Supplementary Material 9: Computational Modeling

9.1 Description of the model for studying the coupling between Ca²⁺ influx and exocytosis

9.1.1 AZ topography: positioning of RRP vesicles and presynaptic Ca²⁺ channels

We considered three main scenarios M1-3 of the mature synapse AZ topography (Fig. 8A in the Results section). There, presynaptic Ca²⁺ channels were treated as 15 nm diameter[8] disks and positioned in the plane of the plasma membrane within the area of the presynaptic density – assumed as a 420×80 nm rectangular stripe. This choice of the presynaptic density area was based on an exemplary reconstruction of the synapse from serial EM sections and is consistent with the STED imaging of $Ca_V1.3$ immunofluorescence (e.g. Fig 3D,G in the Results section). In scenario M1, 36 channels were distributed completely randomly within the stripe. In scenarios M2 and M3, 14 Ca²⁺ channels were placed in contact with the vesicular Ca²⁺ sensors ("private channels"), simulating molecular coupling. In scenario M2, additional 36 channels were distributed randomly within the area of the presynaptic density. We also considered several modifications of M2-3: M2b-d and M3b (Fig. S9). M2b differed from M2 in that there were 76 instead of 36 channels distributed randomly within the presynaptic density, in addition to the 14 "private channels". M2c differed from M2 in that the "private channels" were surrounded by exclusion zones of a channel's diameter wide. M2d differed from M2 in that the Ca²⁺ sensors were shifted by 20 nm towards the center of the vesicles in the plane of the plasma membrane. M3b was identical to M3 except that each vesicle was assigned two "private channels".

In all aforementioned cases, 14 vesicles (40 nm diameter disks) were randomly distributed at the longer sides of the presynaptic density stripe (7 vesicles per side), approximating the number of membrane-proximal, ribbon-associated vesicles, which we assume to represent the readily releasable pool (RRP). In the top view projection of the active zone, vesicles were in contact with the presynaptic density stripe and their Ca²⁺ sensors, located at the level of the plasma membrane, coincided with those contact points. Positions of the channels and the vesicular release sites were fixed throughout a particular simulation. 100 realizations with different positions of the vesicles and randomly placed channels of all those scenarios were used.

For the immature synapse, two scenarios of the AZ, IM1 (Fig. 8) and IM1b (Fig. S9), were considered. In these cases, a presynaptic density area similar to that of M1-3 was distributed within a 1 μm diameter circular area in the form of separate patches, wherein 60 Ca²⁺ channels were distributed with similar average density as in M2. 4 pairs of smaller (45 nm diameter) and 3 single larger (90 nm diameter) density patches were assumed, based on an exemplary reconstruction from serial EM sections. In addition, 60 channels were distributed outside the presynaptic density patches within the 1 μm circular area. In the case of IM1, all additional channels were distributed randomly outside the presynaptic density while in the case of IM1b, 19 of the additional channels were placed in physical contact with the vesicular Ca²⁺ sensors. In both, IM1 and IM1b, vesicles were distributed around the presynaptic density patches and over the remaining AZ as suggested by the EM data. The Ca²⁺ sensors were located at the edges of the vesicles

in such a way that distances from them to the center of mass of the closest presynaptic densities were minimal. 100 realizations with different randomly placed channel positions were considered. Positions of the vesicles and the presynaptic density patches were identical over those 100 realizations.
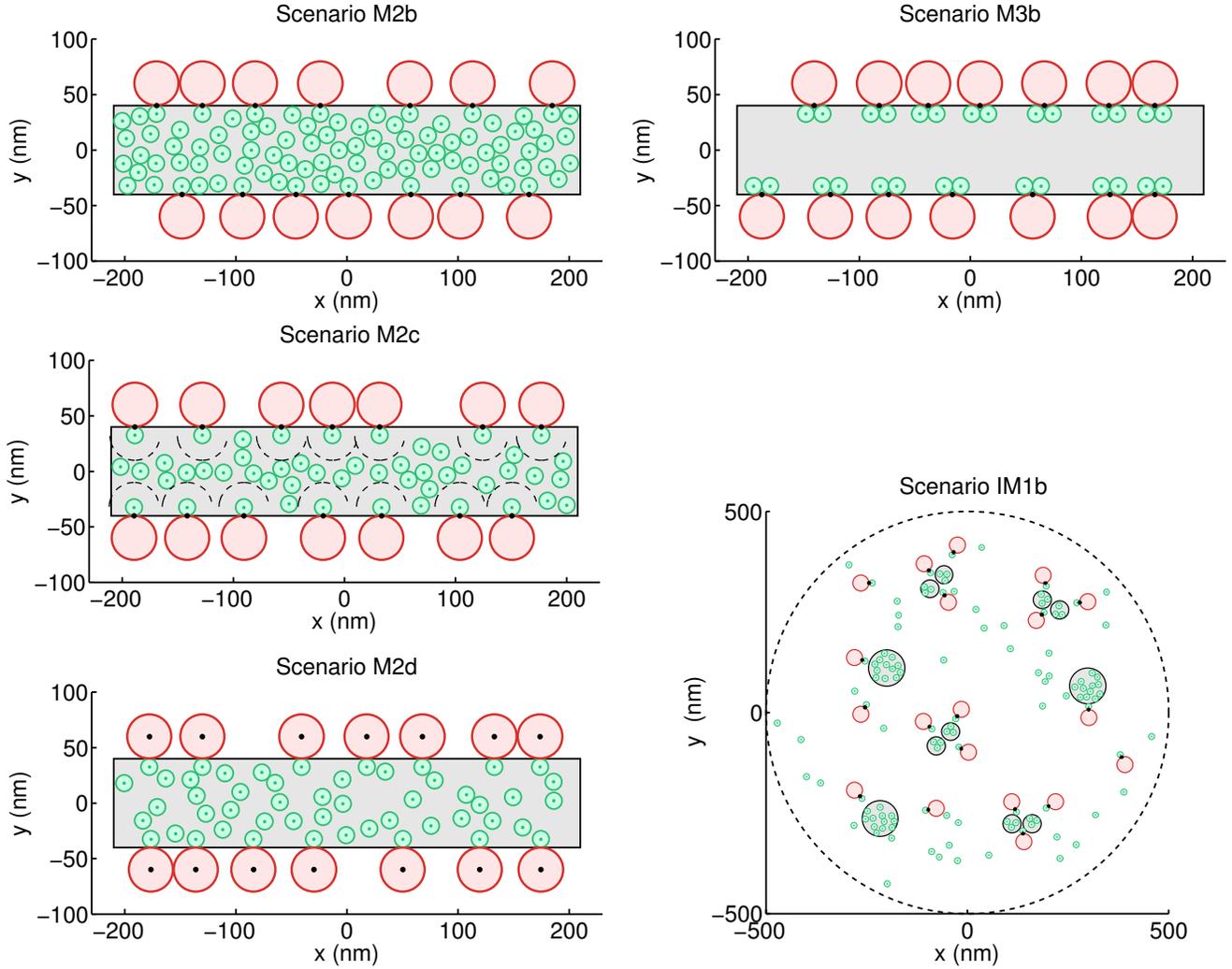

Figure S9: Scenarios M2b, M2c, M2d, M3b of the mature AZ topography and scenario IM1b of the immature AZ topography. Grey area: presynaptic density, red discs: RRP vesicles, green discs: Ca^{2+} channels, black spots: Ca^{2+} sensors of exocytosis.

9.1.2 Ca^{2+} channel gating dynamics

A three state Markov chain model $C_1 \underset{1 \cdot k_{-1}}{\overset{2 \cdot k_{+1}}{\rightleftarrows}} C_2 \underset{2 \cdot k_{-1}}{\overset{1 \cdot k_{+1}}{\rightleftarrows}} O$ was used to describe the gating dynamics of the presynaptic Ca^{2+} channels: $C_{1,2}$ – closed channel states, O – open channel state. Values of the gating rates k_{+1} and k_{-1} were calculated from the activation kinetics reported by ref.9. In particular, it follows from Fig. 1F,G in ref.9 that $\tau = (k_{+1} + k_{-1})^{-1} = 0.45$ ms and $P_o/P_{o,max} = (k_{+1}/(k_{+1} + k_{-1}))^2/P_{o,max} = 0.8$ at membrane potential $V_m = -17$ mV, which is the potential of the depolarizing pulse in our experiments.

16

We assumed $P_{o,max} = 0.8$, as estimated from from the fluctuation analysis of Ca^{2+}-tail currents. These values result in $k_{+1} = 1.78\,\text{ms}^{-1}$ and $k_{-1} = 0.44\,\text{ms}^{-1}$. It should be noted, however, that BayK8644, which increases the channel open probability[10], was used by ref.9 while our aim was to model the channel gating in our current experiments in the absence of BayK8644. Given that the most pronounced effect of BayK8644 is to prolong channel residence times in the open state[10], we increased k_{-1} from $0.44\,\text{ms}^{-1}$ to $1.37\,\text{ms}^{-1}$ such that the open probability was halved. This is consistent with approximately twofold difference of the peak Ca^{2+} current with and without BayK8644[7].

9.1.3 Intracellular $[Ca^{2+}]$ dynamics

$[Ca^{2+}]$ levels in our model were estimated by treating open presynaptic Ca^{2+} channels as Ca^{2+} point sources and by using the linearized buffer approximation[11]. The total flux of Ca^{2+} through an open Ca^{2+} channel at the imposed membrane potential ($V_m = -17\,\text{mV}$) was set to $i_{Ca} = 0.3\,\text{pA}$ for the mature synapse (Fig. 2B, ref.3) and $i_{Ca} = 0.5\,\text{pA}$ for the immature synapse (Fig. 3D, ref.2). We preferred to use directly measured single channel current values rather than their estimate based on the fluctuation analysis of Ca^{2+}-tail currents. 3D Ca^{2+} diffusion was constrained by the reflective plasma membrane but considered unaffected by the synaptic vesicles or the synaptic ribbon. The spatial $[Ca^{2+}]$ profile was assumed to reach steady state instantaneously upon opening and closing of the channels. Overall $[Ca^{2+}]$ at any point of interest was obtained by summing the contributions of separate presynaptic Ca^{2+} channels as is valid in the linearized buffer approximation.

In our model, we considered the following endogeneous Ca^{2+} buffers: calretinin (CR), calbindin (CB), parvalbumin (PV) and ATP. Ca^{2+} and mobile Ca^{2+} buffer concentrations, their binding and unbinding rates, and diffusion coefficients are given in Table S9. Diffusion coefficients of Ca^{2+} bound buffer molecules were assumed to be the same as the diffusion coefficients of the corresponding Ca^{2+}-free buffer molecules. We assumed cooperative Ca^{2+} binding to CR and noncooperative binding to the other three buffers considered. Based on ref.12, we assumed that CR consists of one single binding site (CR1) and two identical, independent pairs of cooperative binding sites (CR2). In Table S9, $CR2_T$ stands for the Ca^{2+} free CR2 while $CR2_R$ stands for CR2 with one Ca^{2+} ion already bound.

Calculation of the steady state levels of Ca^{2+} and Ca^{2+} bound buffer species in our case followed exactly the same path as in ref.11 except that the Jacobian matrix A (Eq. 8 in ref.11) which corresponds to the equations of the Ca^{2+}-buffer reaction dynamics had a different form due to the cooperative Ca^{2+} binding of CR. In our particular case, matrix A was defined as follows:

$$A = \begin{pmatrix} -a_1 & b_1 & 0 & 0 & 0 & 0 & c_1 \\ b_2 & -a_2 & 0 & 0 & 0 & 0 & c_2 \\ 0 & 0 & -a_3 & 0 & 0 & 0 & c_3 \\ 0 & 0 & 0 & -a_4 & 0 & 0 & c_4 \\ 0 & 0 & 0 & 0 & -a_5 & 0 & c_5 \\ 0 & 0 & 0 & 0 & 0 & -a_6 & c_6 \\ a_1 - b_2 & a_2 - b_1 & a_3 & a_4 & a_5 & a_6 & -\sum_{i=1}^{6} c_i \end{pmatrix}. \tag{1}$$

17

Here, index 1 corresponds to $CR2_T$, index 2 corresponds to $CR2_R$, and $a_{1-2}, b_{1-2}, c_{1-2}$ are defined as follows:

$$a_1 = (k_{off,CR2_T} + 2k_{on,CR2_T}[\text{Ca}^{2+}]_{rest}),$$

$$a_2 = 2k_{off,CR2_R},$$

$$b_1 = -2k_{on,CR2_T}[\text{Ca}^{2+}]_{rest},$$

$$b_2 = k_{on,CR2_R}[\text{Ca}^{2+}]_{rest},$$

$$c_1 = \frac{2k_{on,CR2_T}[\text{CR2}]}{1 + 2k_{on,CR2_T}/k_{off,CR2_T}[\text{Ca}^{2+}]_{rest} + (k_{on,CR2_T}k_{on,CR2_R})/(k_{off,CR2_T}k_{off,CR2_R})[\text{Ca}^{2+}]_{rest}^2},$$

$$c_2 = \frac{2(k_{on,CR2_T}k_{on,CR2_R})/k_{off,CR2_T}[\text{CR2}][\text{Ca}^{2+}]_{rest}}{1 + 2k_{on,CR2_T}/k_{off,CR2_T}[\text{Ca}^{2+}]_{rest} + (k_{on,CR2_T}k_{on,CR2_R})/(k_{off,CR2_T}k_{off,CR2_R})[\text{Ca}^{2+}]_{rest}^2}.$$

Indexes 3-6 correspond to CR1, CB, PV, and ATP. a_{3-6}, c_{3-6} correspond to $1/\tau_i$ and κ_i/τ_i defined in ref.11:

$$a_i = 1/\tau_i = (k_{off,i} + k_{on,i}[\text{Ca}^{2+}]_{rest}),$$

$$c_i = \kappa_i/\tau_i = \frac{k_{on,i}k_{off,i}[\text{B}]}{k_{off,i} + k_{on,i}[\text{Ca}^{2+}]_{rest}}.$$

Calculation of $[\text{Ca}^{2+}]$ at any point of interest was based on the diagonalization of the matrix $(\mathbf{D} \cdot \mathbf{A})$, where $\mathbf{D}$ is a diagonal matrix with the diffusion coefficients of the corresponding molecules as in $\mathbf{A}$ put on the diagonal $\mathbf{D} = \mathbf{diag}[D_{CR}, D_{CR}, D_{CR}, D_{CB}, D_{PV}, D_{ATP}, D_{Ca}]$. In particular:

$$[\text{Ca}^{2+}](t) = \sum_{j=1}^{N_{Ca}} s_j(t)/r_j \sum_{i=1}^{7} k_i u_{7,i} e^{-\sqrt{\mu_i}r_j}. \qquad (2)$$

Here, r_j is the distance from the point of interest to the center of the j-th channel and $s_j(t)$ stands for the state of the j-th channel at moment t: $s_j = 0$ if the channel is closed and $s_j = 1$ if the channel is open. $u_{7,i}$ is the seventh component of the i-th eigenvector of $(\mathbf{D} \cdot \mathbf{A})$ and μ_i is the corresponding eigenvalue of $(\mathbf{D} \cdot \mathbf{A})$, while $\{k_i\}_{i=1}^7$ is the solution of $\{\sum_{i=1}^{7} u_{j,i}k_i = i_{Ca}/(4\pi D_{\text{Ca}^{2+}})\delta_{j,7}\}_{j=1}^7$.
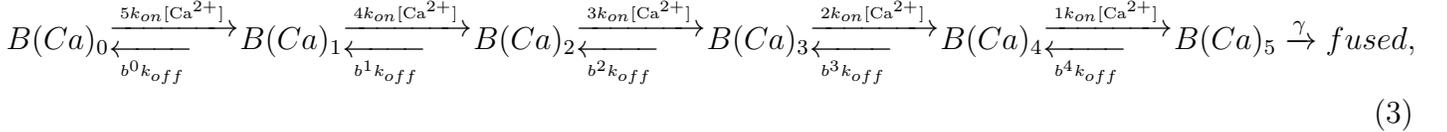
The free ATP concentration was calculated by assuming equilibration of Mg^{2+} to ATP binding with free cellular concentration of Mg^{2+} equal to 0.5 mM[19]. The overall ATP concentration was set to 2 mM, and $K_{d,ATP} = 45\,\mu M$[16]. We did not explicitly take the competition between Mg^{2+} and Ca^{2+} for ATP into account, because $[Mg^{2+}] = 0.5\,mM$ is much higher than $[\text{Ca}^{2+}]$ at the release sites, which leads to only negligible fluctuations of free ATP level around its mean value upon Ca^{2+} channel opening and closing.

	C, mM	k_{on}, $(\mathrm{mM} \cdot \mathrm{ms})^{-1}$	k_{off}, ms^{-1}	D, $\mu\mathrm{m}^2 \cdot \mathrm{ms}^{-1}$
$\mathrm{CR2_T}$	$3.6 \cdot 10^{-2}$ $(4.0 \cdot 10^{-2})$, ref.14	1.8, ref.12	$5.3 \cdot 10^{-2}$, ref.12	$2.0 \cdot 10^{-1}$, ref.12
$\mathrm{CR2_R}$		$3.1 \cdot 10^{2}$, ref.12	$2.0 \cdot 10^{-2}$, ref.12	
CR1	$1.8 \cdot 10^{-2}$ $(2.0 \cdot 10^{-2})$, ref.14	7.3, ref.12	$2.52 \cdot 10^{-1}$, ref.12	$2.0 \cdot 10^{-1}$, ref.12
CB	$2.32 \cdot 10^{-1}$ (1.7), ref.14	$7.5 \cdot 10^{1}$, ref.13	$2.95 \cdot 10^{-2}$, ref.13	$2.0 \cdot 10^{-1}$, ref.17
PV	$1.88 \cdot 10^{-1}$ $(1.38 \cdot 10^{-1})$, ref.14	$1.08 \cdot 10^{2}$, ref.15	$9.8 \cdot 10^{-4}$, ref.15	$4.3 \cdot 10^{-1}$, ref.18
ATP (free)	$1.65 \cdot 10^{-1}$	$1.0 \cdot 10^{3}$, ref.16	$9.0 \cdot 10^{1}$, ref.16	2.0, ref.11
$\mathrm{Ca^{2+}}$ (rest)	$5.0 \cdot 10^{-5}$	$-$	$-$	2.0, ref.11

Table S9 $\mathrm{Ca^{2+}}$ and mobile $\mathrm{Ca^{2+}}$ buffer binding site concentrations (C), their binding (k_{on}) and unbinding (k_{off}) rates, and diffusion coefficients (D) used in the model. Concentration values in the round brackets correspond to immature IHCs, if they differ from the mature IHCs. All concentration values were taken as reported by ref.14 for apical IHCs, which (in mice) were also used for our experiments. Note that the concentration entry for $\mathrm{CR2_T}$ and $\mathrm{CR2_R}$ refers to the concentration of CR2 binding site pairs.

9.1.4 Fusion and replenishment dynamics of RRP vesicles

In our simulations, $\mathrm{Ca^{2+}}$ triggered fusion of the RRP vesicles following the kinetic model proposed by ref.20:

$$B(Ca)_0 \underset{b^0 k_{off}}{\overset{5k_{on}[\mathrm{Ca^{2+}}]}{\rightleftarrows}} B(Ca)_1 \underset{b^1 k_{off}}{\overset{4k_{on}[\mathrm{Ca^{2+}}]}{\rightleftarrows}} B(Ca)_2 \underset{b^2 k_{off}}{\overset{3k_{on}[\mathrm{Ca^{2+}}]}{\rightleftarrows}} B(Ca)_3 \underset{b^3 k_{off}}{\overset{2k_{on}[\mathrm{Ca^{2+}}]}{\rightleftarrows}} B(Ca)_4 \underset{b^4 k_{off}}{\overset{1k_{on}[\mathrm{Ca^{2+}}]}{\rightleftarrows}} B(Ca)_5 \overset{\gamma}{\rightarrow} fused,$$

$$(3)$$

where $B(Ca)_n$ denotes the state of an occupied vesicle release site and n $\mathrm{Ca^{2+}}$ ions bound to the $\mathrm{Ca^{2+}}$ sensor, while $fused$ denotes an empty vesicle release site after a vesicle fusion. Values of the kinetic parameters were taken as reported by ref.20: $k_{on} = 27.6\,\mathrm{mM^{-1}ms^{-1}}$, $k_{off} = 2.150\,\mathrm{ms^{-1}}$, $b = 0.4$, $\gamma = 1.695\,\mathrm{ms^{-1}}$, except that, for the model realizations shown in the Fig. 8 of the Results section, we used two times smaller k_{on} due to reasons discussed below (section 9.2).

Refilling of the empty vesicle release sites during the ongoing stimulation was considered as a single step process with a fixed rate k_{rep}. Mathematically, this was implemented by adding an additional step $fused \overset{k_{rep}}{\rightarrow} B(Ca)_0$ to the kinetic scheme (3). We set $k_{rep} = 0.13\,\mathrm{ms^{-1}}$. This choice was based on the sustained release rate reported in ref.21. According to that study (Table 1, ref.21), each active zone releases 679 vesicles per second in the sustained phase, which is equivalent to 57 vesicles per second per release site (assuming 12 release sites per active zone). This value is the lower limit of the replenishment rate, which cannot be lower than the release rate. Moreover, it is likely that the replenishment rate is higher in a time window before the steady state is reached because number of the RRP surrounding vesicles, which are used to replenish the RRP, is likely to be higher at initial time moments. To take these effects into account, we arbitrarily set k_{rep} two times larger than the vesicle release rate in steady state reported in ref.21. Increase or decrease of k_{rep} by two times from that value did not have significant influence on the main aspects of the modelling results considered in this study.

9.1.5 Propagation of the dynamics of the system across time

Initially, all the channels were set to be in the C_1 closed state and all the vesicle release sites were filled with vesicles, each in state $B(Ca)_0$. This corresponds to the experimental conditions in which cells were kept at $V_m = -87\,\text{mV}$, with negligible $[Ca^{2+}]$ levels at the active zone. After changing V_m to -17 mV at $t = 0$, gating dynamics of each Ca^{2+} channel was propagated across time by using Gillespie's method[22]. Vesicle fusion-replenishment dynamics was propagated by using an extended version of the Gillespie's method, which takes into account the time dependence of the transition rates. Briefly, at iteration i, the amount of time for a vesicle at a particular release site to stay in its current state a was estimated by generating a uniformly distributed random variable $u \in (0, 1)$ and solving the following equation for t_{i+1}

$$\int_{t_i}^{t_{i+1}} \left(\sum_{b=1}^{N_{a\rightarrow}} k_{a\rightarrow b}(t) \right) dt = -\ln(u), \tag{4}$$

where $k_{a\rightarrow b}(t)$ stands for the transition rate from state a to state b at time t and $N_{a\rightarrow}$ is the number of possible transitions from state a. t_i and t_{i+1} are the times when the vesicle entered and left state a. It follows from (2) and (3) that the transition rates $k_{a\rightarrow b}(t)$ are piecewise constant functions of time. (4) was solved numerically exactly. To decide to which one of the possible final states the system moves at t_{i+1}, the unit interval $[0, 1]$ was divided into $N_{a\rightarrow}$ subintervals, each corresponding to one of the possible final states, with lengths $k_{a\rightarrow b}(t_{i+1})/\sum_{b=1}^{N} k_{a\rightarrow b}(t_{i+1})$. The final state was then chosen by generating a uniformly distributed random variable $u \in [0, 1]$ which pointed to one of those subintervals. The vesicle fusion time was defined as the moment when $B(Ca)_5$ transited to *fused*. The spatial position of the vesicle release sites remained constant over time, so that the sensor of a newly replenished vesicle was taking the same position as the sensor of the previously released vesicle.

9.1.6 Estimation of the release vs. Q_{Ca} dependencies

To model the outcome of the channel blocking experiment, we randomly blocked a certain number of channels (in the range from 0 to $N_{Ca} - 1$) for each scenario studied. For each number of blocked channels, we went through 100 different combinations of blocked channels. Simulations with synapse responses for each of those 100 blocked channel combinations were repeated 10 times. The same set of calculations was performed with 100 different realizations of each synapse topography scenarios generated as discussed above (section 9.1.1). Integrated Ca^{2+} influx (Q_{Ca}) and number of released vesicles in response to 20 ms depolarization were calculated and averaged over (# of the synapse realizations) · (# of the blocked channel combinations) · # of repetitions) $= 10^2 \cdot 10^2 \cdot 10 = 10^5$ simulations for each number of blocked channels.

To simulate the i_{Ca} manipulation experiment, we scaled down the single channel current (0.3 pA for mature synapses and 0.5 pA for immature synapses) by a factor of $f_s \in [1, N_{Ca}]$. Q_{Ca} and number of released vesicles in response to 20 ms voltage steps were calculated and averaged over (# of the synapse realizations) · (# of repetitions) $= 10^2 \cdot 10^3 = 10^5$ simulations for each value of the single channel current scaling factor.

The estimates of the exponent m of the release vs. Q_{Ca} relationship under $N_{Ca} \times P_{open}$ manipulation were determined by fitting the relationship in a double logarithmic plot with a line from the maximum

Q_{Ca} down to one-fifth of its value, similar to the that achieved experimentally (20% residual current with $10\,\mu$M isradipine). For i_{Ca} manipulation, the exponent m was determined in the following way. The line fitting started by using the first 5 points corresponding to released vesicle number bigger than 10^{-4}. Then, additional points for fitting were included until the estimated value of the exponent m decreased more than 5% compared to the initial one (based on the first five points). This limitation was required to avoid underestimations of m due to the saturation of the release vs. Q_{Ca} relationship (see section 9.2).

9.1.7 Considerations on the simplifications used in the model

The linearized buffer approximation in general underestimates $[Ca^{2+}]$ as it effectively corresponds to constant free buffer levels independent of Ca^{2+} coming from the sources. The overall mobile buffer concentration is indeed high (0.7 mM to 2 mM binding sites) in our case, which enables substantial concentrations of Ca^{2+} free buffer at distances greater than the Ca^{2+} channel radius. Assuming instantaneous equilibration of $[Ca^{2+}]$ upon the channel opening and closing, on the other hand, leads to an overestimation of the $[Ca^{2+}]$. However, this overestimation is relatively substantial only for those Ca^{2+} sources that are located far from the point of interest (see ref.11). The absolute contribution of these faraway channels to the $[Ca^{2+}]$ at the release sites is altogether small for the AZ topographies considered in our study. Thus, $[Ca^{2+}]$ overestimation due to the assumption of instantaneous $[Ca^{2+}]$ equilibration upon the channel opening and closing can be ignored.

9.2 Further analysis of the release vs. Ca^{2+} influx dependencies

9.2.1 The saturation effect and its origins

A pronounced saturation in the release vs. Q_{Ca} dependencies was observed for some of the mature synapse scenarios considered in our study when strictly adhering to the model parameters as provided in the cited references. The saturation effect was particularly evident for the i_{Ca} manipulation when the k_{on} value reported by ref.20 was used. For example, in the case of scenario M2c, the relationship becomes linear in the a double log plot only when Q_{Ca} drops fivefold below $\max[Q_{Ca}]$ based on i_{Ca} manipulation (red curve Fig. S10, right panel). For Q_{Ca} change based on $N_{Ca} \times P_{open}$ manipulation, the saturation was smaller but still influential. For example, in the case of scenario M2c, the release vs. Q_{Ca} relationship shows saturation for the higher Q_{Ca} and deviates considerably from a linear relationship in a double log plot in the range from $\max[Q_{Ca}]/5$ to $\max[Q_{Ca}]$ (red curve Fig. S10, left panel). This leads to an underestimation of the exponent m which reflects the apparent Ca^{2+} cooperativity.

The origin of the pronounced saturation in the release vs. Q_{Ca} dependencies is fast initial RRP depletion (in respect to the width of the 20 ms observation time window) due to much higher vesicle fusion rate compared to the replenishment rate. Indeed, as shown by the red curve in Fig. S11, which represents scenario M2c, the exocytosis rate is so high that most of the initial RRP vesicles fuse within 6 ms after the stimulus onset. Then, the system enters steady state, during which exocytosis rate ≈ 0.1 vesicle/ms per release site is limited by the replenishment step. During the next 14 ms, around 1.4 vesicles fuse per release site. In other words, around half of all fusion events observed in the 20 ms
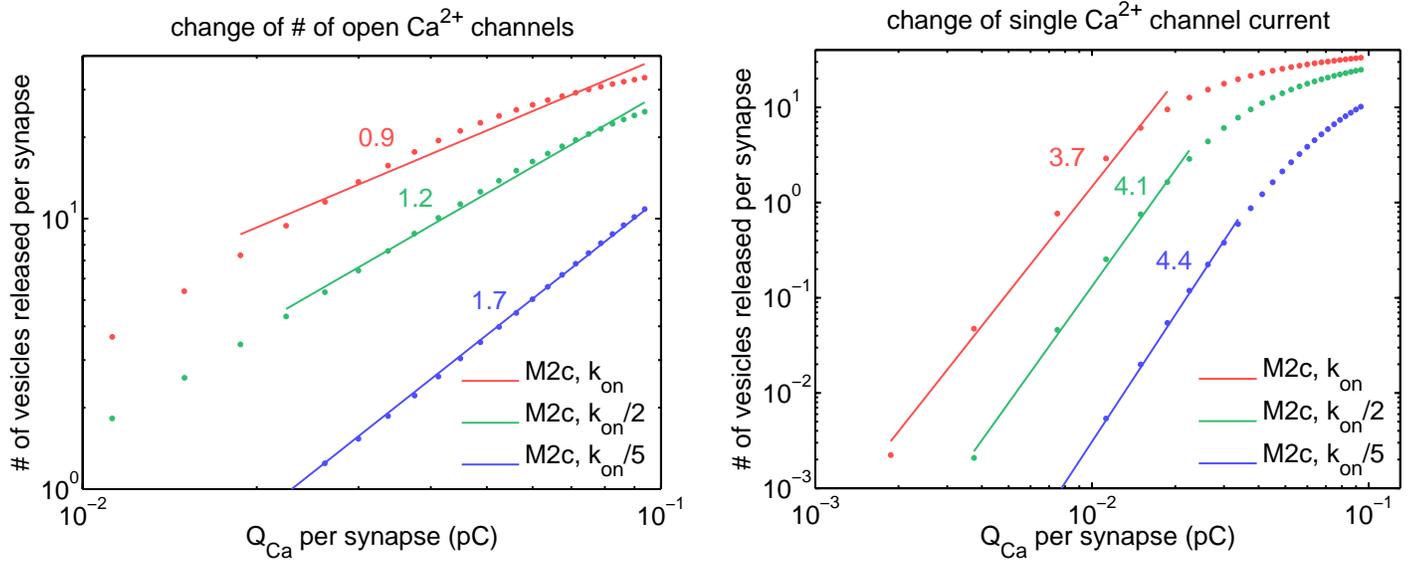
Figure S10: Release vs. Q_{Ca} 20 ms after the stimulus onset for scenario M2c with different scaling of k_{on} values. Left – $N \times P_{open}$ manipulation, right – i_{Ca} manipulation.
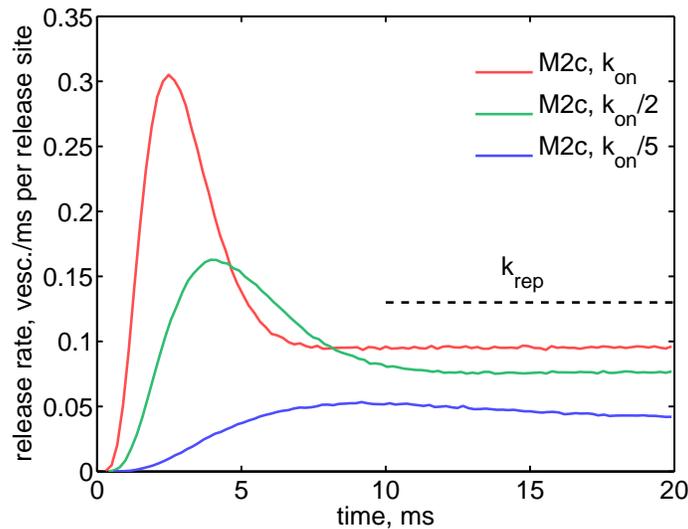

Figure S11: Vesicle release rate evolution after the stimulus application for scenario M2c with different scaling of k_{on} values. Dashed black line indicates to the replenishment rate k_{rep} value on the ordinate axis.

window happen within the first 6 ms. This asymmetry in the exocytosis rate results in a relatively low decrease in overall release upon the reduction of the Q_{Ca}. For example, when Q_{Ca} is reduced two times by reducing i_{Ca} (which is equivalent to reducing k_{on} two times, green curve in Fig. S11), the RRP is depleted in 10 ms on average. During the next 10 ms, 0.8 vesicles fuse per release site on average. Thus, despite the fact that Ca^{2+} is reduced two times, the overall release 20 ms after the stimulus application decreases only $2.4/1.8 \approx 1.3$ times. Note that, in all cases studied here, vesicle fusion rate was not limited by γ, the transition rate from $B(Ca)_5$ to $fused$. Another, only a partial, source of the saturation is based on the fact that vesicle replenishment is not only the rate limiting step, (compare the steady release rates corresponding to the red curve with that showed by the black dashed line) but is also assumed to be Ca^{2+} independent.

When Q_{Ca} is reduced sufficiently, the saturation becomes negligible. Then, the sensitivity of the release to change of Ca^{2+} reflects the real Ca^{2+} influx to exocytosis coupling (Fig. S10). Note, however, that m values based on $N_{Ca} \times P_{open}$ change approach 1 by definition when only a few channels remain unblocked. Thus, data for the lowest Q_{Ca} have to be avoided. In contrast, m values based on i_{Ca} manipulation remain unchanged even at strongly reduced Q_{Ca}, assuming that there is no noticeable saturation in the region of interest. That is because the relative $[Ca^{2+}]$ contribution to the vesicle sensors of a particular Ca^{2+} channel in respect to the rest of the Ca^{2+} channels remains unchanged during this manipulation and, thus, the sensitivity reflects the intrinsic Ca^{2+} dependence of exocytosis.

There are two ways to reduce the saturation effect in the release vs. Q_{Ca} plots without changing the AZ topography or the observation time window. The first way is to increase the replenishment rate so that it becomes much higher than the fusion rate and the vesicle pool stays nearly full throughout the depolarizations. This, however, would strongly deviate from the physiological situation. The replenishment rate should be increased by an order of magnitude or more to achieve a considerable improvement. The second way to reduce the saturation effect is to decrease the fusion rate. This could be achieved by decreasing k_{on} or, equivalently, by decreasing the $[Ca^{2+}]$ levels, which could be, for example, due to a stronger buffering or smaller single channel currents. The effect of the reduction of k_{on} is shown in Fig. S10. We see that the release vs. Q_{Ca} relationship in a double log plot linearises in the range $\max[Q_{Ca}]/5$ to $\max[Q_{Ca}]$ in the case of $N_{Ca} \times P_{open}$ manipulation when k_{on} is decreased (compare the red, green, and blue plots). This leads to a considerable increase in the exponent m estimate. In the case of i_{Ca} manipulation, the range of Q_{Ca} with a constant slope of the release vs. Q_{Ca} relationship in a double log plot increases considerably when k_{on} decreases. Nevertheless, the exponent m estimates based on the fitting to the linear part of the relationship is ~ 4 in all the cases considered. This is how it should be, because release vs. Q_{Ca} relationships based on i_{Ca} manipulation reflect the intrinsic Ca^{2+} dependence of exocytosis, not the Ca^{2+} influx to exocytosis coupling.

It is clearly seen by comparing plots in the left and the right panels in Fig. S10 that the saturation effect is more pronounced in the case of i_{Ca} manipulation than $N_{Ca} \times P_{open}$ manipulation. The explanation of this phenomena is based on disproportionate contributions of different AZ Ca^{2+} channels to vesicle exocytosis rates at different release sites. This is seen the best by considering scenario M3. If we block half of the channels in that case, we effectively make the release nearly absent at half of the sites.

On the contrary, reduction of the single channel current i_{Ca} by two times only decreases the fusion rate, which is otherwise very high, at all sites. This reduction delays the pool depletion only by a few ms which is little compared to 20 ms width observation window. Moreover, the exocytosis rate of the subsequent replenished vesicles is limited by the replenishment, and is considerably lower than the depletion rate of the initial RRP. In this way, the delay of the pool depletion by a few ms due to the reduced i_{Ca} has a smaller effect than the absence of vesicle release at a part of the release sites due to the equivalent $N_{Ca} \times P_{open}$ reduction.

For most of the active zone scenarios considered in our simulations, we chose a two times reduced k_{on} value ($k_{on}/2$) compared to that reported by ref.20. This helped to avoid a very pronounced saturation effect, which would result in unrealistic sublinear release vs. Q_{Ca} relationships (described above). This choice is rather arbitrary and does not necessarily reflect the degree of the saturation observed in the experiments. It is important to note that the $\log(Q_{Ca})$ range available for m estimates based on i_{Ca} change was wide in the case of the model. This guaranteed that these m estimates were not noticeably affected by the saturation. If we limited the analysis of the modeling results to the same width of $\log(Q_{Ca})$ range as in the case of experimental data, the modeling m estimates based on i_{Ca} change (20 ms time window, $k_{on}/2$) would have been even smaller than the experimental ones. This potentially means that the effect of the saturation was more pronounced in the model (20 ms time window, $k_{on}/2$) than the experiment. Note that our modeling results (20 ms time window, $k_{on}/2$) showed that, among all the scenarios considered, experimental estimate $m = 1.4$ based on $N_{Ca} \times P_{open}$ change could be reproduced only in the case of the nanodomain coupling between the Ca^{2+} influx and exocytosis. If this is true with the considered level of saturation in the model, this would be true with a lower level of the saturation too. Below, we provide a more detailed analysis of the effect of the saturation on the exponent m estimates for the considered AZ scenarios to constrain possible interpretations of experimental m estimates in terms of the exocytosis to Q_{Ca} influx coupling.

9.2.2 Assessing the influence of the saturation effect

A parameter-independent way to assess the influence of the saturation effect on the exponent m estimates and reveal the real Ca^{2+} influx to exocytosis coupling is to use very short observation windows. This is not possible to achieve in the experiments due to low signal to noise ratio, but can be conveniently done in the modeling approach. In Fig. S12, release vs. Q_{Ca} dependencies based on 3 ms observation windows are shown for cases k_{on} and $k_{on}/2$ for scenario M2c. The relationships based on $N_{Ca} \times P_{open}$ change are, to a good approximation, linear in double log plots in the range $\max[Q_{Ca}]/5$ to $\max[Q_{Ca}]$ (left panel). The saturation is rather mild for the dependencies based on i_{Ca} change (right panel). Moreover, the estimates of the m exponents are quite similar to those in the case of strongly reduced exocytosis rate ($k_{on}/5$ in Fig. S10), as expected.

It should be noted that one of the possible concerns with using short time windows is that the estimate of the exponent m can be influenced by the gating kinetics of the Ca^{2+} channels. However, in our case, the characteristic equilibration time of the channel gating is 0.45 ms, which is much shorter than the 3 ms length observation window.
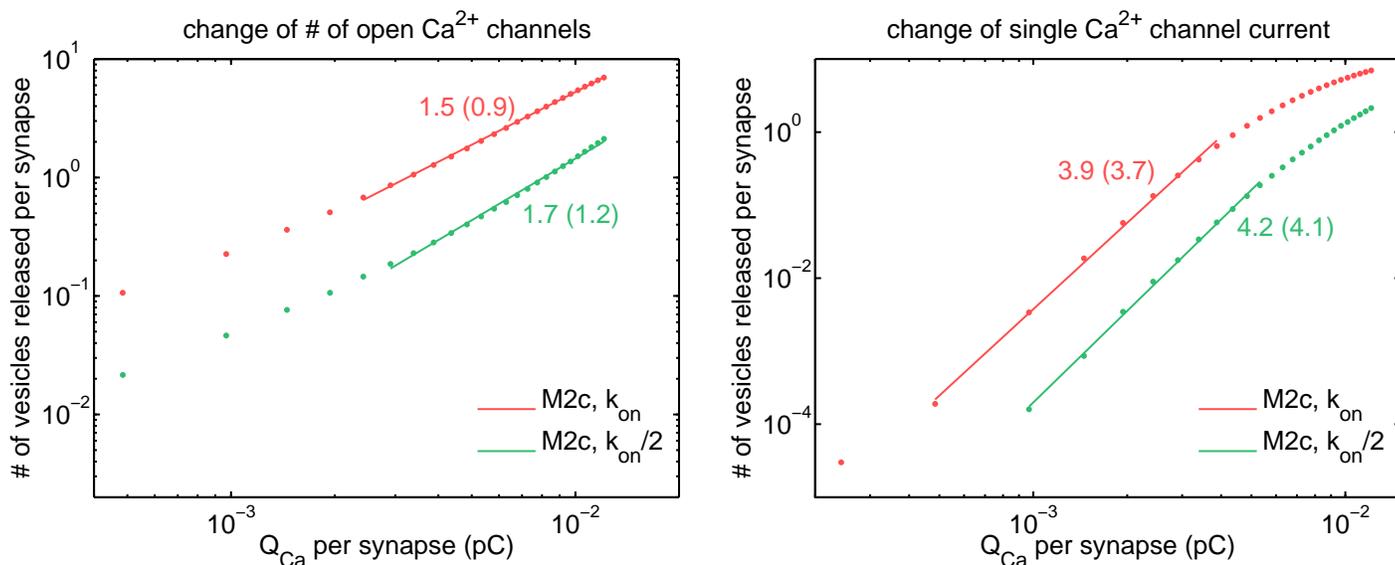
Figure S12: Release vs. Q_{Ca} 3 ms after the stimulus application for scenario M2c with different scaling of k_{on} values. Left – $N \times P_{open}$ manipulation, right – i_{Ca} manipulation. Values in the round brackets correspond to the exponent m estimated from the responses 20 ms (Fig. S11), rather than 3 ms, after the stimulus onset.
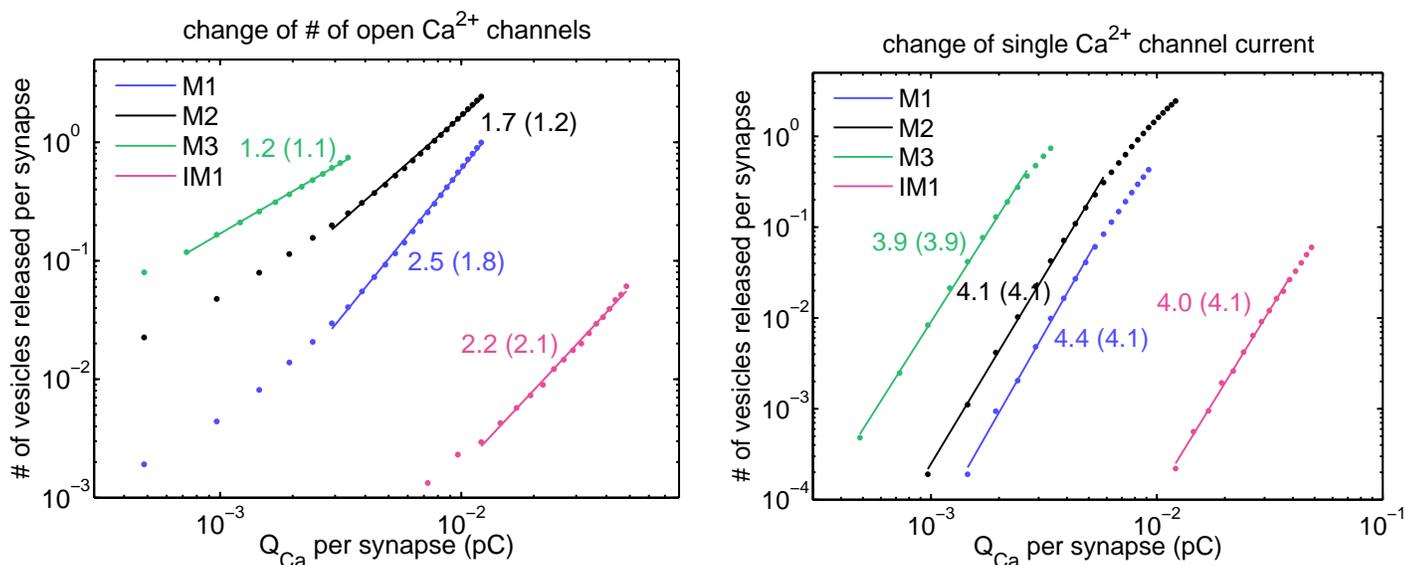

Figure S13: Release vs. Q_{Ca} 3 ms after the stimulus application for scenarios M1-3 and IM1. Left – $N \times P_{open}$ manipulation, right – i_{Ca} manipulation. Values in the round brackets correspond to the exponent m estimated from the responses 20 ms (Fig. 8 E,F), rather than 3 ms, after the stimulus onset. $k_{on}/2$ was assumed.
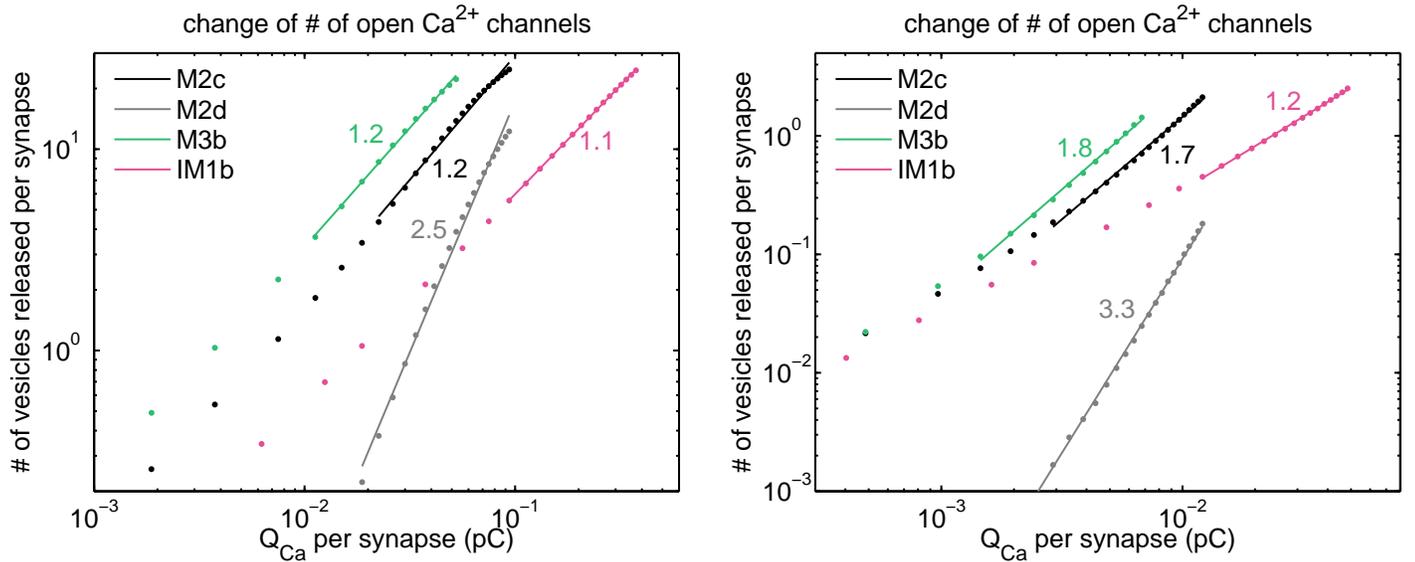
Figure S14: Release vs. Q_{Ca} for scenarios M2c-d, M3b, and IM1b based on $N_{Ca} \times P_{open}$ manipulation. Left – 20 ms after the stimulus onset, right – 3 ms after the stimulus onset. $k_{on}/2$ was assumed.

In Fig. S13, the release vs. Q_{Ca} relationships and exponent m estimates based on 3 ms observation windows for scenarios M1-3 and IM1 are shown and compared with the corresponding estimates based on 20 ms observation windows (values in round brackets) which were discussed in Results section. As can be expected from the previous discussion, m values stayed ~ 4 for the i_{Ca} manipulation for all considered scenarios, though the relationships became linear over a broader range of Q_{Ca} values in the double log plots. In the case of $N_{Ca} \times P_{open}$ manipulation, m values stayed similar for scenarios M3 and IM1. In the case of scenario M3, that happened because this scenario corresponds to almost a pure Ca^{2+} nanodomain controlled vesicle release, when each release site has only one channel contributing to vesicle release at that site. Thus, blocking the channels results in a proportional decrease in the release. In the case of scenario IM1, the m value did not change considerably, because the fusion rate was relatively moderate (due to relatively small $[Ca^{2+}]$ levels) and, thus, the saturation effect was negligible even for the 20 ms observation window. The m value increased from 1.8 to 2.3 for scenario M1 and from 1.2 to 1.7 for scenario M2. These examples show that, despite the fact that m values based on $N_{Ca} \times P_{open}$ change could be underestimated due to the saturation effect, their differences remain considerable among scenarios of different Ca^{2+} influx to exocytosis coupling.

Fig. S14 presents m exponent estimates based on $N_{Ca} \times P_{open}$ change with 20 ms (left) and 3 ms (right) depolarizations for scenarios M2c-d, M3b, and IM1b (Fig. S9). In the case of scenario IM1b, vesicle release at each release site is dominated by a single channel and, thus, m is rather invariant for different time windows as in the case of scenario M3. For scenarios M2c and M3b, the m value increases from 1.2 to 1.7-1.8, similar to what is observed for scenario M2 (Fig. S12). In contrast, in the case of scenario M2d, m increases from 2.5 to 3.3. The main reason why the shift of the sensor by only 20 nm has a substantial effect on the Ca^{2+} to exocytosis coupling (M2c vs. M2d) is that a contribution to $[Ca^{2+}]$ by a particular channel decreases with an increased distance faster at smaller than at larger

distances. Thus, when the sensor is moved away from a "private channel" (Fig. S9), the contribution of this channel decreases more than the contributions of the other channels.