$, while the right panel is for $\Gamma > \Gamma_c$. Note in particular that our extrapolation for $\Gamma > \Gamma_c$ is very likely an overstimate, so one can be fairly confident that $N(\Gamma)$ in the thermodynamic limit drops below $N_{\rm Dyson}(\Gamma)$, ruling out a fit to this form for $\Gamma > \Gamma_c$.

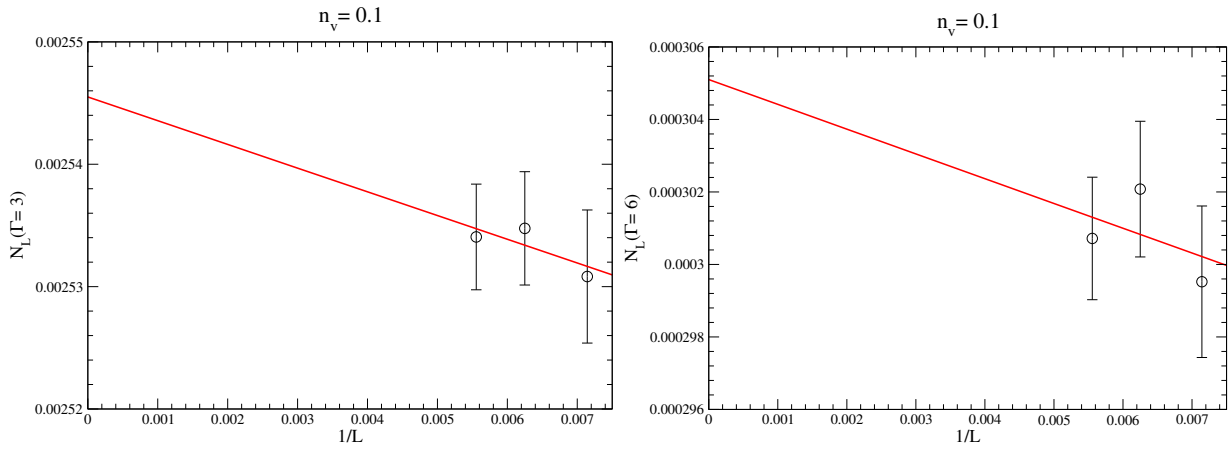


FIG. 11: Examples of extrapolation of $N_L(\Gamma)$ to the thermodynamic limit at $n_v = 0.1$. For this concentration, $\Gamma_c \approx 4$ (see Fig. 4 in the main text), and the left panel illustrates the extrapolation for $\Gamma < \Gamma_c$, while the right panel is for $\Gamma > \Gamma_c$. Note in particular that our extrapolation for $\Gamma > \Gamma_c$ is very likely an overstimate, so one can be fairly confident that $N(\Gamma)$ in the thermodynamic limit drops below $N_{\rm Dyson}(\Gamma)$, ruling out a fit to this form for $\Gamma > \Gamma_c$.

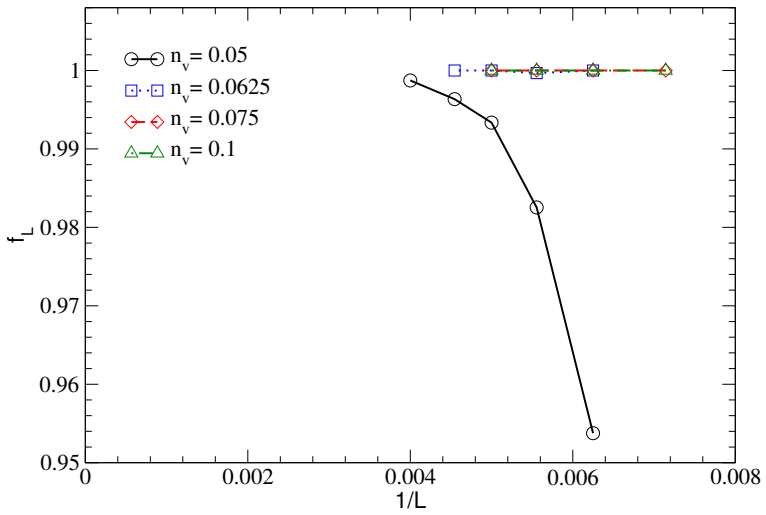


FIG. 12: Probability f_L that an $L \times L$ sample has at least one pair of zero modes tends to 1 in the thermodynamic limit for each concentration studied. Due to our computational constraints, we have been unable to obtain such data at $n_v = 0.02$, where we expect the density of zero modes to be much lower, but f_L to still tend to 1 in the thermodynamic limit (based on the analytical arguments given in the main text).

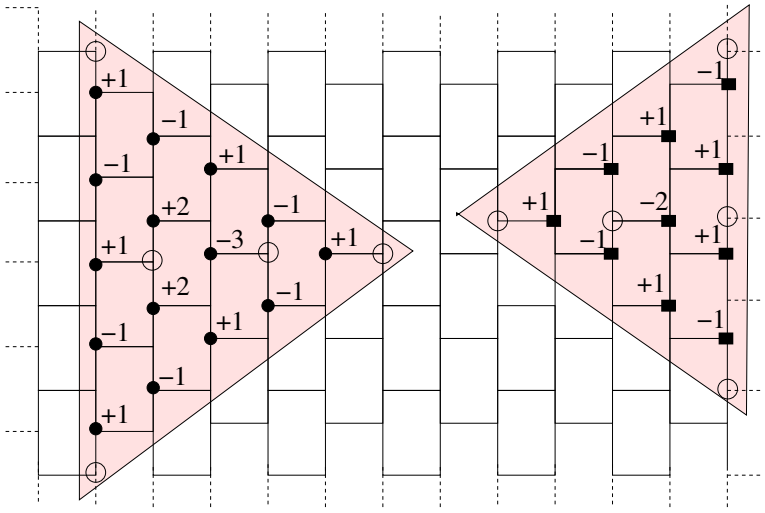


FIG. 13: Two kinds of 5-vacancy clusters ("5-triangles") that host an exact zero mode, with the corresponding wavefunction marked. Open circles correspond to vacancies.

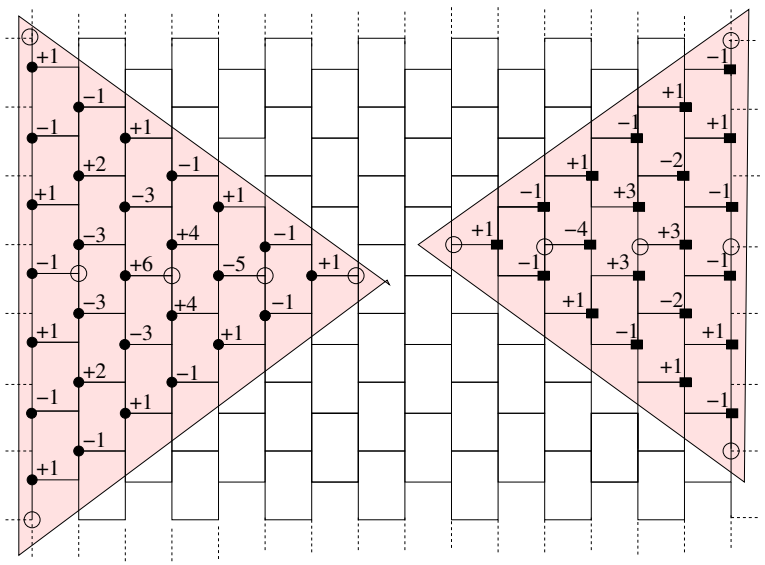


FIG. 14: Two kinds of 6-vacancy clusters ("6-triangles") that host an exact zero mode, with the corresponding wavefunction marked. Open circles correspond to vacancies.

than four impurities and verified that it is not possible to produce a similar zero mode with fewer than four vacancies in a cluster so long as the exclusion constraints outlined in the main text are in place.

The zero mode associated with the $\mathcal{R}_6$ motif described in the main text also generalizes in an obvious way to yield a series of zero modes that all survive the effects of bond disorder in a manner completely analogous to the $\mathcal{R}_6$ zero mode. These $\mathcal{R}_n$ zero modes ($n > 6$) live

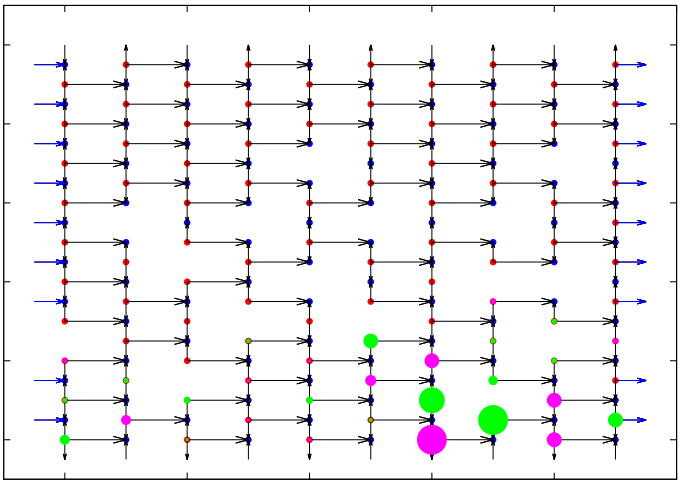


FIG. 15: An $L \times L$ semi-open sample (of the type consistently used in all our numerical work) with $L = 10$, with vacancies represented by missing lattice sites. This sample provides a simple example of a zero mode that does not seem to arise from any of the regular arrangements of vacancies used in our zero mode constructions. The actual wavefunction of this zero mode is represented by color-coded circles. The size of the circle corresponds to the magnitude of the wavefunction at the corresponding site, while the two different colors represent opposite signs for the wavefunction at the corresponding sites.

on larger and larger equilaterial triangles (with zig-zag edges) which are connected to the rest of the lattice only via B (A) sublattice sites but have more undeleted A (B) sublattice sites than B (A) sublattice sites, allowing a zero mode to exist within the triangle for generic realizations of bond-disorder. As in the case of the $\mathcal{R}_6$ zero mode described in the main text, this robustness to disorder follows from the fact that the number of free components of the wavefunction of any such mode is one more than the number of zero-energy equations that they must satisfy.

We have also found other simple examples of such "$\mathcal{R}$-type" zero modes that live near the armchair boundary and are not associated with a specific regular arrangement of vacancies. Instead, as already mentioned earlier, these modes appear to generically live in a region $\mathcal{R}$ which connects to the rest of the lattice only via B (A) sublattice sites belonging to $\mathcal{R}$, although it has more undeleted A (B) sublattice sites than B (A) sublattice sites. In such a region, $T_{AB}T_{AB}^{\dagger}$ ($T_{AB}^{\dagger}T_{AB}$) has a zero mode living on the A (B) sublattice sites, simply

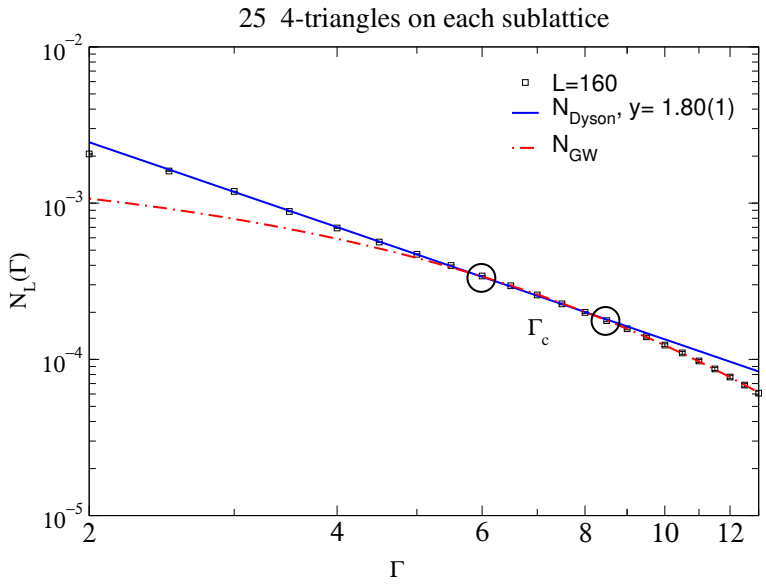


FIG. 16: $N_L(\Gamma)$ in the toy model in which a $L = 160$ sample is diluted with 25 randomly placed 4-triangles on each sublattice. Circles demarcate the crossover region centered at the crossover scale Γ_c. Data for $\Gamma \lesssim \Gamma_c$ fits well to power-law form $N_{\rm Dyson}(\Gamma)$ with the value of y indicated in the figure, while the large-Γ regime fits well to the modified Gade-Wegner form $N_{\rm GW}(\Gamma)$.

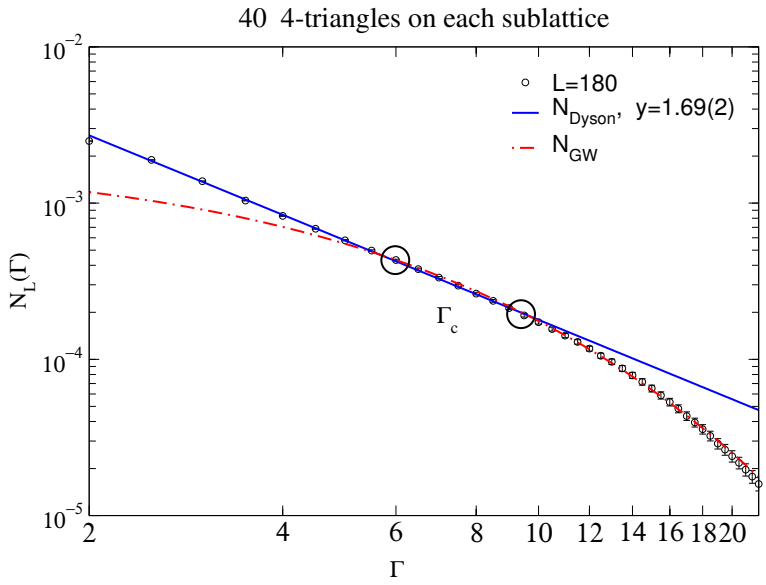


FIG. 17: $N_L(\Gamma)$ in the toy model in which a $L = 180$ sample is diluted with 40 randomly placed 4-triangles on each sublattice. Circles demarcate the crossover region centered at the crossover scale Γ_c. Data for $\Gamma \lesssim \Gamma_c$ fits well to power-law form $N_{\rm Dyson}(\Gamma)$ with the value of y indicated in the figure, while the large-Γ regime fits well to the modified Gade-Wegner form $N_{\rm GW}(\Gamma)$.

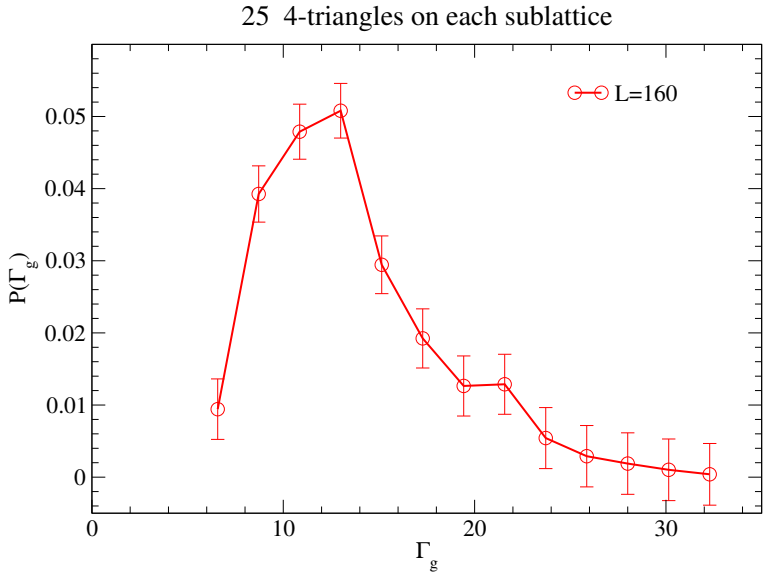


FIG. 18: Histogram of Γ_g, corresponding to the lowest nonzero gap for the $L = 160$ sample diluted with 25 randomly placed 4-triangles on each sublattice.

because the number of constraints that need to be satisfied by this zero mode wavefunction is smaller than the number of A (B) sublattice sites on which this zero mode lives. As already noted, this feature also guarantees that such zero modes survive the effects of disorder in the nearest-neighbour hopping amplitudes. One example of such a mode is shown in Fig. 15. We believe that bulk versions of such more general $\mathcal{R}$-type zero modes provide the dominant contribution to w for the values of n_v studied by us, which is why our lower-bound on w (obtained by thinking in terms of Fig. 2 in the main text) substantially underestimates w at such not-too-small values of n_v. Clearly, no additional local correlations among impurities can entirely eliminate such more general $\mathcal{R}$-type zero modes . Therefore, a non-zero density of zero-energy modes is expected to be a generic feature of such systems. However, we have been unable to convert this observation into an improved lower-bound.

III. DILUTION BY 4-TRIANGLES

Finally, we provide an illustration of the importance of spatial correlations between vacancies via a simple toy model in which vacancies enter the sample only in groups of four, arranged as a 4-triangle at random locations in the sample (as in Fig. 3 of the main text). In Figs. 16 and 17, we respectively display the density of states of $L \times L$ samples with $L = 160$

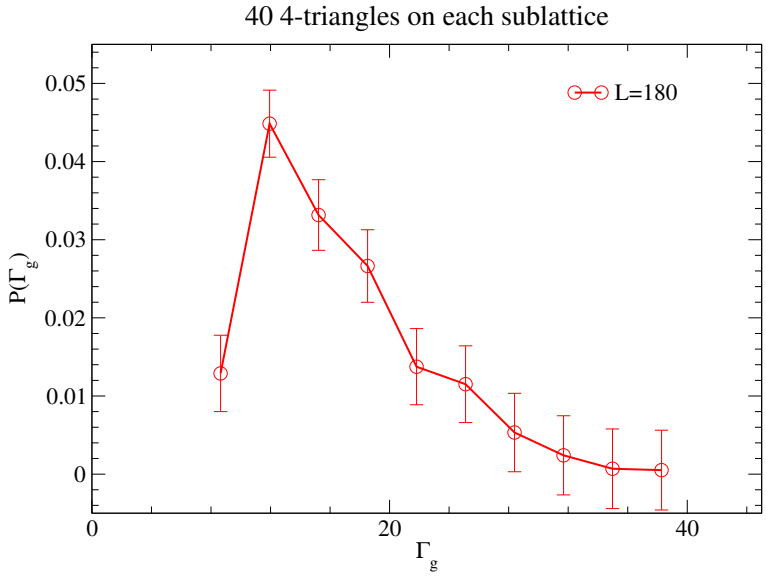


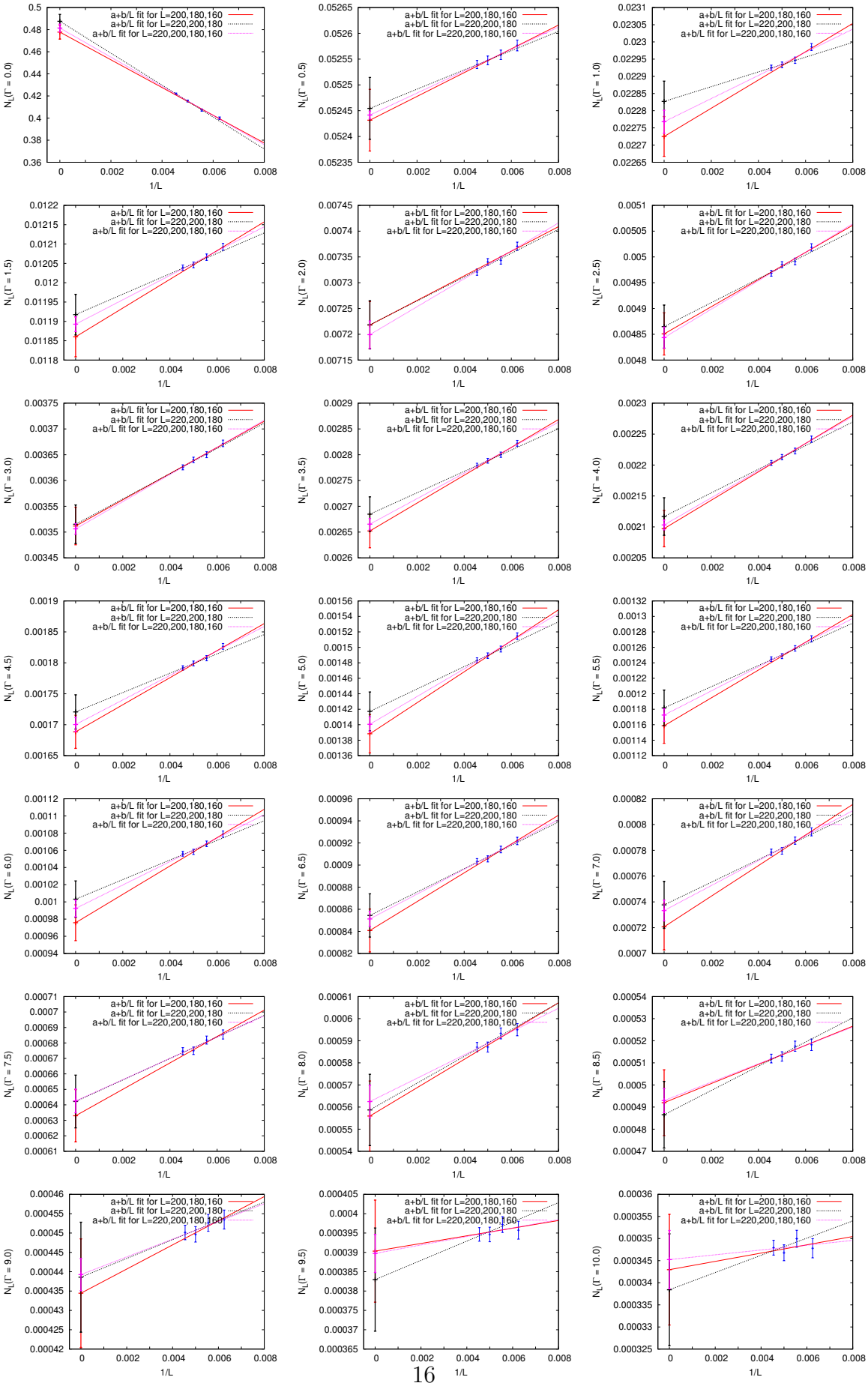
FIG. 19: Histogram of Γ_g, corresponding to the lowest nonzero gap for the $L = 180$ sample diluted with 40 randomly placed 4-triangles on each sublattice.

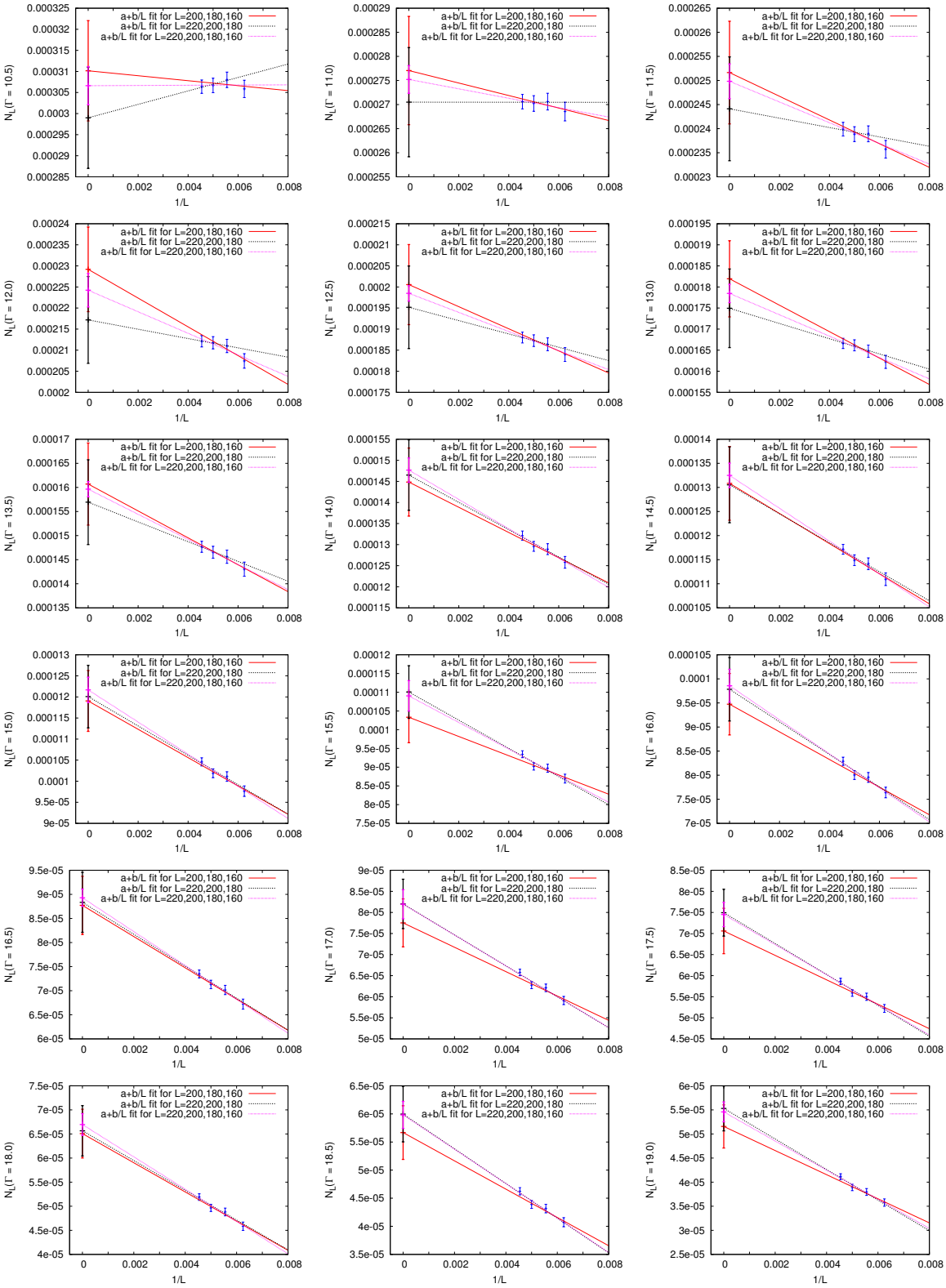
and $L = 180$. The $L = 160$ sample is diluted by 25 4-triangles placed at random on each sublattice, while the $L = 180$ sample is diluted with 40 4-triangles placed at random on each sublattice. The former sample corresponds to a "bare" value of $n_v \approx 0.0039$, while the latter sample corresponds to a bare value of $n_v \approx 0.0049$. These values of n_v are an order of magnitude different from the values of n_v studied by us in the main part of our work (in which the impurities are uncorrelated except for exclusion constraints designed to prevent the occurrence of "trivial" zero modes). However, since all vacancies go in as part of a 4-triangle, $w \approx 9.76 \times 10^{-4}$ for the $L = 160$ sample and $w \approx 1.23 \times 10^{-3}$ for the $L = 180$ sample. The values of w are thus very similar to those obtained in our independently diluted samples with n_v in the range 0.05—0.06.

From Figs. 16 and 17, we see that the density of states again undergoes a crossover that is qualitatively the same as the crossover identified in our main study. However, the corresponding Γ_c is much smaller (*i.e.*, the energy scale $|\epsilon_c|$ is much larger) than one would have expected based on the value of the overall vacancy concentration n_v (had the vacancies been independent as in the main study). Similarly, the value of y is also very different from the (extrapolated) value of y one would have expected at such small n_v. The corresponding histograms of Γ_g are shown in Figs. 18 and 19. From these figures, we see that Γ_g^*, corresponding to the position of the peak in the histogram of Γ_g, is significantly smaller than one

would have expected based on the overall vacancy concentration n_v (had the vacancies been independent, as in the main study). This provides a simple illustration of the importance of spatial correlations between vacancies in setting the lowest gap scale Γ_g^*, and the density of zero modes w. It also emphasizes that the crossover identified by us is a robust and generic aspect of the low-energy physics of vacancy-disorder.

Finally, we note that the values of Γ_c and y in the case of dilution by 4-triangles are apparently predicted much better by the value of w (as opposed to the n_v). This raises the interesting questions already alluded to in the main text: Are Γ_c and y determined in a "universal" way (*i.e.*, independent of short-ranged correlations between vacancies and other such microscopic details) by the value of the zero-mode density w in the limit of small but nonzero w? Can this dependence be understood in terms of a low-energy effective theory or renormalization group approach?


a+b/L fit for L=200,180,160
a+b/L fit for L=220,200,180
a+b/L fit for L=220,200,180,160
N_L(Γ = 0.0)
N_L(Γ = 0.5)
N_L(Γ = 1.0)
N_L(Γ = 1.5)
N_L(Γ = 2.0)
N_L(Γ = 2.5)
N_L(Γ = 3.0)
N_L(Γ = 3.5)
N_L(Γ = 4.0)
N_L(Γ = 4.5)
N_L(Γ = 5.0)
N_L(Γ = 5.5)
N_L(Γ = 6.0)
N_L(Γ = 6.5)
N_L(Γ = 7.0)
N_L(Γ = 7.5)
N_L(Γ = 8.0)
N_L(Γ = 8.5)
N_L(Γ = 9.0)
N_L(Γ = 9.5)
N_L(Γ = 10.0)
1/L


a+b/L fit for L=200,180,160
a+b/L fit for L=220,200,180
a+b/L fit for L=220,200,180,160
1/L
N_L(Γ = 10.5)
N_L(Γ = 11.0)
N_L(Γ = 11.5)
N_L(Γ = 12.0)
N_L(Γ = 12.5)
N_L(Γ = 13.0)
N_L(Γ = 13.5)
N_L(Γ = 14.0)
N_L(Γ = 14.5)
N_L(Γ = 15.0)
N_L(Γ = 15.5)
N_L(Γ = 16.0)
N_L(Γ = 16.5)
N_L(Γ = 17.0)
N_L(Γ = 17.5)
N_L(Γ = 18.0)
N_L(Γ = 18.5)
N_L(Γ = 19.0)

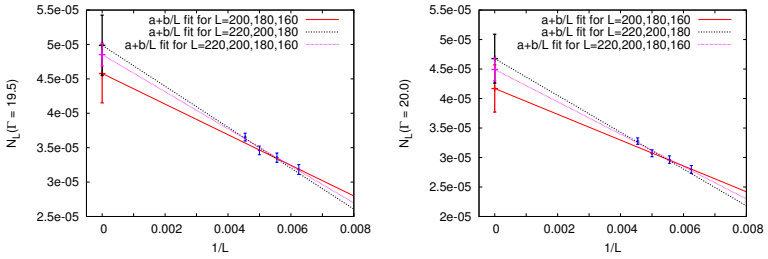


FIG. 20: Comparison of thermodynamic limit of $N_L(\Gamma)$ at $n_v = 0.0625$, taken with and without data at a larger size. At $n_v = 0.0625$, the estimated value of Γ_c is $\Gamma_c \approx 7$. Note that some correlation in the relative ordering of the three extrapolated values is expected over short intervals of Γ since these data are correlated (obtained from the same set of random samples). However, over the range of Γ from $\Gamma = 7$ to $\Gamma = 18$, we already see that there is no consistent ordering of the three extrapolated values, *i.e.*, the red points are not always higher than the black points or vice-versa. Additionally, the three different extrapolations fall within (or lie at the edge of) each other's error bars. Also, the results quoted in the main text (values of Γ_c and y in fits to $N_{\rm Dyson}$ and the quality of the different fits) do not change significantly if our analysis is performed on the thermodynamic limit $N(\Gamma)$ obtained by including data at the larger size. All this, taken together, provides compelling evidence that our extrapolations to the thermodynamic limit are reliable.